Charmless hadronic three-body decays of neutral B mesons with a KS in the final state in the LHCb experiment: branching fractions and an amplitude analysis

Louis Henry

Charmless hadronic three-body decays of neutral B mesons with a K_s^0 in the final state in the LHCb experiment: branching fractions and an amplitude analysis

Désintégrations hadroniques à trois corps sans charme de mésons beaux avec un K_s^0 dans l'état final dans l'expérience LHCb: mesure de rapports d'embranchement et une analyse en amplitude

par

Louis Henry

Jury composé de:

Sébastien Descotes-Genon　Rapporteur
Directeur de recherche (LPT)

Maurizio Pierini　Rapporteur
Researcher (CERN)

Tim Gershon　Membre du jury
Professor (University of Warwick)

Sandrine Laplace　Membre du jury
Directeur de recherche (LPNHE)

François Le Diberder　Membre du jury
Professeur (Université Paris-Diderot)

Guy Wilkinson　Membre du jury
Professor (Oxford University)

Jean-Bernard Zuber　Membre du jury
Professeur (Université Pierre et Marie Curie)

Eli Ben Haim　Directeur de thèse
Maître de conférences (Université Pierre et Marie Curie)

"All changed, changed utterly: a terrible beauty is born."

Yeats, *Easter 1916*

"Ô Mort, vieux capitaine, il est temps! levons l'ancre!
Ce pays nous ennuie, ô Mort! Appareillons!
Si le ciel et la mer sont noirs comme de l'encre,
Nos coeurs que tu connais sont remplis de rayons!

Verse-nous ton poison pour qu'il nous réconforte!
Nous voulons, tant ce feu nous brûle le cerveau,
Plonger au fond du gouffre, Enfer ou Ciel, qu'importe?
Au fond de l'Inconnu pour trouver du *nouveau*!"

Charles Baudelaire, *Le Voyage, Les Fleurs du Mal*

"L'inspiration, c'est une invention des gens qui n'ont jamais rien créé. Nous entretenons
la légende pour nous faire valoir, mais entre nous, c'est un bluff. Le poète ne connaît
que la commande."

Jean Anouilh

Remerciements

Les dernières corrections se mettent en place, la soutenance paraît déjà loin, et j'appose ces derniers mots à ce manuscrit. Curieusement c'est ceux-là, auxquels je pense depuis le début, qui me prennent le plus de temps. C'est qu'il s'agit de n'oublier personne.

Tout d'abord, je souhaite remercier une fois de plus mon directeur de thèse, Eli Ben Haim, qui m'aura accompagné de près tout au long de ces trois années. Grâce à sa capacité à canaliser des pensées pas toujours très ordonnées, j'ai pu profondément évoluer en tant que personne et en tant que chercheur. J'ajoute par ailleurs une mention spéciale à sa patience sans limites pour corriger au mot près toutes mes productions écrites, corrections qui donnaient l'impression qu'il aurait été plus rapide de repeindre ma page entière en rouge. J'espère sincèrement pouvoir continuer à profiter de sa justesse d'analyse et de ses qualités humaines tout au long de ma carrière.

Je tiens à exprimer toute ma gratitude l'ensemble du groupe LHCb du LPNHE pour leur gentillesse et leur disponibilité. Merci aussi à mon parrain, José Ocariz, qui m'a accompagné et supporté mes râleries (surtout en troisième année), toujours avec le sourire. Je remercie l'encadrement du laboratoire, et notamment sa grande ouverture aux étudiants, qu'ils viennent de la licence ou de l'extérieur. Ce n'est pas un hasard que mon premier contact avec le monde de la physique des particules après mon école se soit fait dans ces locaux. Enfin, merci à tous mes collègues doctorants: nos discussions interminables sur tous les sujets au cours de pauses café à rallonge me manquent déjà. Petite pensée spéciale pour Mathilde, qui aura marqué profondément tout le début de ma thèse, et pour l'ensemble de mes prédecesseurs. Je leur souhaite bonne chance pour le futur, dans la recherche et en-dehors.

Je remercie l'ensemble des membres de mon jury, et notamment mes deux rapporteurs, Sébastien Descotes-Genon et Maurizio Pierini, qui ont pu relire et commenter mon manuscrit en moins d'un mois. Le présent document doit beaucoup à l'ensemble des remarques du jury et à leurs questions incisives. Je remercie Jean-Bernard Zuber d'avoir accepté de présider ce jury, et de m'avoir conseillé dans mes choix de carrière en M1. Merci aussi à l'ensemble du groupe KShh, et plus particulièrement à Tom, Stéphane et Rafael. Je remercie sincèrement Diego Milanes, pour son aide, sa constante sympathie, et nos nombreuses soirées et excursions. Ses conseils, prodigués autour d'un café, d'une bière, ou dans un stade de base-ball, m'ont été extrêmement précieux.

Je pense bien entendu à ma famille, qui m'a forgé et a tenté depuis 27 ans de canaliser la boule d'énergie dont ils avaient hérité. J'essaie toujours de me montrer à la hauteur de ce que vous m'avez donné. Cette thèse, vous le savez, c'est beaucoup de vous. J'adresse

aussi une pensée à mes amis, anciens ou nouveaux, et plus particulièrement à Franck et Thibaut, qui m'accompagnent depuis tant d'années. On peut difficilement trouver chemins plus différents, mais votre présence m'est toujours aussi chère dans ces moments importants.

Enfin, toute ma gratitude et ma tendresse à Camille, qui aura eu la malchance de croiser un thésard au début de son voyage et la grâce de le supporter. Son soutien et sa patience m'ont portés tout au long de ces trois années qui changèrent ma vie.

Résumé

Ce manuscrit présente plusieurs études des désintégrations de mésons B^0 et B_s^0 en trois corps non-charmés, dont un méson K_S^0. Ces études portent sur les données enregistrées par l'expérience LHCb pendant le Run I du LHC, correspondant à une luminosité intégrée de $\int \mathcal{L} = 3$ fb^{-1}.

Une première analyse consiste en une mesure des rapports d'embranchement des modes $B_{d,s}^0 \to K_S^0 h^+ h'^-$, où $h^{(')}$ désigne un kaon ou un pion. Les précédentes mesures par LHCb des rapports d'embranchements de ces modes de désintégration, rapportés à celui du mode $B^0 \to K_S^0 \pi^+ \pi^-$, sont mis à jour. De plus, le but principal de cette analyse est de rechercher le mode $B_s^0 \to K_S^0 K^+ K^-$, pas encore observé par les analyses précédentes. Les rapports d'embranchement relatifs sont mesurés :

$$\frac{\mathcal{B}\left(B_s^0 \to K_S^0 \pi^+ \pi^-\right)}{\mathcal{B}\left(B^0 \to K_S^0 \pi^+ \pi^-\right)} = 0.26 \pm 0.04(\text{stat.}) \pm 0.02(\text{syst.}) \pm 0.01(\text{f}_\text{s}/\text{f}_\text{d}),$$

$$\frac{\mathcal{B}\left(B^0 \to K_S^0 K^\pm \pi^\mp\right)}{\mathcal{B}\left(B^0 \to K_S^0 \pi^+ \pi^-\right)} = 0.17 \pm 0.02(\text{stat.}) \pm 0.00(\text{syst.}),$$

$$\frac{\mathcal{B}\left(B_s^0 \to K_S^0 K^\pm \pi^\mp\right)}{\mathcal{B}\left(B^0 \to K_S^0 \pi^+ \pi^-\right)} = 1.84 \pm 0.07(\text{stat.}) \pm 0.02(\text{syst.}) \pm 0.04(\text{f}_\text{s}/\text{f}_\text{d}),$$

$$\frac{\mathcal{B}\left(B^0 \to K_S^0 K^+ K^-\right)}{\mathcal{B}\left(B^0 \to K_S^0 \pi^+ \pi^-\right)} = 0.59 \pm 0.02(\text{stat.}) \pm 0.01(\text{syst.}),$$

(1)

Une première observation de $B_s^0 \to K_S^0 K^+ K^-$ est rapportée, avec une significance globale de $3.7\,\sigma$.

Une analyse non-étiquetée de saveur et indépendante du temps du plan de Dalitz de la désintégration $B^0 \to K_S^0 K^+ K^-$ est présentée, en utilisant l'approche isobare. Les rapports d'embranchement quasi-deux-corps des désintégrations $B^0 \to K_S^0 \phi^0$, $B^0 \to K_S^0 f_2'(1525)$, $B^0 \to K_S^0 f_0(1710)$, et $B^0 \to K_S^0 \chi_{c0}$ sont mesurés. Ils sont compatibles avec les mesures précédentes de BaBar, à l'exception de $B^0 \to K_S^0 f_0(1710)$.

Abstract

This dissertation presents several studies of the decays of both B^0 and B_s^0 mesons to charmless three-body final states including a K_S^0 meson. They use the data recorded by the LHCb experiment during Run I of LHC, corresponding to an integrated luminosity of $\int \mathcal{L} = 3 \text{ fb}^{-1}$.

A first analysis consists of the measurement of the branching fractions of $B_{d,s}^0 \to K_S^0 h^+ h'^-$ decays, where $h^{(')}$ designates a kaon or a pion. Preceding LHCb measurements of branching fractions for all decay channels, relative to that of $B^0 \to K_S^0 \pi^+ \pi^-$, are updated. Furthermore, the primary goal of this analysis is to search for the, as yet, unobserved decay $B_s^0 \to K_S^0 K^+ K^-$. The relative branching fractions are measured to be:

$$\frac{\mathcal{B}\left(B_s^0 \to K_S^0 \pi^+ \pi^-\right)}{\mathcal{B}\left(B^0 \to K_S^0 \pi^+ \pi^-\right)} = 0.26 \pm 0.04(\text{stat.}) \pm 0.02(\text{syst.}) \pm 0.01(\text{f}_\text{s}/\text{f}_\text{d}),$$

$$\frac{\mathcal{B}\left(B^0 \to K_S^0 K^\pm \pi^\mp\right)}{\mathcal{B}\left(B^0 \to K_S^0 \pi^+ \pi^-\right)} = 0.17 \pm 0.02(\text{stat.}) \pm 0.00(\text{syst.}),$$

$$\frac{\mathcal{B}\left(B_s^0 \to K_S^0 K^\pm \pi^\mp\right)}{\mathcal{B}\left(B^0 \to K_S^0 \pi^+ \pi^-\right)} = 1.84 \pm 0.07(\text{stat.}) \pm 0.02(\text{syst.}) \pm 0.04(\text{f}_\text{s}/\text{f}_\text{d}),$$

$$\frac{\mathcal{B}\left(B^0 \to K_S^0 K^+ K^-\right)}{\mathcal{B}\left(B^0 \to K_S^0 \pi^+ \pi^-\right)} = 0.59 \pm 0.02(\text{stat.}) \pm 0.01(\text{syst.}),$$

$$\tag{2}$$

A first observation of $B_s^0 \to K_S^0 K^+ K^-$ is reported with a global significance of $3.7\,\sigma$.

A flavour-untagged, time-independent Dalitz-plot analysis of $B^0 \to K_S^0 K^+ K^-$ is presented, using the isobar approach. The quasi-two-body branching fractions of $B^0 \to K_S^0 \phi^0$, $B^0 \to K_S^0 f_2'(1525)$, $B^0 \to K_S^0 f_0(1710)$, and $B^0 \to K_S^0 \chi_{c0}$ are measured. They are compatible with previous measurements from BaBar, except for $B^0 \to K_S^0 f_0(1710)$.

Contents

Introduction

The study of of b-hadron decays to hadronic final states with no charmed particles allow for a rich array of studies. A few examples are the measurements of branching fractions, CP asymmetries, weak and strong phases; they probe the dynamics of weak and strong interactions. The typical branching fractions of these modes are below 10^{-5} and thus their analyses are feasible only with large data samples and the use of powerful tools to reject background. The LHCb experiment at the CERN Large Hadron Collider (LHC) is an adequate experimental environment for these analyses, offering the possibility to study decays of light B mesons, B_s mesons and b baryons.

This dissertation describes two analyses of $B^0_{d,s} \rightarrow K^0_S h^+ h^{'-}$ decays, where $h^{(')}$ represents a kaon or a pion, that were performed with the $3\,\mathrm{fb}^{-1}$ dataset collected by the LHCb experiment during the years 2011 and 2012, at centre-of-mass energies of 7 and 8 TeV, respectively. The decays under study are dominated by loop transitions, that may have contributions from particles beyond the standard model. The measured observables are therefore probes for new physics. A first analysis consists in the measurement of the six branching fractions of these modes, relative to that of $B^0 \rightarrow K^0_S \pi^+ \pi^-$. This includes a search for the mode $B^0_s \rightarrow K^0_S K^+ K^-$, that has never been observed before. A second study is the first amplitude analysis (or Dalitz-plot analysis) of the mode $B^0 \rightarrow K^0_S K^+ K^-$ from LHCb. It contains a measurement the branching fractions of intermediate states that intervene in the decay, using the isobar approximation. This is the first such study of this mode in LHCb; it will be pursued in steps of increasing complexity with the growing dataset, and will become more and more sensitive to new physics observables.

This dissertation is organized as follows. Section 1 shortly reviews the theoretical framework, as well as concepts related to the amplitude analysis. It also gives a short overview of existing results. Section 2 then describes the LHCb experiment and the related concepts that are useful for the understanding of the analysis work. The presentation of my work is then separated into three parts. Firstly, Sec. 3 presents an alternative procedure to simulate background events. This procedure is used in the measurement of $B^0_{d,s} \rightarrow K^0_S h^+ h^{'-}$ branching fractions. Secondly, Sec. 4 describes the measurement of the branching fractions of $B^0_{d,s} \rightarrow K^0_S h^+ h^{'-}$ modes, along with the search for the missing $B^0_s \rightarrow K^0_S K^+ K^-$ mode. Finally, Sec. 5 presents the untagged, time-independent Dalitz-plot analysis of the $B^0 \rightarrow K^0_S K^+ K^-$ decay.

Chapter 1

Theory

1.1 Introduction

The Standard Model (SM) of particle physics describes the interaction of fundamental particles through the strong and electroweak interactions [1–3]. It is an outstandingly successful theory that predicts nearly all the measurements ever performed with great precision. There are however some hints that point at a larger theory, the SM being an effective model of that theory at lower energies:

- the SM does not explain the number of fermion generations nor their highly hierarchical structure in terms of mass. Instead, masses of particles form the bulk of free parameters of the SM (13 out of 18);

- the SM does not include gravity. In fact, general relativity is even mathematically incompatible with quantum field theory (QFT). The SM has then to be an effective theory that cannot be valid at the Planck energy scale;

- the SM does not provide a candidate for cold dark matter, whose contribution to the mass content of the Universe is found to be about fives times larger than that of ordinary matter [4];

- there is no mechanism in the SM that explains the smallness of the mass of the Higgs boson. Indeed, quantum contributions to the Higgs boson mass from Grand Unification or Planck-scale particles would make the mass huge, unless there is a fine-tuning cancellation between the radiative corrections and the bare mass [5]. This problem may be solved by the presence of physics beyond the SM at low mass scale (1 TeV), which would provide a more natural cancellation;

- the SM fails to account for the matter-antimatter asymmetry observed in the Universe.

These issues motivate the search for new physics (NP), and also provide some hints that it should be accessible at energies close to the TeV scale.

Searches for new physics can be classified in two categories: direct and indirect searches. Direct searches look for the production of on-shell particles beyond the SM, such as supersymmetric particles (squarks, gluinos) [4]. Indirect searches focus on deviations of measurements of observables from a theoretical SM prediction due to the effect of off-shell NP particles. These searches require both a clean theoretical prediction and a clean experimental measurement so that possible deviations can be attributed to the effects of NP; they are better performed on decays where a contribution from NP is expected. In general, direct searches need an accurate description of the background, whereas features of the background can be usually inferred from data in indirect searches.

The violation of the CP symmetry, described in Sec. 1.2, is a feature of the Standard Model which is strongly related to the matter-antimatter asymmetry in the Universe.[1] It depends on few parameters of the Standard Model, thus its predictive power is rather high. The study of the violation of this symmetry in $B_{d,s}^0 \to K_S^0 h^+ h'^-$ decays provides opportunities to perform indirect searches for NP. Indeed, decays of the type $B_{d,s}^0 \to K_S^0 h^+ h'^-$, where $h^{(')}$ are kaons or pions, are dominated by so-called penguin diagrams that include a loop of virtual particles. Particles of NP could contribute inside of that loop and cause a deviation of some observables from the SM prediction. Additionally, these decays also provide a relatively clean experimental context in the LHCb experiment, where sample purities larger than 90% can be achieved.

Section 1.2 details the Standard Model description of the CP violation, and Sec. 1.3 presents some general concepts of amplitude analysis. Finally, Section 1.4 presents an overview of the motivations and experimental context of the study of $B_{d,s}^0 \to K_S^0 h^+ h'^-$ decays.

1.2 Violation of the CP symmetry

The violation of the CP symmetry, described in Sec. 1.2.2 is a key factor to understand the matter-antimatter asymmetry of the Universe. Indeed, the required conditions so that a model could allow for a matter-antimatter asymmetry, denoted Sakharov conditions [7], are

- the existence of an interaction that does not conserve the baryon number;

- the existence of an interaction that violates both the C and CP symmetries;

- non-thermal equilibrium.

The baryon number is not conserved in some non-perturbative electroweak processes, for instance the processes called sphalerons [8]. The existence of such processes relies however on the existence of a CP violation at the perturbative scale.

---

[1]As described in the following, CP violation is a key ingredient to explain this asymmetry, but this CP violation is too small by 9 orders of magnitude to explain the matter-antimatter asymmetry of the Universe [6].

1.2.1 Introduction to symmetries

Symmetries play a fundamental role in modern physics, as they constitute the building blocks of any Lagrangian theory. They can be continuous or discrete. Continuous symmetries are families of symmetries that depend on a continuous parameter. For instance, $U(1)$ is a group of the continuous, global symmetries that describe rotations in a plane. It can be defined as

$$\{t_\alpha \in U(1); y \to y \times e^{i\alpha}\}, \tag{1.1}$$

where y is a complex number and α is a real number.

Continuous, global symmetries can be extended into gauge symmetries, where the parameter is itself a function of the position in space and time. For instance, the gauged version of the global $U(1)$ symmetry would be

$$\{t_\alpha \in U(1); y(\mathbf{x}) \to y(\mathbf{x}) \times e^{i\alpha(\mathbf{x})}\}, \tag{1.2}$$

where $\mathbf{x}$ is a position in space-time, $y(\mathbf{x})$ is a complex operator, and $\alpha(\mathbf{x})$ is a real function. The Standard Model is a gauge theory of the $SU(3)_\mathrm{C} \otimes SU(2)_\mathrm{L} \otimes U(1)_\mathrm{Y}$ group. This underlying structure constrains the particle content of the theory and the interactions between these particles.

The strong interaction is described by the underlying $SU(3)_\mathrm{C}$ symmetry, where the C stands for "colour" charge of the interaction. Properties of that symmetry group naturally yield the gluon self-interaction, which is the underlying cause for the confinement of quarks into colourless hadrons.

The electromagnetic and weak interactions are described by the underlying $SU(2)_\mathrm{L} \otimes U(1)_\mathrm{Y}$ symmetry, where the L stands for "left-handed" and the Y for the hypercharge. The left-handed aspect of the $SU(2)_\mathrm{L}$ symmetry is what explains the nonexistence of right-handed neutrinos, and thus the violation of parity (see Sec. 1.2.2) by the weak interactions. The $SU(2)_\mathrm{L} \otimes U(1)_\mathrm{Y}$ symmetry is spontaneously broken at the current Universe energy density, leaving only the residual $U(1)_Q$ symmetry that is responsible for electromagnetic interaction and whose mediator is the massless photon γ. The mechanism of that symmetry breaking, where the vacuum expectation value of one of the scalar fields of the theory is nonzero, is known as the Higgs mechanism. This mechanism gives rise to the masses of fermions and of the gauge bosons of the weak interaction, $W^\pm$ and Z^0, and has been confirmed by the discovery of the Higgs boson by the ATLAS and CMS experiments in 2011 [9, 10]. The weak interaction is the only one known to couple different flavours. In the quark sector, it couples up-type quarks (u,c,t) and down-type quarks (d,s,b).

Discrete symmetries do not depend on a continuous parameter, and cannot be gauged. They are however interesting in the building of a model as they correspond to conserved

quantum numbers, through the Noether theorem (also valid for continuous symmetries). The conservation of these quantum numbers in a process governed by an interaction that is invariant under the corresponding symmetry allows to build selection rules. The following section describes three of these discrete symmetries, C, P, T, as well as the CP and CPT products.

1.2.2 The C, P, and T symmetries

The charge-conjugation operator C

The charge-conjugation operator C transforms a particle to the corresponding antiparticle. This antiparticle shares all the properties of the original particle, except for reversed electric, flavour, and colour charges. The Lagrangians of the electromagnetic and strong interaction are invariant under C, unlike the Lagrangian of the weak interaction.

The parity operator P

The parity operator is defined as the reversal of all the spatial coordinates of a 4-vector, while the time component is conserved. It conserves all the charges of the particle and its spin. The angular momentum $\mathbf{L}$ is conserved, which means that the sign of the helicity of the particle, defined as

$$H = \frac{\mathbf{L}.\mathbf{p}}{|\mathbf{p}|}, \tag{1.3}$$

is reversed. Hence P transforms left-handed ($H = -1$) particles into right-handed ($H = 1$) particles, and inversely. The helicity is strongly related to the chirality of the particle, which defines its transformation under P.[2] In contrary to the helicity, however, chirality does not depend on the reference frame in the case of massive particles.

Following the observation that parity is conserved by the electromagnetic and strong interactions, weak interaction was initially thought to conserve that symmetry. However, Lee and Yang [11] raised concern that the weak interaction could be sensitive to the chirality of particles ("chiral interaction"). This was confirmed by the observation that β decays only emit left-handed neutrinos [12]. More generally, only left-handed particles (and right-handed antiparticles) interact via the weak interaction.

The T operator

The time-reversal operator T is complementary to the parity operator P, as it transforms $(t, \mathbf{x})$ into $(-t, \mathbf{x})$. It is conserved by the electromagnetic and the strong interactions. The first direct observation of the violation of the T symmetry by the weak interaction has been performed in the study of the B^0 system [13].

---

[2] The helicity and the chirality are equal for massless particles.

The *CP* and *CPT* operators

The previous results on *C* and *P* operators could mean that the product of the *C* and *P* operator, denoted *CP*, is conserved by weak interactions as this operator transforms left-handed neutrinos into right-handed antineutrinos [14]. The first demonstration of *CP* violation in nature has been obtained through the study of the mixing of neutral mesons such as the K^0 [15] and the B^0 [16]

The *CPT* theorem states that the Lagrangian of the SM must be invariant under the *CPT* product. This is related to Lorentz invariance and locality. Searches for *CPT* violation have for now not found any significant violation.

Under the assumption of the *CPT* theorem, any observation of violation of the *T* or the *CP* symmetry results in the violation of *CP* or *T*, respectively. This has led to the first observation of time-reversal symmetry violation in the neutral kaon system, under the assumption of *CPT* [17]. Additionally, measurements of the *T*-violation have constrained the violation of *CP* by the strong interaction to smaller than 10^{-10} [18]. This constitutes the strong *CP* problem, as the strong interaction could in principle violate *CP*. We consider in the rest of this dissertation that the strong interaction if *CP*-conserving.

Section 1.2.3 describes the mechanism of *CP* violation in the mixing of neutral mesons, along with the different types of *CP* violation in the Standard Model.

1.2.3 Neutral mesons mixing and *CP* violation

We consider a neutral meson $|P^0\rangle$ such that $|P^0\rangle \neq |\overline{P^0}\rangle$, decaying to a final state f. There are three different bases that can be used to describe the $|P^0\rangle$-$|\overline{P^0}\rangle$ system:

- $|P^0\rangle$ and $|\overline{P^0}\rangle$ (flavour eigenstates);

- $\frac{1}{\sqrt{2}}(|P^0\rangle + |\overline{P^0}\rangle)$ and $\frac{1}{\sqrt{2}}(|P^0\rangle - |\overline{P^0}\rangle)$ (*CP*-eigenstates);

- $|P_\mathrm{L}\rangle$ and $|P_\mathrm{H}\rangle$ (eigenstates of the Hamiltonian).

In the two eigenstates of the Hamiltonian, L and H stand for "light" and "heavy", respectively. The weak Hamiltonian conserves *CP* if and only if the eigenstates of the Hamiltonian are also eigenstates of *CP*.

The effective Hamiltonian H, describing the evolution of an initial state containing a mixture of $|P^0\rangle$ and $|\overline{P^0}\rangle$ (and ignoring final states), can be written as

$$H = M - \frac{i}{2}\Gamma, \tag{1.4}$$

where M and Γ are hermitian matrices defined as

$$M = \begin{pmatrix} m_{11} & m_{12} \\ m_{12}^* & m_{22} \end{pmatrix}, \Gamma = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma_{22} \end{pmatrix}. \tag{1.5}$$

The *CPT* invariance requires that the diagonal terms of these matrices are equal. The introduction of the matrix Γ in the Hamiltonian removes its property of hermiticity, which is linked to the conservation of probability. This allows to introduce the lifetime of the states described by this Hamiltonian, as the square of the wave-function that describes them is decreasing exponentially with time.

The Schrödinger equation that governs the time-evolution of a wave-function is

$$i\frac{\mathrm{d}|\Psi(t)\rangle}{\mathrm{d}t} = H|\Psi(t)\rangle. \tag{1.6}$$

The integration of this equation applied to the $|P_{\mathrm{L,H}}\rangle$ states yields

$$|P_{\mathrm{L,H}}(t)\rangle = |P_{\mathrm{L,H}}\rangle e^{-i\left(M_{\mathrm{L,H}} - \frac{i}{2}\Gamma_{\mathrm{L,H}}\right)t} \tag{1.7}$$

where $\left(M_{\mathrm{L,H}} - \frac{i}{2}\Gamma_{\mathrm{L,H}}\right)$ are the corresponding eigenvalues of the Hamiltonian. The terms p and q are defined as the (nonvanishing) coefficients that allow to change the basis

$$\begin{aligned}
|P_{\mathrm{L}}\rangle &= p|P^0\rangle + q|\overline{P^0}\rangle, \\
|P_{\mathrm{H}}\rangle &= p|P^0\rangle - q|\overline{P^0}\rangle,
\end{aligned} \tag{1.8}$$

where $|p|^2 + |q|^2 = 1$. Conversely, these coefficients can be used to write

$$\begin{aligned}
|P^0\rangle &= \frac{1}{2p}(|P_{\mathrm{L}}\rangle + |P_{\mathrm{H}}\rangle), \\
|\overline{P^0}\rangle &= \frac{1}{2q}(|P_{\mathrm{L}}\rangle - |P_{\mathrm{H}}\rangle).
\end{aligned} \tag{1.9}$$

We remark that if $p = q = \frac{1}{\sqrt{2}}$, $|P_{\mathrm{L}}\rangle$ and $|P_{\mathrm{H}}\rangle$ are exactly equal to $\frac{1}{\sqrt{2}}(|P^0\rangle + |\overline{P^0}\rangle)$ and $\frac{1}{\sqrt{2}}(|P^0\rangle - |\overline{P^0}\rangle)$, and *CP* is conserved.

Finally, combining Eq. 1.9 and 1.7, the time-evolution of $|P^0\rangle$ and $|\overline{P^0}\rangle$ states writes

$$\begin{aligned}
|P^0\rangle(t) &= f_+(t)|P^0\rangle + \frac{q}{p}f_-(t)|\overline{P^0}\rangle, \\
|\overline{P^0}\rangle(t) &= f_+(t)|\overline{P^0}\rangle + \frac{p}{q}f_-(t)|P^0\rangle,
\end{aligned} \tag{1.10}$$

Table 1.1 – Experimental average for Δm and $\Delta \Gamma$ in different neutral-meson systems from [19].

B^0 mixing parameters	
$\Delta m_d (\,\mathrm{ps}^{-1})$	0.510 ± 0.003
$\Delta \Gamma_d / \Gamma_d$	0.001 ± 0.010
$|q/p|$	1.0009 ± 0.0013
B_s^0 mixing parameters	
$\Delta m_s (\,\mathrm{ps}^{-1})$	$17.757 \pm 0.020 \pm 0.007$
$\Delta \Gamma_s / \Gamma_s$	0.124 ± 0.009
$|q/p|$	1.0038 ± 0.0021

where

$$f_{\pm}(t) = \frac{1}{2} \left(e^{-i\left(M_{\mathrm{L}} - \frac{i}{2}\Gamma_{\mathrm{L}}\right)t} \pm e^{-i\left(M_{\mathrm{H}} - \frac{i}{2}\Gamma_{\mathrm{H}}\right)t} \right). \tag{1.11}$$

We define the quantities

$$\Delta m = m_{\mathrm{H}} - m_{\mathrm{L}}, \Delta \Gamma = \Gamma_{\mathrm{L}} - \Gamma_{\mathrm{H}}, \tag{1.12}$$

and obtain

$$f_{\pm}(t) = \frac{1}{2} \left(e^{-im_{\mathrm{L}}t} e^{-\frac{1}{2}\Gamma_{\mathrm{L}}t} \left[1 \pm e^{-i\Delta m t} e^{-\frac{1}{2}\Delta \Gamma t} \right] \right). \tag{1.13}$$

This function governs the mixing in the $|P^0\rangle$–$|\overline{P^0}\rangle$ system.

The Δm and $\Delta \Gamma$ parameters can be predicted from SM calculations, and experimentally measured. Table 1.1 summarizes the current world averages for the B^0 and B_s^0 meson systems [19].

We consider the decay of the $|P^0\rangle$ meson to a final state f, associated with the amplitude $\mathcal{A}_f$.[3] The parameter

$$\lambda_f = \frac{q}{p} \frac{\overline{\mathcal{A}_{\overline{f}}}}{\mathcal{A}_f} \tag{1.14}$$

contains the information about CP violation in that decay. Indeed, if the modulus of λ_f is not 1, or if its imaginary part is not vanishing, CP violation in the $|P^0\rangle \to f$ decay occurs. Defining the three observables

---

[3]In the following, the conjugate decay of $|\overline{P^0}\rangle$ to $\overline{f}$ is associated with the amplitude $\overline{\mathcal{A}_{\overline{f}}}$.

$$C_f = \frac{1 - \lambda_f^2}{1 + \lambda_f^2}, \tag{1.15}$$

$$S_f = \frac{2\Im(\lambda_f)}{1 + \lambda_f^2}, \tag{1.16}$$

$$\mathcal{A}_f^{\Delta\Gamma} = -\frac{2\Re(\lambda_f)}{1 + \lambda_f^2}, \tag{1.17}$$

$$\tag{1.18}$$

the decay rate of $|P^0\rangle$ as a function of time writes

$$\Gamma(t) \propto \frac{e^{-\Gamma_{|P^0\rangle}t}}{2\tau} \left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) + \mathcal{A}_f^{\Delta\Gamma} \sinh\left(\frac{\Delta\Gamma t}{2}\right) + \left(C_f \cos\left(\Delta m t\right) - S_f \sin\left(\Delta m t\right)\right) \right], \tag{1.19}$$

where $\tau = (\frac{\Gamma_L + \Gamma_H}{2})^{-1}$, $\Gamma_{|P^0\rangle} = \frac{\Gamma_L + \Gamma_H}{2}$, and $\Delta\Gamma = \Gamma_L - \Gamma_H$. It is necessary to perform a time-dependent analysis of a decay in order to measure all the CP-violation effects with precision, as well as to determine the flavour of the neutral meson that decays ("tagging").

In the case where several channels contribute to the total amplitude, the amplitudes $\mathcal{A}$ and $\overline{\mathcal{A}}$ of the total decay can be written

$$\mathcal{A} = \sum_i \mathcal{A}_i e^{i(\phi_i - \delta_i)}, \overline{\mathcal{A}} = \sum_i \mathcal{A}_i e^{i(\phi_i + \delta_i)}, \tag{1.20}$$

where the sum runs over the channels contributing to the amplitude and $\mathcal{A}_i$ is the magnitude of the contribution of each channel. The phases ϕ_i and δ_i are the CP-conserving and CP-violating components of the phase corresponding the each channel, respectively. The effect of the CP symmetry can only induce a difference in phase, not magnitude, in each channel taken separately. However, in the presence of two or more contributing channels, the difference in the pattern of interference induced by the CP-violating phase can result in CP violation in the decay.

Three types of CP violation sources can be distinguished, with different physical interpretations.

CP violation in decays

In presence of several contributions to the amplitude that both have a relative CP-conserving phase and different CP-violating phases, the rate of a decay and its conjugate may be different. Indeed, in the case of two contributing diagrams,

$$\mathcal{A} = A_1 e^{i(\phi + \delta_1)} + A_2 e^{i(-\phi + \delta_2)}, \overline{\mathcal{A}} = A_1 e^{i(\phi - \delta_1)} + A_2 e^{i(-\phi - \delta_2)}, \tag{1.21}$$

where the relative CP-conserving phase between diagrams 1 and 2 is 2ϕ, and $\delta_{1,2}$ is the CP-violating phase between these diagrams. If both the CP-conserving and the CP-violating phases are not 0, the decay rates related to $\mathcal{A}$ and $\overline{\mathcal{A}}$ are different.

This is the only possible type of CP violation in decays of charged mesons or baryons.

CP violation through mixing

As underlined before, CP violation can be induced by the mixing of neutral mesons. Considering for instance Eq. 1.8, CP is violated in the mixing of neutral mesons if and only if $|\frac{p}{q}| \neq 1$. As shown in Tab. 1.1, this ratio is consistent with 1 in the case of the B^0 and B_s^0 mesons.

CP violation in interference between mixing and decay

Another type of CP violation is associated to the interference between mixing and decay processes of neutral mesons to the same CP-eigen state. Contrary to the CP violation in decay, it does not require several channels to contribute to the amplitude, as the interference happens between the mixed and unmixed amplitudes. This type of CP violation occurs in case that the imaginary part of λ takes a nonzero value. The parameter that outlines this measurement is contained in the term S_f.

Direct and indirect CP violation

CP violation can be alternatively classified into direct or indirect CP violation. Direct CP violation corresponds to CP violation through decay, whereas indirect CP violation refers to CP violation through mixing or through the interference between mixing and decay. As shown in Tab. 1.1, the CP violation in the mixing of the B^0 meson can be neglected in most cases, and thus "indirect CP violation" often refers to interference between mixing and decay when considering decays of the B^0 meson.

1.2.4 The CKM matrix and the KM mechanism

As described in Sec. 1.1, flavour eigenstates are eigenstates of the electroweak interaction. They are however not necessarily eigenstates of the strong interaction, or of the Hamiltonian. This section describes how the change of basis between eigenstates of the electroweak interaction and of the Hamiltonian introduces an irreducible phase in the SM, and thus to CP violation, when three or more quark generations exist.

We consider the change of basis between the quark eigenstates of flavour and of the Hamiltonian by the matrices U_{L}^f and U_{R}^f, defined such as

$$M_{\mathrm{mass}} = \left(U_{\mathrm{L}}^f\right)^\dagger M_{\mathrm{flavour}} U_{\mathrm{R}}^f, \tag{1.22}$$

where M_{mass} and M_{flavour} are the matrices that describe quark currents in the mass basis and the flavour basis, respectively. The idea of different bases to describe the mass and the weak eigenstates was first proposed by Cabibbo [20]. The motivation was to explain the suppression of the decay of strange particles, and thus the long lifetime of these particles. The GIM mechanism is an extension of this concept that requires the existence of a second-generation up-type quark, the c quark [21]. It allows to forbid any flavour-changing neutral current at tree-level in the Standard Model.

A 2×2 unitary matrix V can be described by a single real parameter. Starting from the original 2×4 real parameters (*e.g.* magnitudes and phases), unitarity relations state that

$$\forall(i,j), \sum_k V_{ik}V_{jk}^* = \delta_{ij}, \forall(i,j), \sum_k V_{ki}V_{kj}^* = \delta_{ij}, \tag{1.23}$$

which removes 4 parameters. Finally, phases between quark currents are physically meaningless, thus removing $2N - 1 = 3$ parameters, leaving only one real parameter. The comparison with real orthogonal matrices leads to defining this parameter as an angle θ_{C}, and so

$$V = \begin{pmatrix} \cos\theta_{\text{C}} & \sin\theta_{\text{C}} \\ -\sin\theta_{\text{C}} & \cos\theta_{\text{C}} \end{pmatrix}. \tag{1.24}$$

This idea has first been proposed with the two lightest quark generations, this angle θ_{C} being named the Cabibbo angle. Kobayashi and Maskawa have proposed to extend this idea to three quark generations and showed how this resulted in the introduction of a physical phase in the SM, responsible for *CP* violation [22].[4] Indeed, an extension of the discussion above shows that a 3×3 unitary matrix can be described by 4 real parameters, one of which being an irreducible phase. The 3×3 basis-changing matrix in the case of three quark generations is called the Cabibbo-Kobayashi-Maskawa (CKM) matrix. The 2008 Nobel prize of physics was awarded to Kobayashi and Maskawa after precise measurements of *CP* violation showed that it was indeed consistent with their description.

The CKM matrix is written as

$$V_{CKM} = (U_{\text{L}}^u)^\dagger U_{\text{R}}^d = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{1.25}$$

It is important to note that, due to the fact that flavour-changing neutral currents (FCNC) are forbidden at tree-level in the SM, up-type quarks are only paired with down-type quarks, and inversely. Following the discussion on the number of degrees of freedom, this matrix can be parameterized by three real parameters and one imaginary parameter. These three angles are defined as $\theta_{12}(= \theta_{\text{C}})$, θ_{13}, and θ_{23}. For each angle θ_{ij}, its cosine and sine are noted c_{ij} and s_{ij}, respectively, and the CKM matrix may be written as

---

[4]This irreducible phase is equivalent to a violation of *CP*, as this symmetry is anti-linear.

$$V_{\mathrm{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \qquad (1.26)$$

where δ is the irreducible phase. Since the term s_{12} is small, this form of the CKM matrix can be written as an expansion of $\lambda = s_{12} \approx 0.22$, and three parameters that are close to unity: $A = \frac{s_{23}}{\lambda^2}$, $\rho = \frac{s_{13}}{\lambda s_{23}}\cos\delta$, and $\eta = \frac{s_{13}}{\lambda s_{23}}\sin\delta$. This yields the Wolfenstein parameterization [23]

$$V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \qquad (1.27)$$

Finally, the

$$\overline{\rho} = \rho\left(1 - \frac{\lambda^2}{2}\right), \overline{\eta} = \eta\left(1 - \frac{\lambda^2}{2}\right) \qquad (1.28)$$

terms can be defined to yield the Buras parameterization [24] which is is valid at $\mathcal{O}(\lambda^5)$

$$V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 - \lambda^4/8 & \lambda + \mathcal{O}(\lambda^7) & A\lambda^3(\rho - i\eta) \\ -\lambda + A^2\lambda^5\left[1 - 2(\rho + i\eta)\right]/2 & 1 - \lambda^2/2 - \lambda^4\left(1 + 4A^2\right)/8 & A\lambda^2 + \mathcal{O}(\lambda^8) \\ A\lambda^3(1 - \overline{\rho} - i\overline{\eta}) & -A\lambda^2 + A\lambda^4\left[1 - 2(\rho + i\eta)\right]/2 & 1 - A^2\lambda^4/2 \end{pmatrix} \qquad (1.29)$$

1.2.5   The unitarity triangles

The unitarity of the CKM matrix can be formulated as

$$L_i^* L_j = \sum_i V_{ik}^* V_{jk} = \delta_{ij},$$
$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (1.30)$$
$$C_i^* C_j = \sum_i V_{ki}^* V_{kj} = \delta_{ij},$$

where $L_{i(j)}$ and $C_{i(j)}$ are the ith (j^{th}) line and column, respectively. These unitarity constraints yield 9 equations, among which six involve different lines or columns[5]:

---

[5]As the CKM matrix can be written using only four terms, these equations are highly redundant.

$$V_{ud}^* V_{us} + V_{cd}^* V_{cs} + V_{td}^* V_{ts} = 0, \tag{1.31}$$

$$V_{ud}^* V_{ub} + V_{cd}^* V_{cb} + V_{td}^* V_{tb} = 0, \tag{1.32}$$

$$V_{us}^* V_{ub} + V_{cs}^* V_{cb} + V_{ts}^* V_{tb} = 0, \tag{1.33}$$

$$V_{cd}^* V_{ud} + V_{cs}^* V_{us} + V_{cb}^* V_{ub} = 0, \tag{1.34}$$

$$V_{td}^* V_{ud} + V_{ts}^* V_{us} + V_{tb}^* V_{ub} = 0, \tag{1.35}$$

$$V_{td}^* V_{cd} + V_{ts}^* V_{cs} + V_{tb}^* V_{cb} = 0. \tag{1.36}$$

These constraints can be represented by triangles in the complex plane, denoted by "unitarity triangles". Most of them include terms of different orders in λ, thus corresponding to flat triangles. Equation 1.32 and 1.35, however, only include terms that are proportional to λ^3.

The triangle defined by Eq. 1.32 is often called "the" unitarity triangle, as it has been the focus of many measurements. Indeed, the three sides of this triangle are all of order λ^3, compared to other triangles that are flatter. Alternatively, it is referred to as the B^0 unitarity triangle. Its sides are normalized by $V_{cd}^* V_{cb}$, and its internal angles are thus defined as:

$$\alpha = \arg\left(-\frac{V_{tb}^* V_{td}}{V_{ub}^* V_{ud}}\right) = \arg\left(-\frac{1 - \rho - i\eta}{\rho + i\eta}\right) + \mathcal{O}(\lambda^2), \tag{1.37}$$

$$\beta = \arg\left(-\frac{V_{cb}^* V_{cd}}{V_{tb}^* V_{td}}\right) = \arg\left(\frac{1}{1 - \rho - i\eta}\right) + \mathcal{O}(\lambda^4), \tag{1.38}$$

$$\gamma = \arg\left(-\frac{V_{ub}^* V_{ud}}{V_{cb}^* V_{cd}}\right) = \arg\left(\rho + i\eta\right) + \mathcal{O}(\lambda^2). \tag{1.39}$$

Figure 1.1 shows a sketch of this unitarity triangle specifying the angles and the expressions of the lengths of its sides [19].

The angles and the sides of the triangle can be measured experimentally, to constrain the location of its apex.[6] These different constraints set by the measurements must overlap in at least one region of space so that the unitarity of the CKM matrix is respected. Figure 1.2 shows the status of the constraints on this unitarity triangle, from the CKMfitter collaboration [19]. These constraints arise from the measurement of physics observables by several experiments. They include

- the measurement of ε_K and ε_K' (CP-violating parameters of the neutral kaon system) [26];

- the constraint on Δm_d, measured first by the UA1 [27] and ARGUS [28] collaborations; current world average is dominated by B-factories and LHCb;

---

[6]The freedom to set the origin of the referential and its orientation can be used to set two of the tree apexes of the triangle to 0 and 1, leaving only one apex to be determined.
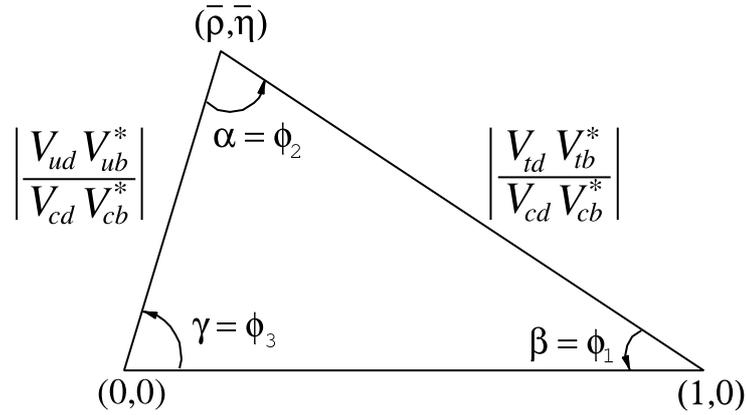
Figure 1.1 – Sketch of the unitarity triangle defined by Eq. 1.32, from [19].

- the constraint on Δm_s, firstly measured by CDF [29]; LHCb [30] is dominating the current world average[7];

- the measurement of β performed in $b \to c\bar{c}s$ modes by BaBar [31], Belle [32], and LHCb [33];

- the measurement of the angle α, measured in time-dependent analyses of $b \to u\bar{u}d$ decays such as $B \to \pi\pi$, $B \to \rho\rho$, and $B \to \rho\pi$;

- the constraint on γ, set with the best precision in charmed B tree decays, and measured by CDF, BaBar, Belle, and LHCb. It is one of the least known parameters of the B^0 unitarity triangle.

The mixing phase between the B_s^0 and the $\overline{B}_s^0$ is noted ϕ_s, and is equal to

$$\phi_s = -2\beta_s = arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right), \tag{1.40}$$

where β_s is one of the angles of the B_s^0 unitarity triangle defined by Eq. 1.33. The LHCb experiment disposes of a large sample of B_s^0 mesons that allows it to improve constraints on this triangle.

1.2.6 B^0 oscillations and the β angle

As discussed in Sec. 1.2.3, flavoured neutral mesons (K^0, D^0, B^0, and B_s^0) oscillate when they propagate. The short-range terms related to these oscillations can be described at first order by box diagrams like those shown in Fig. 1.3. Long-range terms and upper or-

---

[7]The ratio $\Delta m_d / \Delta m_s$ is cleaner than the individual observables, as it cancels some hadronic uncertainties.
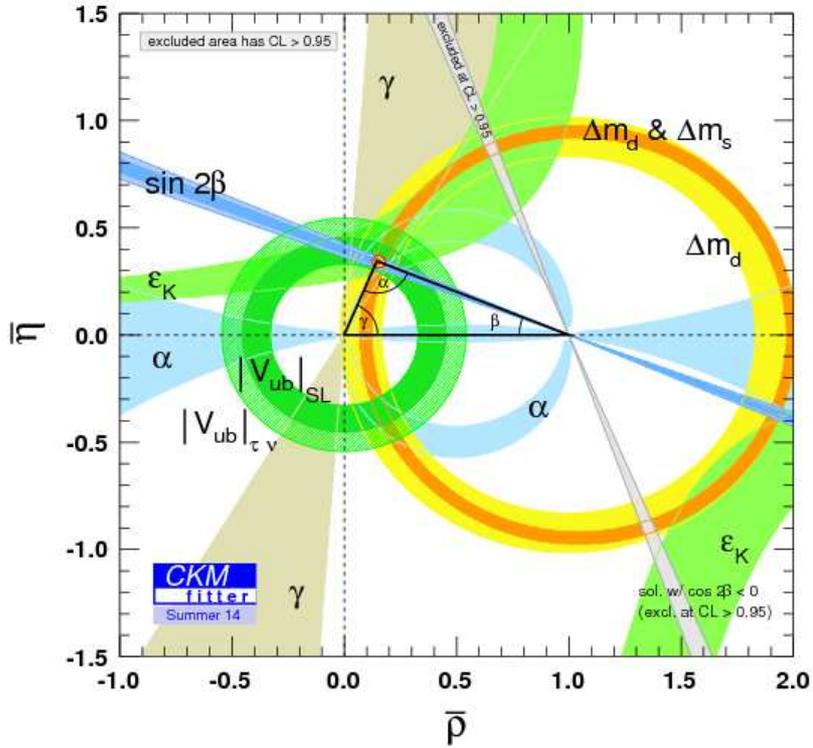
Figure 1.2 – Constraints on the apex of the unitarity triangle defined by Eq. 1.32 from the CKMfitter collaboration [25]
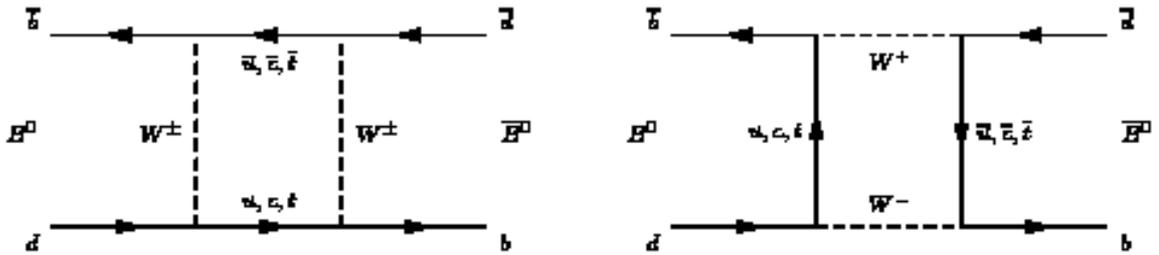
Figure 1.3 – Second-order weak interaction Feynman diagrams that give rise to the mixing of the B^0 meson. The virtual loop in both diagrams is dominated by the top-quark.

ders are neglected. The contribution from virtual quarks inside of the loop are dominated by the top-quark. It is then a very good approximation to consider the amplitude to be proportional to $V_{tb}V_{td}^*/V_{tb}^*V_{td}$, whose phase is equal to -2β at $\mathcal{O}(\lambda^4)$. This expression also yields that $|q/p| = 1 + \mathcal{O}(\lambda^4)$, thus strongly suppressing CP violation in the mixing of B^0

mesons.

The angle β can be extracted from various channels that allow to measure the interference between the mixing and the decay of B^0 mesons. Considering a $B^0 \to f$ decay, where f is a *CP* eigenstate and only one process contributes to the amplitude, no direct *CP* violation is possible and

$$S_f = \sin\left(\arg\left(\lambda_f\right)\right) = \sin\left(\arg\left(\frac{q}{p}\frac{\overline{\mathcal{A}}_f}{\mathcal{A}_f}\right)\right) = \eta_f \sin 2\beta, \qquad (1.41)$$

where $\eta_f = \pm 1$ is the eigenvalue of the f final state. The observable S_f can be extracted from an analysis that measures $\Gamma(t)$ (time-dependent analysis).

Decays of the form $B^0 \to K_{\mathrm{S}}^0(K_{\mathrm{L}}^0)(c\bar{c})$ are dominated by the tree-level transition $b \to c\bar{c}s$ and thus allow for a clean measurement of the angle β by means of a time-dependent analysis. This allows to extract a clean measurement of β in modes where no significant contribution from NP processes is expected. This value can then be compared to the value of β from modes that include a virtual loop.

Charmless B^0 decays involve an underlying $b \to q\bar{q}s$ transition. They are strongly suppressed at tree level as the only tree-level contribution involves a $b \to u$ transition, that is suppressed by a factor λ^2 in branching fractions compared to a $b \to c$ transition. Figure 1.4 shows a compilation of the CKM angle β and of β_{eff} as of 2014 [19], in the $b \to c\bar{c}s$ and the $b \to q\bar{q}s$ transitions, respectively. These two averages are compatible, but most of the $b \to q\bar{q}s$ measurements are smaller than measurements in $b \to c\bar{c}s$ modes.

1.3 Amplitude analyses concepts

1.3.1 Three-body particle decays and the Dalitz plot

The differential cross-section associated with the decay of a particle of mass M and momentum P into n particles of momenta p_i and energies E_i is

$$d\Gamma = \frac{(2\pi)^4}{2M}\left|\mathcal{M}\right|^2 d\Phi_n(P; p_1...p_n), \qquad (1.42)$$

where

$$d\Phi_n(P; p_1, ...p_n) = \delta^4(P - \sum_{i=1}^n p_i) \prod_{i=1}^n \frac{d^3p_i}{(2\pi)^3 2E_i} \qquad (1.43)$$

is the phase-space element of volume, and the scattering matrix $\mathcal{M}$ contains all the information related to underlying dynamics (such as resonances or hadronic factors). Conservation of momentum is ensured by the Dirac function δ.

In the case of three-body decays, the previous equation becomes

$$d\Gamma = \frac{1}{(2\pi)^5}\frac{1}{16M^2}\left|\mathcal{M}\right|^2 dE_1 dE_3 d\alpha d(\cos\beta)d\gamma \qquad (1.44)$$
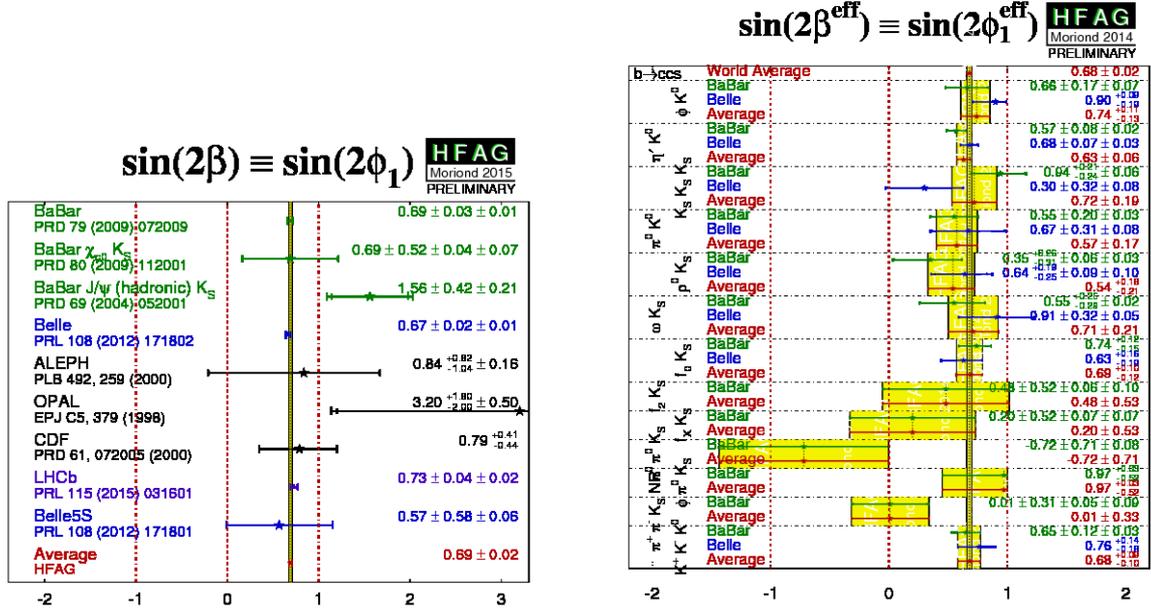
Figure 1.4 – World average of β from [19], extracted from $b \to c\bar{c}s$ decays (left) and $b \to q\bar{q}s$ decays (right). The world average from $b \to c\bar{c}d$ is also indicated in the right hand-side figure.

where E_1 and E_3 are the energy of particles 1 and 3 in the rest frame of the mother particle. The angles α, β, and γ are the Euler angles that define the plane where momenta of the daughters are contained. Here, the initial twelve degrees of freedoms are reduced to five when the conservation of momentum and the masses of the three final-state particles is taken into account.

In the case of the decay of a (pseudo-)scalar particle into three (pseudo-)scalar particles, the process is isotropic. This means that the dependency on angles can be integrated out, further reducing the number of degrees of freedom from five to two. Equation 1.44 becomes

$$d\Gamma = \frac{1}{(2\pi)^3} \frac{1}{8M} |\mathcal{M}|^2 dE_1 dE_3. \tag{1.45}$$

This equation can be rewritten as

$$d\Gamma = \frac{1}{(2\pi)^3} \frac{1}{32M^3} |\mathcal{M}|^2 dm_{12}^2 dm_{13}^2, \tag{1.46}$$

where the m_{ij} masses are the invariant masses of the particle pair ij. This amplitude only depends on two variables, which allows to represent the whole phase-space on a single plane. A graphic representation of this plane is called a Dalitz plot [34].

The conservation of momentum and the mass of the mother particle set constraints on the Dalitz plot. Figure 1.5 shows a typical Dalitz plot along with kinematical boundaries.
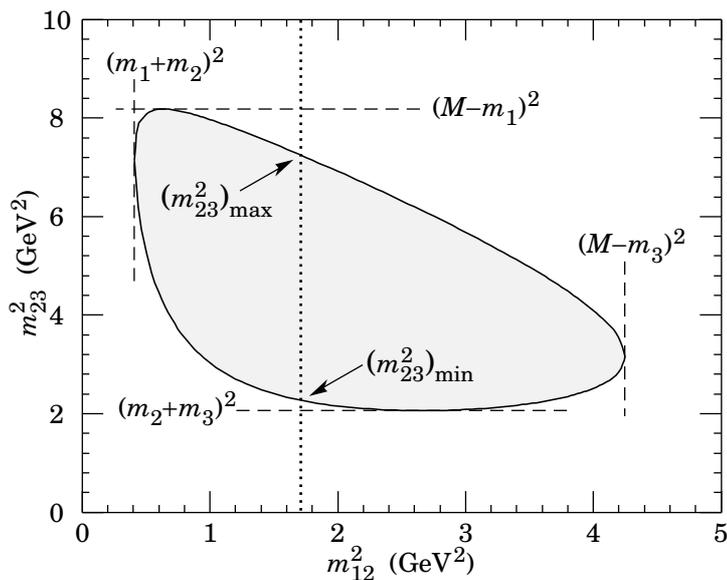
Figure 1.5 – Typical Dalitz plot, along with kinematical boundaries [4].

The physical region corresponds to the gray area, limited by boundaries where all particles are collinear. The corners of the physical region correspond to the case where one of the particles is at rest.

As shown in Eq. 1.46, the only possible source of non-uniformities over the Dalitz plot is the scattering matrix $\mathcal{M}$. Such non-uniformities typically arise in the presence of quasi-two body (Q2B) decays (see Sec. 1.3.3).

1.3.2 The square Dalitz plot

An alternative representation of events that is sometimes easier to manipulate is the square Dalitz plot [35]. Its coordinates m' and θ' are defined as

$$\mathrm{m}' = \frac{1}{\pi}\arccos\left(2\frac{\mathrm{m}_{\mathrm{ij}} - \mathrm{m}_{\mathrm{ij}}^{\mathrm{min}}}{\mathrm{m}_{\mathrm{ij}}^{\mathrm{max}} - \mathrm{m}_{\mathrm{ij}}^{\mathrm{min}}} - 1\right), \tag{1.47}$$

$$\theta' = \frac{1}{\pi}\theta_{ij}, \tag{1.48}$$

where $m_{ij}^{\mathrm{max(min)}}$ designates the maximum (minimum) mass of the ij pair

$$m_{ij}^{\mathrm{max}} = M - m_k \tag{1.49}$$

$$m_{ij}^{\mathrm{min}} = (m_i + m_j), \tag{1.50}$$

18
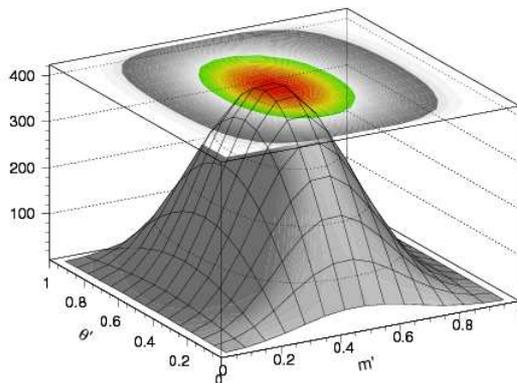
Figure 1.6 – Jacobian of the transformation from the usual Dalitz plot to the square Dalitz plot. [36]

where M is the mass of the mother particle and $m_{i,j,k}$ is the mass of the daughter i, j, or k. The angle θ_{ij} is the helicity angle of a given ij system, which is defined between the momenta of the particles k and i in the ij rest frame. These coordinates are defined between 0 and 1, and the change of coordinates between the regular Dalitz plot and the square Dalitz plot is defined as

$$dm_{ij}^2 dm_{jk}^2 \to |\det J| d\,\mathrm{m}'d\theta', \tag{1.51}$$

$$|\det J| = 4 \left|p_{ij}^*\right| \left|p_k^*\right| \frac{\delta m_{ij}}{\delta \,\mathrm{m}'} \frac{\delta \cos\theta_{ij}}{\delta\theta'}, \tag{1.52}$$

J being the Jacobian of the transformation. The momenta $p_{ij}^* = \sqrt{E_{ij}^2 - m_{ij}^2}$ and $p_k^* = \sqrt{E_k^2 - m_k^2}$ are defined in the ij rest frame. Figure 1.6 shows the distribution of this Jacobian over the square Dalitz plot.

This representation is especially useful in charmless B decays, as they populate areas of the Dalitz plot close to its boundaries, due to the small mass of intermediate resonances compared to the mass of the B meson. Additionally, from a technical point of view, the square shape of this plot allows to bin the plane more easily.

A major difference between the usual Dalitz plot and the square Dalitz plot is that the square Dalitz plot area is not proportional to the element of phase space. This means that structures over the square Dalitz plot are not necessarily related to any dynamics, unlike in the usual Dalitz plane. This is illustrated by Fig. 1.6, as the Jacobian can be interpreted as the shape over the square Dalitz plot of a flat, phase-space, component.
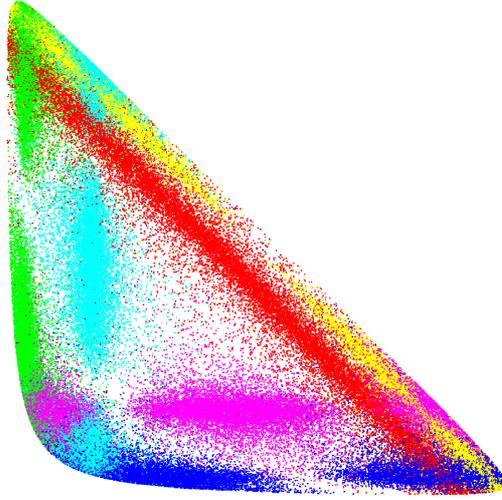
Figure 1.7 – Sketch of a Dalitz plot including several Q2B decays, represented in different colours. The red strip corresponds to a scalar resonance. The green and dark blue points correspond to vector resonances, while the magenta and light blue correspond to tensor resonances. A spin 3 resonance is also shown, in yellow. Interference regions are clearly visible where resonances overlap, such as the red and green ones.

1.3.3 Quasi-two body decays

A decay $A \to B + C + D$ can proceed via an intermediate state R that decays for instance into B and C. The decay $A \to (R \to B + C)D$ is called a "quasi-two-body" decay, where D is sometimes denoted as the "bachelor" particle.

A Q2B decay appears as a strip over the Dalitz plot with a mean and a width that are related to the mass and the width of the resonance, respectively. The variations of the magnitude along the strip provide information about the spin of the resonance. Figure 1.7 shows an example of a Dalitz plot with several Q2B decays with different spins of the resonances.

1.3.4 The isobar model

The amplitude of a three-body decay can be modelled in different ways. The isobar approach approximates the total amplitude as

$$\overset{(-)}{\mathcal{A}} = \sum_n^N \overset{(-)}{\mathcal{A}_n}, \qquad (1.53)$$

where the sum runs over N coherent contributions and

$$\overset{(-)}{\mathcal{A}_n} = \overset{(-)}{a_n} F_n \left(m_{ij}^2, m_{jk}^2 \right), \qquad (1.54)$$

are the partial amplitudes that are characterized by their lineshapes $F_n\left(m_{ij}^2, m_{jk}^2\right)$ [37–39]. The $\overset{(-)}{a_n}$ coefficients are complex numbers.

Under the assumption of heavy-quark factorization, the lineshape $F_n\left(m_{ij}^2, m_{jk}^2\right)$ is only related to strong dynamics. Hence this does not contain any information about CP violation[8], and is not decay-dependent.

The isobar coefficients $\overset{(-)}{a_n}$ can be parameterized in several ways. In the analyses that are detailed in this dissertation, they are parameterized as

$$a_n = c_n(1 \pm b_n)e^{i(\phi_n \pm \delta_n)}, \tag{1.55}$$

where ϕ_n and δ_n are the strong and weak phases, respectively. This parameterization has been proposed by the BaBar experiment in Ref. [40].

The coefficients a_n are not physical as, for instance, all the magnitudes in a given model could be multiplied by a factor without changing the description of data. Fit fractions of resonances are defined as

$$FF_n = \frac{\underset{\text{DP}}{\iint} \left(|\mathcal{A}_n|^2 + \left|\overline{\mathcal{A}_n}\right|^2\right) dm_{12}^2 dm_{13}^2}{\underset{\text{DP}}{\iint} \left(|\mathcal{A}|^2 + \left|\overline{\mathcal{A}}\right|^2\right) dm_{12}^2 dm_{13}^2} \tag{1.56}$$

Each Q2B branching fractions is related to the fit fraction of the corresponding resonance by

$$\mathcal{B}(A \to RD) = FF_R \times \mathcal{B}(A \to BCD). \tag{1.57}$$

Similarly, interference fit fractions between two resonances i and j can be defined as

$$FF_{nm} = 2\,\mathrm{Re}\left(\frac{\underset{\text{DP}}{\iint} \left(\mathcal{A}_n\mathcal{A}_m^* + \overline{\mathcal{A}_n\mathcal{A}_m}^*\right) dm_{12}^2 dm_{13}^2}{\underset{\text{DP}}{\iint} \left(|\mathcal{A}|^2 + \left|\overline{\mathcal{A}}\right|^2\right) dm_{12}^2 dm_{13}^2}\right). \tag{1.58}$$

The sum of fit fractions FF_n is not necessarily unity, because of interference. In the contrary, the relation

$$\sum_{n \le m} FF_{nm} = 1 \tag{1.59}$$

is fulfilled.

The parameters of the coefficients a_n can also be used to define CP-violating observables $\mathcal{C}_n$ and $\mathcal{S}_n$ for resonance n as

$$\mathcal{C}_n = \frac{2b_n}{1 + b_n^2} \tag{1.60}$$

and

$$\mathcal{S}_n = \frac{1 - b_n^2}{1 + b_n^2} \sin(2\beta_{\text{eff},n}), \tag{1.61}$$

---

[8]CP violation by the strong interaction is negligible.

where
$$\beta_{\text{eff},n} = \beta + \delta_n. \tag{1.62}$$
The angle $\beta_{\text{eff},n}$ quantifies the interference between mixing and decay for a given Q2B decay. The observable $\mathcal{S}_n$ is meaningful only for CP-eigen Q2B decays such as $B^0 \to K_{\text{s}}^0 \phi$.

1.3.5 Resonance dynamics

Resonance dynamics are contained in the $F_n\left(m_{ij}^2, m_{jk}^2\right)$ terms of the isobar decomposition. This term can be decomposed as
$$F_n\left(m_{ij}^2, m_{jk}^2\right) = X_L(|\mathbf{p}^*|r')X_L(|\mathbf{q}|r)T_n(L, \mathbf{p}, \mathbf{q})R_n(m_{ij}), \tag{1.63}$$
where:

- i and j are the daughters of the resonance;

- m_{ij} is the mass of the decay products of the resonance;

- L is the angular momentum of the resonance;

- $\mathbf{p}^*$ is the momentum of the bachelor particle, evaluated in the B rest frame;

- r and r' are the Blatt-Weisskopf barrier radii;

- $\mathbf{p}$ and $\mathbf{q}$ are the momenta of one of the resonance daughters and of the bachelor particle, respectively. They are both evaluated in the rest frame of the resonance.

We review in the following the definition and the physical meaning of the terms X_L, T_n, and R_n.

Blatt-Weisskopf momentum barrier factors X_L

The maximum angular momentum L of a strong decay is limited by the momentum and by a distance that is comparable to the "radius" of the resonance. The Blatt-Weisskopf momentum barrier factor ([41]) depends on these two variables, and reweights the amplitudes in order to enforce the global conservation of angular momentum. The value of this factor depending on the angular momentum L of the resonance is

$$L = 0; X_L(z = |\mathbf{p}|r) = 1;$$

$$L = 1; X_L(z = |\mathbf{p}|r) = \sqrt{\frac{1 + z_0^2}{1 + z^2}};$$

$$L = 2; X_L(z = |\mathbf{p}|r) = \sqrt{\frac{9 + 3z_0^2 + z_0^4}{9 + 3z^2 + z^4}};$$

where z_0 is the value of the $z = |\mathbf{p}|r$ variable when the invariant mass of the two daughter particles is exactly the mass of the resonance. In the following of this dissertation, we take the radii values as $r' = 0$ and $r = (4 \pm 2.5)\,(\text{GeV}/\text{c}^2)^{-1}$ from Ref. [40].

Zemach tensor T_n

The Zemach tensor T_n describes the angular distribution of the resonance daughters [42]. It depends on the resonance angular momentum L following:

$$L = 0; T_n = 1; \tag{1.64}$$

$$L = 1; T_n = 4\mathbf{p.q}; \tag{1.65}$$

$$L = 2; T_n = \frac{16}{3} \left[3(\mathbf{p.q})^2 - (|\mathbf{p}||\mathbf{q}|)^2) \right]. \tag{1.66}$$

$$\tag{1.67}$$

This terms explains the variations of the amplitude along a resonance, as seen in Fig. 1.7.

The resonance lineshape R_n

Hadronic resonances are defined as poles of the scattering matrix S, which describes the unitary operator that relates the asymptotic initial and final states. They appear in several ways, for instance as an increase in the total cross-section when s approaches the square of the mass of the resonance. The amplitude can be expanded in several ways around such a pole.

The Breit-Wigner formalism is well-suited to model the amplitude near an isolated pole that is far from the opening of any threshold. It is a first-order Taylor expansion of the amplitude around the pole. The associated lineshape is defined as

$$R_n(m_{ij}) = \frac{1}{m_r^2 - m_{ij}^2 - im_r\Gamma_{ij}(q)}, \tag{1.68}$$

where r is a resonance decaying into the particles i and j, and q is the momentum of the resonance in the mother rest frame. The mass-dependent width Γ is

$$\Gamma_{ij}(q) = \Gamma_r \left(\frac{|\mathbf{q}|}{|\mathbf{q_r}|} \right)^{2L+1} \left(\frac{m_r}{m_{ij}} \right) X_L^2(q, q_0), \tag{1.69}$$

where Γ_r is the intrinsic width of the resonance and $\mathbf{q_r}$ is the value of $\mathbf{q}$ when $m = m_r$. It is worth mentioning that a sum of close Breit-Wigner distributions breaks the unitarity of the S matrix. In the case of overlapping or broad resonances, other parameterizations can be used, such as the Gounaris-Sakurai for the ρ^0 resonance [43]. One of these parameterizations is the Flatté formula ([44]) that describes the amplitude of a resonance close to a threshold, such as the $f^0(980)$ (close to the KK threshold)

$$R_n(m_{ij}) = \frac{1}{m_r^2 - m_{ij}^2 - i(\rho_1 g_1^2 + \rho_2 g_2^2)}, \tag{1.70}$$

where $g_1^2 + g_2^2 = m_r\Gamma_r$. The g_i are coupling constants that are measured experimentally. The ρ_i factors are phase-space terms that contain the information about the different

masses of the final states. In the case of the $f_0(980)$, they are written

$$\rho_1 = \rho_{\pi\pi} = \sqrt{\left(1 - \left(\frac{2m_{pi}}{m_{ij}}\right)^2\right)}, \tag{1.71}$$

$$\rho_2 = \rho_{KK} = \frac{1}{2}\sqrt{\left(1 - \left(\frac{2m_{K^\pm}}{m_{ij}}\right)^2\right)} + \sqrt{\left(1 - \left(\frac{2m_{K^0}}{m_{ij}}\right)^2\right)}. \tag{1.72}$$

The $\mathcal{K}$-matrix formalism ([45]) describes the scattering process by decomposing the $\hat{T}$ matrix as

$$\hat{T} = (I - i\rho\hat{\mathcal{K}})^{-1}\hat{\mathcal{K}}, \tag{1.73}$$

where ρ is the phase-space matrix, and $\hat{\mathcal{K}}$ is a Lorentz-invariant matrix defined as

$$\hat{\mathcal{K}}_{ij} = \sum_\alpha \alpha \frac{\sqrt{m_\alpha \Gamma_{\alpha,i}(m) m_\alpha \Gamma_{\alpha,j}(m)}}{(m_\alpha^2 - m^2)\sqrt{\rho_i \rho_j}}, \tag{1.74}$$

where the sum runs overs all resonances α. This construction explicitly enforces the unitarity of the $\hat{T}$ operator. Additionally, this expression yields the same result as a Breit-Wigner in the case of a single resonance in a single channel. The $\mathcal{K}$-matrix formalism is best defined in the case of scattering. It can be transposed to the case of three-body decays under the assumption that there are no interactions between the bachelor particle and the daughters of the resonance.

1.3.6 Nonresonant amplitude

The nonresonant amplitude is not related to any pole of the S matrix, and covers the whole phase space. It is especially important to consider in the case of B decays as the phase space is large, and as resonances cover a small portion of it (even more so in the case of charmless B decays, as charmless resonances have a small mass compared to the B mass). As a result, while the typical nonresonant contribution to charmed decays is of the order 10%, it can be as large as 90% in $B \to KKK$ decays ([46]). This nonresonant amplitude is poorly understood theoretically, and may even be the result of the presence of several broad resonances. Several parameterizations of the nonresonant amplitude have been used by different analyses from different collaboration, including a flat distribution, an exponential distribution, and a polynomial. These parameterizations have usually been defined in an ad hoc manner. For instance the use of a flat nonresonant distribution was motivated by the presence of signal events in the centre of the Dalitz plot.

In the context of the factorization approach, a large scalar contribution is expected in $B \to KKK$ [47]. However, as discussed in the following, a large additional P-wave contribution has been observed by the BaBar experiment in these modes.

Table 1.2 – Summary of favoured ("Fav.") and suppressed ("Sup.") $B_{d,s}^0 \rightarrow K_S^0 h^+ h'^-$ decays. The suppression factor is equal to λ^2, where $\lambda \approx 0.22$ is the sine of the Cabibbo angle.

$B_{d,s}$ meson	Final state			
	$K_S^0 \pi^+ \pi^-$	$K_S^0 \pi^+ K^-$	$K_S^0 K^+ \pi^-$	$K_S^0 K^+ K^-$
B^0	Fav.	Sup.	Sup.	Fav.
B_s^0	Sup.	Fav.	Fav.	Sup.

1.4 The study of $B_{d,s}^0 \rightarrow K_S^0 h^+ h^-$ decays

Decays of B^0 and B_s^0 mesons to $K_S^0 h^{\pm} h'^{\mp}$ are a privileged sector to perform indirect searches for NP. They have been studied for years in different experiments such as BaBar, Belle, and now LHCb. This section presents the general properties of these decays, with an emphasis on the $B^0 \rightarrow K_S^0 K^+ K^-$ decay, along with a state of the art.

1.4.1 $B_{d,s}^0 \rightarrow K_S^0 h^+ h'^-$ decay amplitudes

The $B_{d,s}^0 \rightarrow K_S^0 h^+ h'^-$ decays proceed through $b \rightarrow q\bar{q}u$ tree-level transitions, as well as $q \rightarrow q\bar{q}d$ and $b \rightarrow q\bar{q}s$ penguin transitions, where $q = d$ or s. Figure 1.8 shows all the possible dominant diagrams that contribute to a three-body decay of a heavy meson, where Q denotes the heavy quark, $\mathcal{T}$, $\mathcal{C}$, $\mathcal{A}$, $\mathcal{E}$, and $\mathcal{P}$ stand for "tree", "colour-suppressed tree", "annihilation", "exchange", and "penguin", respectively. As discussed in Sec. 1.2.4, $b \rightarrow u$ transitions are suppressed with respect to $b \rightarrow c$ transitions by a factor λ^2, where $\lambda \approx 0.22$ is the sine of the Cabibbo angle. This results in the suppression of tree-level diagrams in these decays, relative to the penguin amplitudes. The following discussion thus focuses on the properties of penguin amplitudes.

Depending on the nature of the mother particle and on the number of kaons in the final state, a $B_{d,s}^0 \rightarrow K_S^0 h^+ h'^-$ decay proceeds via the Cabibbo-favoured $b \rightarrow q\bar{q}s$ transition or the Cabibbo-suppressed $b \rightarrow q\bar{q}d$ transition, as shown in Table 1.2.

In B-meson decays, it is a good approximation to factorize the weak and strong parts of the decay, due to the large mass of the b quark compared to Λ_{QCD}.

The $B^0 \rightarrow K_S^0 K^+ K^-$ decay contains the $B^0 \rightarrow K_S^0 \phi(1020)$ contribution, which is a particularly good channel to study time-dependent CP violation. Indeed, it is dominated by a $b \rightarrow s s\bar{s}$ transition that proceeds via a gluonic $b \rightarrow s$ penguin. There is no tree contribution to this channel ("tree pollution"), which means that deviations of the measured value of the CKM angle β compared to that performed in $b \rightarrow c\bar{c}s$ transitions such as $B^0 \rightarrow J/\psi K_S^0$ can be an indication to NP.

The $K_S^0 K^+ K^-$ final state is not a CP-eigenstate. Indeed, for a given orbital angular momentum L of the $K^+ K^-$ system,

$$CP|K_S^0 K^+ K^-\rangle = (-1)^L |K_S^0 K^+ K^-\rangle. \tag{1.75}$$
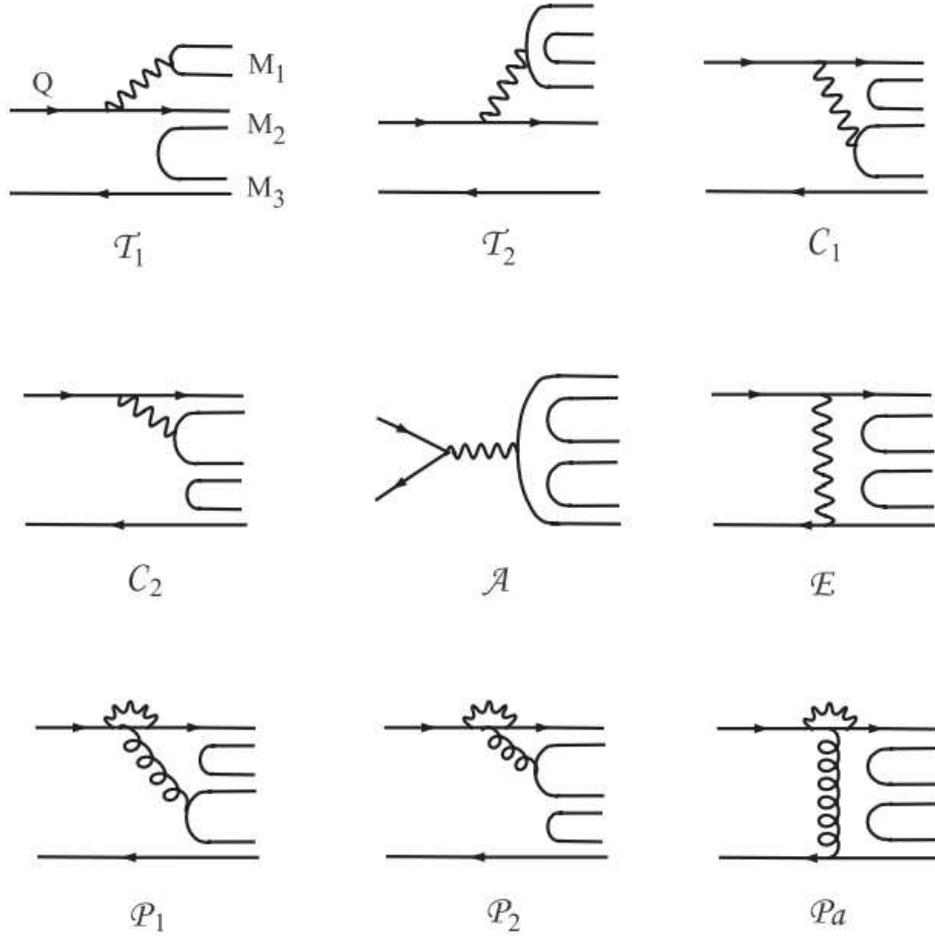
Figure 1.8 – Quark diagrams for three-body decays of a heavy meson. Q denotes the heavy quark. The dominant diagrams in $B^0 \to K_\mathrm{S}^0 K^+ K^-$ are P_1 and P_2.

One of the uses of a Dalitz plot analysis of this mode is to separate the different partial-wave contributions in order to measure CP-violating observables in a CP-eigen final state.

1.4.2 Previous studies of $B_{d,s}^0 \to K_\mathrm{S}^0 h^+ h^{'-}$ and $B_{d,s}^0 \to K_\mathrm{S}^0 K^+ K^-$ decays

The work described in this document is part of decade-long efforts by several collaborations to study $B_{d,s}^0 \to K_\mathrm{S}^0 h^+ h^{'-}$ decays. The main goal of the analyses described in the next section is to refine previous measurements and to gather more information in the specific LHCb environment. It lays the groundwork for future flavour-tagged, time-dependent, analyses of these modes in LHCb.

Legacy from B factories

The $B^0 \to K_S^0 \pi^+ \pi^-$ and $B^0 \to K_S^0 K^+ K^-$ decays have been studied by BaBar ([40,48]) and Belle ([49,50]). Both experiments have measured the branching fractions of these decays and performed a flavour-tagged time-dependent analysis that extracted β_{eff} in several Q2B modes. Additionally, BaBar has reported the observation of the $B^0 \to K_S^0 K^{\pm} \pi^{\mp}$ decay in [51] with a total significance of 5.2σ.

Studies of three kaons final states from BaBar and Belle have shown that they are dominated by a large nonresonant component. This nonresonant amplitude cannot be described with a flat phase-space shape, and BaBar has shown that it can be described as a sum of S-wave and P-wave contributions. One of the main goals of the time-integrated amplitude analysis described in Sec. 5 is to provide more insight on this component.

Analysis with LHCb data ($1\,\text{fb}^{-1}$)

The LHCb experiment has reported in [52] a measurement of the $B_{d,s}^0 \to K_S^0 h^+ h'^-$ branching fractions, relative to the $B^0 \to K_S^0 \pi^+ \pi^-$ branching fraction, as well as a glimpse of the Dalitz-plot distribution of signal events. The observation of the $B^0 \to K_S^0 K^{\pm} \pi^{\mp}$ mode from BaBar was confirmed, and the $B_s^0 \to K_S^0 \pi^+ \pi^-$ and $B_s^0 \to K_S^0 K^{\pm} \pi^{\mp}$ modes were observed for the first time.

Chapter 2

Description of the LHCb experiment

The LHCb detector is placed at one of the interaction points of the Large Hadron Collider (LHC), presented in Sec. 2.1. The detector, described in Sec. 2.2, is designed in order to take advantage of the large amount of $b\bar{b}$ and $c\bar{c}$ pairs produced near the beam axis in the conditions of the LHC. The system of particle identification is shortly discussed in Sec. 2.3.

The large number of collisions and their short spacing in time (50 ns in 2010–2012, then 25 ns) requires a trigger system, described in Sec. 2.4. Finally, modern high-energy physics relies on accurate simulations of physics and detector response, and I describe in Sec. 2.5 the software environment for Monte-Carlo productions in the LHCb experiment.

2.1 The Large Hadron Collider

The LHC is the largest and most powerful accelerator in the world in terms of centre-of-mass energy, and is located at the Conseil Européen pour la Recherche Nucléaire (CERN), in Geneva. It is the final point of a chain of accelerators located at CERN, shown in Fig. 2.2. It accelerates bunches of protons from 450 GeV to energies of 3.5, 4, or 6.5 TeV, depending on the data-taking period. This acceleration is performed using 16 radio-frequencies (RF) cavities located along the 27 km tunnel. The accelerator is also designed to accelerate beams of lead ions during dedicated runs.

The programme of the LHC is separated into several parts called "Runs" by long shutdowns (LS), during which the characteristics of the accelerator remain rather stable. The LHC accelerator physics programme is divided in Runs separated by long shut-downs (LSD) during which both the accelerator and the detectors can be maintained and/or upgraded. The data-taking during 2011 and 2012 is designated as "Run I", and the data-taking period starting from 2015 is designated as "Run II". Data used in this thesis was entirely acquired during Run I. Figure 2.1 shows the running plan for the LHC in the next few years, including planned upgrades for the experiments.

Nominal proton beams are composed of bunches of 1.2–$1.4.10^{11}$ protons separated by $50(25)$ ns in Run I(II). A beam can contain up to 2,808 bunches of protons, and can remain stable for over 8 hours. The beams are steered by 1,232 superconducting (1.3 K)
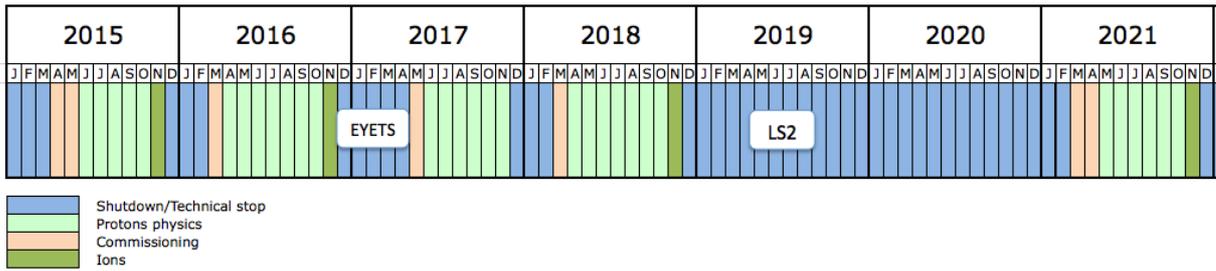
Figure 2.1 – Plan for the LHC in the next few years. Long shutdowns are indicated as "LS". The upgrade of the LHCb experiment is planned during "LS2".
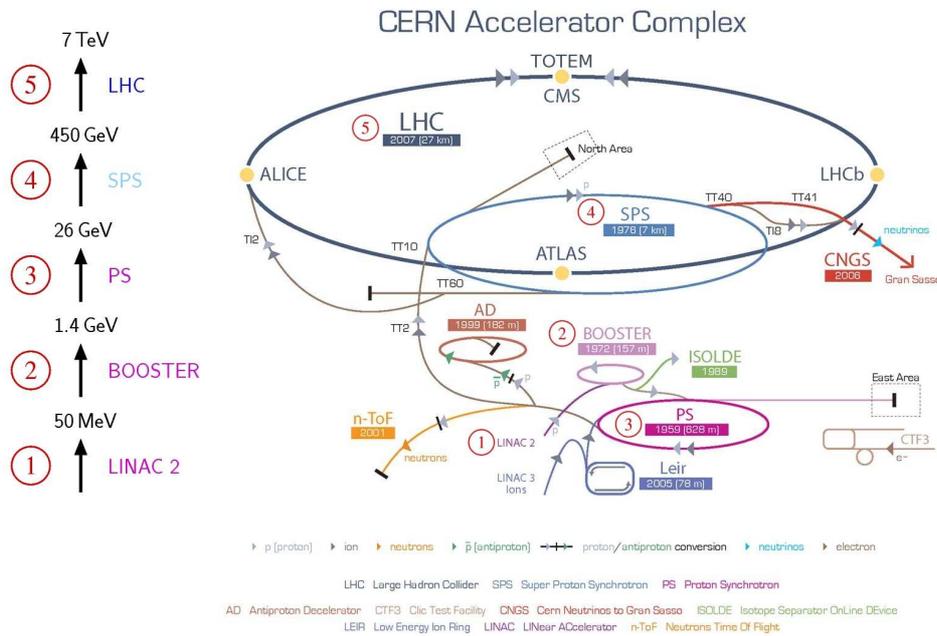
Figure 2.2 – Acceleration complex of the LHC. The four main experiments are also shown along the LHC.

Niobium-Titanium dipole magnets, cooled by superfluid helium, each of them creating a field of up to 8.3 T. Quadrupole and octupole magnets are also used to focus the beam and correct chromatic aberrations.

The two beams collide in 4 interaction points along the LHC, and seven experiments are located at these points. The ATLAS and CMS experiments use giant general-purpose detectors (GPDs) with a barrel-like geometry to study the product of collisions that have a large transverse momentum p_T. This physics programme includes the study of the
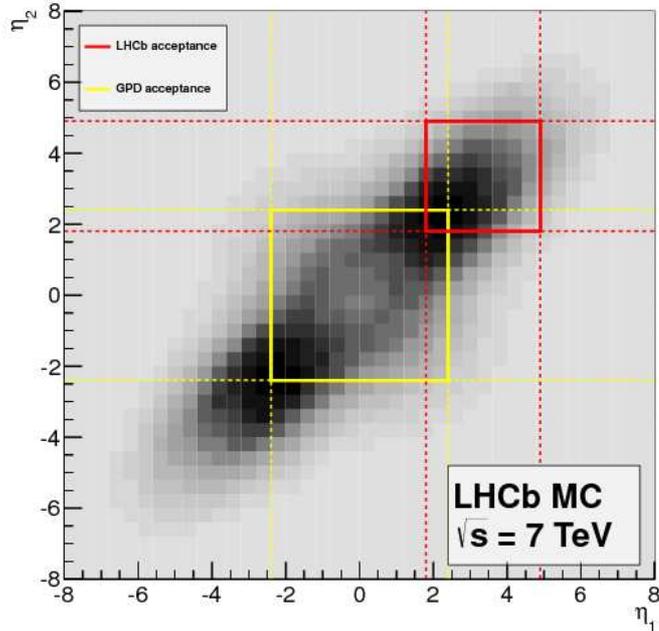
Figure 2.3 – Differential branching fraction of $b\overline{b}$ pair production at the energies of the LHC during the 2011 data-taking, and angular acceptance of GPDs (CMS and ATLAS, marked in yellow) and LHCb (marked in red).

Higgs boson that these experiments discovered in 2012 [9, 10], of the top quark, and searches for New Physics (NP) particles produced on shell. These two experiments, and CMS in particular, also have sensitivity to processes relevant for flavour physics, such as $B_s^0 \rightarrow \mu^+\mu^-$ [53, 54].

The other experiments being operated at the LHC are LHCb (described in Sec. 2.2), ALICE, TOTEM, LHCf, and MoEDAL. Figure 2.3 compares the angular coverage of the ATLAS, CMS, and LHCb detectors, illustrating the different purposes of these experiments.

2.2 The LHCb detector

The LHC accelerator is the world most intense source of b and c quark pairs. The LHCb detector is designed to take advantage of the localization of these pairs by covering only the forward regions near the beam axis. Figure 2.4 shows the diagrams responsible for heavy-quark pair production at the LHC.

The LHCb detector, shown in Fig. 2.5, is designed as a single-arm forward spectrom-
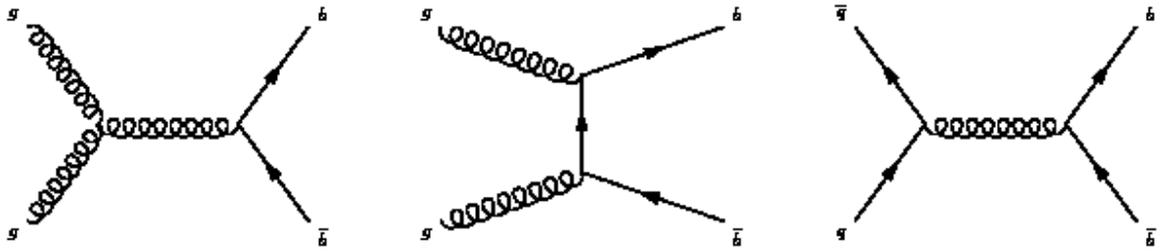
Figure 2.4 – Dominant diagrams for $b\bar{b}$ and $c\bar{c}$ pair production at the energies of the LHC: (left) $q\bar{q}$ annihilation; (middle) gluon separation; (right) gluon fusion.

eter. This geometry covers an angular acceptance of 15–300(250) mrad in the bending (non-bending) plane of the magnet.[1] A right-handed coordinate system is defined with the z-axis parallel to the beam axis in the direction from the VELO towards the muon stations, and the y-axis pointing upwards. In this arrangement the magnetic field bends trajectories in the xz plane. Additionally, the terms "upstream" and "downstream" are often used to designate the relative position of two points with respect to the interaction point.

---

[1]This is equivalent to a pseudorapidity coverage of $2 < \eta < 5$. The pseudorapidity η is defined as $\eta = -\log(\tan\theta/2)$, where θ is the polar angle with respect to the beam axis.
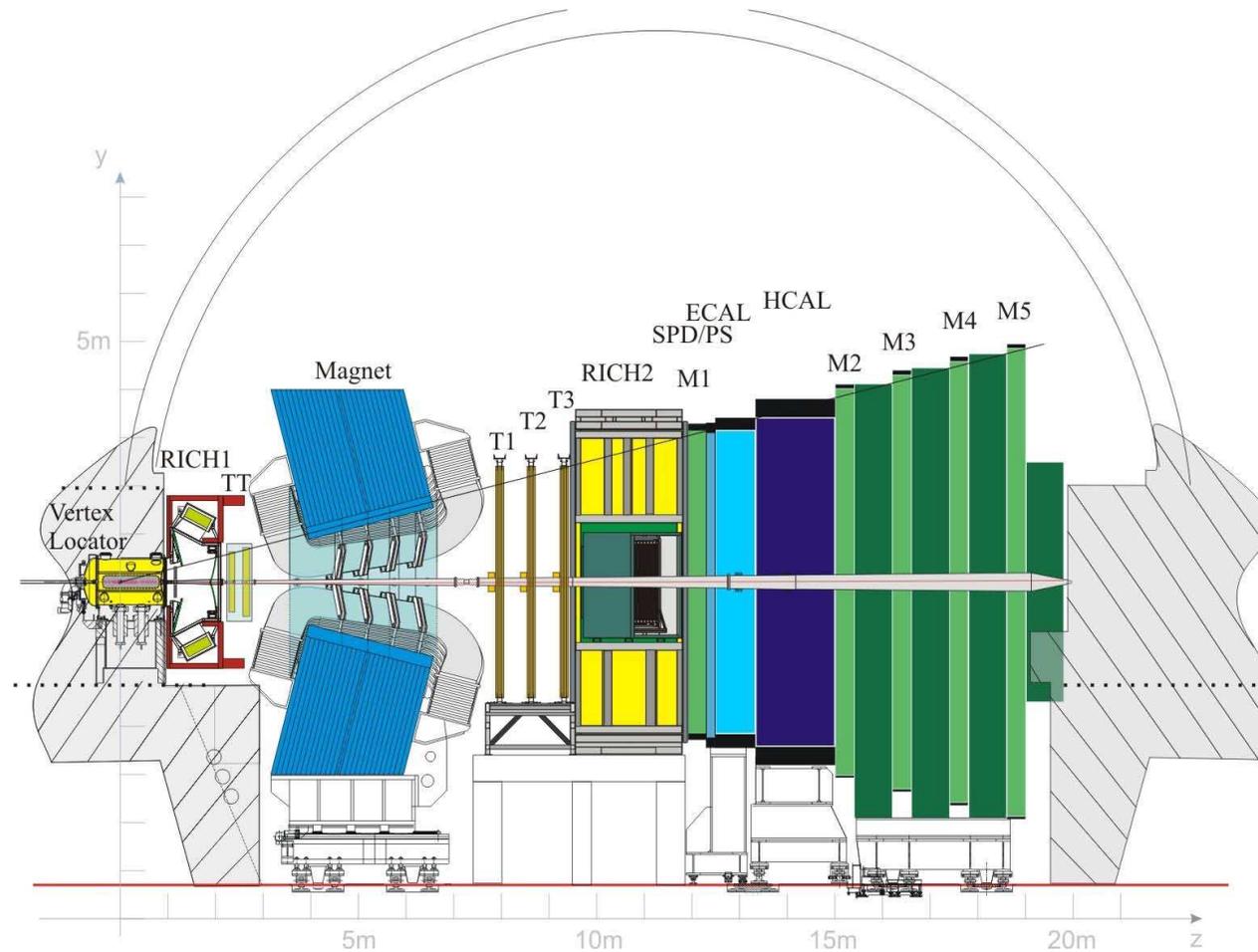
Figure 2.5 – Overview of the LHCb detector. The z axis is along the horizontal, while the y axis is along the vertical. The interaction point is located on the left, inside of the VELO subdetector. The beam pipe and the contours of the pit are filled with gray. Upstream and downstream directions correspond to the left and the right, respectively.
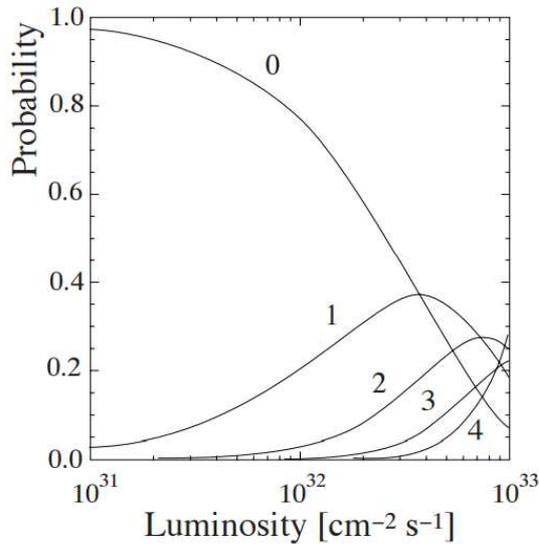
Figure 2.6 – Number of primary vertices as a function of the luminosity.

2.2.1 Beam conditions at the LHCb interaction point

The LHCb detector has been designed to operate at a nominal luminosity of $2 \times 10^{32} \, \text{cm}^2 \, \text{s}^{-1}$, which is lower than the maximum that can be provided by the LHC. Indeed, as shown in Fig. 2.6, larger instantaneous luminosities induce an increased number of multiple pp inelastic collisions. These multiple pp inelastic collisions increase the amount of data recorded by the detector, but induce larger occupancies and thus less accurate reconstruction. They also increase the amount of radiations absorbed by the detector. The luminosity for the LHCb experiment can be tuned by changing the beam focus at its interaction point independently from the other interaction points. This allows LHCb to maintain its optimal luminosity for the whole duration of a fill, as shown in Fig. 2.7. The luminosity has been increased to $3.5 \times 10^{32} \, \text{cm}^2 \, \text{s}^{-1}$ and $4.5 \times 10^{32} \, \text{cm}^2 \, \text{s}^{-1}$ in 2011 and 2012, respectively.

2.2.2 The magnet

The LHCb dipole magnet [55] is located between the TT and tracking stations. It generates a magnetic field that is perpendicular to the beam axis, so that the trajectory of all charged particles that pass through is curved. The curvature radius of the trajectory allows for a measurement of the track momentum. In order to achieve a 0.5% relative precision on p up to $200 \, \text{GeV}/c$ momenta, the integrated bending power is equal to $4 \, \text{Tm}$ for tracks of $10 \, \text{m}$ length.

The magnet is composed of two saddle-shaped aluminium coils maintained by an iron yoke, as shown in Fig. 2.8. An important feature of the LHCb magnet is the ability to
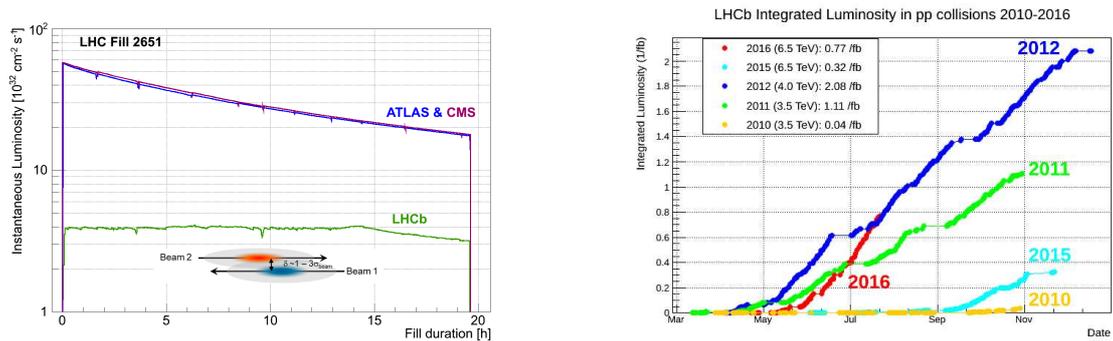
Figure 2.7 – (Left) Example of the evolution of instantaneous luminosity of the lifetime of a fill for different experiments. (Right) Integrated luminosity in fb^{-1} per year of data taking.
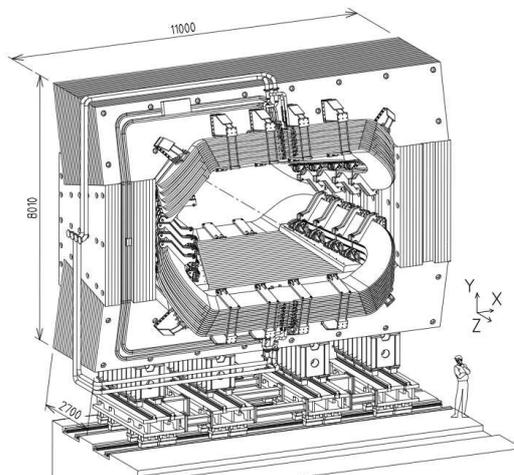
Figure 2.8 – Layout of the LHCb magnet.

reverse the polarity of its magnetic field.[2] This allows to cancel out detection asymmetries such as the charge detection asymmetry [56].

2.2.3   The tracking system

The goal of the tracking system is to measure the trajectories and momenta of charged particles ("tracks") in the detector acceptance. It is composed of the VELO and of two ensembles of stations located upstream and downstream from the magnet. These ensembles are the TT and the T1–3 stations. The T1–3 stations are composed of two distinct subdetectors: the Inner Tracker (IT) and the Outer Tracker (OT).

---

[2]The two polarities are referred to as *MagUp* and *MagDown*.
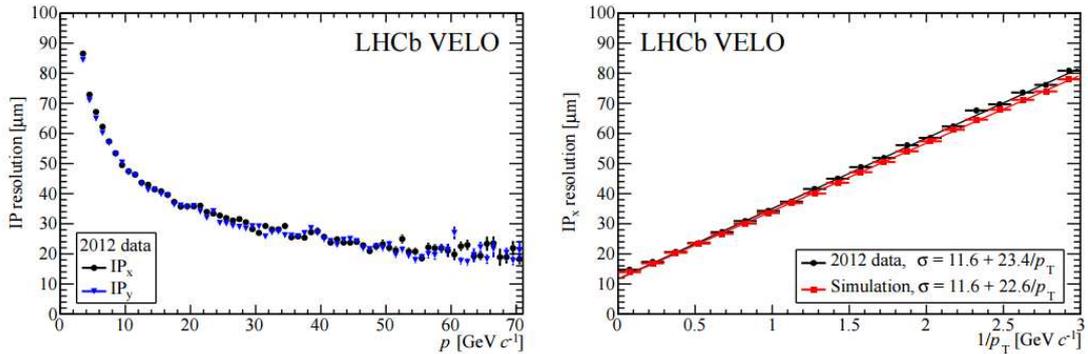
Figure 2.9 – (Left) Resolution of the impact parameter of a track with respect to a vertex as a function of the momentum p of that track. (Right) Resolution of the impact parameter of a track with respect to a vertex as a function of the inverse of the transverse momentum p_T of that track.

The vertex locator (VELO)

The VErtex LOcator (VELO) provides precise measurements of charged track coordinates close to the interaction region, which are used to identify the primary vertices and the displaced (secondary) vertices.[3]

The information on detached vertices is used to enrich the b-hadron content of the data written to tape, as well as in the LHCb offline analysis in order to measure particles lifetimes and to reject backgrounds. Indeed, the main source of background for most analyses is the combinatorial background, where one or several tracks are matched to the wrong decay vertex.

The ability of the VELO to differentiate between the multiple primary and secondary vertices is strongly related to its resolution of the impact parameter (IP) of tracks with respect to these vertices. This parameter is defined as the smallest distance of approach of a track to the vertex, and is expected to be zero for tracks originating from this vertex. The resolution of the VELO on the impact parameter of a track relative to a vertex depends on the transverse momentum p_T of that track. Figure 2.9 shows the performances of the VELO as a function of the p and p_T of a track [57].

The detector is divided in two halves, each consisting of 21 modules mounted around and downstream of the interaction point and perpendicular to the beam as shown in Fig. 2.10. The number of modules is chosen such that tracks that are inside the acceptance of the rest of the tracking system (and originate up to 10 cm downstream of the interaction point) traverse at least 3 modules. Each module is equipped with silicon strips oriented in the r and ϕ directions to measure the azimuthal and radial coordinates of charged

---

[3]Primary vertices are the vertices of the pp interaction. Conversely, vertices formed by the decay of particles ($e.g.B^0$ mesons) are called secondary vertices.
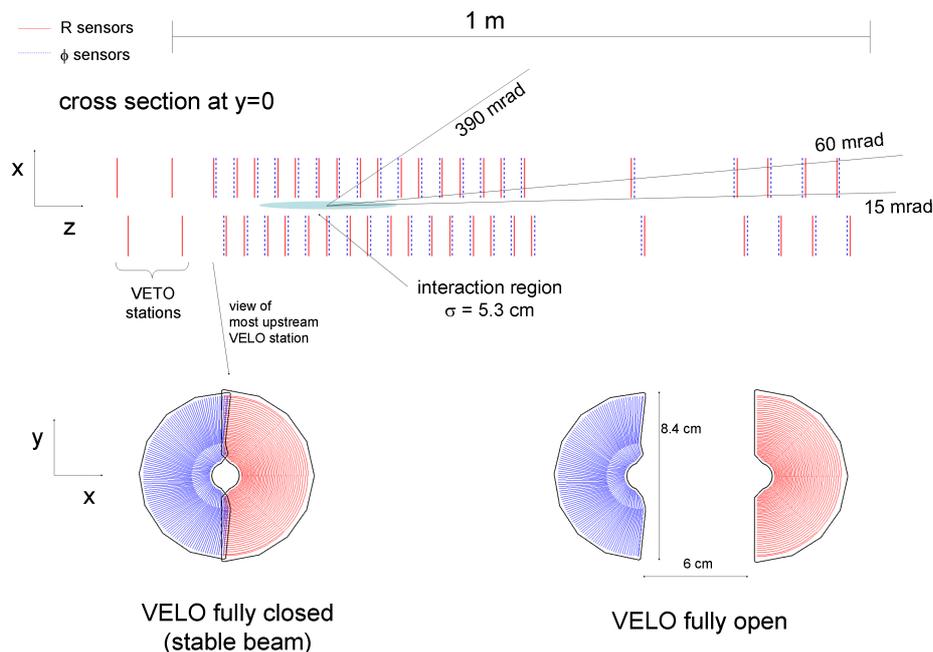
Figure 2.10 – Top: cut view of the VELO; the rest of the LHCb detector is located downstream (on the right). Bottom: view of a VELO module in closed (left) and open (right) positions.

particles.[4] The pitch within a module varies from 38 μm at the inner radius of 8.2 mm, increasing linearly to 102 μm at the outer radius of 42 mm. Figure 2.11 shows a projection of a module with its r and a ϕ silicon strips.

Two additional stations are placed upstream of the interaction point. They are used to aid the instantaneous measurement of luminosity. To protect the detector while LHC beams are not squeezed at the IP, the two VELO halves are retracted 35 mm from the beam axis, as shown in Fig. 2.10.

Tracking Turicensis (TT) stations

The Tracker Turicensis[5] (TT) detector is located upstream from the magnet, after the RICH1 subdetector. This station is composed of four planar layers 150 cm wide and 130 cm high, covering an active area of 8.4 m². These layers are arranged in a "x–u–v–x" layout, with vertical (x-layers) and rotated by stereo angles of +5° and -5° (u and v-layers, respectively) readout strips. The structure of these planes is illustrated in Fig. 2.12. This layout allows the TT to resolve the x and y position of the hits in the stations.

---

[4]This geometry is chosen to enable fast pattern recognition in the trigger.
[5]The Tracker Turicensis was formerly known as the Trigger Tracker.
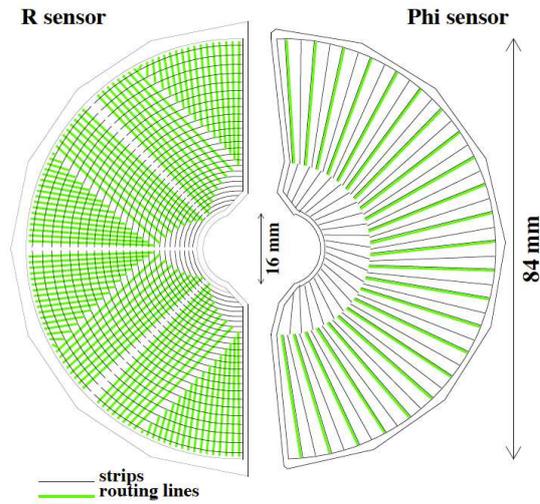
Figure 2.11 – Silicon strips used to measure the r (left) and the ϕ (right) coordinates in the VELO.
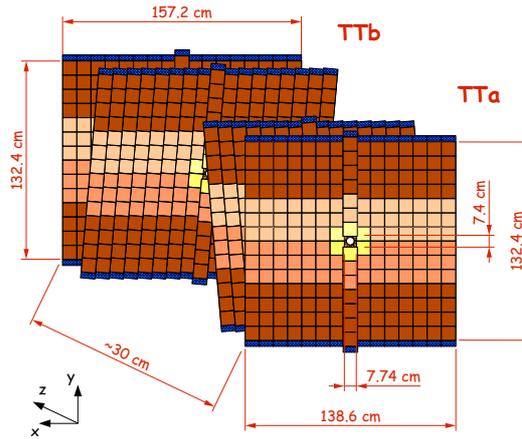
Figure 2.12 – Sketch of the TT subdetector. The two inner layers are tilted by a stereo angle of $\pm 5°$ (u/v-layers) in order to provide information on the y coordinate.

The planes which comprise the TT are manufactured using silicon micro-strip technology similar to that used in the VELO, with a strip pitch of $183\,\mu m$ and $500\,\mu m$ thick p+-on-n sensors.
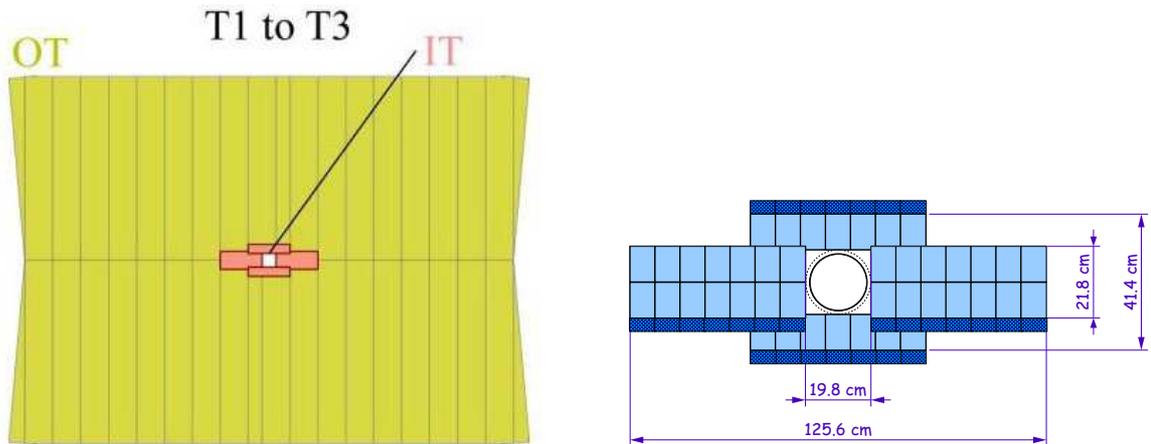
Figure 2.13 – (Left) View of a tracking module. The inner tracker covers the central region, while the outer tracker covers the rest of the angular acceptance. (Right) View of an IT module. Dimensions are given in cm and refer to the sensitive surface covered by the IT.

Tracking stations T1–3

The tracking stations T1–3 are located downstream from the magnet. Each of them is composed of four substations organized in a x–u–v–x layout, described in Sec. 2.2.3. In order to avoid uncovered regions in the acceptance, the top and bottom modules are staggered 4 mm in the z-axis and 3 mm in the x-axis, with respect to the lateral ladders. Each of these substations is divided into an inner tracker (IT) and an outer tracker OT, as shown in Fig. 2.13.

The central regions near the beam pipe feature large occupancies and require a fine granularity. The IT is positioned in the three downstream tracking stations T1–T3, and uses a silicon micro-strip technology. It is separated into single and double lines of seven staggered silicon ladders. Figure 2.13 shows a projection of this subdetector.

The remaining area has a significant reduction in the occupancy, allowing a coarser granularity. Therefore, the OT detector covers this large acceptance (total area of $5 \times 6 \, \text{m}^2$) utilizing a drift-tube technology. The OT acceptance extends from the outer boundaries of the inner tracker up to the nominal LHCb coverage. It is designed in four layers of arrays of gaseous straw tubes 2.4 m long and 4.9 mm in diameter. Each of these modules contains two monolayers of drift tube as shown in Fig. 2.14. The gas is composed of a mixture of Ar (70%) and CO_2 (30%). Both these characteristics enable the detector to achieve a fast drift-time across the drift-tubes under 50 ns, which is the performance required for the tracking algorithm.

Types of tracks in LHCb

The quality of a track in LHCb depends on the subdetectors used in its reconstruction. Four types of tracks are defined in LHCb: Long, Down, T, and Muon. Figure 2.15 shows
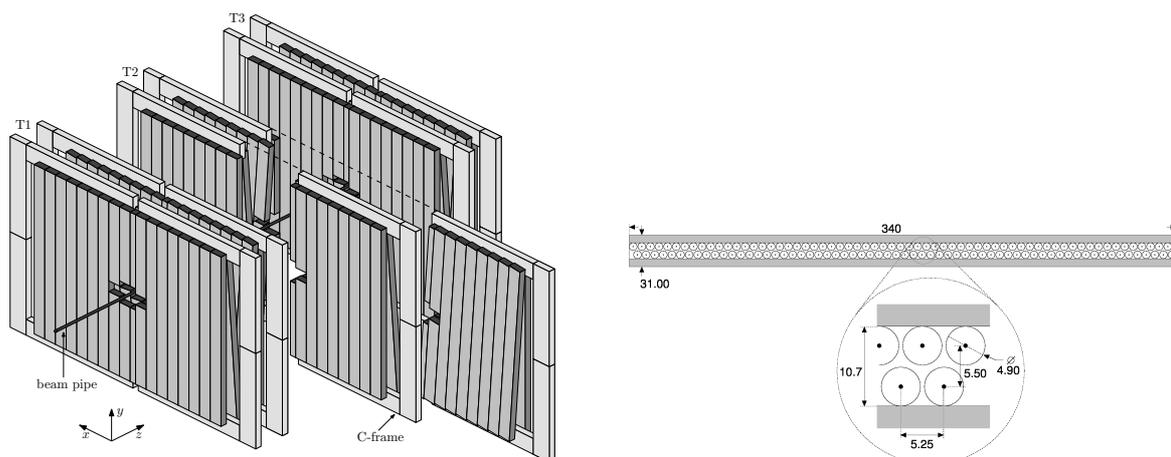
Figure 2.14 – a (Left) View of the OT. The T2 station is in open position. (Right) Illustration of the OT layer and zoom on the arrangement of straw tubes (distances are given in cm). [58]

these different types, with the exception of Muon tracks, which also have hits in the muon chambers.

A track is reconstructed as Long if it crosses at least three VELO stations. As K_S^0 mesons fly a typical distance of 1 m, some of them do not decay inside of the VELO acceptance and thus their pion daughters are reconstructed as Down tracks. Three types of K_S^0 mesons are defined in LHCb, depending on the track type of their daughters: Down-Down, Long-Long, and Long-Down. As the resolution of the momenta of Down tracks is worse than the resolution of Long tracks, the resolution of the mass of Down-Down K_S^0 mesons is worse than for Long-Long K_S^0 mesons.

2.2.4 The RICH1 and RICH2

The ring imaging Cherenkov (RICH) stations are located upstream (RICH1) and downstream (RICH2) of the magnet. These stations are filled with a radiative material of refractive index n that emits a ring of Cherenkov light whenever a high-energy particle traverses them. All photons are emitted at an angle

$$\theta_c = \frac{1}{n\beta},\tag{2.1}$$

where β is the ratio between the particle velocity and the speed of light. A precise measurement of this ratio and of the particle momentum (performed by the tracking system in the case of charged tracks) allows to extract the mass of the particle, and then to identify it.

The choice of the refractive index of the material determines the momentum range in which the detector efficiently determines the mass of the particle. Figure 2.16 shows

39
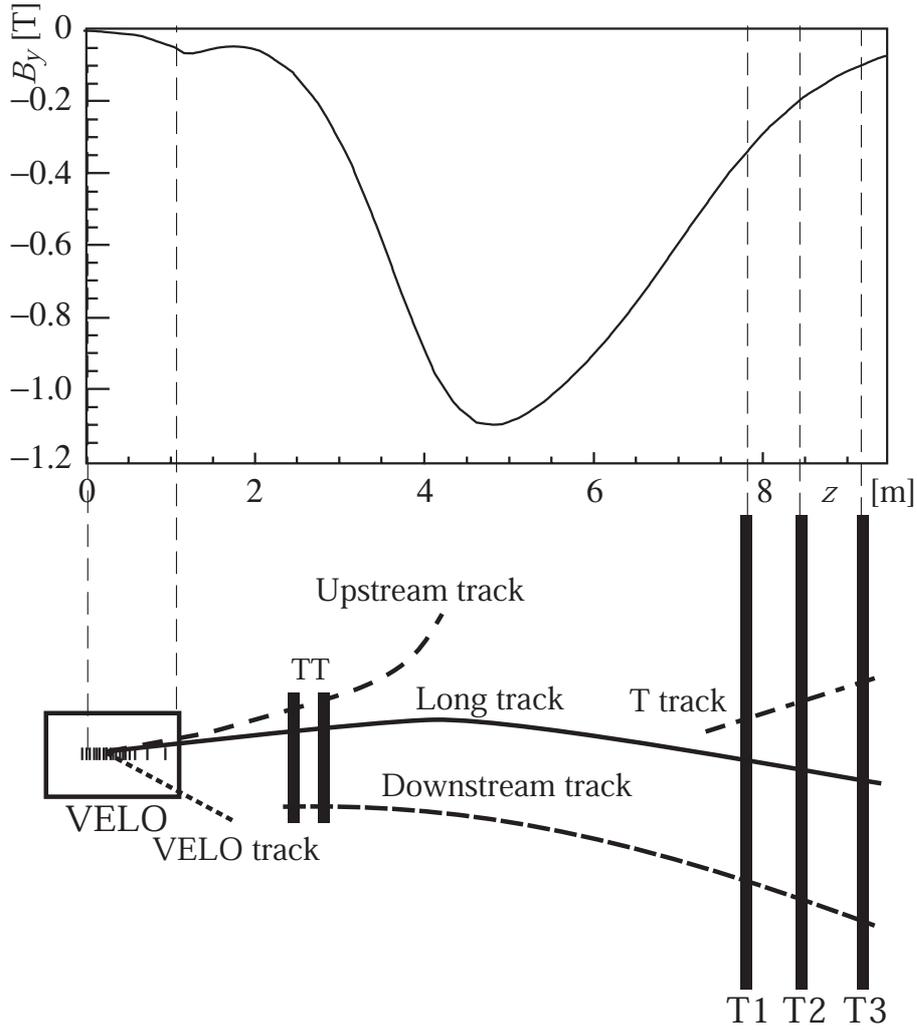
Figure 2.15 – Types of track in LHCb, along with the magnetic field as a function of z [58]. Muon tracks that only leave hits in the muon stations are not shown.

the Cherenkov angle as a function of momentum for two different refractive indexes. The RICH1 subdetector uses two different refractors, SO_2 (n=1.03) and C_4F_{10} (n=1.0014). The RICH2 subdetector uses only CF_4 (n=1.0005) as a refractor.

The full coverage of the nominal momentum range 2–100 GeV/c is achieved through the use of different technologies in the RICH1 and the RICH2. Upstream from the magnet, low-momentum particles associated to a large angular aperture are covered by the RICH1 detector momentum acceptance 2–60 GeV/c. Particles that have a larger momentum or a smaller aperture are covered by the RICH2 momentum acceptance of 15–100 GeV/c. Different momentum ranges correspond to the choice of different refractive indexes. While RICH1 covers the LHCb tracking acceptance, RICH2 has a reduced angular acceptance of 120 mrad (horizontal) and 100 mrad (vertical), as it is dedicated to the PID of particles
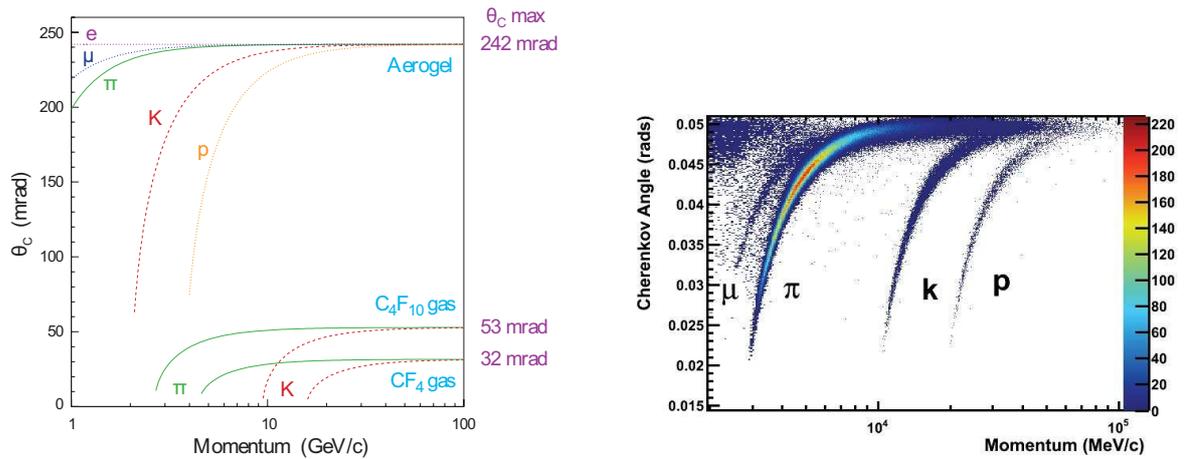
Figure 2.16 – (Left) Cherenkov angle as a function of the particle momentum for different radiators and particles. (Right) Reconstructed Cherenkov angle for isolated tracks, as a function of track momentum in the C_4F_{10} radiator [59]. The Cherenkov bands for muons, pions, kaons and protons are clearly visible.

with high momenta, which are mainly at small angles.

Figure 2.17 shows the layout of the two RICH stations. Particles pass through the middle of the detector, and emits Cherenkov light that is reflected by the mirrors located on the sides. The photons are finally collected by hybrid photo-detectors (HPDs) located outside of the LHCb acceptance.

The RICH system provides good particle identification over the entire momentum range. The average efficiency for kaon identification for momenta in the 2–100 GeV/c interval is 95%, with a corresponding average pion misidentification rate of 5%. Around 30 GeV/c the identification probability is close to 97% and the misidentification probability roughly 5%.

2.2.5 Calorimeters

The LHCb calorimeter is located downstream of the T1–3 and RICH2 stations, and consists of the electromagnetic calorimeter ECAL and the hadronic calorimeter HCAL. It provides information about the energy and the position of all particles, including neutral particles (π^0,γ) that do not leave a trace until that point. Two additional systems, PS and SPD, are dedicated to the detection of neutral particles. This strategy is designed in order to separate electrons and pions (charged or neutral), which requires a longitudinal separation of the showers. The electromagnetic calorimeter (ECAL) is designed to stop electrons and photons, and the hadronic calorimeter (HCAL) is designed to stop hadrons.

Calorimeter systems perform a *destructive* detection of the particle. This detection is performed by converting the energy of the incoming particle in a shower of particles that excite a radiator medium, whose nature depends on the type of particles that are detected.

41
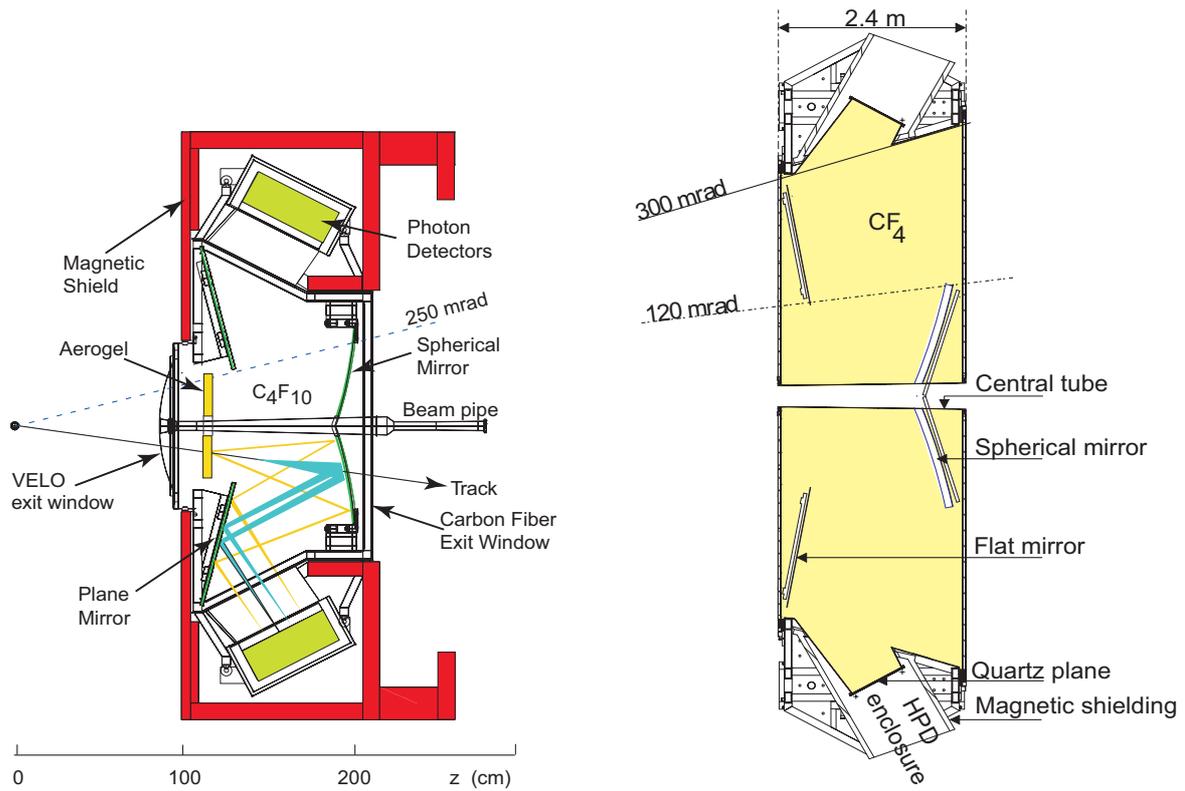
Figure 2.17 – Schematic cross section of the RICH1 (left) and RICH2 (right) detectors. [60]

This radiator medium then emits UV photons that are collected by the detector. The total amount of collected light is proportional to the energy of the incoming particle.

The SPD and the PS

The PS and SPD consist of two identical planes of scintillator pads with a 15 mm thick lead plane in between that corresponds to $2.5X_0$ for electrons and photons, but only to 6% hadronic interaction lengths. This allows to gain information about the nature of incoming particles, in particular on the γ/e separation. Figure 2.18 shows the typical longitudinal shower profile for different kind of particles.

Additionally, the SPD hit multiplicity information is used in the hardware trigger as it is correlated to the multiplicity of the event.

The electromagnetic calorimeter ECAL

The electromagnetic calorimeter is responsible for measuring the energy of incoming electrons and photons. It is a shasklik-type sampling calorimeter of thickness $25X_0$, composed of 66 layers of lead plates and scintillating tiles; the scintillating light is collected by photomultipliers. The cell size varies from 4×4 cm in the inner part of the detector, to 6×6 cm
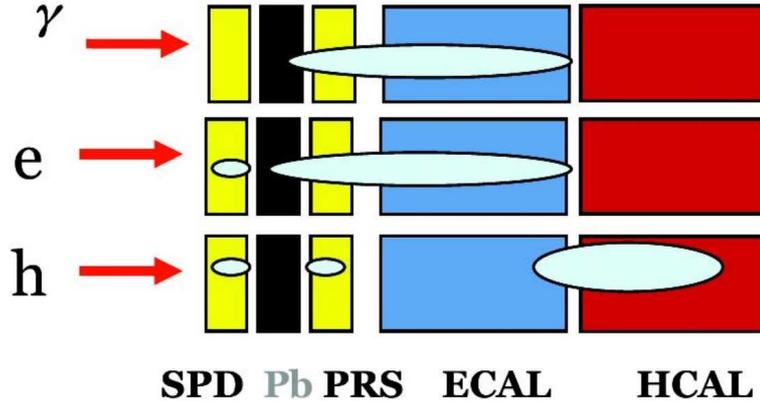
Figure 2.18 – Sketch of the typical repartition of showers depending on the nature of the incoming particle. "Pb" refers to a lead plate that converts photons to cascades of charged particles.
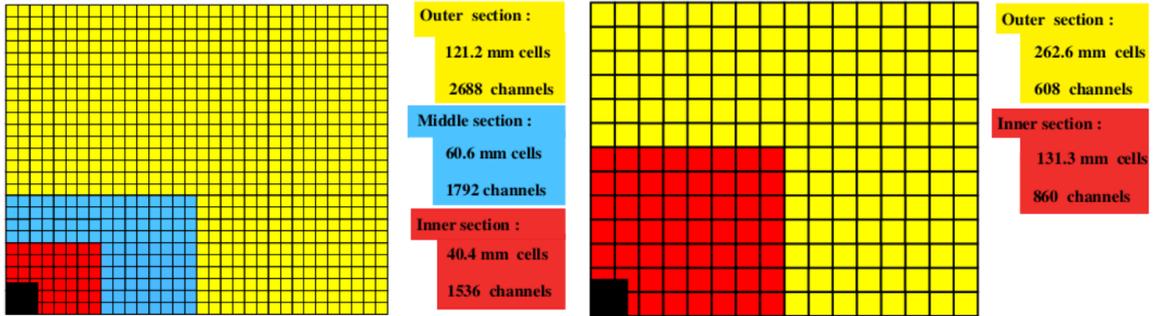
Figure 2.19 – (Left) Segmentation of the ECAL. (Right) Segmentation of the HCAL. The beam pipe region is filled with black, and is outside of the acceptance of both calorimeters.

and 12×12 cm in the middle and outer parts. Figure 2.19 shows these different segmentation schemes. The cell granularity corresponds to that of the SPD and the PS, aiming at a combined use in γ/e separation.

The hadronic calorimeter HCAL

The hadronic calorimeter is designed to absorb the entire energy of incoming hadrons. It is organized as a succession of 26 layers of thin iron plates and scintillating tiles arranged parallel to the beam pipe. Figure 2.19 shows the segmentation of the two calorimeters.

2.2.6   Muon chamber

The five muon chambers are responsible for the identification of muons and for providing a fast-response detection of high-p_T muons in the trigger system. Four of them consist of
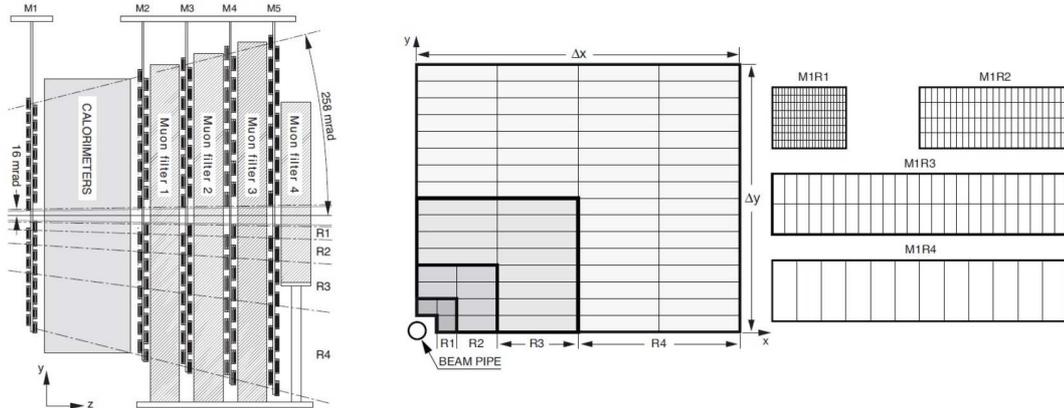
Figure 2.20 – Schematic view of the (left) side of the LHCb muon detector and the (right) two mechanically independent station halves with the four regions (R1-R4) indicated.

multi-wire projection chambers (MWPC). The first muon chamber is located upstream from the calorimeter systems in order to improve the resolution of the transverse momentum of muons, as calorimeters induce multiple scattering. It is equipped with a triple gas electron multiplier (GEM) that is more resistant to the increased radiation in this region. The other muon chambers (M2–5) are located downstream from the calorimeter systems. A 80 cm-thick layer of iron absorber is placed in front of each of these chambers in order to reduce backgrounds, and another is placed in front of M5 for the same purposes.

After the hadronic calorimeter, most hadrons and electromagnetic (e, γ) particles have been absorbed. Inversely, muons mostly pass through the whole detector without being absorbed, due to their low rate of energy loss dE/dx. Hence, the matching of a track to a deposit in the muon chambers increases its probability to be associated with a muon. This information is used in particle identification.

Figure 2.20 shows the layout of the muon chambers. The segmentation of the readout is finer in the regions near the beam axis, as these correspond to higher occupancies.

The information from muon chambers is used as a veto in the analyses described in further parts of the document. Indeed, the muon identification provided by the information from these chambers allows to veto out muon misidentification with a large efficiency.

2.3   Particle identification in LHCb

The identification of particles in the LHCb experiment relies on information from most of its subdetectors, such as:

- Cherenkov radiation angle from the RICH (charged particles only);

- track measurement in the muon chambers (muons);

- deposited energy in the ECAL, associated with the track momentum (electrons);

- information from the PS and SPD (neutral particles).

Additionally, a cluster formed in the calorimeters that is not associated with a track can be attributed to a γ or a π^0. As all analyses described in this document use charged pions and kaons as final-state particles, we only describe in the following the techniques relative to these particles.

Two approaches are used in LHCb for the identification of particles. The first method, named `DLL`, computes the difference of likelihood between a mass hypothesis and the pion hypothesis for each subsystem, and combines them linearly. The second method, named `ProbNN`, uses information from all subdetectors as inputs to a multivariate method that outputs a single probability for each hypothesis. It takes the correlations between subdetectors responses into account, as well as additional information. The training of this multivariate method is performed on inclusive simulated B events. Its performance depends on the blending of MC samples used (tune).

2.4   Trigger system in the LHCb experiment

The LHC accelerator operates at a bunch-crossing frequency of 40 MHz. Due to the lower luminosity settings at the LHCb interaction point and to the detector geometry, the rate of visible interactions was 15 MHz for 2012 data-taking conditions.[6] At a luminosity of 2×10^{32} the bunch crossings with visible pp interactions are expected to contain a rate of about 100 kHz of $b\bar{b}$-pairs. However, only about 15% of these events will include at least one B meson with all its decay products contained in the spectrometer acceptance. Furthermore the branching ratios of interesting B meson decays used to study for instance CP violation are typically less than 10^{-3}. The role of the LHCb trigger system is to reduce the rate down to 5 kHz while enriching the samples with events that are interesting for LHCb analyses. The trigger is also required not to bias interesting observables too much, which is especially challenging in the case of particle lifetimes.

Figure 2.21 shows the overall structure of the LHCb trigger system during Run I, along with the rates associated to each level. The trigger is divided in two levels: the hardware trigger (level-0 trigger or L0), and the software trigger (high-level trigger or HLT). The structure of the trigger system in LHCb has been overhauled for Run II, with the suppression of the hardware trigger.

2.4.1   The hardware trigger (L0 trigger)

The purpose of the L0 trigger is to reduce the LHC beam crossing rate of 40 MHz to the rate of 1 MHz with which the entire detector can be read out. The logic of the L0 trigger takes advantage of the fact that the dominant source of transverse momentum and energy

---

[6]An interaction is defined to be visible if it produces at least two charged particles with sufficient hits in the VELO and T1–T3 to allow them to be reconstructible.
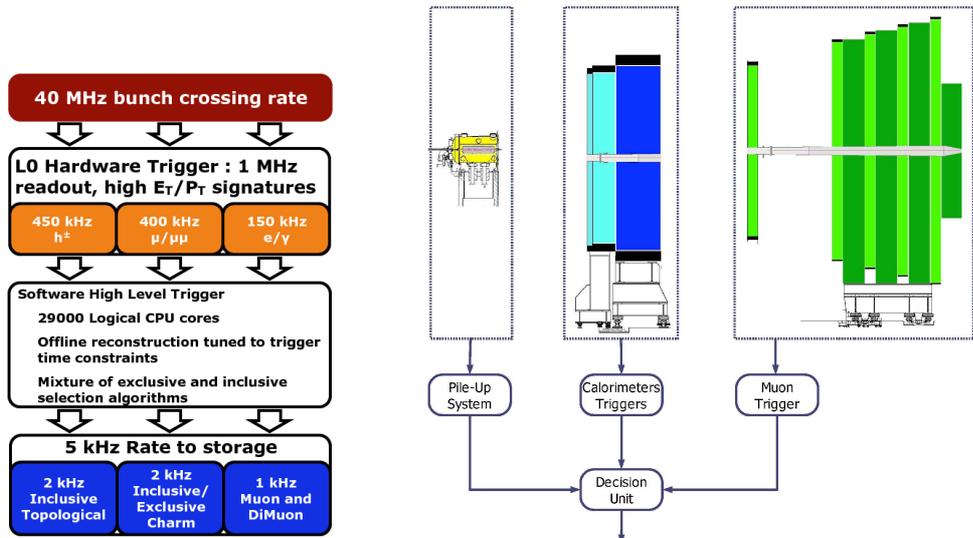
Figure 2.21 – (Left) Overview of the LHCb trigger system in 2012, along with the allowed rates for each line. (Right) Detector subsystems involved in the L0 trigger decision.

in the LHCb acceptance comes from the decays of heavy particles and resonances. A partial read-out of the muon chamber information and of the calorimeter information is processed in order to estimate the transverse momentum of single/di-muon candidates and the highest E_T cluster, respectively. Both quantities are then compared to a pre-defined threshold. The rates of each decision line in 2012 data-taking conditions are shown in Fig. 2.21, along with a sketch showing the subdetectors involved in the decision.

In parallel with this filtering, the information from the SPD is also read in order to estimate the total number of tracks in the event, respectively. This allows to veto events that would have too many tracks and that would be triggered due to large combinatorics. These events would also occupy a disproportionate fraction of the data-flow bandwidth or available processing power in the HLT.

The L0 trigger is operated synchronously with the 40 MHz bunch-crossing frequency, using custom-made electronics.

2.4.2 The software trigger (HLT trigger)

The High Level Trigger (HLT) is designed as a series of C++ algorithms that reduces the output rate to approximately 3.5 kHz and 5 kHz in 2011 and 2012, respectively. This corresponds to the nominal event rate for being permanently stored. The HLT architecture is divided in two stages: fast partial event reconstruction with an inclusive selection (HLT1) in order to reduce the rate to 40 and 80 kHz for 2011 and 2012, respectively; complete event reconstruction with final trigger selection (HLT2).

With the additional information available, the strategy of a single track trigger is implemented in HLT1 using information on the quality of the track and the displacement

from the primary vertex. Further improvements on the tracking search are obtained by reducing the phase-space boundaries to consider only tracks with transverse momentum above the required threshold conditions. In particular, an inclusive approach for beauty decays has been designed, which comprises a large fraction of the output bandwidth. The strategy of this approach is to select a high transverse momentum, significantly displaced track, and a significantly displaced vertex containing this track and 1–3 other tracks [61]. This design triggers efficiently on B decays with at least two charged daughters.

2.4.3 Trigger conventions

Another important consideration is the association of a trigger object with a signal track. An event is classified as trigger-on-signal (TOS) if the signal under study triggers the event, whereas trigger-independent-of-signal (TIS) categorizes the trigger objects not associated to the signal. This separation is especially relevant when the main trigger line for an analysis introduces a bias on the variables of interest that has to be studied.

2.5 Monte-Carlo simulations in LHCb

The simulation of a physics event in LHCb is divided in several phases, integrated in the GAUDI framework. Firstly, the underlying physical event and its interaction with the LHCb detector is simulated inside the GAUSS framework. The digitization of hits in the subdetectors and the building of the raw dataset is then modelled by the BOOLE package. The reconstruction of tracks from this raw dataset is then modelled by the BRUNEL package. Finally, the DAVINCI package simulates the further offline analysis steps, such as the building of physical variables from tracks. Figure 2.22 shows the data flow of simulated events in LHCb.

In this section, I focus on the GAUSS and DAVINCI parts of the framework. I also detail several sources of data/MC discrepancies that are relevant to the analyses discussed in this document.

2.5.1 The GAUSS framework

The GAUSS package simulates pp collisions in LHCb and the detector response to the products of the collision. It operates in two phases that can be run sequentially or independently.

The first phase consists of the event generation of pp collisions and the decay of the B-mesons in channels of interest for the LHCb physics programme. It is interfaced to PYTHIA for the event production and to a specialized decay package, EVTGEN, for the B-meson decay. The generator phase of GAUSS also handles the simulation of the running conditions, the smearing of the interaction region due to the transverse and longitudinal sizes of the proton bunches and the change of luminosity during a fill due to the finite beam lifetime. Single and multiple pp collisions are produced according to the chosen running luminosity. Other event generator engines can be interfaced in this phase if required.
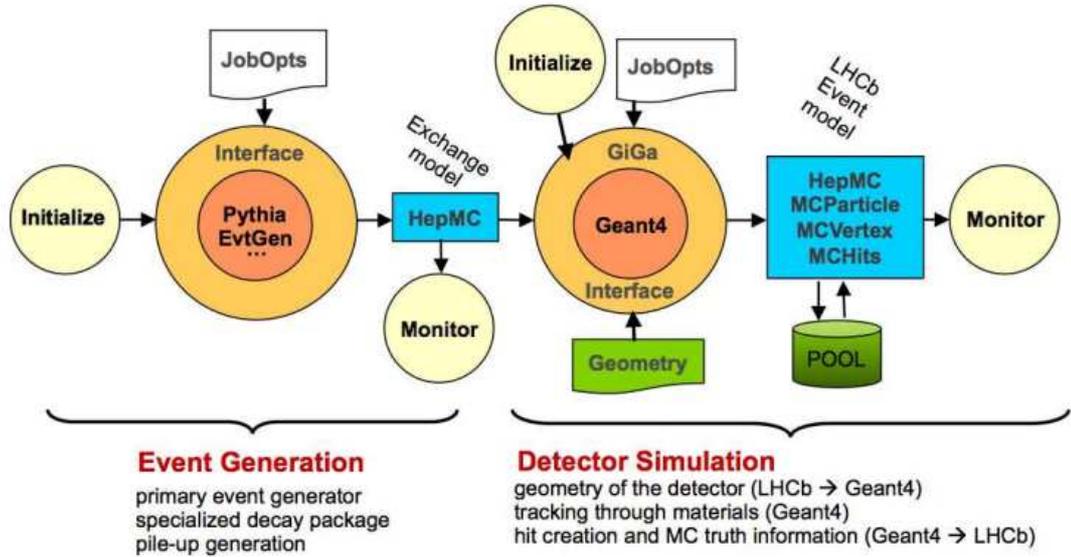
Figure 2.22 – The LHCb data processing applications and data flow. Underlying all of the applications is the Gaudi framework and the event model describes the data expected. The arrows represent input/output data

The second phase of GAUSS consists in the propagation in the LHCb detector of the particles produced by the generator phase. The simulation of the physics processes that the particles undergo when traveling through the experimental setup is delegated to the GEANT4 toolkit. The behaviour of the GEANT4 simulation engine in terms of detectors to simulate, physics models to use, details of the Monte-Carlo truth to be provided, is controlled at run time via job options configuration.

The behaviour of the generator phase and of the detector response is regularly studied using reference decay channels for which LHC disposes of a large high-purity dataset. The correction of some variables in the Monte-Carlo production in order to match data better is called "tuning". It is especially relevant to the generator phase, where the output of PYTHIA and other hadronization tools is closely scrutinized.

As the second phase is the most CPU-intensive, it is possible to specify a set of requirements ("generator-level cuts") to the first phase, in order to veto out events that have no chance to be reconstructed by LHCb, for instance, a signal event with a charged track outside of the LHCb acceptance.

2.5.2 The DAVINCI framework

The DaVinci package manages the creation of physical objects such as tracks from the output of GAUSS or from the detector response to real data-taking. Additionally, DaVinci contains tools to tag the flavour of particles, or to refit events taking constraints such as masses or vertices into account. This allows, for instance, to constrain the masses of all
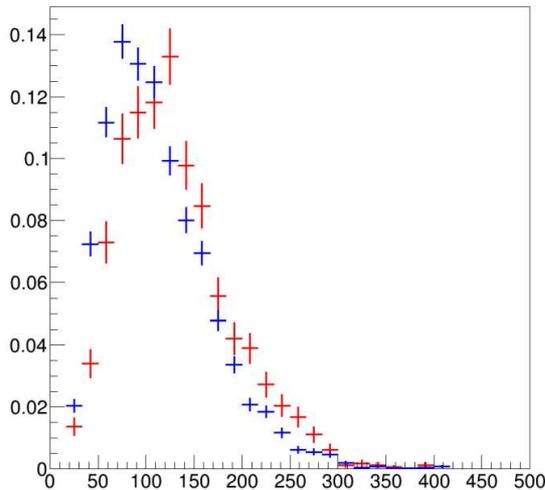
Figure 2.23 – Distributions of the number of tracks in $B^0 \to K_S^0 \pi^+ \pi^-$ events. Simulated events are represented in blue, while signal events from data are represented in red.

particles to be the "true" mass when considering a Dalitz plot.

The DaVinci package is implemented in such a way that Monte-Carlo productions and real data are treated the same way.

2.5.3 Data/MC discrepancies

Despite regular tuning and an overall excellent performance, the LHCb Monte-Carlo production does not match data perfectly, for several reasons. I describe in the following several sources of data/MC discrepancies that are relevant to the analyses described in this dissertation.

Our understanding of strong interaction and hadronic physics is limited, and the simulation of the underlying event is only an approximation. As a result, the kinematical spectrum of the products of pp interactions and the number of these products is different in Monte-Carlo and data. The response of several subdetectors, such as the RICH and the calorimeters, is correlated to the track multiplicity in the event. Figure 2.23 shows the different distributions in track multiplicity of events in data and Monte-Carlo. The track-finding efficiency of LHCb is correlated to momenta and to the track multiplicity, and so differs between data and simulation. The kinematical dependency of that difference means that it depends a priori on the Dalitz plot.

As mentioned in Sec. 2.2.4, performances of the RICH detectors depend on the refractive index n of their radiator. This index is highly sensitive to temperature and pressure changes through the whole year, a change that is impossible to match perfectly in Monte-Carlo. Additionally, the performances of these subdetectors depend on the track multiplicity in the event and on the kinematics of particles. The `PIDCalib` tool is used to reweight Monte-Carlo productions to match the efficiency of a given selection on PID

variables. The weights are calculated using a reference data sample and the dependence on track multiplicity and the kinematics of the given particle is taken into account.

Chapter 3

Fast Monte-Carlo method for background studies

In this chapter, I describe a fast simulation method that I have developed to model the partially reconstructed backgrounds in $B^0_{d,s} \to K^0_S h^+ h^-$ modes. I present the strategy of the method in Sec. 3.1, and then present the study of a fully simulated sample of one of these backgrounds in Sec. 3.2. I show the results of exporting this study to another sample of partially reconstructed background in Sec. 3.3. I discuss in Sec. 3.4 the modelling of acceptance effects by means of selection criteria, and finally present the results of a full fast MC simulation in Sec. 3.5.

3.1 Strategy of the fast MC method

Partially reconstructed background in $B^0_{d,s} \to K^0_S h^+ h^-$ modes consists of events such as $X \to K^0_S h^+ h^- Y$, where Y is not being reconstructed (*e.g.* it is soft or out of the acceptance). It can originate from a variety of channels with different mother particles, missed particles, or intermediate resonances.

Studying and modelling the partially reconstructed background is usually done by generating large samples of fully reconstructed Monte-Carlo (MC) events, which is CPU-consuming. As described in Sec. 2.5, the simulation of the detector is the most expensive part of the generation of simulation samples in terms of CPU. We thus aim at modelling the effects of the detector on the distribution of invariant masses without simulating the whole detector.

The detector affects the distributions of physical variables because of its finite resolution, and because of its finite acceptance. The proposed fast Monte-Carlo method consists of smearing generator-level variables event by event, to account for resolution effects on invariant masses, after applying some requirements on the generator-level distributions in order to account for the acceptance effects.

For most kinematic variables, the distributions of variables at generator-level and reconstruction-level are barely distinguishable. We study the resolution of a variable X by means of the distribution of
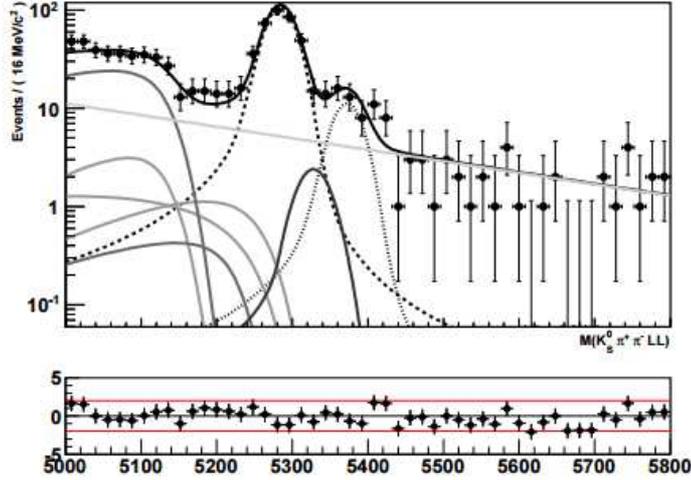
Figure 3.1 – Mass spectrum of $B^0_{d,s} \to K^0_S \pi^+ \pi^-$ (LHCb data) from [52]. The partially reconstructed background can be seen to the left of the $B_{d,s}$ signal peaks.

$$\Delta_X = X - X_{\text{true}}. \tag{3.1}$$

In the case of a variable at reconstruction level X_{rec}, the resolution quantifies the detector effect on the variable. We attempt to obtain a choice of variables to smear with a parametric function modelling their resolutions, in a way that ensures that resolution effects on all the invariant masses are fairly well taken into account.

For obvious reasons, we prefer having a set of variables for which the resolutions are not correlated. Resolutions of the components p_x, p_y and p_z of the three final-state-particles momenta do not have these properties (x, y, and z are the usual LHCb coordinates defined in Sec. 2.2). We therefore use the resolutions of $1/p_z$, the polar angle θ, and the azimuthal angle ϕ, along with the resolution of $m_{K^0_S}$. Once we extract the resolution distributions, we use them to smear the corresponding variables, and compute the resulting invariant masses.

We investigate the results of extracting resolution functions from one channel, and applying them on another. Indeed, the partially reconstructed background of $B^0_{d,s} \to K^0_S h^+ h'^-$ is composed of channels of various types. They can differ by their topologies, or by the missing particle. For instance, the following modes contribute to the partially reconstructed background of $B^0_{d,s} \to K^0_S \pi^+ \pi^-$:

- $B^0 \to (K^{*0} \to K^0_S \pi^0)(\rho^0 \to \pi^+ \pi^-)$, P→VV topology, missing π^0 (massive calorimetric object);[1]

---

[1] The P,V, and S letters stand for pseudo-scalar, vector, and scalar, respectively.

- $B^0 \to K_S^0(\eta \to \pi^+\pi^-\pi^0)$, P $\to$ PS, S$\to$PPP topology, missing π^0;

- $B^0 \to K_S^0\pi^+\pi^-\gamma$, nonresonant, missing γ (massless calorimetric object);

- $B^0 \to K_S^0(\eta' \to \pi^+\pi^-\gamma)$, resonant, missing γ;

- $B^0 \to (K^{*0} \to K_S^0\pi^+)\pi^+\pi^-$, P$\to$VPP topology, missing π^+ (massive track).

The invariant-mass distribution of each of these modes is modelled by an ARGUS distribution convoluted with a Gaussian. The ARGUS distribution is parameterized as

$$f(m; c, s, m_0) = \mathcal{N}\frac{m}{m_0}\left(1 - \left(\frac{m}{m_0}\right)^2\right)^c . e^{-\frac{1}{2}s^2\left(1 - \frac{m^2}{m_0^2}\right)}, \tag{3.2}$$

where the parameters c, s, and m_0 are the curvature, slope, and threshold mass, respectively, and $\mathcal{N}$ is a normalization factor.

3.2 Study of a $B^0 \to (K^{*0} \to K_S^0\pi^0)(\rho^0 \to \pi^+\pi^-)$ sample

We extract the resolutions from a fully simulated sample of roughly 20,000 Monte-Carlo events of $B^0 \to (K^{*0} \to K_S^0\pi^0)(\rho^0 \to \pi^+\pi^-)$. This section only presents the study of events with Down-DownK_S^0 reconstruction, as defined in Sec.2.2.3. The results also hold for events with a Long-Long K_S^0 reconstruction.

For each particle in each event, we extract Δ_θ, Δ_ϕ, and Δ_{1/p_z}, along with $\Delta_{m_{K_S^0}}$, and fit them with analytic functions. These are then used to smear the generator-level variables, to obtain fast Monte-Carlo distributions. These distributions are not complete fast Monte-Carlo distributions, as we apply here the smearing procedure on the generator level of a fully reconstructed sample. We control the results obtained by comparing the fast Monte-Carlo and reconstruction-level distributions of invariant masses. We also compare the distributions of Δ_X for the reconstructed sample X_{rec} and for the fast Monte-Carlo sample X_{fastMC}.

Figure 3.2 shows the distributions of Δ_θ, Δ_ϕ, and Δ_{1/p_z} for the π^+ with respect to its p_z momentum. The three variables show a strong dependence on p_z. We account for this dependence by handling twelve intervals of p_z that contain roughly the same number of events, and fitting Δ_θ, Δ_ϕ, and Δ_{1/p_z} in each of them. Figure 3.2 shows the distribution of $\Delta_{m_{K_S^0}}$ along the $p_z(K_S^0)$ axis. The resolution of $m(K_S^0)$ depends on the momentum of the K_S^0, but this dependence is ignored in the following as it is smaller than the dependence of other resolutions with respect to the momentum.

We model the distributions of Δ_θ, Δ_ϕ, and Δ_{1/p_z} by a sum of two Gaussians in each interval. Figure 3.3 shows the results of a fit to the resolution distributions obtained for π^+ in a single p_z interval. The fit is overall satisfactory.

To generate fast Monte-Carlo distributions, we proceed as following:
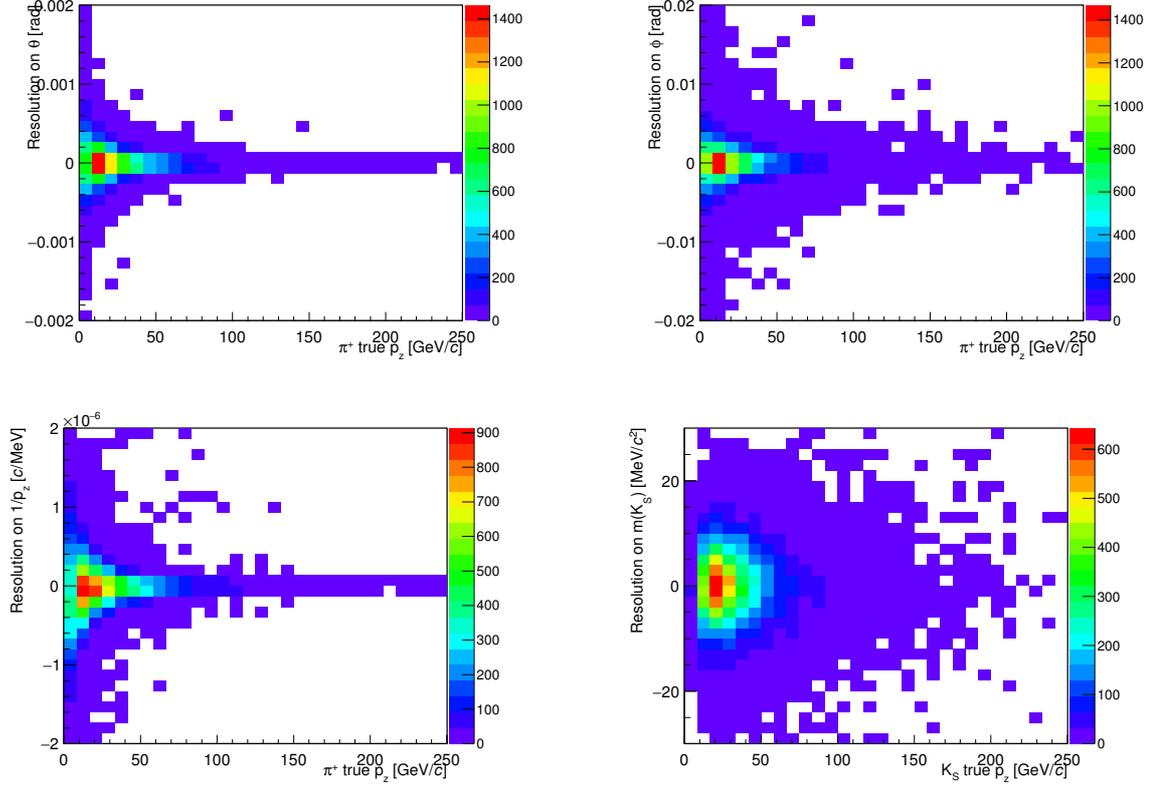
Figure 3.2 – Distributions versus p_z of Δ_θ (top, left), Δ_ϕ (top, right), and Δ_{1/p_z} (bottom, left) for π^+, and distribution of $\Delta_{m_{K_S^0}}$ versus p_z (bottom, right).
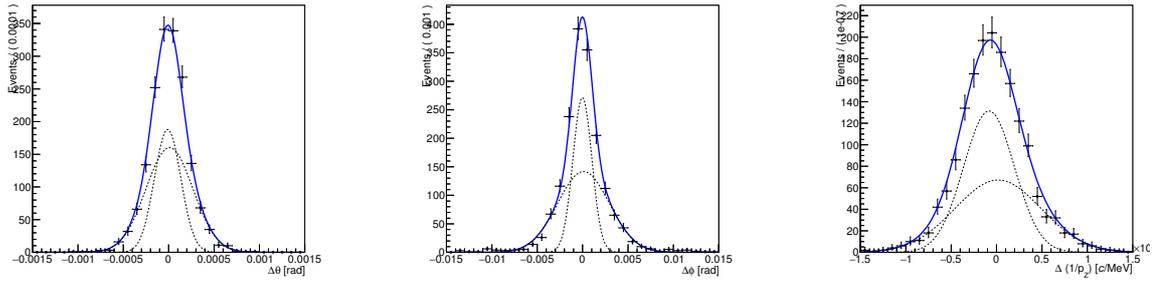
Figure 3.3 – Fit results for p_z (π^+) between 8.5 GeV/c^2 and 11.5 GeV/c^2. Left: $\Delta_\theta(\pi^+)$. Middle: $\Delta_\phi(\pi^+)$. Right: $\Delta_{1/p_z}(\pi^+)$.

- for each event i, generate a random value according to the PDF of $\Delta_{m_{K_S^0}}$. We

obtain $m^i_{K^0_\mathrm{S},\mathrm{fastMC}} = m^i_{K^0_\mathrm{S},\mathrm{true}} + \Delta m_{K^0_\mathrm{S}}{}^i$;

- for each particle, determine in which p_z interval it lies, and generate a random value according to the corresponding PDFs to get $\Delta_\theta{}^i$, $\Delta_\phi{}^i$, and $\Delta_{1/p_z}{}^i$; we then obtain $\theta^i_\mathrm{fastMC} = \theta^i_\mathrm{true} + \Delta_\theta{}^i$, $\phi^i_\mathrm{fastMC} = \phi^i_\mathrm{true} + \Delta_\phi{}^i$, and $1/p^i_{z,\mathrm{fastMC}} = 1/p^i_{z,\mathrm{true}} + \Delta_{1/p_z}{}^i$;

- deduce the fast MC values of momenta, $m_{K^0_\mathrm{S}h^+h'}$, $m_{K^0_\mathrm{S}h}$, $m_{K^0_\mathrm{S}h'}$, and $m_{hh'}$.

Figure 3.4 shows a comparison of the reconstructed and fast Monte-Carlo distributions of $m_{K^0_\mathrm{S}h^+h'}$. These do not agree perfectly well, but given the small amount of partially reconstructed background events in our modes of interest, it is good enough for our purpose. In the same figure, we also compare the reconstructed and fast Monte-Carlo resolution for $m_{K^0_\mathrm{S}h^+h'}$. The shape is sensibly the same, but the fast Monte-Carlo distribution is slightly narrower. Possible explanations include imperfections of the used resolution functions or missed correlations between variables.

Figure 3.5 shows the distributions of background events over the Dalitz plane in reconstruction level and fast MC. It also shows the distribution of the difference between the two former distributions, divided by the standard error on the difference. In the following, this distribution is referred to as the distribution of pulls between reconstructed and fast Monte-Carlo distributions. These pulls are small and show no overall structure, thus showing that the agreement between the reconstructed level and the fast Monte-Carlo is rather good.
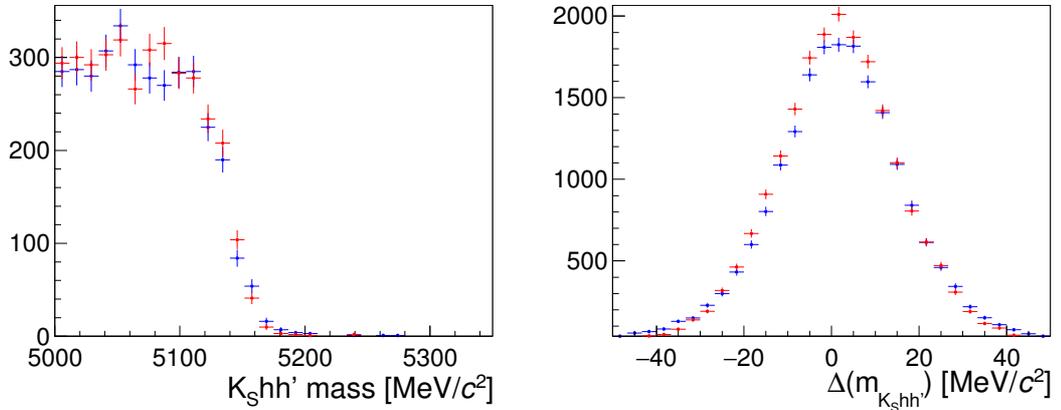
Figure 3.4 – Left: Distributions of $m_{K^0_\mathrm{S}h^+h'}$ for reconstructed (blue) and fast MC (red) $B^0 \to K^{*0}\rho^0$ events. Right: Resolutions of $m_{K^0_\mathrm{S}h^+h'}$ for reconstructed (blue) and fast MC (red) $B^0 \to K^{*0}\rho^0$ events.
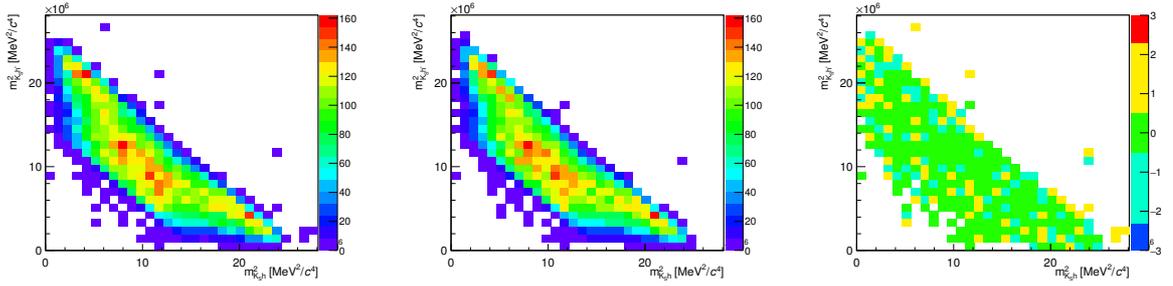
Figure 3.5 – Distributions of $m^2_{K^0_S h}$ versus $m^2_{K^0_S h'}$ for the $B^0 \to K^{*0}\rho^0$ mode, with DD reconstruction of the K^0_S. Left: Reconstructed events. Middle: Fast MC events. Right: Pulls between the two distributions.

3.3 Study of the resolution model applied to other channels

We apply the resolution functions extracted from $B^0 \to K^{*0}\rho^0$ decays, using the same procedure and p_z intervals as in Sec. 3.2, to a variety of partially reconstructed samples. The fully simulated and fast Monte-Carlo distributions of $m_{K^0_S h + h'}$ and $\Delta_{m_{K^0_S h + h'}}$ are then compared, as well as the distributions of events over the Dalitz plane. As in Sec. 3.2, this is not a complete fast Monte-Carlo distribution, as we still use the generator level of a fully reconstructed sample.

In this section I discuss the results of this procedure applied to a sample of $B^0 \to K^0_S(\eta' \to \pi^+\pi^-\gamma)$ decays. This channel appears as a background in the $K^0_S\pi^+\pi^-$ spectrum, and the missed particle is massless, which makes this contribution dangerous to our analysis. In Annex C, I summarize the results of the same procedure applied to other channels of partially reconstructed decays.

Figure 3.6 shows the distribution of $m_{K^0_S h + h'}$ near the threshold for reconstructed and fast MC events, as well as the resolution distributions for $m_{K^0_S h + h'}$. The distributions agree quite well for $m_{K^0_S h + h'}$, with a well-reproduced behaviour at the threshold. As before, the distribution of $\Delta_{m_{K^0_S h + h'}}$ is slightly narrower in the fast MC case. Figure 3.7 shows the distributions of events over the Dalitz plane for reconstructed and fast MC events, along with the distribution of the pulls between these two distributions. The distributions are similar, and the pulls are rather small and show no structure. Overall, the result we obtain with this channel are satisfactory.

3.4 Study of generator-level reconstruction effects

The generator-level distributions that we smeared in Sec. 3.2 and Sec. 3.3 are not those that are directly produced by the GAUSS generation. They contains events that pass the
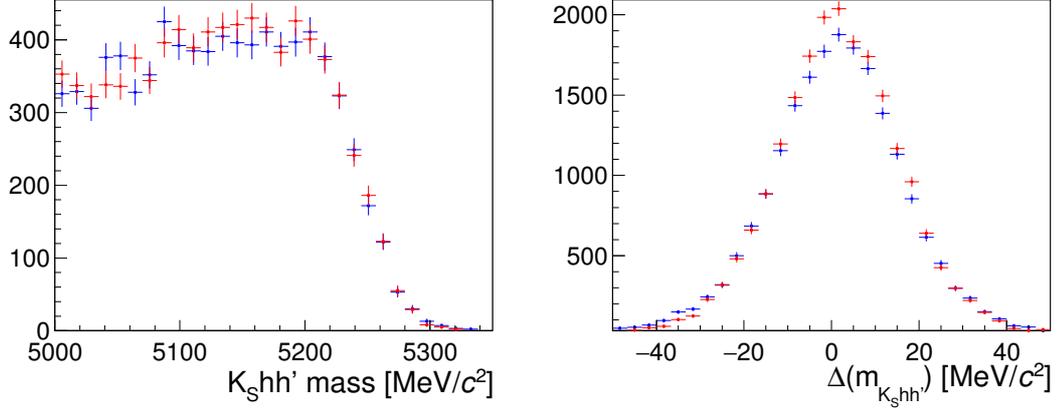
Figure 3.6 – Left: Distributions of $m_{K_S^0 h + h'}$ for reconstructed (blue) and fast MC (red) $B^0 \to K_S^0 \eta'$ events. Right: Resolutions of $m_{K_S^0 h + h'}$ for reconstructed (blue) and fast MC (red) $B^0 \to K_S^0 \eta'$ events.
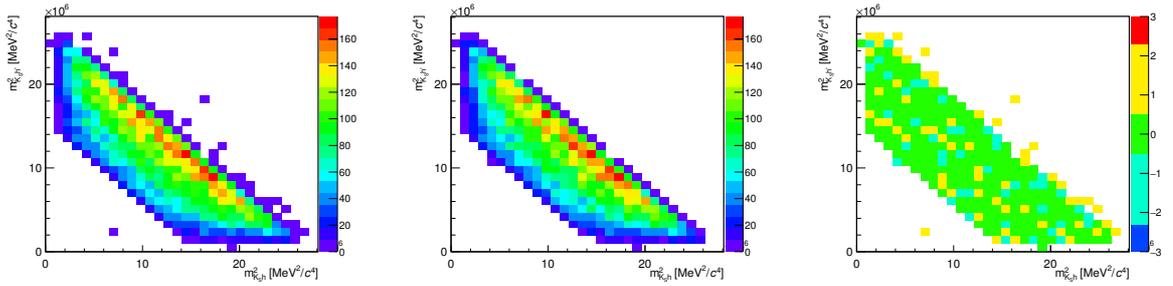
Figure 3.7 – Distributions of $m^2_{K_S^0 h}$ versus $m^2_{K_S^0 h'}$ for the $B^0 \to K_S^0 \eta'$ mode, with DD reconstruction of the K_S^0. Left: Reconstructed events. Middle: Fast MC events. Right: Pulls between the two distributions.

trigger, reconstruction, and a step called "stripping", all of which modify the generator-level distributions of invariant masses and momenta. In this section I discuss the possibility of accounting for these effects by applying selection criteria on variables available at generator level. To do so, we use the same samples as in Sec. 3.2 and check the consistency of our procedure on the samples used in Sec. 3.3.

3.4.1 K_S^0 reconstruction mode

Firstly, we have to determine if a K_S^0 in the generator level would be reconstructed as Down-Down or as Long-Long. For this purpose, we study four variables: the z position of the K_S^0 end vertex, the radial coordinate r of this vertex in the LHCb usual coordinate

Table 3.1 – Selection criteria applied on events to assign them to Long-Long reconstruction.

A	$\min_z > 30\,\text{mm}$ & $z(K_\mathrm{S}^0\text{ end vertex}) < 650\,\text{mm}$
B	$r(K_\mathrm{S}^0\text{ end vertex}) < 35\,\text{mm}$
C	$z(K_\mathrm{S}^0\text{ end vertex}) < 250\,\text{mm}$
D	$\theta(K_\mathrm{S}^0) < 0.07\,\text{rad}$
E	$z(K_\mathrm{S}^0) < 300\,\text{mm} \;||\; (z(K_\mathrm{S}^0) + 1.2{\times}\min_z) > 550\,\text{mm}$
Total cut	A & B & (C || (!C & D)) & E

Table 3.2 – Summary of DD-LL requirements on all channels under study. Efficiency is defined as the portion of events from the corresponding K_S^0 reconstruction mode that pass the criterion; power is defined as the proportion of events from the other K_S^0 reconstruction mode that do not pass the criteion.

Channel	Efficiency (%)	Power(%)
$B^0 \to K^{*0}\rho^0$	98.7	99.7
$B^0 \to K_\mathrm{S}^0\eta$	98.9	99.5
$B^0 \to K_\mathrm{S}^0\pi^+\pi^-\gamma$	98.4	99.7
$B^0 \to K_\mathrm{S}^0\eta'$	98.8	99.7
$B^+ \to K^{*+}\pi^+\pi^-$	98.0	99.6
$B^0 \to K^{*0}\phi$	98.5	99.7
$B^+ \to K^{*+}\phi$	98.6	99.8
$B_s^0 \to K^{*0}\phi$	98.4	99.6

system, the polar angle θ of the K_S^0, and a variable named $\min_z$. This variable accounts for the fact that a track can be reconstructed as Long only if it crosses at least three VELO stations (see Sec. 2.2.3). To design the selection criteria on these variables, we consider the VELO as a cylinder of radius $r = 35\,\text{mm}$.

Figure 3.8 shows the distribution of the K_S^0 end vertex, θ, and $\min_z$ for Down-Down and Long-Long reconstruction in fully simulated $B^0 \to K^{*0}\rho^0$ events. We can achieve a good separation between the two samples with rather simple requirements presented in Table 3.1.

These requirements are then tested on all other available channels, and the results are shown in Table 3.2. Here, the efficiency is defined as the percentage of Long-Long events that pass these requirements, and the power is defined as the percentage of Down-Down events that do not pass these requirements. Both the efficiency and the power are consistently high for all channels, and do not vary significantly between them.
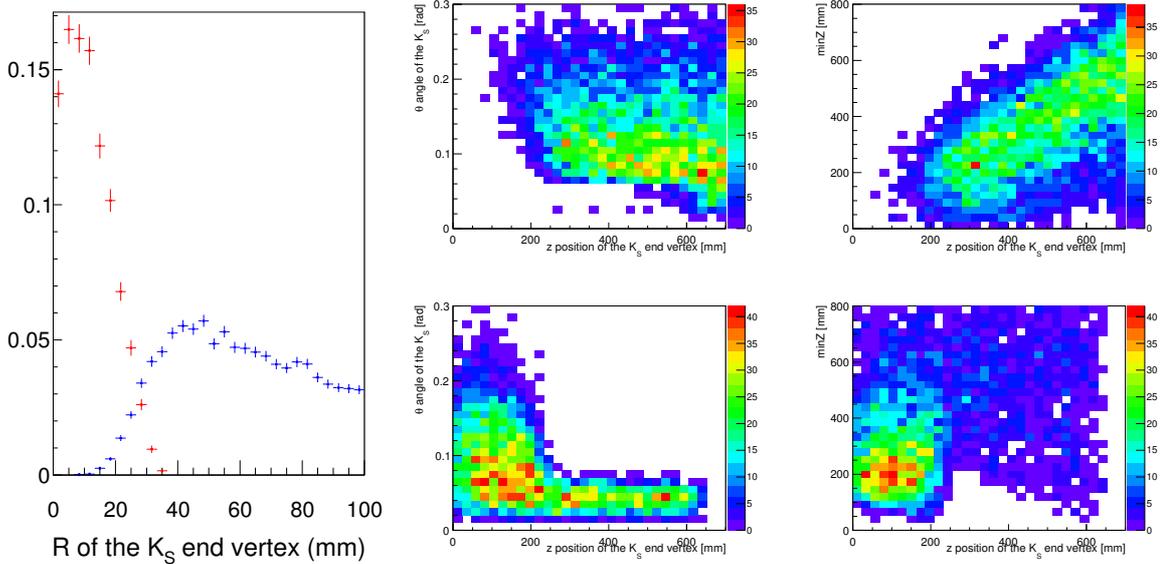
Figure 3.8 – Left: Distributions of the r coordinate of the K_S^0 end vertex for Down-Down events (blue) and Long-Long events (red), in the $r < 100\,\mathrm{mm}$ region. Middle: Distributions of the θ angle of the K_S^0 versus the K_S^0 end vertex z position in the $z < 700\,\mathrm{mm}$ region (top: DD, bottom: LL). Right: Distributions of $\min_z$ with respect to the z position of the K_S^0 end vertex (top: DD, bottom: LL), in the $z < 700\,\mathrm{mm}$ region.

3.4.2 Reproduction of the generator-level distributions

We generate with GAUSS 20,000 events of each of the following modes: $B^0 \to K^{*0}\rho^0$, $B^0 \to K_S^0\eta$, $B^0 \to K_S^0\eta'$, $B^0 \to K_S^0\pi^+\pi^-\gamma$, $B^0 \to K^{*0}\phi$, and $B^+ \to K^{*+}\phi$. The selection applied to these GAUSS samples are presented in Table 3.3. Most of them are described in Ref. [62], but some values are changed to obtain a higher efficiency of the fully simulated samples.[2] We also add requirements on $\theta(B)$ and on the B-meson flight distance. These do not affect the efficiency of fully simulated samples, and improve the agreement between $p_{z,\mathrm{GAUSS}}(B)$ and $p_{z,\mathrm{full}}(B)$.

Figure 3.9 shows a comparison between the distributions of $m_{K_S^0 h}$, $m_{K_S^0 h'}$, $m_{hh'}$, and $p_z(B)$ at generator level for fully simulated samples and our GAUSS samples of $B^0 \to K^{*0}\rho^0$ events. The invariant mass distributions are similar, while the $p_z(B)$ fully reconstructed distribution is shifted towards higher values. This could come from a χ^2 selection on the B vertex, as this variable is correlated with the momentum of the B meson. The momentum of the B meson affects our smearing procedure, as resolutions depend on p_z.

We account for this effect by reweighting the $p_z(B)$ distributions. Figure 3.10 shows the distributions of the weights calculated for all channels of which we generated a GAUSS

---

[2]This readjustment is necessary because the selection criteria quoted in Ref. [62] are applied to reconstruction-level variables, which are affected by the resolution.

Table 3.3 – List of generator-level selections

Down-Down	Long-Long
$p_z(B) > 25\,\mathrm{GeV}/c$	
$p_T(B) > 1.5\,\mathrm{GeV}/c$	
$\sum\limits_{\mathrm{daughters}} p_T > 3\,\mathrm{GeV}/c$	
$p_{T,daugh} > 800\ \mathrm{GeV}/c$ for at least 2 B-meson daughters	
$cos(DIRA_B) > 0.999$	
$p_T(h1) > 250\,\mathrm{GeV}/c$	
$p_T(h2) > 250\,\mathrm{GeV}/c$	
$p_{K^0_\mathrm{S}\mathrm{daughters}} > 2\,\mathrm{GeV}/c$	
$\mathrm{p}(K^0_\mathrm{S}) > 6\ \mathrm{GeV}/c$	n/a
$\theta(K^0_\mathrm{S}) > 0.01$ rad	
$\theta(K^0_\mathrm{S}) < 0.35$ rad	
$\theta(h^{(')}) > 0.01$ rad	
$\theta(h^{(')}) > 0.4$ rad	
$z_{K^0_\mathrm{S}\mathrm{endvertex}} < 2400\,\mathrm{mm}$	n/a
$z_{K^0_\mathrm{S}\mathrm{endvertex}} > 100\,\mathrm{mm}$	n/a
$R_{K^0_\mathrm{S}\mathrm{endvertex}} > 15\,\mathrm{mm}$	n/a
B-meson flight distance $> 1.5\,\mathrm{mm}$	B-meson flight distance $> 1\,\mathrm{mm}$

sample. For a given bin i and channel j, we define the weight $\mathrm{w_j}^i$ as following:

$$\mathrm{w_j}^i = \frac{\mathrm{N_{j,full}}^i}{\mathrm{N_{j,Gauss}}^i} \times \frac{\mathrm{N_{j,Gauss}}}{\mathrm{N_{j,full}}}, \text{with } \mathrm{N_{j,(Gauss,full)}} = \sum_i \mathrm{N^i_{j,(Gauss,full)}}. \tag{3.3}$$

Within uncertainties, all these weights are compatible. We can then reweight our generator-level distributions using weights from all our fully simulated samples.

3.5 Complete fast Monte-Carlo test on $B^0 \to K^{*0}\rho^0$

In this section I present a comparison between fast Monte-Carlo events and the reconstruction-level in fully reconstructed Monte-Carlo events. We generate 50,000 $B^0 \to K^{*0}\rho^0$ events using GAUSS first stage (see Sec. 2.5.1), with generator-level cuts on the production angle of the daughters, to ensure that they are in the LHCb acceptance. We also apply a selection on $m_{K^0_\mathrm{S}h^+h'}$, forced to be larger than $4800\,\mathrm{MeV}/c^2$. We apply the generator-level cuts described in Sec. 3.4.1 and in Table 3.3 to our generated sample; we apply weights as described in Sec. 3.4.2.

Prior to the smearing, we combine all the fully reconstructed samples used in Sec. 3.2 and Sec. 3.3 to extract new resolution functions. The larger number of events allows to
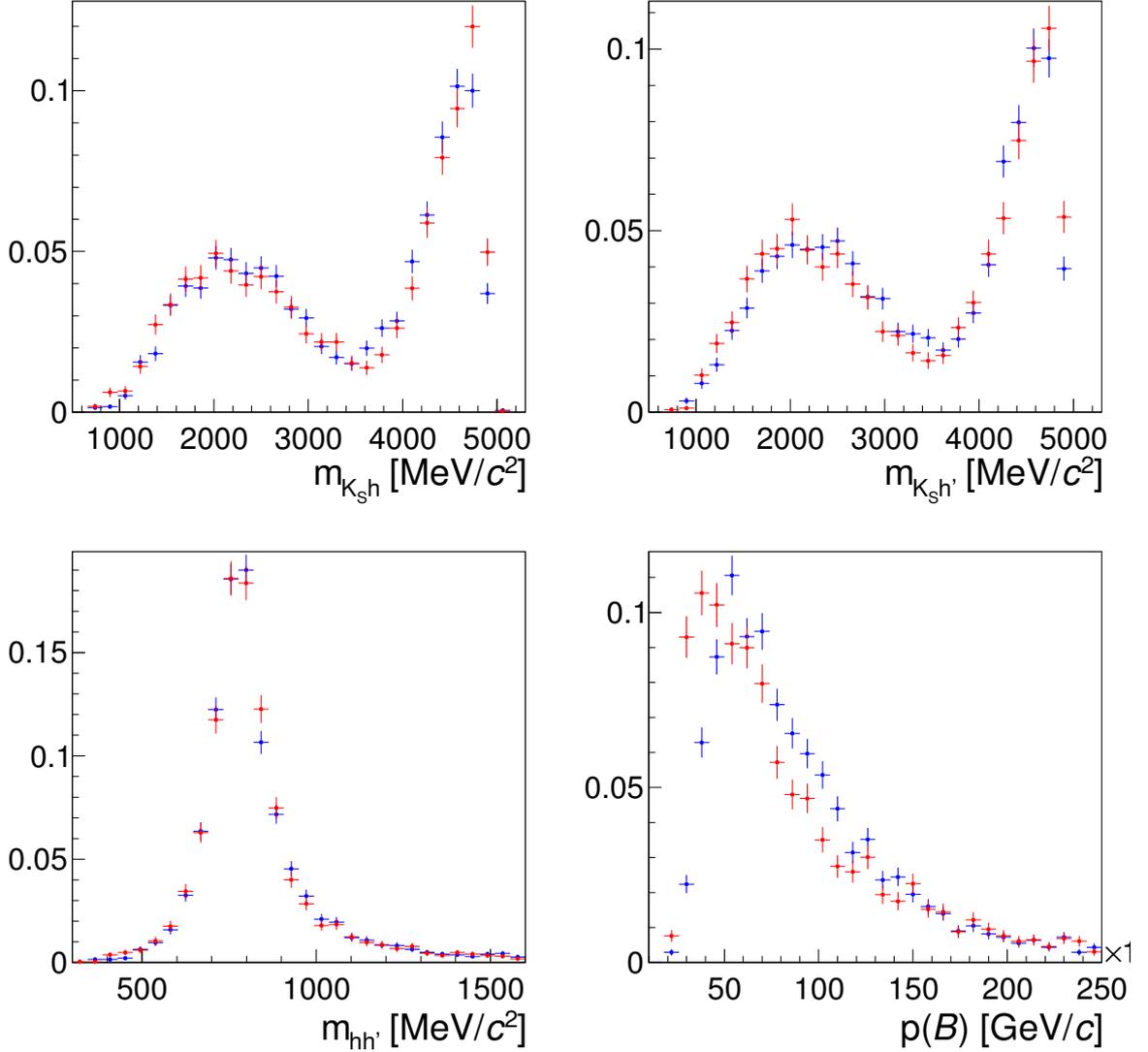
Figure 3.9 – Distributions of $m_{K_S^0 h}$ (top, left), $m_{K_S^0 h'}$ (top, right), $m_{hh'}$ (bottom, left), and $p_z(B)$ (bottom, right) for fully simulated (blue) and generated (red) events of $B^0 \to K^{*0}\rho^0$ at generation-level. The agreement between centre-of-mass distributions is satisfactory, but the z momentum distributions do not agree well between fully simulated samples and generated samples.

divide the p_z axis into 24 intervals and the fit is of better quality. We then smear the generator-level variables as described in Sec. 3.2.

The resulting $m_{K_S^0 h^+ h'}$ and $\Delta_{m_{K_S^0 h^+ h'}}$ distributions are shown and compared to the fully simulated distributions for the same channel in Fig. 3.11 with Down-Down and Long-Long K_S^0 reconstruction. The level of agreement between the two distributions is
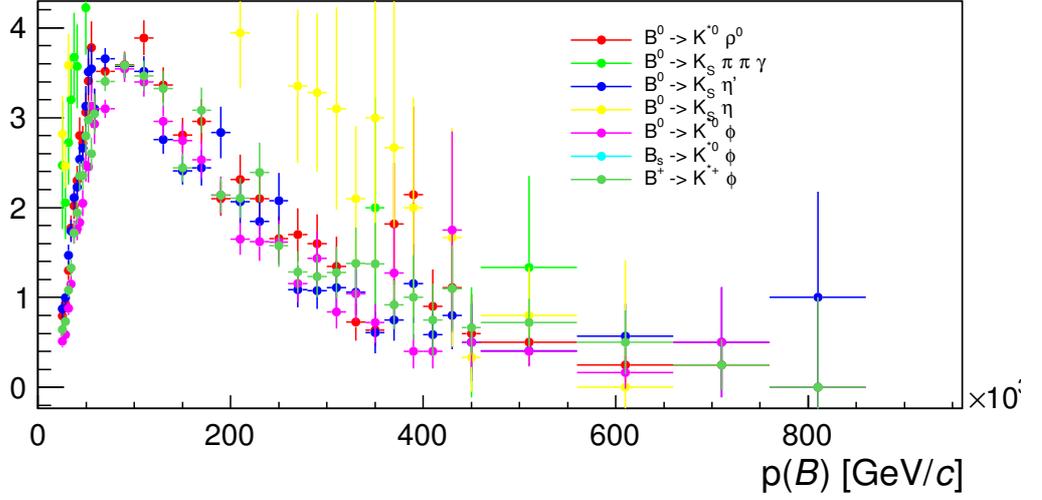
Figure 3.10 – Distribution of the weights calculated using Eq. 3.3 for different channels of partially reconstructed background.

good enough for our purposes. Figure 3.12 shows the distribution of the Dalitz plane in fast Monte-Carlo events, and in fully simulated events, as well as the pulls between these two distributions. The pulls are larger than in Sec. 3.3, but overall satisfactory given the small amount of partially reconstructed background events that we expect.

3.6 Conclusion

In the case of the partially reconstructed background of $B^0_{d,s} \rightarrow K^0_S h^+ h'^-$, it is possible to account for resolution effects on invariant masses by smearing event-by-event the θ, ϕ, $1/p_z$ of each reconstructed particle, along with the K^0_S mass. The functions we use to smear these variables can be extracted from only one Monte-Carlo sample, and exported from one channel to another, regardless of the missed particle or the type of reconstructed hadron.

We also demonstrated that we can emulate the acceptance effects on our samples by using selection criteria on variables available at generator level only, and by reweighting the resulting sample. This opens the possibility to generate a large amount of events with GAUSS, and to obtain sensible distributions of invariant masses both in Down-Down and Long-Long K^0_S reconstruction modes. This procedure can be useful in modelling backgrounds coming from a large variety of channels, such as the partially reconstructed background.

However, this procedure still suffers from inaccuracies in several levels. Firstly, the functions we use to fit resolution distributions are not perfect, and there are dependencies that we did not take into account, for instance between the distributions of Δ_ϕ for the
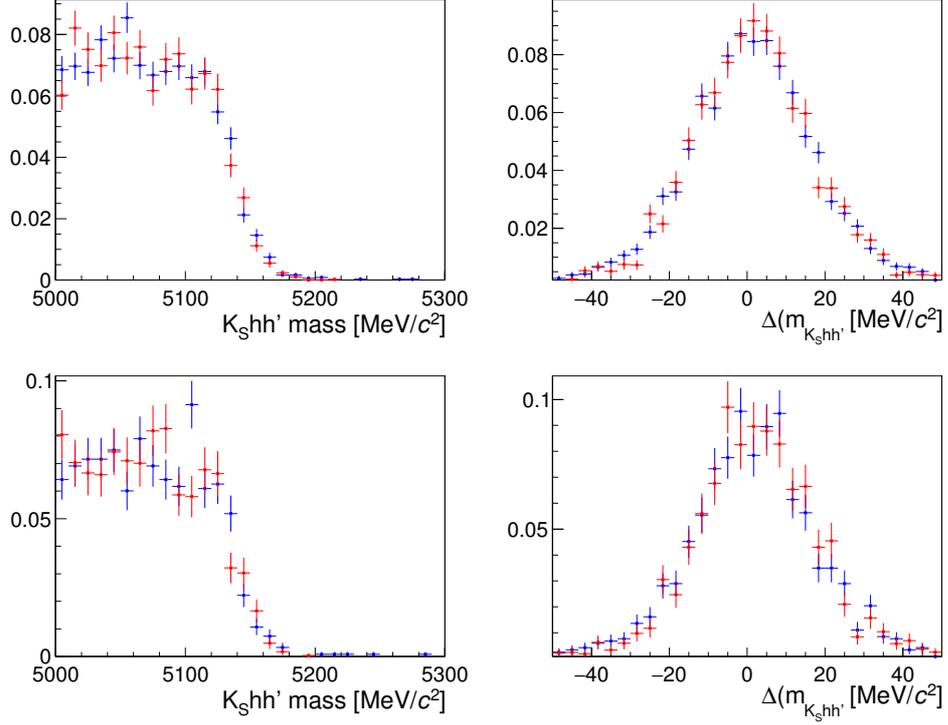
Figure 3.11 – Distributions of $m_{K_S^0 h^+ h'}$ (left) and $\Delta_{m_{K_S^0 h^+ h'}}$ (right) for full (blue) and fast (red) Monte-Carlo methods, with Down-Down (top) and Long-Long (bottom) K_S^0 reconstruction mode.

two charged tracks, which could explain the $\Delta_{m_{K_S^0 h^+ h'}}$ behaviour in Sec. 3.2 and Sec. 3.3. The generator-level cuts could be improved using Monte-Carlo samples disposing of the whole generator-level information, which would improve the agreement between invariant-masses distributions shown in Fig. 3.11 and in Fig . 3.12. The weighting procedure would greatly benefit from a careful study and larger samples. Finally, this procedure has only been tested on charmless backgrounds. The different topology of open-charm decays could require another set of generator-level cuts. Overall, this method provides sufficient modelling power for reconstruction effects on invariant masses for the analyses described in this dissertation.
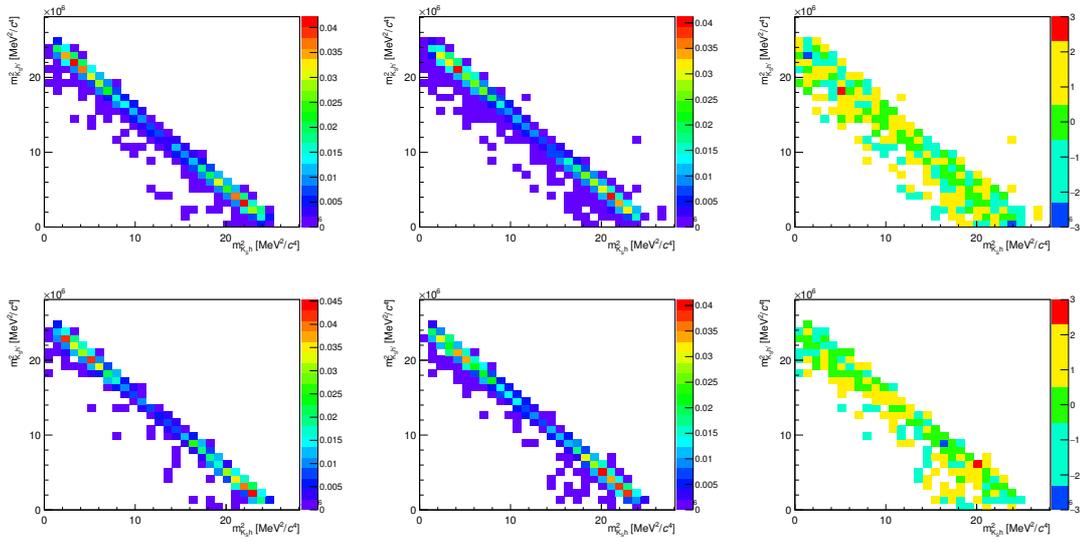
Figure 3.12 – Left: Distributions of fully simulated $B^0 \to K^{*0}\rho^0$ events with Down-Down (top) and Long-Long (bottom) K^0_S reconstruction mode over the Dalitz plane. Middle: Distributions of fast Monte-Carlo $B^0 \to K^{*0}\rho^0$ events with Down-Down (top) and Long-Long (bottom) K^0_S reconstruction mode over the Dalitz plane. Right: Pulls between the two distributions with Down-Down (top) and Long-Long (bottom) K^0_S reconstruction mode.

Chapter 4

Measurement of the branching fractions of the $B^0_{d,s} \to K^0_{\mathrm{S}} h^+ h'^-$ modes

In this chapter I describe the measurement of the branching fractions of the $B^0_{d,s} \to K^0_{\mathrm{S}} h^+ h'^-$ modes. In Sec. 4.1, I shortly present the analysis strategy. I then review in Sec. 4.2 the formalism and methods employed in the mass fit and in the extraction of the signal shapes over the Dalitz plot. The different event species considered in each mass spectrum and their modelling are described in Sec. 4.3. The results of the mass fit to data are shown in Sec. 4.4, and the results of the toy studies used to validate the model are shown in Sec. 4.5. I then present the different sources of systematic uncertainties originating from the mass fit in Sec. 4.6. Finally, I discuss the extraction of the distribution of signal events over the Dalitz plot in Sec. 4.7, and present the measurements of the branching fractions in Sec. 4.8.

4.1 Analysis strategy

The first LHCb analysis of the $B^0_{d,s} \to K^0_{\mathrm{S}} h^+ h'^-$ modes, performed with $1\,\mathrm{fb}^{-1}$ of 2011 data, was published in 2012 [63] and updated in 2013 [52]. The present analysis integrates the additional $2\,\mathrm{fb}^{-1}$ of data from 2012, disposes of more simulated samples, and makes use of more refined analysis techniques on several points. Our aim is to update the measurements of the branching fractions of the modes previously observed, along with observing the $B^0_s \to K^0_{\mathrm{S}} K^+ K^-$ decay for the first time. In order to avoid any experimenter bias in this search, we blind the region of the $B^0_s \to K^0_{\mathrm{S}} K^+ K^-$ signal in the mass fit. Finally, the result of the mass fit performed in this analysis is one of the key inputs to the three Dalitz-plot analyses performed on the Cabibbo-favoured signal modes.

We consider separately four different final states: $K^0_{\mathrm{S}} K^+ K^-$, $K^0_{\mathrm{S}} K^+ \pi^-$, $K^0_{\mathrm{S}} \pi^+ K^-$, and $K^0_{\mathrm{S}} \pi^+ \pi^-$. However, due to experimental differences, we have to simultaneously fit 24 different spectra in total. Firstly, we have to split our data between the Down-Down and

Long-Long K_s^0 reconstruction modes[1], described in Sec. 2.2.3, as the shapes and yields of the different event species are expected to differ between these two configurations. Secondly, we have to consider separately the 2011 data and two different data-taking periods in 2012, due to the difference in the trigger configuration between these periods described in Sec. 2.4.

Each of these mass spectra contains signals events from B^0 and B_s^0 decays, as well as several species of backgrounds. In the case of components that are well separated from the signal, such as the charmed background described in Sec. 4.3.2, we use a veto to remove most of their contributions. We describe in the following the different strategies adopted for the remaining backgrounds.

The combinatorial background, described in Sec. 4.3.5, is first suppressed using some preselection criteria that have a high efficiency on signal. It is then further suppressed using Boosted Decision Tree (BDT optimization) methods, trained using Monte-Carlo as the signal reference and events from the upper-mass sideband ($m(K_\mathrm{s}^0 h^\pm h'^\mp) > 5450 \,\mathrm{MeV}/c^2$) as the combinatorial background reference. In order not to bias the Dalitz plot, the variables used as an input to the BDT method are mainly topological variables. These methods produce an output variable for which the signal and combinatorial background distributions are well separated, as shown in Fig. 4.1. We then apply a selection on this variable so that the resulting signal and combinatorial background yields maximize the figure of merit

$$FoM = \frac{N(\mathrm{Sig})}{\sqrt{N(\mathrm{Sig}) + \mathrm{N(Bkg)}}} \qquad (4.1)$$

for all observed signal modes. $N(\mathrm{Sig})$ and $N(\mathrm{Bkg})$ are the number of signal and background events after the selection is applied. For the unobserved $B_s^0 \to K_\mathrm{s}^0 K^+ K^-$ mode, we use the Punzi figure of merit [64]

$$FoM = \frac{\epsilon_\mathrm{sig}}{\sqrt{\frac{a}{2} + N(\mathrm{Bkg})}} \qquad (4.2)$$

with $a = 2$, where ϵ_sig is the signal efficiency, estimated by means of Monte-Carlo samples. We use two different sets of requirements on the BDT output variable for each spectrum, depending on the signal component that we use to calculate the figure of merit. Indeed, considering different signal components for the parameter $N(\mathrm{Sig})$ will result in different cut values as cross-sections are different between B^0 and B_s^0 signals. Selection cuts optimized using a figure of merit calculated with the Cabibbo-favoured mode will be referred to as "loose", whereas those calculated with the Cabibbo-suppressed mode will be refereed to as "tight".

---

[1]We do not dispose of a dedicated stripping line for Long-Down K_s^0 candidates.
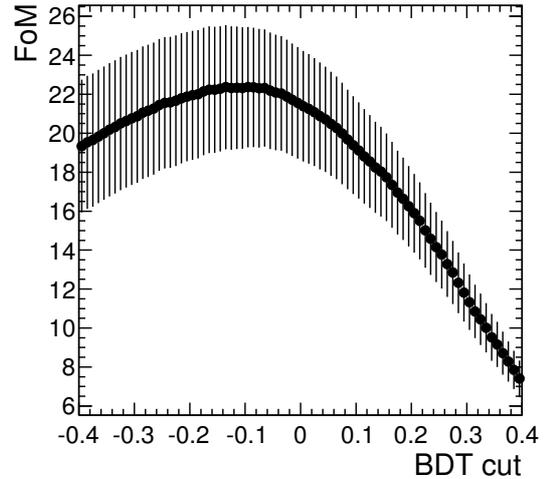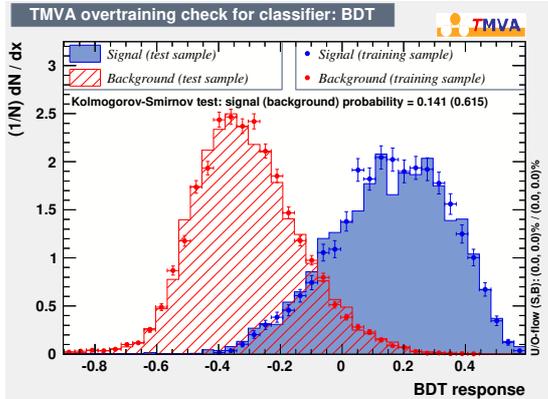
Figure 4.1 – Left: distribution of the output variable of a BDT method for signal Monte-Carlo events (red) and upper-mass sideband events (blue). Dashed histograms represent the distribution of the BDT output variable in training samples, while dots represent the distribution of the BDT output variable in test samples. Right: value of the figure-of-merit calculated on $B^0 \to K_S^0 K^+ K^-$, 2011, Down-Down signal for different BDT selection cut values. The maximum is chosen as the cut value for the loose BDT selection in that sample.

We apply a selection on particle identification variables (PID cuts) to reduce the contributions from misidentified signal events, or cross-feeds, with criteria of the type

$$(\mathrm{PROBNN(Pi/K)} - \mathrm{PROBNN(K/Pi)} > \alpha). \tag{4.3}$$

These likelihood-based PID classifiers are described in Sec. 2.2.4. The threshold α is optimized for each spectrum, using the same figure of merit as in the corresponding BDT cut optimization. The value of α is therefore different between the loose and the tight optimizations. The values of α are also chosen in such a way that no event can contribute to two different spectra.

Partially reconstructed background events, already discussed in Sec. 3, peak at a lower reconstructed $K_S^0 h^\pm h'^\mp$ invariant mass than signal events. In order to reduce the number of partially reconstructed background events, we include in the fit events with a reconstructed $K_S^0 h^\pm h'^\mp$ mass between $5150\,\mathrm{MeV}/c^2$ and $5800\,\mathrm{MeV}/c^2$.

The efficiency of the trigger, stripping, and selection criteria is not constant over the Dalitz plane. We estimate its distribution using Monte-Carlo simulations, corrected for data/MC discrepancies in the tracking and the trigger efficiencies. The total efficiency is taken as the factorized product of three components

$$\epsilon^{\mathrm{tot}} = \epsilon^{gen} \epsilon^{\mathrm{sel|gen}} \epsilon^{\mathrm{PID|(sel|gen)}}, \tag{4.4}$$

where ϵ^{gen} is the efficiency of the generator-level cuts in the Monte-Carlo, $\epsilon^{\mathrm{sel|gen}}$ is the trigger, selection, and stripping efficiency, and $\epsilon^{\mathrm{PID|(sel|gen)}}$ is the PID efficiency, estimated using the `PIDcalib` tool discussed in Sec. 2.5.3. The "|" symbol states that the efficiency is calculated using samples that passed the selection corresponding to the phases on the right of the symbol. These efficiencies are corrected for differences between data and Monte-Carlo in tracking and trigger efficiencies. This is done in LHCb by a standard recipe based on kinematics-dependent correction tables. Unfortunately, a problem was very recently discovered in the tables used in the correction of the trigger efficiency and is now under study in the collaboration. The results presented in this dissertation will be consequently updated before the final publication.

We perform a simultaneous unbinned maximum likelihood fit to the 24 reconstructed B mass spectra in order to extract the signal yields. We also perform a second fit, letting only the signal and combinatorial background yields to vary, in order to extract sWeights. These sWeights allow to estimate the distribution $f_{B^0_{d,s} \to K^0_S h^+ h'^-}(\mathrm{m}', \theta')$ of signal events over the Dalitz plane as

$$\epsilon = \frac{\sum\limits_{e \in \mathrm{data}} sW(e)}{\sum\limits_{e \in \mathrm{data}} \frac{sW(e)}{\epsilon(e)}}, \tag{4.5}$$

where $sW(e)$ and $\epsilon(e)$ are the sWeight associated to the event e and the signal efficiency of the event e, respectively. We estimate the total efficiency of signal events

$$\epsilon^{\mathrm{tot}} = \iint_{sqDP} f_{B^0_{d,s} \to K^0_S h^+ h'^-}(\mathrm{m}', \theta') \epsilon_{B^0_{d,s} \to K^0_S h^+ h'^-}(\mathrm{m}', \theta') \mathrm{d}\,\mathrm{m}'\,\mathrm{d}\theta', \tag{4.6}$$

where the integration is performed over the square Dalitz plot variables m' and θ', described in Sec. 1.3.2. The efficiency-corrected signal yield of a particular channel is then

$$N^{\mathrm{corr}}_{B^0_{d,s} \to K^0_S h^+ h'^-} = \epsilon^{\mathrm{tot}} N_{B^0_{d,s} \to K^0_S h^+ h'^-}, \tag{4.7}$$

where $N_{B^0_{d,s} \to K^0_S h^+ h'^-}$ is the yield from the signal fit. The branching fraction of each signal mode is then

$$\mathcal{B}(B^0_{d,s} \to K^0_S h^+ h'^-) = \frac{N^{\mathrm{corr}}_{B^0_{d,s} \to K^0_S h^+ h'^-}}{\mathcal{L}.\sigma_{pp \to b\bar{b}}.f_{d,s}} \tag{4.8}$$

where $f_{d,s}$ is the fraction of b quarks hadronising to B^0 or B_s^0 mesons, $\mathcal{L}$ is the integrated luminosity, and $\sigma_{pp\to b\bar{b}}$ is the cross-section of $b\bar{b}$ pair production in LHCb. The two last parameters cancel out when we consider ratios of branching fractions of $K_S^0 h^\pm h'^\mp$ modes, and the uncertainty on the f_s/f_d ratio is smaller than the uncertainty on f_d and f_s individually. Therefore, we aim at measuring the ratios

$$\frac{\mathcal{B}(B_{d,s}^0 \to K_S^0 h^+ h'^-)}{\mathcal{B}(B^0 \to K_S^0 \pi^+ \pi^-)} = \frac{f_{d,s}}{f_d} \frac{N_{B_{d,s}^0 \to K_S^0 h^+ h'^-}^{\text{corr}}}{N_{B^0 \to K_S^0 \pi^+ \pi^-}^{\text{corr}}}. \tag{4.9}$$

My personal contribution to this analysis is detailed in the following, and is focused on the extended maximum likelihood fit to data, and on the extraction of the signal distribution over the Dalitz plane.

4.2 Tools and formalism of the B-meson invariant mass fit

4.2.1 The unbinned maximum extended likelihood fit

Maximum-likelihood estimation is a widely used method of fitting parameters of a model to some data. For a variable x, we consider a model f, function of a parameter θ.[2] Given a set of measurements x_i, the likelihood of the model is

$$L(\theta) = \prod_{i=1}^{N} f(x_i, \theta), \tag{4.10}$$

which is a function of θ. The maximum-likelihood estimator $\hat{\theta}$ for θ is then the value of θ that maximizes the likelihood.[3] Maximum likelihood estimators are generally asymptotically unbiased and efficient for large data samples.

In the case where several event species are present in the model, the number N_i of events in each event species is itself a random variable. In the general case, it follows a Poisson distribution with the observed number of event parameter $N_{i,0}$

$$f(N_i) = \frac{N_i^{N_{i,0}}}{N_{i,0}!} e^{-N_i}. \tag{4.11}$$

---

[2]All the following assertions and formulae extend naturally to the case where there are several parameters and/or variables.

[3]As this value also maximizes the logarithm of the likelihood, it is often preferred to work with the logarithm.

Including this term in the model $f(x, \theta)$, and defining $N_0 = \sum_i N_{i,0}$ and $N = \sum_i N_i$, the extended likelihood is defined as

$$L(\theta; N_i) = \frac{e^{-N}}{N_0!} \prod_{i=1}^{N_0} f(x_i, \theta; N_i), \; -\ln(L) = N + \sum_{i=1}^{N_0} ln\left(f(x_i, \theta; N_i)\right). \tag{4.12}$$

In this expression, the normalization term $N_0!$ of the Poisson law was dropped, as it does not change the estimator for θ.

4.2.2 Gaussian constraints

One of the advantages of the likelihood estimator is that it is possible to "plug in" an external knowledge about some parameters by adding a term to the log-likelihood function. This effectively constrains the parameter by adding a penalty to the likelihood. We often choose so-called "Gaussian constraints" that result in the likelihood

$$L'(\theta) = L(\theta) \times e^{-\frac{(\theta - \theta_0)^2}{2\sigma_\theta^2}}, \tag{4.13}$$

where θ_0 is the central value of the constraint and σ_θ is its uncertainty. Gaussian constraints are often used to allow a proper convergence of a fit where the sensitivity to one or several parameters is poor.

4.2.3 The $_s\mathcal{P}lots$ method

Subtracting background from distributions in physics analyses can be performed in several ways. The $_s\mathcal{P}lots$ method [65] uses the covariance matrix extracted from a fit to a discriminating variable X to disentangle the signal and background distributions of some control variables Y_i. This covariance matrix is extracted from a fit in which only the yields of the different event species present in the dataset are varied.

Let a model with N_S event species, each with a yield noted N_k and a normalized PDF noted f_k. The $_s\mathcal{P}lots$ method defines for each event e and event species n the weight

$$sP_n(e) = \frac{\sum_{j=1}^{N_S} V_{nj} f_j(X_e)}{\sum_{k=1}^{N_S} N_k f_k(X_e)}, \tag{4.14}$$

where V_{nj} is the covariance between the yields of species n and j. The estimated distribution of each control variable Y_i for the event species n is denoted $sM_n(\overline{Y_i})$ and is defined by

$$N_n s M_n(\overline{Y_i}) \delta Y_i = \sum_{e \subset [\overline{Y_i} - \delta Y_i, \overline{Y_i} + \delta Y_i]} s P_n(e). \tag{4.15}$$

In the presence of species with fixed yields, Eq. 4.15 becomes

$$N_n.s M_n(Y).\delta Y = \sum_{e \subset [Y - \delta Y, Y + \delta Y]} s P_n(e) + c_n.M_0(Y), \tag{4.16}$$

where $M_0(Y)$ is the distribution of the variable Y for the ensemble of species with fixed yields. The parameter c_n is extracted from the covariance matrix of the fit and is

$$c_n = N_n - \sum_{j=1}^{N_S} V_{nj}, \tag{4.17}$$

where N_n and V_{nj} are defined as in Eq. 4.14. The parameter N_S is the number of species with varying yields.

The `RooFit` implementation of the $_s\mathcal{P}lot$ method does not allow to fix a part of the yields in the fit. We discuss in Annex A the pitfalls of the current implementation and propose a new implementation that we use in the following.

4.3 The B-meson invariant mass fit model

In this section, we review the different event species present in our dataset and their models in the mass fit.

4.3.1 B^0 and B_s^0 signal

The signal is modelled by a double Crystal-Ball distribution, which is the sum of two Crystal-Ball distributions [66], defined by

$$t = m - \mu,$$
$$F(m) = \begin{cases} \exp(-t^2/2\sigma^2) & \text{if } t/\sigma > -\alpha \\ (\frac{n}{|\alpha|})^n \exp(-\alpha^2/2)(\frac{n-\alpha^2}{|\alpha|} - \frac{t}{\sigma})^{-n} & \text{if } t/\sigma \le -\alpha \end{cases} \tag{4.18}$$

This distribution combines a Gaussian-type core, parameterized by μ and σ, and a radiative tail, parameterized by α and n. Depending on the sign of α, the tail can be on the left or on the right of the Gaussian core. Figure 4.2 shows an example of a Crystal-Ball distribution with a tail on the left, superimposed with a Gaussian distribution for comparison.
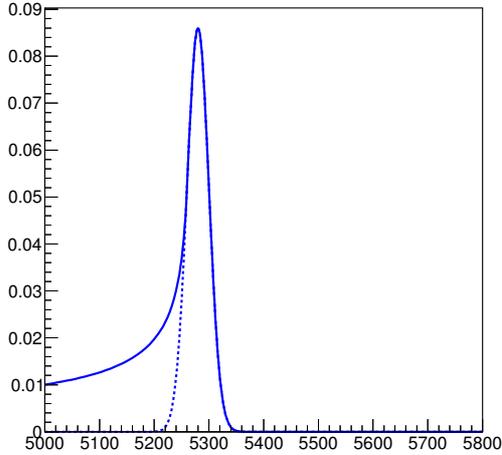
Figure 4.2 – Example of a Crystal-Ball distribution centered around $5280 \, \text{MeV}/c^2$, with a tail on the left (solid line). The distribution drawn with a dashed line is a Gaussian with the same μ and σ (normalization has been changed accordingly).

The two Crystal-Ball distributions of each signal component share their Gaussian parameters, and have their tails on opposite sides. The left-hand side tail accounts for radiative energy loss, while the right-hand side tail accounts for small stochastic dispersion. The fraction of the distribution with the tail on the left is denoted f, for a total of 7 parameters per signal component. Considering that there are two signal contributions per spectrum (B^0 and B_s^0 mesons), we have to constrain some of the parameters in the fit to data. For that purpose, we first perform a simultaneous fit to fully simulated signal Monte-Carlo samples, using the following fit model:

- The turnover point α_0 and the tail parameter n_0 of the left-hand side tail are different for each reconstruction mode and data-taking period, but are the same for the B^0 and B_s^0 mesons, as well as for Down-Down and Long-Long candidates.

- The parameters α_1 and n_1 of the right tail are the same in all the modes, data-taking periods, and B meson types. They are thought to be related to tracking effects, and all the modes under study have similar kinematics at first order.

- The fraction f is assumed to be the same for the B^0 and B_s^0 mesons, and the Down-Down and Long-Long reconstruction modes, but is taken as different in the different reconstructed modes.

- The parameters μ of both B^0 and B_s^0 are free to vary in the fit and are the same for all reconstruction modes and invariant masses.

- The width of the $K_S^0 \pi^+ \pi^-$ decays is varied in the fit and two (multiplicative) scale factors for the widths of the $K_S^0 K^\pm \pi^\mp$ and $K_S^0 K^+ K^-$ decays are also free to vary in the fit. The ratio of the widths of Down-Down and Long-Long candidates is assumed to be the same, in order to reduce the number of free parameters. The ratio of the widths of B^0 and B_s^0 is also assumed to be common between all reconstruction modes and data-taking periods.

Fits to Monte-Carlo samples are shown in Fig. 4.3, and the results are satisfactory. In the fit to data, we fix all the tail parameters and the fractions, but let the Gaussian parameters μ and σ vary, as well as the multiplicative factors.

4.3.2 Charmed contributions

Sub-decays with charmed intermediate states, such as $B^0 \to (\overline{D}^0 \to K^+\pi^-)K_S^0$, have different physics properties than the signal and are backgrounds to our analysis. Furthermore, they do not interfere with our signal because of the long lifetime of the charmed mesons. As such decays generally have larger branching fractions than our signal we veto them out. We thus apply a selection on the invariant mass of the daughters of the charmed hadron. Figure 4.4 shows this contribution in data events reconstructed according to the $K_S^0 K^+ \pi^-$ mass hypothesis and that passed the trigger requirements, along with the distribution obtained from simulated $B^0 \to K_S^0 K^\pm \pi^\mp$ signal events.

We also have to take into account misidentification of one of the daughters. For instance, $B \to (\overline{D}^0 \to K^+\pi^-)K_S^0$ decays can also contribute to the $K_S^0 \pi^+ \pi^-$ spectrum, and are also vetoed there.

4.3.3 Λ background

Another source of background comes from Λ baryons misidentified as K_S^0 mesons, as the proton from the Λ decay has been wrongly identified as a pion. Figure 4.5 shows the distribution of $K_S^0 \pi^+ \pi^-$ data events on the proton PID of one of the K_S^0 daughters and the mass of the K_S^0 using a proton hypothesis for this K_S^0 daughter. A clear peak near the Λ mass is present, indicating the presence of the Λ baryon background in data.

We veto out this contribution by imposing that an event is either outside the $|m(K_S^0 \text{ as } \Lambda) - \mathrm{m}(\Lambda)| < 10\,\mathrm{MeV}/c^2$ window, or the PID variable `ProbNNp` of each of the pion daughters is inferior to 0.05 unities. The efficiency of this requirement is estimated on MC, and shown to be around 99%.

4.3.4 Beauty baryons backgrounds

The mass of beauty baryons is larger than $5600\,\mathrm{MeV}/c^2$ and if a proton is misidentified as a pion or a kaon, may fall into our considered mass range. We apply an additional selection criterion on the proton PID, required to be inferior to 0.5 unities for both h^+ and h'^-, in order to veto these backgrounds out.
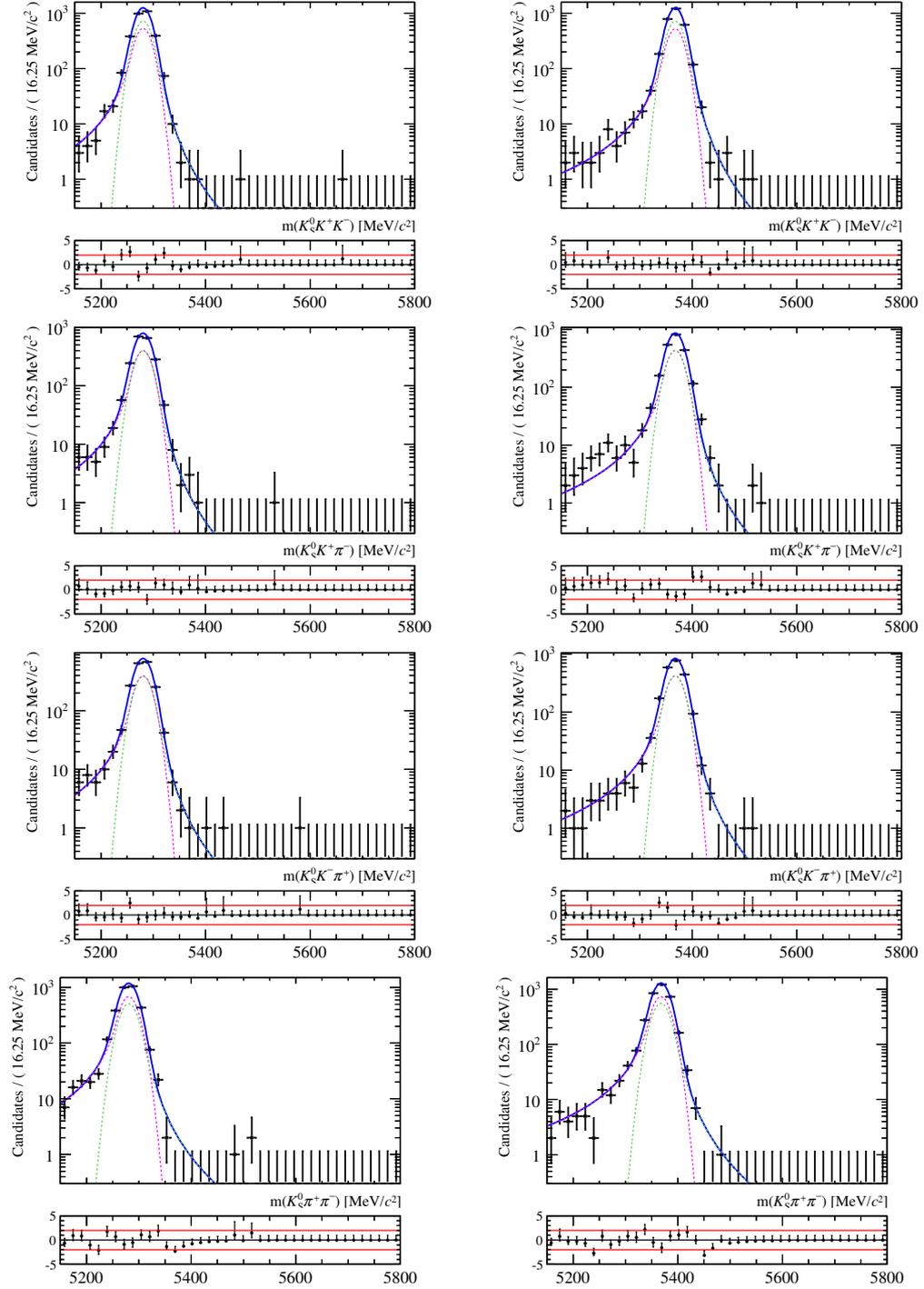
Figure 4.3 – Results of mass fits on simulated signal samples (2011)(Down-Down), using the loose BDT optimization, shown in logarithmic scale. $K_S^0 K^+ K^-$, $K_S^0 K^+ \pi^-$, $K_S^0 \pi^+ K^-$, and $K_S^0 \pi^+ \pi^-$ are shown from top to bottom, while B^0 decays are shown on the left and B_s^0 decays on the right.
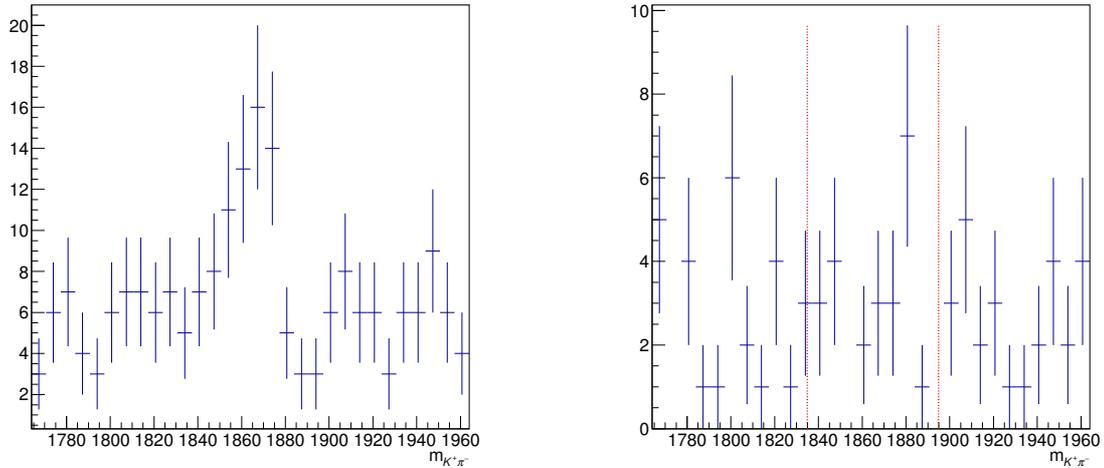
Figure 4.4 – Reconstructed $K^+\pi^-$ mass (in MeV/c^2) from trigged Down-Down $B^0 \to K^0_S K^\pm \pi^\mp$ data events before selection (left) and from trigged Down-Down $B^0 \to K^0_S K^\pm \pi^\mp$ simulated events before selection (right).
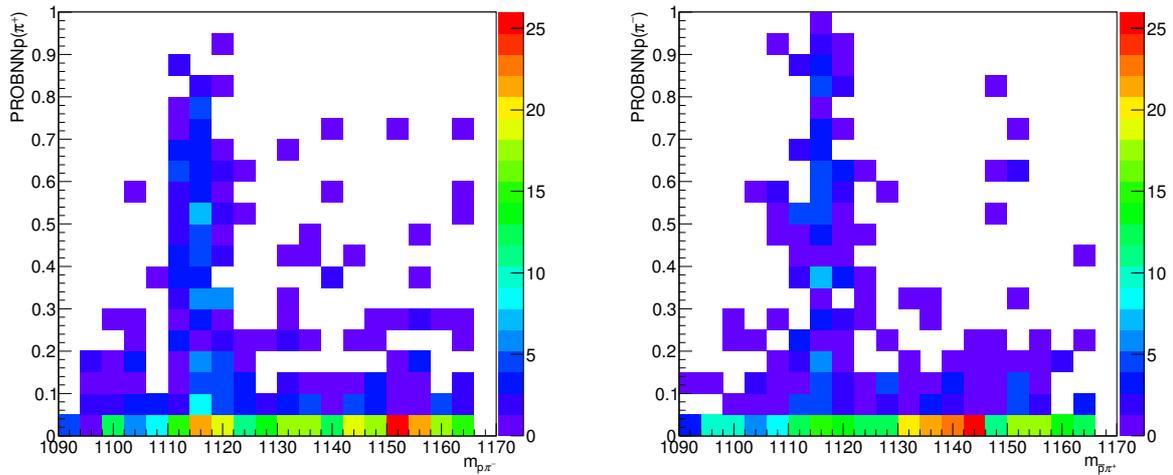
Figure 4.5 – Left (Right): Distributions of selected data events on the reconstructed K^0_S mass, calculated using proton mass hypothesis on the π^+ (π^-) with respect to the proton PID of the same particle.

4.3.5 Combinatorial backgrounds

The dominant source of background in the analysis is the random combination of tracks from several decays. In the LHCb experiment, we do not dispose of a dedicated Monte-

Carlo simulation for such contributions. However, we could use specific stripping lines looking for $K_\mathrm{S}^0 h^+ h'^+$ candidates. These candidates can only come from random combination of tracks or other sources of background that we did not consider here, and are expected to be kinematically close to the combinatorial background. These stripping lines have been prepared but could not be included in this iteration of the analysis.

We consider two different shapes for the combinatorial background: exponential and linear. Both yield similar results, and we choose the linear shape as a baseline for our analysis. We add two multiplicative factors to the linear shape to account for differences between Long-Long and Down-Down reconstruction modes and between invariant mass spectra. The slopes for different data-taking periods are considered independent.

4.3.6 Cross-feeds

Cross-feeds are the contributions to a mass spectrum $K_\mathrm{S}^0 h^\pm h'^\mp$ originating from a signal decay $B_{d,s}^0 \to K_\mathrm{S}^0 h^\pm h''^\mp$, where h'' is misidentified as an h'. These contributions typically lie near a signal peak, and thus are dangerous to the fit as they can be absorbed in the tails of signal distributions. We model them by double Crystal-Ball distributions, and fix all of their parameters (including the parameters μ and σ) to their value extracted from the fit to Monte-Carlo. Figure 4.6 shows the results of some of these fits.

Even with fixed shape, these contributions are too close to the signal to be properly accounted for by an unconstrained fit. We thus constraint their yields using known efficiencies and the yield parameters of the signal yield from which they originate

$$N(B^0 \to K_\mathrm{S}^0 h^\pm h'^\mp \text{ as } K_\mathrm{S}^0 hh'') = N(B^0 \to K_\mathrm{S}^0 h^\pm h'^\mp) f(B^0 \to K_\mathrm{S}^0 h^\pm h'^\mp \text{ as } K_\mathrm{S}^0 hh''), \quad (4.19)$$

$$f(B^0 \to K_\mathrm{S}^0 h^\pm h'^\mp \text{ as } K_\mathrm{S}^0 hh'') = \epsilon^{\mathrm{PID|sel\&gen}}(B^0 \to K_\mathrm{S}^0 h^\pm h'^\mp \text{ as } K_\mathrm{S}^0 hh'') \frac{\epsilon^{\mathrm{sel|gen}}(B^0 \to K_\mathrm{S}^0 h^\pm h'^\mp)}{\epsilon^{\mathrm{sel|gen}}(B^0 \to K_\mathrm{S}^0 h^+ h''^-)}. \quad (4.20)$$

The width of the Gaussian constraint on the parameter $f(B^0 \to K_\mathrm{S}^0 h^\pm h'^\mp \text{ as } K_\mathrm{S}^0 hh'')$ is derived from uncertainties on the relevant efficiencies.

4.3.7 Partially reconstructed backgrounds

We already discussed the nature and general properties of partially reconstructed backgrounds in $B_{d,s}^0 \to K_\mathrm{S}^0 h^+ h'^-$ decays in Sec. 3.1. Table 4.1 shows the categories of partially reconstructed backgrounds that contribute to each reconstructed invariant mass. We model the shapes of these contributions using the fast Monte-Carlo method described in Chapter 3, and fix all their shape parameters in the fit to data.

As these contributions are expected to be small and as their distributions overlap, we constrain the yields of each partially reconstructed background category using
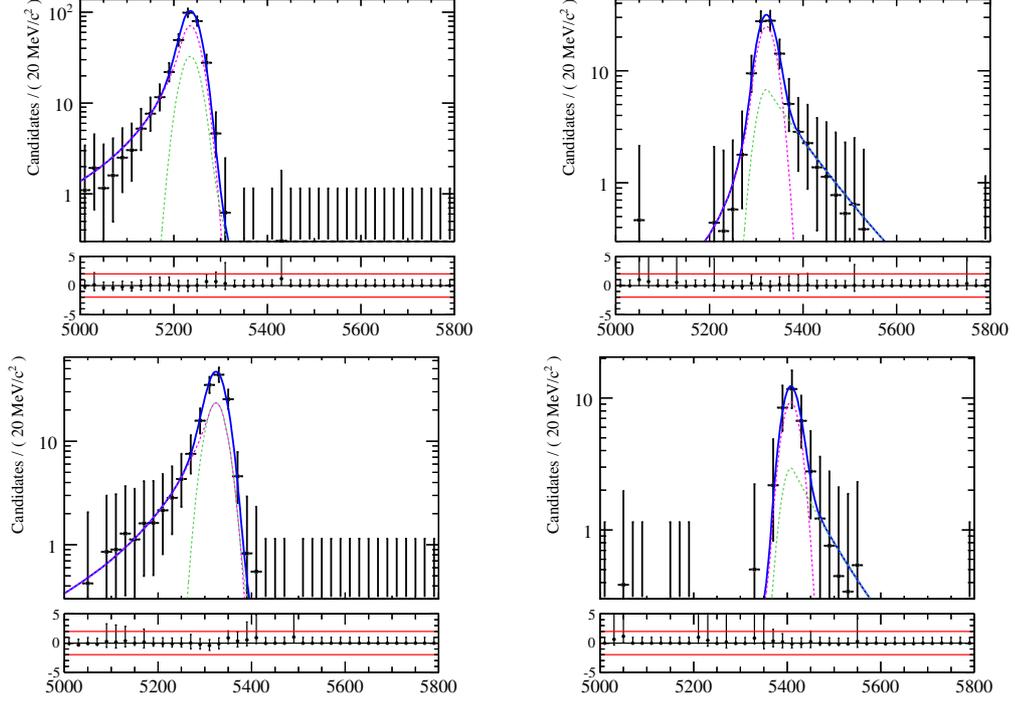
Figure 4.6 – Results of the fit of the reconstructed invariant mass of the misidentified signal decays, using the loose optimization of the BDT on 2011, Down-Down simulated signal samples. Top: on the left, $B^0 \to K_S^0 K^+ K^-$ as $K_S^0 K^\pm \pi^\mp$; on the right, $B^0 \to K_S^0 \pi^+ \pi^-$ as $K_S^0 K^\pm \pi^\mp$. Bottom: on the left, $B_s^0 \to K_S^0 K^\pm \pi^\mp$ as $K_S^0 \pi^+ \pi^-$; on the right, $B_s^0 \to K_S^0 K^\pm \pi^\mp$ as $K_S^0 K^+ K^-$.

Table 4.1 – Categories of partially reconstructed backgrounds included in each invariant mass spectrum.

Category	$K_S^0 \pi^+ \pi^-$	$K_S^0 K^\pm \pi^\mp$	$K_S^0 K^+ K^-$
Charmed(B^0)	Yes	Yes	Yes
Charmless(B^0)	Yes	Yes	Yes
Charmed(B_s^0)	Yes	Yes	No
Charmless(B_s^0)	Yes	Yes	No
Missing γ(B^0, NR)	Yes	No	No
Missing γ(B^0, resonant)	Yes	No	No

$$N(\mathrm{Cat}) = N(\mathrm{Sig}) h(\mathrm{Cat}) \mathcal{B}(\mathrm{Cat}), \qquad (4.21)$$

$$h(\mathrm{Cat}) = \frac{1}{\mathcal{B}(\mathrm{Sig})} \frac{\epsilon(\mathrm{Sig})}{\epsilon(\mathrm{Cat})}, \qquad (4.22)$$

Table 4.2 – Values and statistical uncertainties on yield parameters extracted from the fit to data (loose BDT optimization)

	2011	2012(pre-June)	2012(post-June)
$N(B^0 \to K_S^0 K^+ K^-)(DD)$	281 ± 19	181 ± 15	671 ± 30
$N(B^0 \to K_S^0 K^+ K^-)(LL)$	222 ± 17	119 ± 12	344 ± 20
$N(B_s^0 \to K_S^0 K^+ K^-)(DD)$	23 ± 9	2 ± 6	25 ± 13
$N(B_s^0 \to K_S^0 K^+ K^-)(LL)$	7 ± 8	6 ± 5	8 ± 7
$N(B^0 \to K_S^0 K^+ \pi^-)(DD)$	52 ± 12	44 ± 11	73 ± 14
$N(B^0 \to K_S^0 K^+ \pi^-)(LL)$	37 ± 8	29 ± 8	30 ± 9
$N(B_s^0 \to K_S^0 K^+ \pi^-)(DD)$	152 ± 15	92 ± 12	255 ± 19
$N(B_s^0 \to K_S^0 K^+ \pi^-)(LL)$	91 ± 11	51 ± 8	118 ± 13
$N(B^0 \to K_S^0 K^- \pi^+)(DD)$	52 ± 12	47 ± 11	91 ± 16
$N(B^0 \to K_S^0 K^- \pi^+)(LL)$	26 ± 7	21 ± 8	56 ± 10
$N(B_s^0 \to K_S^0 K^- \pi^+)(DD)$	181 ± 17	113 ± 14	307 ± 22
$N(B_s^0 \to K_S^0 K^- \pi^+)(LL)$	115 ± 12	49 ± 9	143 ± 14
$N(B^0 \to K_S^0 \pi^+ \pi^-)(DD)$	803 ± 36	553 ± 30	1410 ± 46
$N(B^0 \to K_S^0 \pi^+ \pi^-)(LL)$	471 ± 27	286 ± 19	654 ± 30
$N(B_s^0 \to K_S^0 \pi^+ \pi^-)(DD)$	65 ± 18	16 ± 15	83 ± 22
$N(B_s^0 \to K_S^0 \pi^+ \pi^-)(LL)$	23 ± 12	15 ± 8	42 ± 14

where $\mathcal{B}(\text{Cat})$ is the estimated inclusive branching fraction of the category, and Sig refers to the Cabibbo-favoured signal mode of the corresponding $K_S^0 h^\pm h'^\mp$ spectrum. The parameter $h(\text{Cat})$ is Gaussian-constrained using information from Monte-Carlo simulation, under the assumption that the efficiencies of all the decays within a category are roughly equal. In order to account for the lack of precise knowledge of the efficiencies on partially reconstructed backgrounds, we multiply the width of the constraint by a factor two.

Except for radiative decays $B^0 \to K_S^0 \pi^+ \pi^- \gamma$ and $B^0 \to K_S^0 \eta'$, the branching fractions $\mathcal{B}(\text{Cat})$ are not known. We use information from the PDG to obtain an estimate of the minimum of each inclusive branching fraction. We then perform a fit to data while fixing the parameter $h(\text{Cat as } K_S^0 h^\pm h'^\mp)$ in order to extract an estimate of this branching fraction, and then fix it to the value obtained from this fit.

4.4 Results of the mass fit

Table 4.2 and Table 4.3 show the results of the fit to data on the loose and tight BDT optimizations, respectively. Figure 4.7-4.10 show the corresponding fits for the 2011 data-taking period. All spectra are satisfyingly modelled.

A naive, statistical only significance of the $B_s^0 \to K_S^0 K^+ K^-$ observation can be obtained by performing a fit fixing all $B_s^0 \to K_S^0 K^+ K^-$ yields to 0, and mesuring the differ-

Table 4.3 – Values and statistical uncertainties on yield parameters extracted from the fit to data (tight BDT optimization)

	2011	2012(pre-June)	2012(post-June)
$N(B^0 \to K_S^0 K^+ K^-)(DD)$	122 ± 11	129 ± 12	299 ± 17
$N(B^0 \to K_S^0 K^+ K^-)(LL)$	149 ± 12	71 ± 8	140 ± 11
$N(B_s^0 \to K_S^0 K^+ K^-)(DD)$	5 ± 3	2 ± 3	5 ± 4
$N(B_s^0 \to K_S^0 K^+ K^-)(LL)$	4 ± 3	1 ± 2	1 ± 2
$N(B^0 \to K_S^0 K^+ \pi^-)(DD)$	34 ± 9	29 ± 8	48 ± 11
$N(B^0 \to K_S^0 K^+ \pi^-)(LL)$	28 ± 7	23 ± 6	24 ± 7
$N(B_s^0 \to K_S^0 K^+ \pi^-)(DD)$	118 ± 12	65 ± 10	222 ± 18
$N(B_s^0 \to K_S^0 K^+ \pi^-)(LL)$	78 ± 10	40 ± 7	73 ± 9
$N(B^0 \to K_S^0 K^- \pi^+)(DD)$	42 ± 10	34 ± 8	74 ± 13
$N(B^0 \to K_S^0 K^- \pi^+)(LL)$	22 ± 6	22 ± 7	41 ± 8
$N(B_s^0 \to K_S^0 K^- \pi^+)(DD)$	139 ± 14	90 ± 11	268 ± 19
$N(B_s^0 \to K_S^0 K^- \pi^+)(LL)$	91 ± 10	42 ± 8	102 ± 11
$N(B^0 \to K_S^0 \pi^+ \pi^-)(DD)$	514 ± 25	392 ± 23	898 ± 34
$N(B^0 \to K_S^0 \pi^+ \pi^-)(LL)$	386 ± 23	239 ± 17	441 ± 23
$N(B_s^0 \to K_S^0 \pi^+ \pi^-)(DD)$	43 ± 10	16 ± 8	86 ± 14
$N(B_s^0 \to K_S^0 \pi^+ \pi^-)(LL)$	21 ± 8	15 ± 6	38 ± 8

ence in NLL. This method yields

$$\Delta_{\mathrm{NLL}} = -7.70 \text{ (loose)}$$
$$\Delta_{\mathrm{NLL}} = -4.65 \text{ (tight)}$$

which correspond to a significance of 3.9 and 3.1σ, respectively. This difference in significance may be related to the fact that crossfeed events are relatively more abundant in the tight spectra than in the loose spectra.
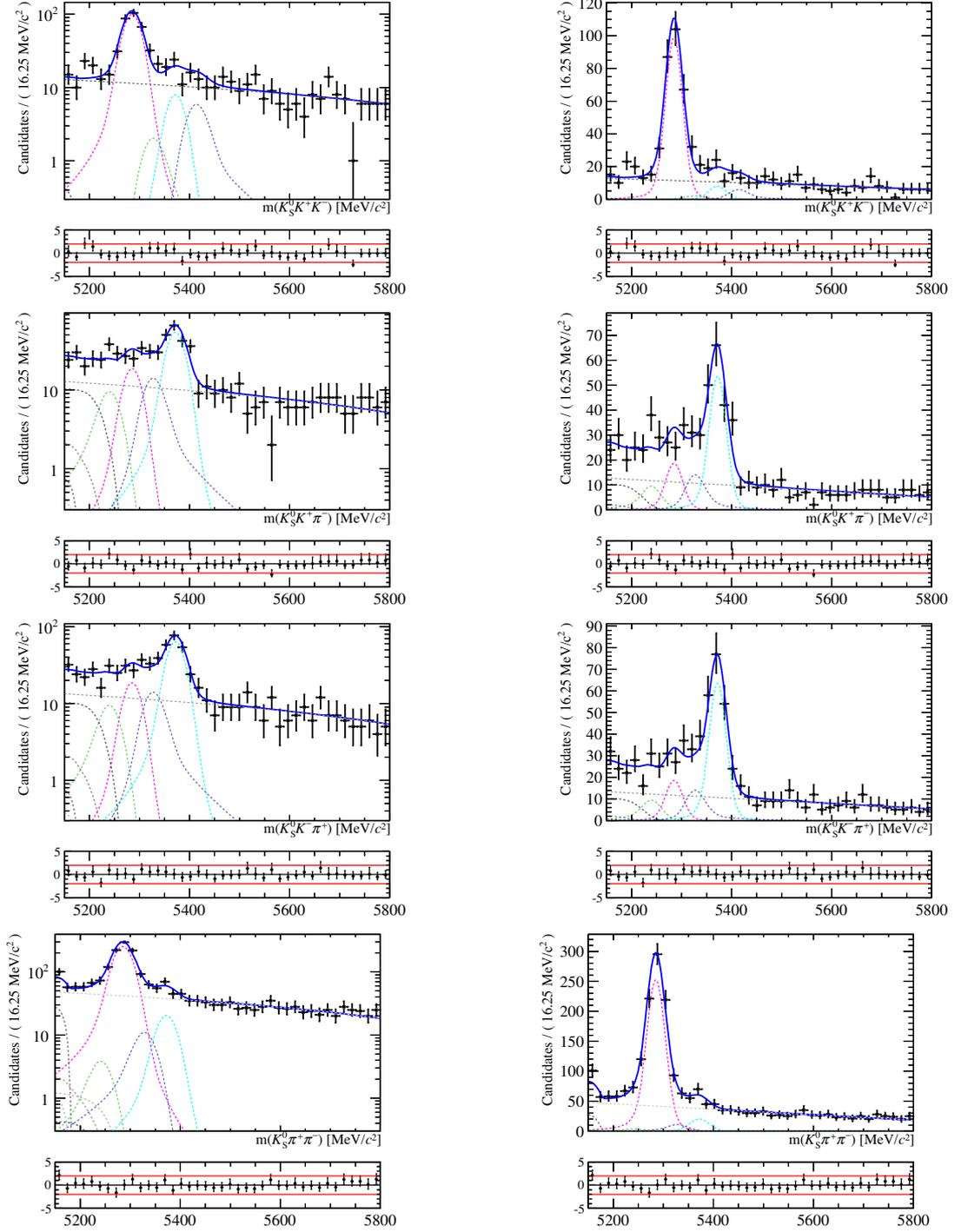
Figure 4.7 – Result of the simultaneous fit to data (Down-Down, 2011) with the loose BDT optimisation. $K_S^0 K^+ K^-$, $K_S^0 K^+ \pi^-$, $K_S^0 \pi^+ K^-$ and $K_S^0 \pi^+ \pi^-$ are shown from top to bottom, while the left plots show the result on a linear scale and the right on a logarithmic scale.
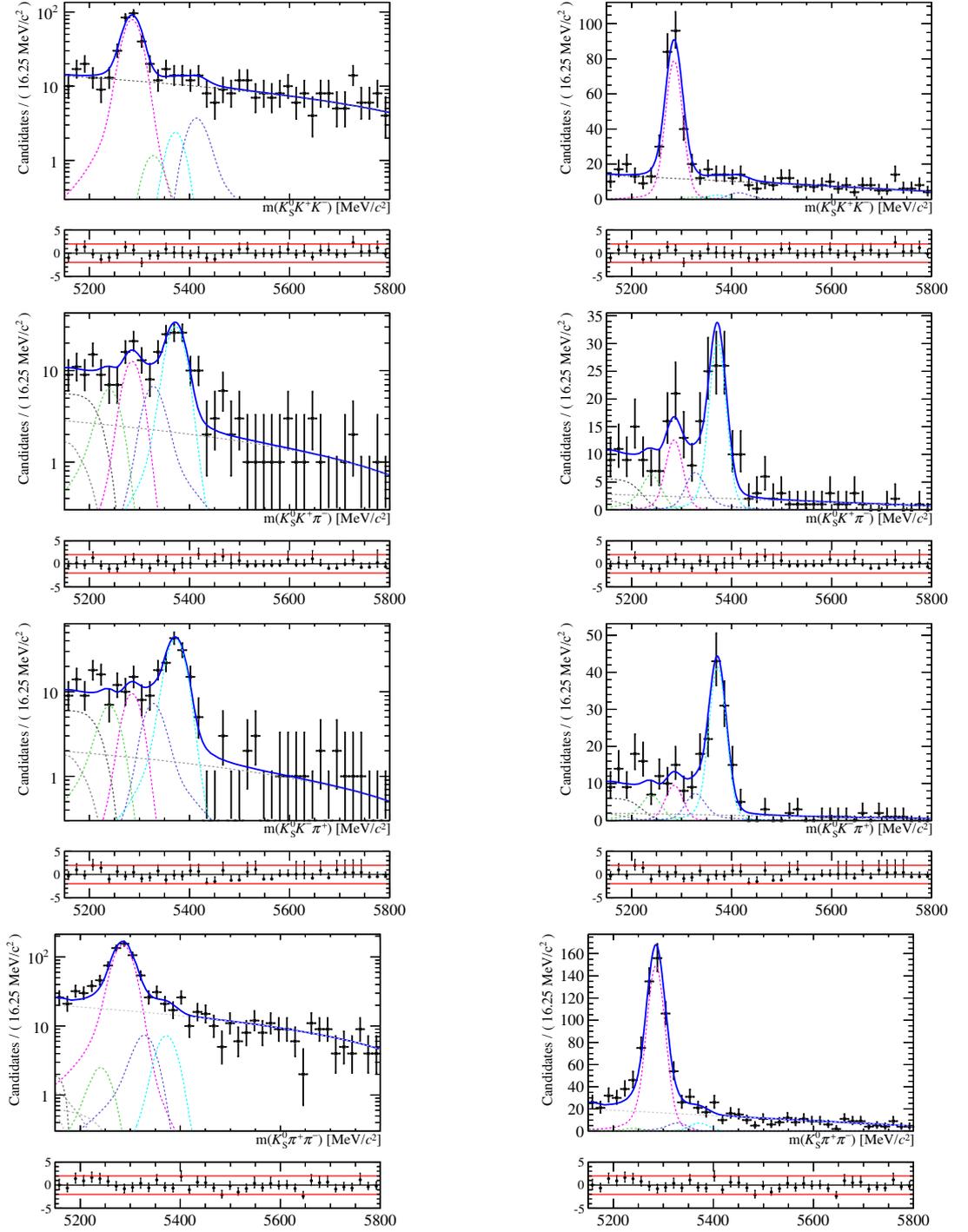
Figure 4.8 – Result of the simultaneous fit to data (Long-Long, 2011) with the loose BDT optimisation. $K_S^0 K^+ K^-$, $K_S^0 K^+ \pi^-$, $K_S^0 \pi^+ K^-$ and $K_S^0 \pi^+ \pi^-$ are shown from top to bottom, while the left plots show the result on a linear scale and the right on a logarithmic scale.

81
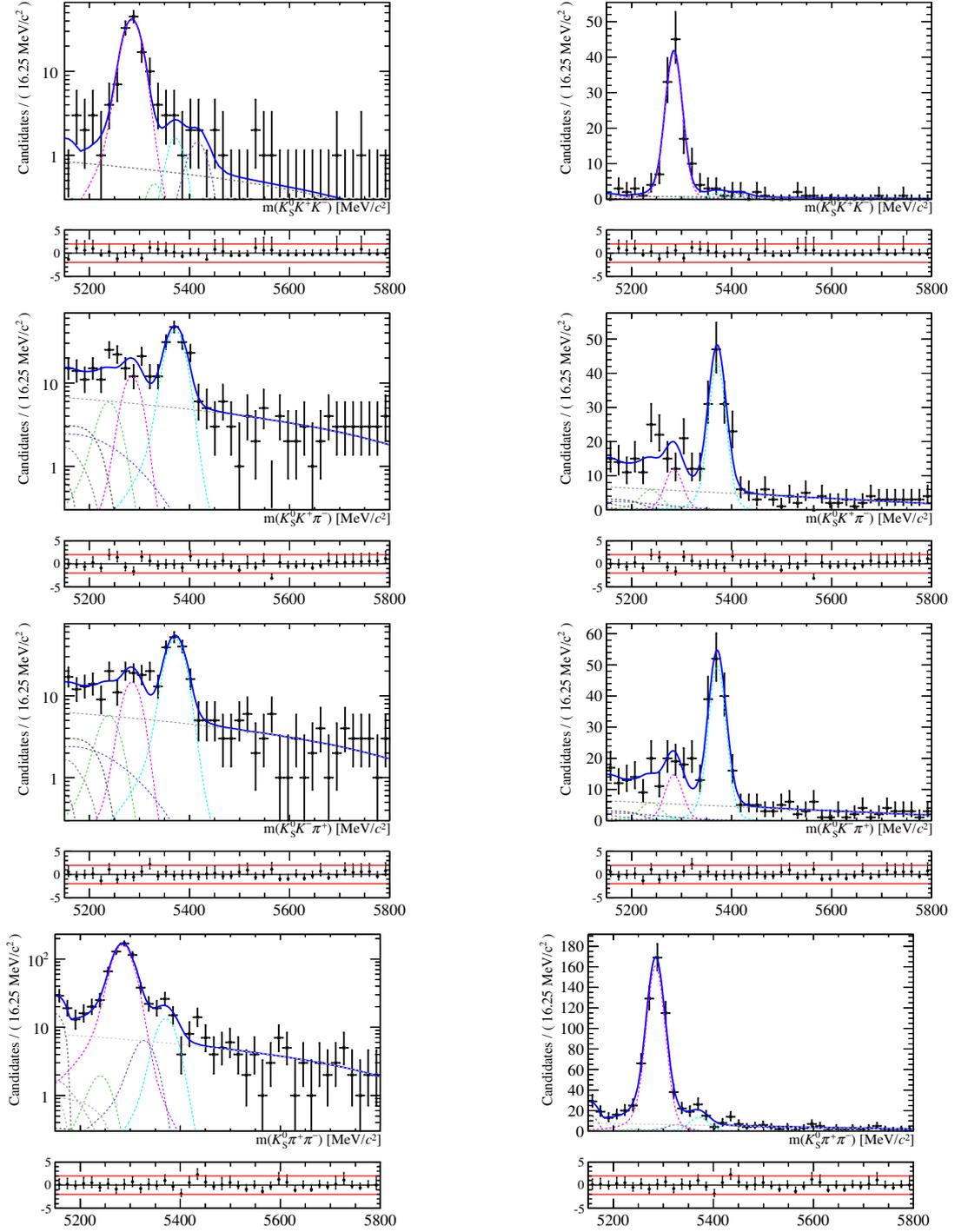
Figure 4.9 – Result of the simultaneous fit to data (Down-Down, 2011) with the tight BDT optimisation. $K_S^0 K^+ K^-$, $K_S^0 K^+ \pi^-$, $K_S^0 \pi^+ K^-$ and $K_S^0 \pi^+ \pi^-$ are shown from top to bottom, while the left plots show the result on a linear scale and the right on a logarithmic scale.
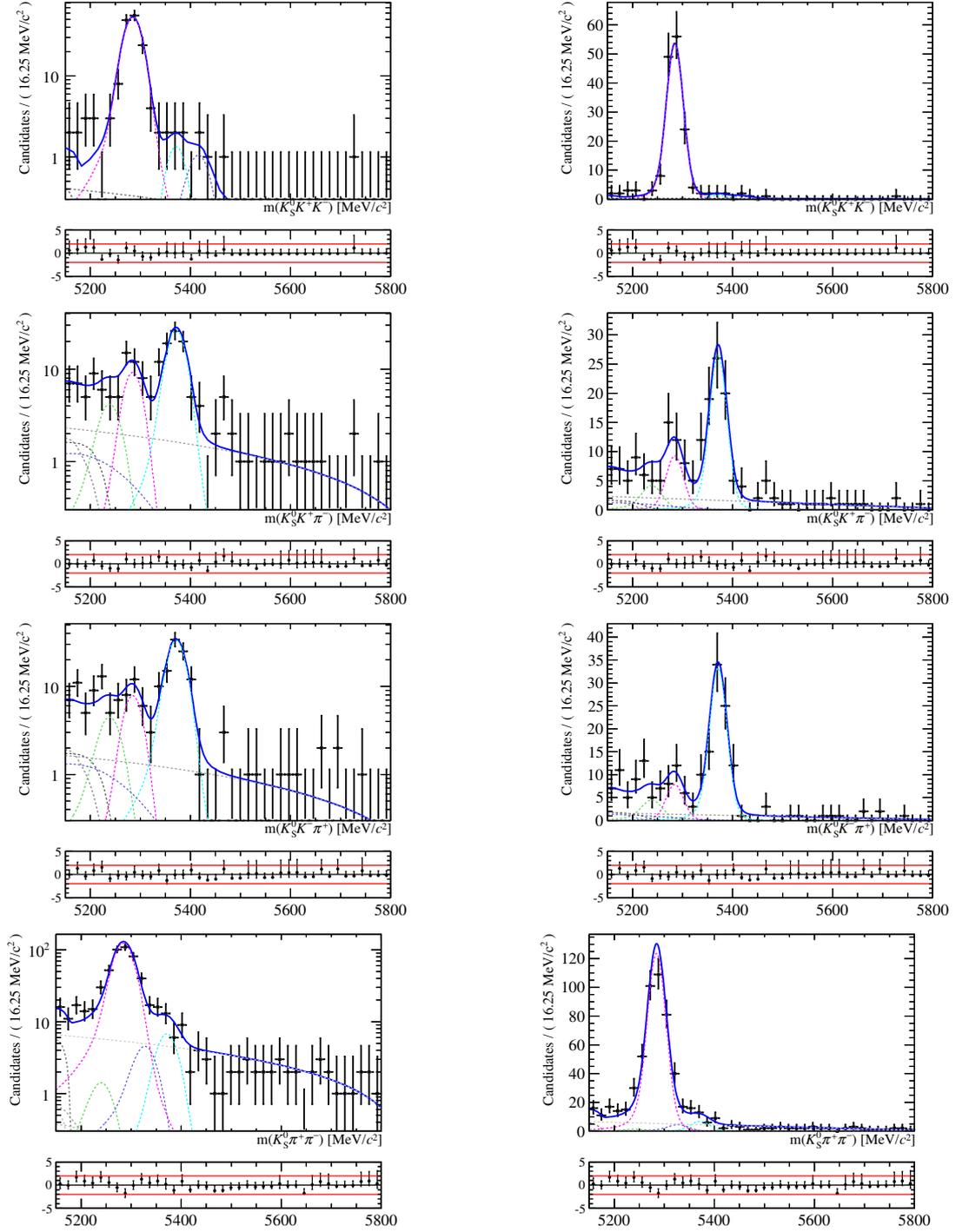
Figure 4.10 – Result of the simultaneous fit to data (Long-Long, 2011) with the tight BDT optimisation. $K_S^0 K^+ K^-$, $K_S^0 K^+ \pi^-$, $K_S^0 \pi^+ K^-$ and $K_S^0 \pi^+ \pi^-$ are shown from top to bottom, while the left plots show the result on a linear scale and the right on a logarithmic scale.

4.5 Validation of the mass fit model

A fit model can be tested by means of pseudo-experiments. Performing the fit to these samples allows to check the sensitivity to certain parameters, their covariance, and the convergence properties of the model. For each parameter θ and each pseudo-experiment i, we define the bias

$$\Delta_i(\theta) = \theta_t - \theta_i \tag{4.23}$$

and the pull statistic

$$p_i(\theta) = \frac{\Delta_i(\theta)}{\sigma_i(\theta)}, \tag{4.24}$$

where θ_t is the value used to generate the toys, θ_i is the value extracted from the fit to pseudo-experiment i, and $\sigma_i(\theta)$ is the uncertainty on θ_i. In the case of an unbiased estimator that properly covers uncertainties, the pull statistic is expected to follow a standard Gaussian with mean 0 and width 1. Deviations of the mean from 0 indicate a bias in the fit, while deviations of the width from 1 indicate an incorrect coverage of the uncertainty.

We test the invariant-mass fit model by means of 500 pseudo-experiments. They are generated using the fit model with all parameters set to the value extracted from the fit, except for the yields, which are varied according to their Poisson distribution. We summarise in Table 4.6 and in Table 4.7 the results of the fit validation procedure. When the deviation from the standard Gaussian is significant, we correct the measured value of a parameter θ with

$$\theta_{corr} = \theta_0 - \frac{\Delta(\theta)}{2} \tag{4.25}$$

and its uncertainty $\delta(\theta)$ with

$$\delta(\theta)_{corr} = \frac{\delta(\theta)}{\sigma(p(\theta))}. \tag{4.26}$$

A systematic uncertainty

$$\delta_{\mathrm{syst}}(\theta) = \frac{\Delta(\theta)}{2} \tag{4.27}$$

is associated to this correction.

Figure 4.14 shows the pull distributions for 2011 $K_{\mathrm{S}}^0 K^+ K^-$ signal yields using the loose BDT optimization. Tables 4.4 and 4.5 detail the signal yields for which a bias larger than 2σ was observed in the loose and tight BDT optimizations, respectively. The width of the pulls distribution is always consistent with the unity expectation.
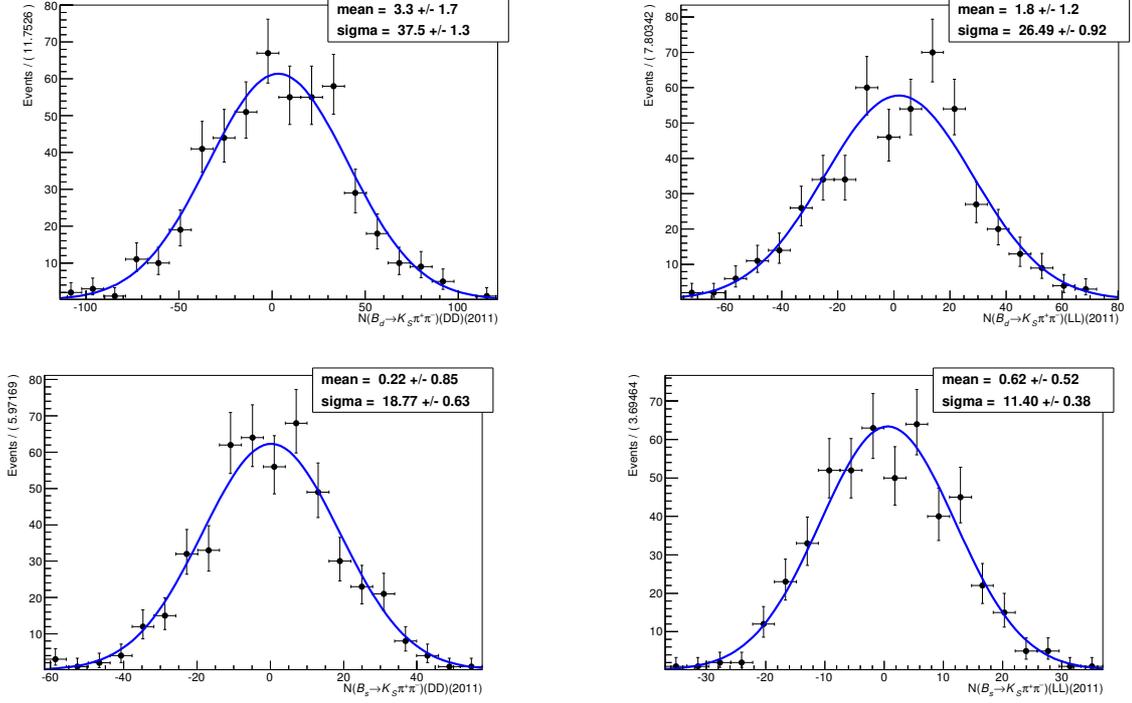
Figure 4.11 – Residuals distributions for yields of $K_S^0\pi^+\pi^-$ 2011 signals. Top: $B^0 \to K_S^0\pi^+\pi^-$ signal, bottom: $B_s^0 \to K_S^0\pi^+\pi^-$ signal. Left: Down-Down, right: Long-Long.

Table 4.4 – Signal yields for which a bias was observed in the loose BDT optimization, along with the bias. They are corrected for in the final results and accounted for in the uncertainties.

Signal yield	Bias
$B^0 \to K_S^0 K^+\pi^-$ (Down-Down)(2011)	1.7 ± 0.5
$B^0 \to K_S^0\pi^+\pi^-$ (Down-Down)(2012a)	3.3 ± 1.4
$B^0 \to K_S^0\pi^+\pi^-$ (Down-Down)(2012b)	5.2 ± 2.0
$B^0 \to K_S^0\pi^+\pi^-$ (Long-Long)(2012a)	1.7 ± 0.8
$B^0 \to K_S^0\pi^+\pi^-$ (Long-Long)(2012b)	3.1 ± 1.4
$B_s^0 \to K_S^0 K^+\pi^-$ (Down-Down)(2011)	1.7 ± 0.7
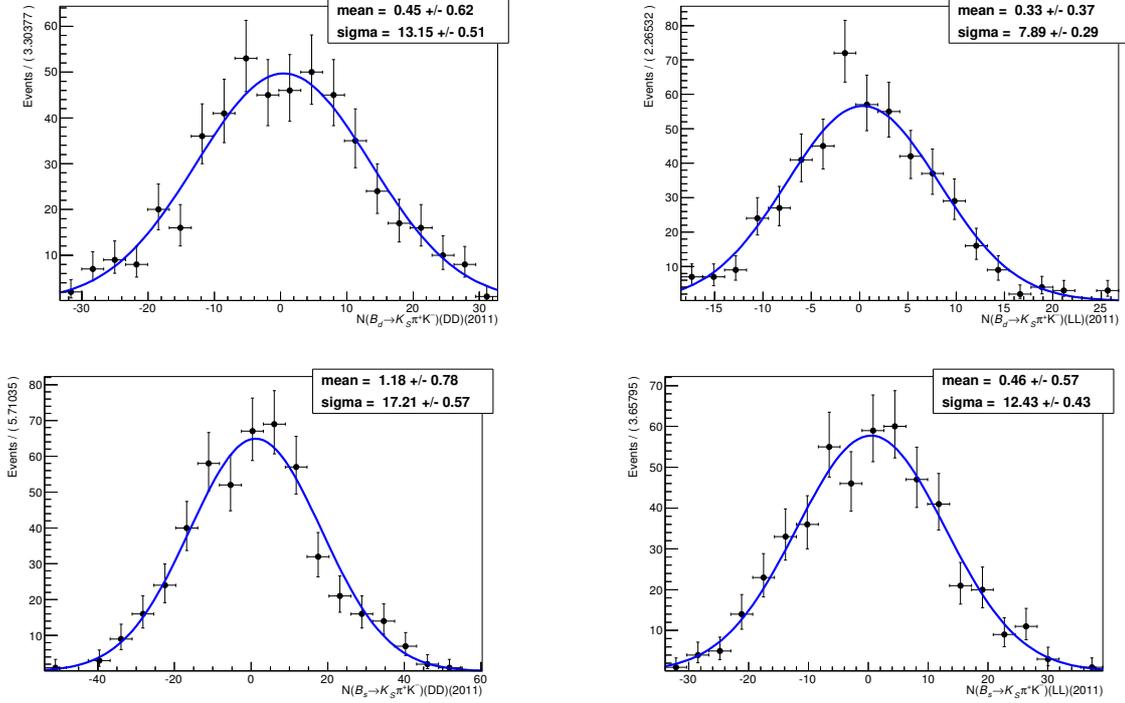
Figure 4.12 – Residuals distributions for yields of $K_S^0\pi^+K^-$ 2011 signals. Top: $B^0 \to K_S^0K^-\pi^+$ signal, bottom: $B_s^0 \to K_S^0K^-\pi^+$ signal. Left: Down-Down, right: Long-Long.

Table 4.5 – Signal yields for which a bias was observed in the tight BDT optimization, along with the bias. They are corrected for in the final results and accounted for in the uncertainties.

Signal yield	Bias
$B^0 \to K_S^0\pi^+\pi^-$ (Down-Down)(2012b)	3.3 ± 1.6
$B_s^0 \to K_S^0K^+K^-$ (Down-Down)(2012b)	2.6 ± 0.7
$B_s^0 \to K_S^0K^+\pi^-$ (Long-Long)(2011)	0.9 ± 0.5
$B_s^0 \to K_S^0K^+\pi^-$ (Long-Long)(2012a)	0.9 ± 0.3
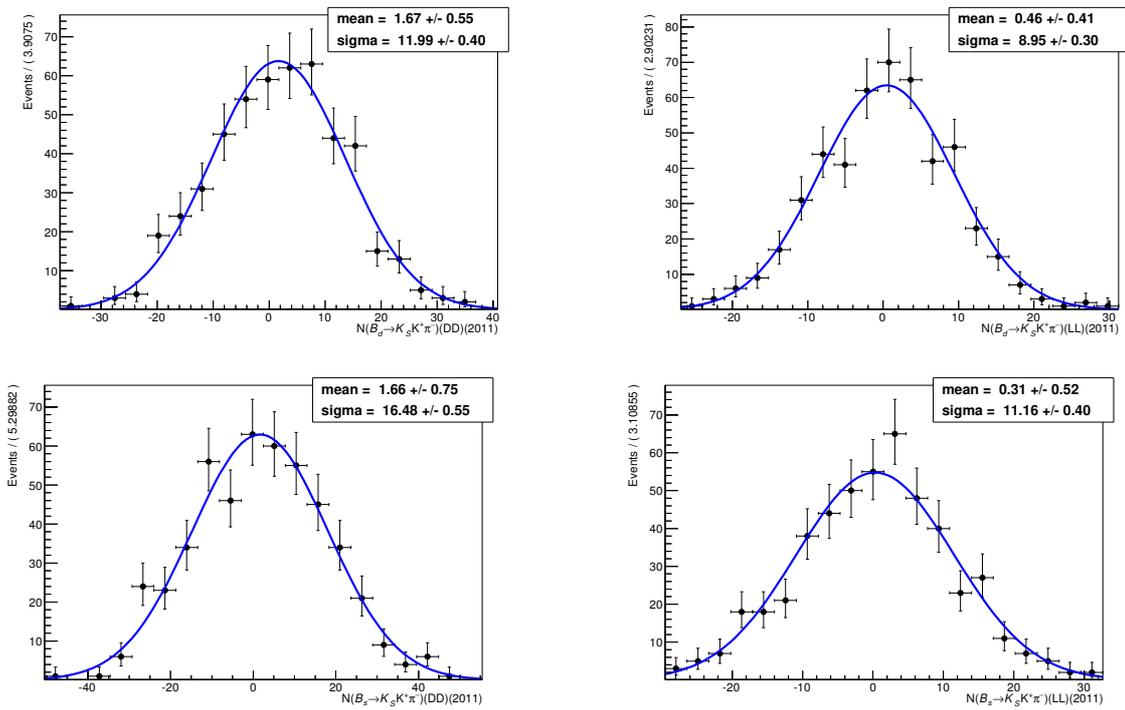
Figure 4.13 – Residuals distributions for yields of $K_S^0 K^+ \pi^-$ 2011 signals. Top: $B^0 \to K_S^0 K^+ \pi^-$ signal, bottom: $B_s^0 \to K_S^0 K^+ \pi^-$ signal. Left: Down-Down, right: Long-Long.
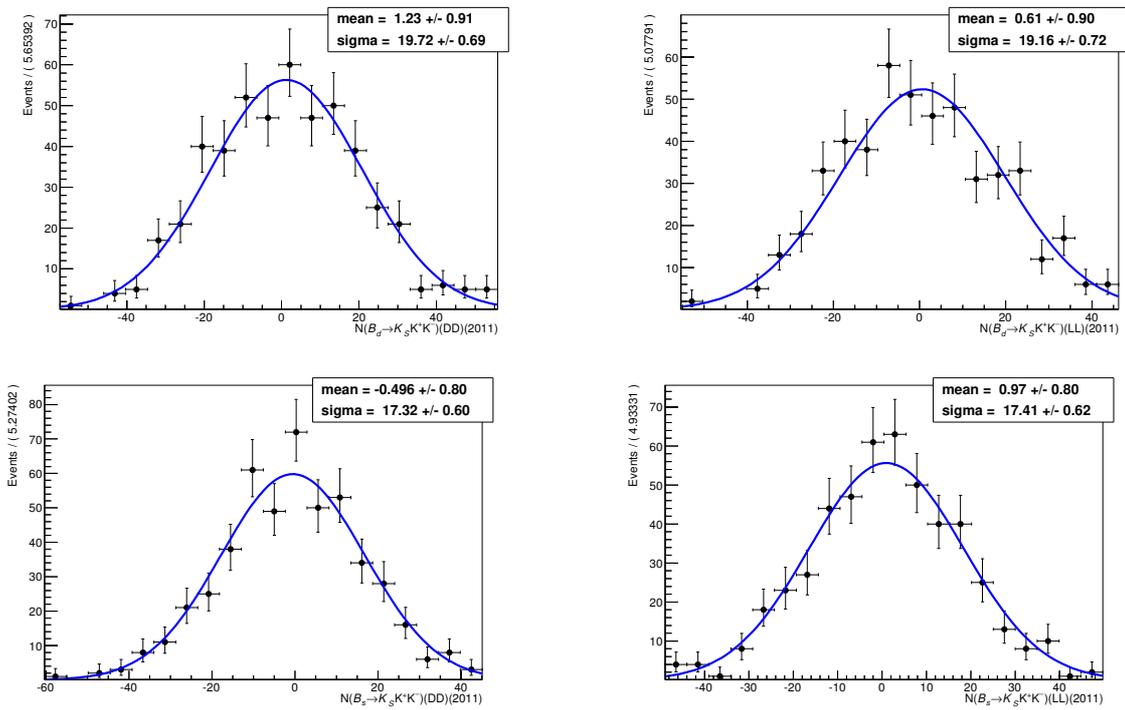
Figure 4.14 – Residuals distributions for yields of $K_S^0 K^+ K^-$ 2011 signals. Top: $B^0 \to K_S^0 K^+ K^-$ signal, bottom: $B_s^0 \to K_S^0 K^+ K^-$ signal. Left: Down-Down, right: Long-Long.

4.6 Estimation of systematic uncertainties

Systematic uncertainties can arise from different sources, such as biases or assumptions made on model parameters. We consider two types of systematic uncertainties related to the fit model.

The first type is related to parameters that are fixed to values determined from simulated events. We extract these systematic uncertainties by performing several hundreds of alternative fits to data, varying all the fixed parameters according to the correlation matrix of the fit to simulated samples. The distribution of differences between yields in the nominal fit and alternative fits is fitted using a Gaussian distribution. The systematic uncertainty on a yield X is then

$$\Delta_X = \sqrt{\left(\frac{\mu}{2}\right)^2 + \sigma^2},\tag{4.28}$$

where μ and σ are the mean and the width of the Gaussian. The fixed parameters of the fit model are:

- signal model: the tail parameters of the CB functions (α_0, n_0, $\frac{\alpha_1}{\alpha_0}$, $\frac{n_1}{n_0}$), and the fraction of the two functions, f;

- partially reconstructed background model: the two parameters of all the ARGUS functions. The threshold is varied within $1\,\mathrm{MeV}/c^2$ of its nominal value;

- cross-feeds model: all the parameters for each considered event species.

The second type of systematic uncertainties related to the fit model originates from the choice of the models used in the nominal fit. Toy experiments are used to estimate the systematics due to these effects: a pseudo-dataset is generated according to the result of the fit of an alternative model to data; the pseudo-dataset is then fitted with both the nominal model and the alternative model. The distribution of the differences of the yields of the two fits is fitted with a Gaussian function. The associated systematic uncertainty is then estimated as in Eq. 4.28.

Both the partially reconstructed background and the cross-feed shapes suffer from a large statistical uncertainty due to small Monte-Carlo samples, and it is believed that the toy exercise described above covers any reasonable variation of the shapes. Hence, the uncertainty due to the choice of the model will be estimated for the signal and combinatorial background models only. We consider the following alternative models:

- signal: the Cruijff distribution, defined as

$$t = m - \mu,$$
$$F(m) = \begin{cases} \exp(-t^2/(2\sigma_{\mathrm{L}}^2 + \alpha_{\mathrm{L}}^2 t^2)) & \text{if } t/\sigma \leq 0 \\ \exp(-t^2/(2\sigma_{\mathrm{R}}^2 + \alpha_{\mathrm{R}}^2 t^2)) & \text{if } t/\sigma > 0 \end{cases}\tag{4.29}$$

is taken as an alternative description of the signal;

Table 4.6 – Systematic uncertainties on signal yields related to fixed parameters of the signal shapes (loose BDT optimization).

	2011		2012 pre-June		2012 post-June	
	DD	LL	DD	LL	DD	LL
$B^0 \to K_S^0 K^+ K^-$	0.3	0.2	0.2	0.1	0.8	0.5
$B^0 \to K_S^0 K^+ \pi^-$	0.1	0.1	0.1	0.1	0.2	0.1
$B^0 \to K_S^0 K^- \pi^+$	0.1	0.1	0.1	0.1	0.2	0.1
$B^0 \to K_S^0 \pi^+ \pi^-$	1.1	0.9	0.8	0.4	1.7	0.9
$B_s^0 \to K_S^0 K^+ K^-$	0.1	0.1	0.0	0.0	0.2	0.2
$B_s^0 \to K_S^0 K^+ \pi^-$	0.3	0.2	0.2	0.1	0.4	0.2
$B_s^0 \to K_S^0 K^- \pi^+$	0.3	0.2	0.2	0.1	0.5	0.3
$B_s^0 \to K_S^0 \pi^+ \pi^-$	0.4	0.3	0.2	0.1	0.6	0.3

Table 4.7 – Systematic uncertainties on signal yields related to fixed parameters of the signal shapes (tight BDT optimization).

	2011		2012 pre-June		2012 post-June	
	DD	LL	DD	LL	DD	LL
$B^0 \to K_S^0 K^+ K^-$	0.2	0.2	0.1	0.1	0.4	0.1
$B^0 \to K_S^0 K^+ \pi^-$	0.1	0.1	0.1	0.0	0.1	0.1
$B^0 \to K_S^0 K^- \pi^+$	0.1	0.1	0.1	0.1	0.2	0.1
$B^0 \to K_S^0 \pi^+ \pi^-$	0.6	0.7	0.6	0.3	1.1	0.5
$B_s^0 \to K_S^0 K^+ K^-$	0.0	0.0	0.0	0.0	0.1	0.0
$B_s^0 \to K_S^0 K^+ \pi^-$	0.2	0.1	0.1	0.1	0.4	0.1
$B_s^0 \to K_S^0 K^- \pi^+$	0.2	0.1	0.1	0.1	0.4	0.1
$B_s^0 \to K_S^0 \pi^+ \pi^-$	0.2	0.2	0.2	0.1	0.5	0.2

- combinatorial background: the exponential distribution as taken as an alternative to the linear shape;

- common parameters in the combinatorial background model: in the nominal model, the ratios of the slopes between data-taking periods and K_S^0 reconstruction modes are constrained. We consider an alternative model where all these constraints are removed.

Tables 4.6 and 4.7 show the estimated systematic uncertainties on signal yields that originate from fixed parameters in the signal shapes. Tables 4.8 and 4.9 show the estimated systematic uncertainties on signal yields that originate from fixed parameters

Table 4.8 – Systematic uncertainties on signal yields related to fixed parameters of the cross-feed shapes (loose BDT optimization).

	2011		2012 pre-June		2012 post-June	
	DD	LL	DD	LL	DD	LL
$B^0 \to K_{\mathrm{S}}^0 K^+ K^-$	0.2	0.4	0.1	0.2	0.7	0.4
$B^0 \to K_{\mathrm{S}}^0 K^+ \pi^-$	0.5	0.3	0.4	0.2	0.8	0.4
$B^0 \to K_{\mathrm{S}}^0 K^- \pi^+$	0.5	0.2	0.4	0.2	0.8	0.4
$B^0 \to K_{\mathrm{S}}^0 \pi^+ \pi^-$	0.9	0.8	1.0	0.4	2.0	0.9
$B_s^0 \to K_{\mathrm{S}}^0 K^+ K^-$	0.1	0.1	0.1	0.1	0.3	0.1
$B_s^0 \to K_{\mathrm{S}}^0 K^+ \pi^-$	0.4	0.2	0.3	0.1	0.6	0.3
$B_s^0 \to K_{\mathrm{S}}^0 K^- \pi^+$	0.4	0.2	0.3	0.1	0.6	0.3
$B_s^0 \to K_{\mathrm{S}}^0 \pi^+ \pi^-$	0.3	0.2	0.2	0.1	0.4	0.2

Table 4.9 – Systematic uncertainties on signal yields related to fixed parameters of the cross-feed shapes (tight BDT optimization).

	2011		2012 pre-June		2012 post-June	
	DD	LL	DD	LL	DD	LL
$B^0 \to K_{\mathrm{S}}^0 K^+ K^-$	0.1	0.2	0.1	0.1	0.6	0.2
$B^0 \to K_{\mathrm{S}}^0 K^+ \pi^-$	0.1	0.1	0.1	0.1	0.3	0.1
$B^0 \to K_{\mathrm{S}}^0 K^- \pi^+$	0.1	0.1	0.1	0.1	0.3	0.1
$B^0 \to K_{\mathrm{S}}^0 \pi^+ \pi^-$	0.2	0.2	0.2	0.1	0.3	0.1
$B_s^0 \to K_{\mathrm{S}}^0 K^+ K^-$	0.1	0.1	0.1	0.0	0.2	0.1
$B_s^0 \to K_{\mathrm{S}}^0 K^+ \pi^-$	0.1	0.1	0.2	0.1	0.2	0.1
$B_s^0 \to K_{\mathrm{S}}^0 K^- \pi^+$	0.1	0.1	0.2	0.1	0.2	0.1
$B_s^0 \to K_{\mathrm{S}}^0 \pi^+ \pi^-$	0.2	0.1	0.1	0.1	0.4	0.1

in the cross-feeds shapes.

Tables 4.10 and 4.11 show the estimated systematic uncertainties on signal yields that originate from fixed parameters in the partially reconstructed background shapes.

Figure 4.15 shows fits to simulated signal events using the Cruijff distribution, and Tables 4.12 and 4.13 show the systematic uncertainties evaluated using this distribution as an alternative signal model.

Figure 4.16 shows fits to 2011, Down-Down data using the exponential shape to model the combinatorial background and the loose BDT optimization. Tables 4.14 and 4.15 summarise the systematic uncertainties associated with the choice of combinatorial shape.

In the fit to data, we factorize the parameters of the combinatorial background shapes
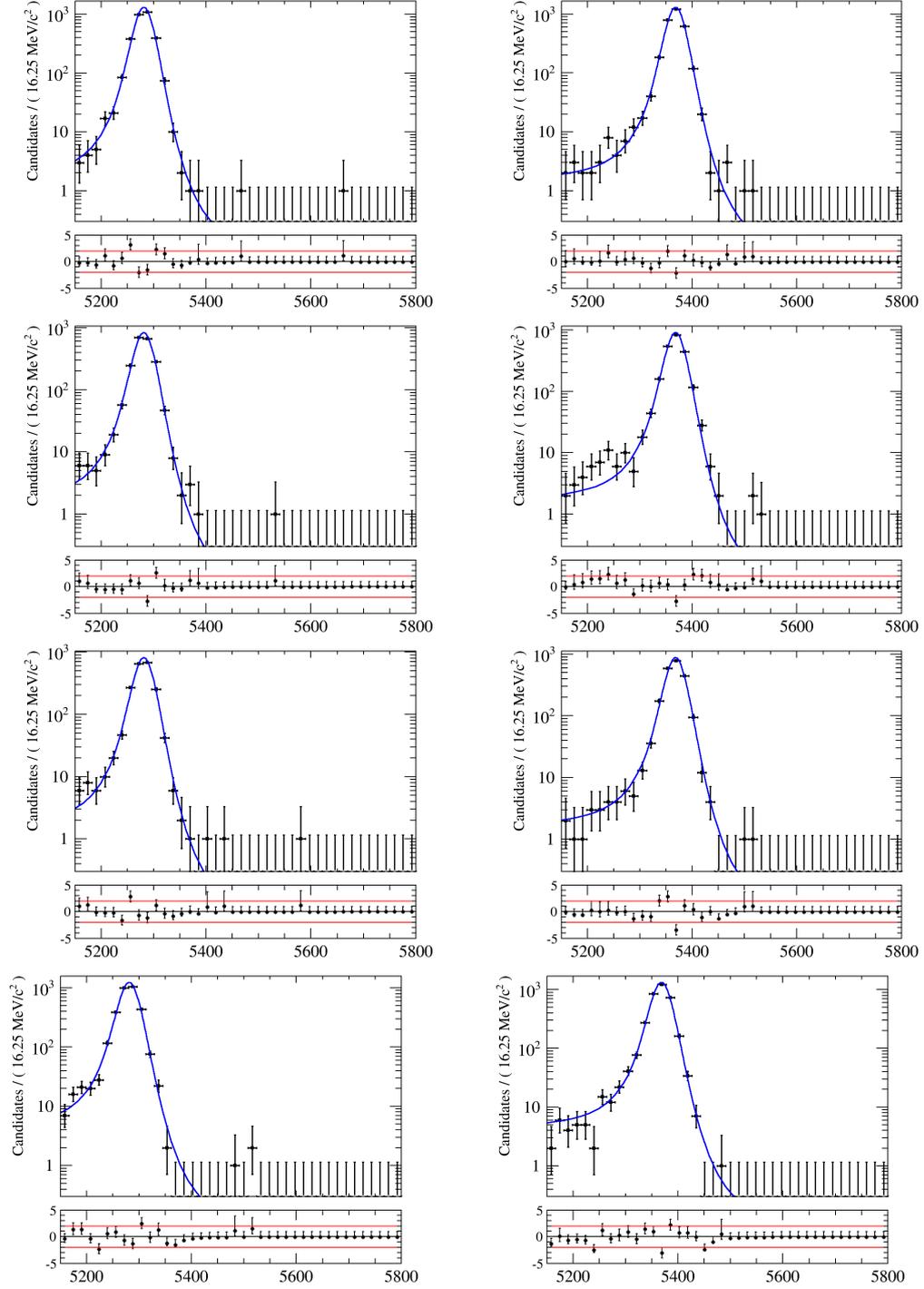
Figure 4.15 – Result of the simultaneous fit of the reconstructed invariant mass on simulated samples of the signal decays (Down-Down), using the loose optimization of the BDT and a Cruijff distribution (logarithmic scale). $K_S^0 K^+ K^-$, $K_S^0 K^\pm \pi^\mp$ and $K_S^0 \pi^+ \pi^-$ are shown from top to bottom, while B^0 decays are shown on the left and B_s^0 decays on the right.

Table 4.10 – Systematic uncertainties on signal yields related to fixed parameters of the partially reconstructed backgrounds shapes (loose BDT optimization).

	2011		2012 pre-June		2012 post-June	
	DD	LL	DD	LL	DD	LL
$B^0 \to K_\mathrm{S}^0 K^+ K^-$	0.1	0.1	0.1	0.1	0.2	0.3
$B^0 \to K_\mathrm{S}^0 K^+ \pi^-$	0.1	0.0	0.0	0.0	0.0	0.0
$B^0 \to K_\mathrm{S}^0 K^- \pi^+$	0.0	0.0	0.0	0.0	0.1	0.0
$B^0 \to K_\mathrm{S}^0 \pi^+ \pi^-$	0.1	0.0	0.1	0.1	0.2	0.1
$B_s^0 \to K_\mathrm{S}^0 K^+ K^-$	0.0	0.0	0.0	0.0	0.0	0.1
$B_s^0 \to K_\mathrm{S}^0 K^+ \pi^-$	0.1	0.1	0.1	0.0	0.1	0.1
$B_s^0 \to K_\mathrm{S}^0 K^- \pi^+$	0.1	0.0	0.0	0.0	0.1	0.1
$B_s^0 \to K_\mathrm{S}^0 \pi^+ \pi^-$	0.0	0.0	0.0	0.0	0.0	0.0

Table 4.11 – Systematic uncertainties on signal yields related to fixed parameters of the partially reconstructed background shapes (tight BDT optimization).

	2011		2012 pre-June		2012 post-June	
	DD	LL	DD	LL	DD	LL
$B^0 \to K_\mathrm{S}^0 K^+ K^-$	0.1	0.0	0.1	0.0	0.1	0.0
$B^0 \to K_\mathrm{S}^0 K^+ \pi^-$	0.0	0.0	0.0	0.0	0.0	0.0
$B^0 \to K_\mathrm{S}^0 K^- \pi^+$	0.0	0.0	0.0	0.0	0.0	0.0
$B^0 \to K_\mathrm{S}^0 \pi^+ \pi^-$	0.1	0.0	0.0	0.0	0.1	0.0
$B_s^0 \to K_\mathrm{S}^0 K^+ K^-$	0.0	0.0	0.0	0.0	0.0	0.0
$B_s^0 \to K_\mathrm{S}^0 K^+ \pi^-$	0.0	0.0	0.0	0.0	0.1	0.1
$B_s^0 \to K_\mathrm{S}^0 K^- \pi^+$	0.0	0.0	0.0	0.0	0.0	0.0
$B_s^0 \to K_\mathrm{S}^0 \pi^+ \pi^-$	0.0	0.0	0.0	0.0	0.0	0.0

in order to constrain the ratio of their value between years and between K_S^0 reconstruction modes

$$\alpha(K_\mathrm{S}^0 h^{\pm} h^{'\mp})(\text{period})(K_\mathrm{S}^0 \text{ mode}) = k_{\mathrm{mode}} k_{\mathrm{period}} \alpha(K_\mathrm{S}^0 h^{\pm} h^{'\mp})(2011)(\text{DD}). \qquad (4.30)$$

We evaluate the systematic uncertainty associated with this choice by removing these constraints. Tables 4.16 and 4.17 show this systematic uncertainty on each signal yield.

Table 4.12 – Systematic uncertainties related to the choice of the shape of the signal distribution (loose BDT optimization).

	2011		2012 pre-June		2012 post-June	
	DD	LL	DD	LL	DD	LL
$B^0 \to K_S^0 K^+ K^-$	2.2	1.3	1.2	0.8	4.2	1.4
$B^0 \to K_S^0 K^+ \pi^-$	4.4	2.3	2.8	1.6	5.5	2.6
$B^0 \to K_S^0 K^- \pi^+$	4.4	1.9	2.7	1.6	5.9	2.7
$B^0 \to K_S^0 \pi^+ \pi^-$	10.8	4.1	5.1	3.0	14.4	5.8
$B_s^0 \to K_S^0 K^+ K^-$	0.4	0.4	0.3	0.2	0.7	0.3
$B_s^0 \to K_S^0 K^+ \pi^-$	6.8	3.8	4.4	2.6	8.0	3.4
$B_s^0 \to K_S^0 K^- \pi^+$	4.9	2.3	2.7	1.6	6.6	2.9
$B_s^0 \to K_S^0 \pi^+ \pi^-$	3.4	1.4	1.8	1.1	5.0	2.3

Table 4.13 – Systematic uncertainties related to the choice of the shape of the signal distribution (tight BDT optimization).

	2011		2012 pre-June		2012 post-June	
	DD	LL	DD	LL	DD	LL
$B^0 \to K_S^0 K^+ K^-$	0.9	2.3	1.1	0.7	1.6	1.5
$B^0 \to K_S^0 K^+ \pi^-$	2.8	1.7	1.4	1.1	4.1	1.2
$B^0 \to K_S^0 K^- \pi^+$	2.8	1.4	1.3	1.2	4.5	1.2
$B^0 \to K_S^0 \pi^+ \pi^-$	7.3	3.9	3.8	2.5	12.1	4.4
$B_s^0 \to K_S^0 K^+ K^-$	0.2	0.8	0.3	0.5	0.3	0.7
$B_s^0 \to K_S^0 K^+ \pi^-$	4.9	2.9	2.8	1.8	5.8	1.7
$B_s^0 \to K_S^0 K^- \pi^+$	3.3	1.8	1.5	1.2	5.4	1.4
$B_s^0 \to K_S^0 \pi^+ \pi^-$	2.6	1.5	1.4	1.0	4.6	2.0

4.6.1 Total uncertainties on yields

Tables 4.18, 4.19, and 4.20 show the uncertainties on yield parameters of the signal for 2011, 2012 pre-June, and 2012 post-June, respectively. The uncertainties are dominated by the statistical uncertainty and the systematic uncertainty related to the combinatorial shape. In parallel to the increase of the size of datasets, next iterations of this analysis will thus have to scrutinise the modelling of combinatorial background. The study of same-sign data samples, formed of $K_S^0 h^+ h'^+$ events, would improve this modelling and reduce the systematic uncertainties related to the combinatorial background.

Table 4.14 – Systematic uncertainties on signal yields related to the choice of combinatorial background shape (loose BDT optimization).

| | 2011 | | 2012 pre-June | | 2012 post-June | |
	DD	LL	DD	LL	DD	LL
$B^0 \to K_S^0 K^+ K^-$	1.8	2.3	1.5	1.3	2.4	1.6
$B^0 \to K_S^0 K^+ \pi^-$	2.5	1.3	1.6	1.1	3.4	1.6
$B^0 \to K_S^0 K^- \pi^+$	2.5	1.2	1.5	1.1	3.6	1.7
$B^0 \to K_S^0 \pi^+ \pi^-$	7.5	4.4	5.1	3.1	11.1	5.3
$B_s^0 \to K_S^0 K^+ K^-$	0.4	0.5	0.3	0.1	0.9	0.3
$B_s^0 \to K_S^0 K^+ \pi^-$	4.8	2.8	3.5	2.0	6.2	2.5
$B_s^0 \to K_S^0 K^- \pi^+$	2.9	1.3	1.8	1.1	4.2	1.8
$B_s^0 \to K_S^0 \pi^+ \pi^-$	3.9	1.9	2.5	1.2	5.6	2.6

Table 4.15 – Systematic uncertainties on signal yields related to the choice of combinatorial background shape (tight BDT optimization).

| | 2011 | | 2012 pre-June | | 2012 post-June | |
	DD	LL	DD	LL	DD	LL
$B^0 \to K_S^0 K^+ K^-$	1.6	2.0	1.9	1.2	2.1	5.0
$B^0 \to K_S^0 K^+ \pi^-$	1.9	1.3	1.0	1.1	2.9	0.9
$B^0 \to K_S^0 K^- \pi^+$	1.8	1.1	0.9	1.2	3.3	0.9
$B^0 \to K_S^0 \pi^+ \pi^-$	5.8	5.7	4.4	3.0	9.0	5.0
$B_s^0 \to K_S^0 K^+ K^-$	0.3	0.3	0.4	0.2	0.4	0.8
$B_s^0 \to K_S^0 K^+ \pi^-$	3.5	2.6	1.9	1.7	5.5	3.0
$B_s^0 \to K_S^0 K^- \pi^+$	1.9	1.1	0.9	0.8	3.7	1.7
$B_s^0 \to K_S^0 \pi^+ \pi^-$	2.4	2.3	1.7	1.0	4.7	2.1
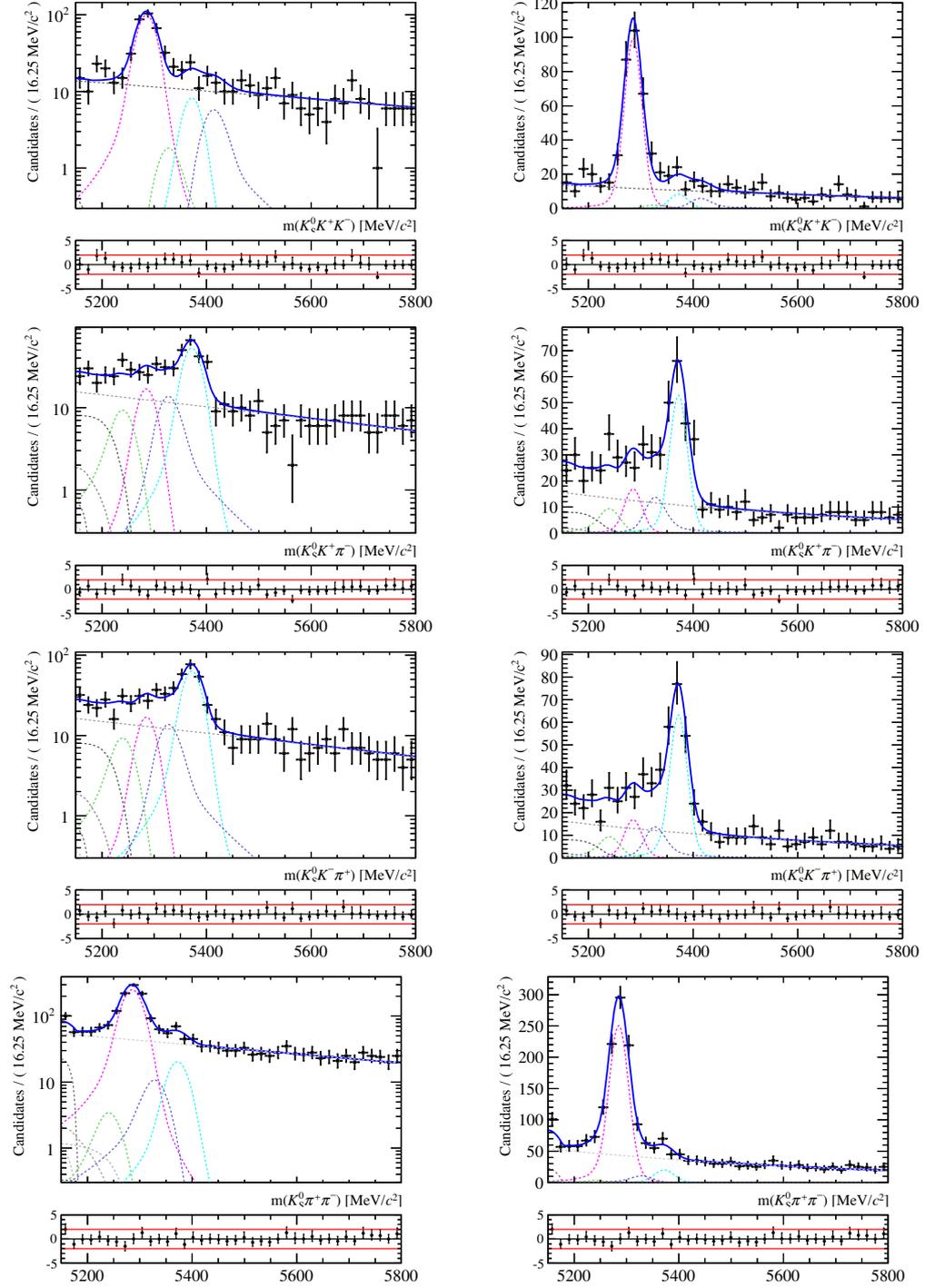
Figure 4.16 – Result of the simultaneous fit of the data (Down-Down, 2011) with the loose BDT optimization. $K_S^0 K^+ K^-$, $K_S^0 K^\pm \pi^\mp$, $K_S^0 \pi^\pm K^\mp$ and $K_S^0 \pi^+ \pi^-$ are shown from top to bottom, while the left plots show the result on a linear scale and the right on a logarithmic scale.

Table 4.16 – Systematic uncertainties on signal yields related to the choice of combinatorial background model (loose BDT optimization).

| | 2011 | | 2012 pre-June | | 2012 post-June | |
	DD	LL	DD	LL	DD	LL
$B^0 \to K_S^0 K^+ K^-$	3.5	2.7	2.5	1.9	7.3	2.6
$B^0 \to K_S^0 K^+ \pi^-$	4.0	1.7	3.6	1.5	4.4	3.2
$B^0 \to K_S^0 K^- \pi^+$	4.1	1.5	3.4	1.6	4.8	3.2
$B^0 \to K_S^0 \pi^+ \pi^-$	13.2	4.4	7.1	4.6	15.2	10.9
$B_s^0 \to K_S^0 K^+ K^-$	1.3	1.2	1.0	0.7	2.7	0.8
$B_s^0 \to K_S^0 K^+ \pi^-$	6.1	3.1	5.1	2.6	7.0	3.7
$B_s^0 \to K_S^0 K^- \pi^+$	4.0	1.7	3.0	1.4	4.7	3.0
$B_s^0 \to K_S^0 \pi^+ \pi^-$	5.4	2.1	3.0	1.9	6.3	5.1

Table 4.17 – Systematic uncertainties on signal yields related to the choice of combinatorial background model (tight BDT optimization).

| | 2011 | | 2012 pre-June | | 2012 post-June | |
	DD	LL	DD	LL	DD	LL
$B^0 \to K_S^0 K^+ K^-$	2.9	10.0	2.3	1.6	3.9	9.8
$B^0 \to K_S^0 K^+ \pi^-$	3.5	1.8	1.5	1.1	4.7	1.4
$B^0 \to K_S^0 K^- \pi^+$	3.3	1.4	1.4	1.1	5.2	1.4
$B^0 \to K_S^0 \pi^+ \pi^-$	8.3	4.0	4.3	3.3	14.3	6.1
$B_s^0 \to K_S^0 K^+ K^-$	1.0	4.4	0.8	0.7	1.2	3.9
$B_s^0 \to K_S^0 K^+ \pi^-$	5.3	2.8	2.7	1.8	6.2	2.1
$B_s^0 \to K_S^0 K^- \pi^+$	3.3	1.5	1.4	1.1	5.3	1.6
$B_s^0 \to K_S^0 \pi^+ \pi^-$	3.4	1.7	1.7	1.4	6.2	2.9

Table 4.18 – Signal yields for 2011 data samples, along with uncertainties (loose BDT optimization). Comb. shape and Comb. model refer to systematic uncertainties related to the combinatorial background shape and to common parameters in the combinatorial background model, respectively.

	Yield	Stat	Bias	Sig. model	CF. model	PartRec	Comb. model	Comb. shape	Sig. shape	Total
$B^0 \to K_{\mathrm{S}}^0 \pi^+ \pi^-$	803.0	35.6	0.0	1.1	0.9	0.1	20.1	1.0	16.0	44.0
$B^0 \to K_{\mathrm{S}}^0 \pi^+ \pi^-$	471.3	26.6	0.0	0.9	0.8	0.0	1.0	4.5	5.1	27.5
$B^0 \to K_{\mathrm{S}}^0 K^- \pi^+$	52.4	12.4	0.0	0.1	0.5	0.0	6.4	1.8	7.9	16.2
$B^0 \to K_{\mathrm{S}}^0 K^- \pi^+$	26.2	7.3	0.0	0.1	0.3	0.0	1.5	0.8	3.2	8.2
$B^0 \to K_{\mathrm{S}}^0 K^+ \pi^-$	52.4	12.3	0.8	0.2	0.5	0.1	6.3	1.7	7.7	16.0
$B^0 \to K_{\mathrm{S}}^0 K^+ \pi^-$	37.5	8.3	0.0	0.1	0.3	0.0	1.6	0.7	3.9	9.4
$B^0 \to K_{\mathrm{S}}^0 K^+ K^-$	281.1	19.2	0.0	0.3	0.2	0.2	5.7	1.7	3.5	20.4
$B^0 \to K_{\mathrm{S}}^0 K^+ K^-$	222.4	16.9	0.0	0.2	0.4	0.2	1.8	3.3	1.0	17.4
$B_s^0 \to K_{\mathrm{S}}^0 \pi^+ \pi^-$	65.4	18.1	0.0	0.4	0.3	0.0	8.0	6.2	4.5	21.2
$B_s^0 \to K_{\mathrm{S}}^0 \pi^+ \pi^-$	23.2	12.0	0.0	0.3	0.2	0.0	1.0	2.6	1.3	12.4
$B_s^0 \to K_{\mathrm{S}}^0 K^- \pi^+$	181.0	16.9	0.0	0.3	0.4	0.1	6.5	4.1	8.9	20.6
$B_s^0 \to K_{\mathrm{S}}^0 K^- \pi^+$	115.5	11.9	0.0	0.2	0.2	0.1	1.8	1.5	3.7	12.7
$B_s^0 \to K_{\mathrm{S}}^0 K^+ \pi^-$	152.2	15.1	0.8	0.3	0.4	0.2	7.6	4.5	10.5	20.4
$B_s^0 \to K_{\mathrm{S}}^0 K^+ \pi^-$	91.2	10.6	0.0	0.2	0.2	0.1	2.7	3.1	5.6	12.7
$B_s^0 \to K_{\mathrm{S}}^0 K^+ K^-$	22.8	9.4	0.0	0.1	0.1	0.1	2.0	0.5	0.2	9.6
$B_s^0 \to K_{\mathrm{S}}^0 K^+ K^-$	6.8	8.1	0.0	0.1	0.1	0.0	1.3	0.7	0.1	8.2

Table 4.19 – Signal yields for 2012 pre-June data samples, along with uncertainties (loose BDT optimization). Comb. shape and Comb. model refer to systematic uncertainties related to the combinatorial background shape and to common parameters in the combinatorial background model, respectively.

	Yield	Stat	Bias	Sig. model	CF. model	PartRec	Comb. model	Comb. shape	Sig. shape	Total
$B^0 \to K_S^0 \pi^+ \pi^-$	553.1	30.3	1.7	0.9	1.0	0.2	3.3	2.5	4.5	31.0
$B^0 \to K_S^0 \pi^+ \pi^-$	286.4	19.5	0.8	0.4	0.4	0.1	3.2	1.5	1.9	19.9
$B^0 \to K_S^0 K^- \pi^+$	46.9	10.7	0.0	0.1	0.4	0.1	5.9	0.5	4.7	13.1
$B^0 \to K_S^0 K^- \pi^+$	21.2	7.8	0.0	0.1	0.2	0.0	2.1	0.3	2.6	8.5
$B^0 \to K_S^0 K^+ \pi^-$	43.9	11.0	0.0	0.1	0.4	0.0	6.3	0.5	4.9	13.6
$B^0 \to K_S^0 K^+ \pi^-$	28.7	7.8	0.0	0.1	0.2	0.1	1.9	0.3	2.5	8.5
$B^0 \to K_S^0 K^+ K^-$	180.7	15.3	0.0	0.2	0.1	0.1	0.8	1.8	1.5	15.5
$B^0 \to K_S^0 K^+ K^-$	119.2	11.8	0.0	0.1	0.2	0.1	1.6	1.9	0.5	12.1
$B_s^0 \to K_S^0 \pi^+ \pi^-$	16.5	15.1	0.0	0.2	0.2	0.1	1.1	3.7	1.2	15.6
$B_s^0 \to K_S^0 \pi^+ \pi^-$	15.4	7.7	0.0	0.1	0.1	0.0	1.4	1.3	0.7	8.0
$B_s^0 \to K_S^0 K^- \pi^+$	112.9	13.7	0.0	0.2	0.3	0.1	5.1	2.5	4.7	15.6
$B_s^0 \to K_S^0 K^- \pi^+$	48.6	9.1	0.0	0.1	0.1	0.0	2.0	1.3	2.5	9.7
$B_s^0 \to K_S^0 K^+ \pi^-$	92.0	12.1	0.0	0.2	0.3	0.1	7.5	2.3	5.7	15.5
$B_s^0 \to K_S^0 K^+ \pi^-$	50.8	8.5	0.0	0.1	0.1	0.0	2.7	0.7	3.1	9.5
$B_s^0 \to K_S^0 K^+ K^-$	1.9	6.4	0.0	0.0	0.1	0.0	0.6	0.3	0.2	6.4
$B_s^0 \to K_S^0 K^+ K^-$	6.0	4.9	0.0	0.0	0.1	0.0	0.8	0.1	0.1	4.9

Table 4.20 – Signal yields for 2012 post-June data samples, along with uncertainties (loose BDT optimization). Comb1 and Comb2 refer to systematic uncertainties related to the combinatorial background shape and to common parameters in the combinatorial background model, respectively.

	Yield	Stat	Bias	Sig. model	CF. model	PartRec	Comb. model	Comb. shape	Sig. shape	Total
$B^0 \to K_{\rm S}^0 \pi^+ \pi^-$	1409.5	46.0	2.6	1.7	2.0	0.2	13.6	7.7	20.7	52.9
$B^0 \to K_{\rm S}^0 \pi^+ \pi^-$	653.6	30.1	1.5	0.9	0.9	0.1	17.9	1.8	7.1	35.9
$B^0 \to K_{\rm S}^0 K^- \pi^+$	90.6	15.6	0.0	0.2	0.8	0.1	5.6	3.7	10.4	19.9
$B^0 \to K_{\rm S}^0 K^- \pi^+$	55.7	10.5	0.0	0.1	0.4	0.1	5.3	0.9	4.6	12.6
$B^0 \to K_{\rm S}^0 K^+ \pi^-$	72.9	14.1	0.0	0.2	0.8	0.1	5.3	3.5	9.5	18.2
$B^0 \to K_{\rm S}^0 K^+ \pi^-$	29.8	9.0	0.0	0.1	0.4	0.0	5.3	1.0	4.4	11.4
$B^0 \to K_{\rm S}^0 K^+ K^-$	671.4	29.9	0.0	0.8	0.7	0.3	10.1	0.2	6.8	32.3
$B^0 \to K_{\rm S}^0 K^+ K^-$	343.8	19.8	0.0	0.7	0.4	0.6	0.4	1.9	1.8	20.0
$B_s^0 \to K_{\rm S}^0 \pi^+ \pi^-$	83.3	21.6	0.0	0.6	0.4	0.0	5.4	8.5	6.7	24.8
$B_s^0 \to K_{\rm S}^0 \pi^+ \pi^-$	41.7	13.9	0.0	0.3	0.2	0.0	8.4	3.3	2.6	16.8
$B_s^0 \to K_{\rm S}^0 K^- \pi^+$	306.8	22.1	0.0	0.5	0.6	0.1	5.7	6.1	11.6	26.3
$B_s^0 \to K_{\rm S}^0 K^- \pi^+$	143.4	13.9	0.0	0.3	0.3	0.2	5.0	2.1	5.0	15.8
$B_s^0 \to K_{\rm S}^0 K^+ \pi^-$	255.3	19.2	0.0	0.4	0.6	0.2	6.1	7.1	12.0	24.5
$B_s^0 \to K_{\rm S}^0 K^+ \pi^-$	118.2	12.8	0.0	0.3	0.3	0.2	5.5	2.3	5.0	15.0
$B_s^0 \to K_{\rm S}^0 K^+ K^-$	25.5	12.9	0.0	0.2	0.3	0.1	3.8	1.3	0.6	13.6
$B_s^0 \to K_{\rm S}^0 K^+ K^-$	7.5	6.5	0.0	0.3	0.1	0.2	0.0	0.3	0.1	6.6

4.7 Modelling the signal distribution over the Dalitz plot using $_s\mathcal{P}lots$

The efficiency of signal events is estimated using Monte-Carlo samples, and varies across the Dalitz plane. As discussed in Sec. 4.1, we need to estimate the distribution of signal events over the Dalitz plane in order to properly correct the signal yields for the efficiencies. For this purpose, we use the $_s\mathcal{P}lots$ method described in Sec. 4.2.3.

A special fit is performed in order to extract these $_s\mathcal{P}lots$. Firstly, the mass interval limit on the left is taken as $5200\,\mathrm{MeV}/c^2$ instead of $5150\,\mathrm{MeV}/c^2$ in order to reduce the impact of the partially reconstructed background. Furthermore, the $_s\mathcal{P}lots$ method does not include cases where Gaussian constraints are present in the model. Yields of cross-feeds and partially reconstructed backgrounds are thus fixed to the value obtained from the nominal fit. As partially reconstructed backgrounds are negligible in the 5200–$5800\,\mathrm{MeV}/c^2$ invariant-mass window, we only consider the effect of cross-feeds when correcting the $_s\mathcal{P}lots$ following Eq. 4.16.

The distributions of signal and cross-feeds events over the square Dalitz plane depend on each other, and must thus be determined simultaneously. Indeed, by definition the distribution $M_n(\mathrm{m}'_m, \theta'_m)$ of events from a signal mode n over the square Dalitz plane corresponding to the signal mode m is

$$M_n(\mathrm{m}'_m, \theta'_m) = sM_n(\mathrm{m}'_m, \theta'_m) = \sum sP_n(e) + c_n . M_{0,n}(\mathrm{m}'_m, \theta'_m), \qquad (4.31)$$

where c_n is the parameter defined in Eq. 4.16, and $M_{0,n}(\mathrm{m}'_m, \theta'_m)$ is the estimated distribution of the cross-feeds events that contribute to the invariant-mass spectrum of the signal n over the square Dalitz plot corresponding to the signal mode m. As there are cross-feeds in each invariant-mass spectrum, we use the following iterative procedure:

- all distributions $M_{0,n}^i(\mathrm{m}'_m, \theta'_m)$ are set to 0 for $i = 0$;

- for each step $i > 0$, we extract the distribution of each signal species n over the square Dalitz plane corresponding to the signal mode m using

$$N_n . sM_n^i(\mathrm{m}'_m, \theta'_m) . \delta\,\mathrm{m}'_m \delta\theta'_m = \sum sP_n(e) + c_n . M_{0,n}^{i-1}(\mathrm{m}'_m, \theta'_m), \qquad (4.32)$$

  where N_n, c_n, and $sP_n(e)$ are the same variables as defined in Eq. 4.16, sM_n^i is the estimated distribution of the signal n over the Dalitz plane for iteration i, and $M_{0,n}^{i-1}$ is the Dalitz-plot distribution estimated in iteration $i - 1$ of all the cross-feeds that contribute to the invariant-mass spectrum of the signal mode n;

- the procedure is stopped when the χ^2 calculated between the sM_n^{i+1} and sM_n^i distributions reaches a predetermined lower threshold for each n and m. In practice, the convergence of the procedure is fast and this threshold is reached at $i = 2$ or $i = 3$.

Figure 4.17 shows examples of distributions of sWeighted signal events over the square Dalitz plot. Table 4.21 shows the total efficiencies of all signal modes for the loose and tight optimizations, except for $B_s^0 \to K_S^0 K^+ K^-$.
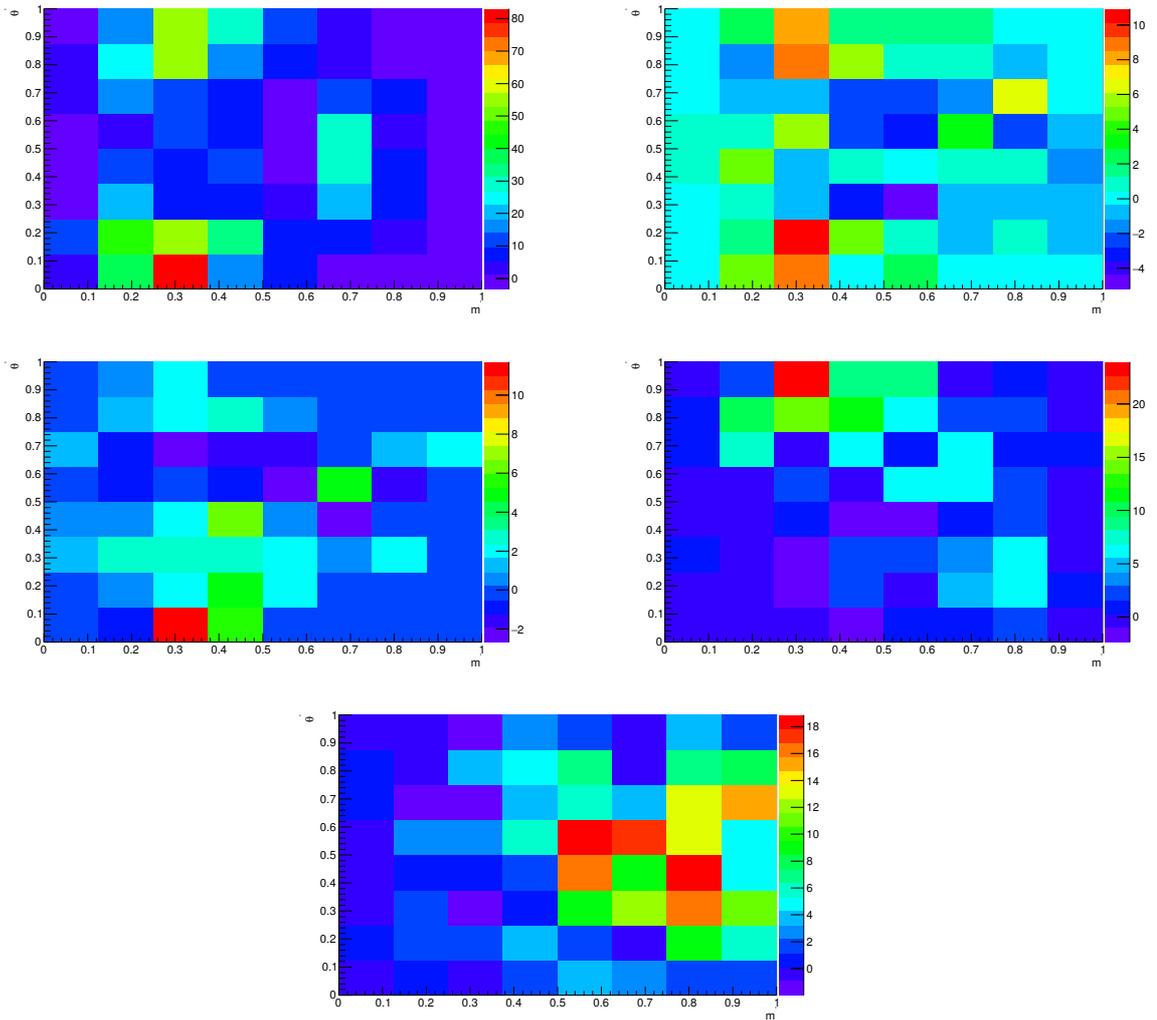
Figure 4.17 – Distributions of sWeighted signal events over the Dalitz plot for $B^0 \to K^0_S \pi^+ \pi^-$ (top,left), $B^0_s \to K^0_S \pi^+ \pi^-$ (top, right), $B^0 \to K^0_S K^+ \pi^-$ (middle, left), $B^0_s \to K^0_S K^+ \pi^-$ (middle, right), and $B^0 \to K^0_S K^+ K^-$ (bottom), in 2011 Down-Down data samples.

4.8 Measurement of the branching fractions

4.8.1 Internal consistency

The measurement of a physical observable such as a branching fraction does not depend on K^0_S reconstruction mode or data-taking period. In order to check the internal consistency of the model, we compare ratios of yields in different categories, corrected for efficiencies. Table 4.22 shows the ratios of the different modes with respect to $B^0 \to K^0_S \pi^+ \pi^-$. They are obtained separately for the two K^0_S reconstruction modes and the three data-taking periods. The agreement between data categories as indicated by the χ^2 of the combination

102

Table 4.21 – Integrated signal efficiencies, using distributions obtained from sWeights in units of 10^{-4}.

Signal channel (selection)	2011		2012 pre-June		2012 post-June	
	DD	LL	DD	LL	DD	LL
$B^0 \to K^0_S \pi^+ \pi^-$ (loose)	4.90	1.89	4.25	1.66	4.57	1.62
$B^0_s \to K^0_S K^\pm \pi^\mp$ (loose)	3.62	1.48	3.23	1.14	4.02	1.39
$B^0 \to K^0_S K^+ K^-$ (loose)	2.93	1.57	2.64	1.50	3.96	1.29
$B^0_s \to K^0_S \pi^+ \pi^-$ (tight)	3.18	1.75	4.98	1.46	3.36	1.08
$B^0 \to K^0_S K^\pm \pi^\mp$ (tight)	3.24	1.60	4.17	1.23	4.05	1.00

Table 4.22 – Measured ratios of branching fractions corresponding to different data categories. The denominator is the branching fraction of the $B^0 \to K^0_S \pi^+ \pi^-$ mode. Quoted uncertainties include statistical uncertainties on yields and efficiencies, along with uncertainties on f_s/f_d. The χ^2 of the combination is indicated for each channel.

Branching fraction	2011		2012 pre-June		2012 post-June		χ^2
	DD	LL	DD	LL	DD	LL	
$B^0 \to K^0_S \pi^+ \pi^-$	1.0	1.0	1.0	1.0	1.0	1.0	
$B^0_s \to K^0_S \pi^+ \pi^-$	$0.49 \pm$ 0.14	$0.23 \pm$ 0.12	$0.10 \pm$ 0.09	$0.23 \pm$ 0.13	$0.31 \pm$ 0.09	$0.38 \pm$ 0.13	6.8
$B^0 \to K^0_S K^\pm \pi^\mp$	$0.20 \pm$ 0.04	$0.15 \pm$ 0.03	$0.23 \pm$ 0.04	$0.20 \pm$ 0.05	$0.13 \pm$ 0.02	$0.22 \pm$ 0.04	9.6
$B^0_s \to K^0_S K^\pm \pi^\mp$	$2.25 \pm$ 0.26	$2.23 \pm$ 0.28	$1.59 \pm$ 0.21	$1.50 \pm$ 0.24	$1.81 \pm$ 0.18	$1.86 \pm$ 0.22	12.5
$B^0 \to K^0_S K^+ K^-$	$0.62 \pm$ 0.05	$0.60 \pm$ 0.06	$0.55 \pm$ 0.06	$0.49 \pm$ 0.06	$0.58 \pm$ 0.03	$0.70 \pm$ 0.05	7.2

shows some tensions, although the global p-value remains above the percent level. The largest χ^2/ndf (2.1) is observed for $B^0_s \to K^0_S K^\pm \pi^\mp$, and there is no clear trend across the different channels and data categories.

4.8.2   Combination of branching fractions

The measurements of branching fractions from each data-taking period and K^0_S reconstruction modes are combined. The central value is the average of all measurements weighted by

$$w_i = \frac{1}{\sigma_i^2},$$
(4.33)

where w_i is the weight associated by the measurement in category i and σ_i is the total uncertainty associated with that measurement.

The total uncertainty of the combination is obtained by propagating the total uncertainties of each measurement, excluding the systematic uncertainty related to f_s/f_d, which is 100% correlated between all data categories. The statistical uncertainty of the combination is computed under the hypothesis of the absence of a systematic uncertainty, and the total systematic uncertainty is evaluated as

$$\Delta_{\mathrm{sys}} = \sqrt{\Delta_{\mathrm{tot}}^2 - \Delta_{\mathrm{stat}}^2}, \tag{4.34}$$

where Δ_{tot} and Δ_{stat} are the total and statistical uncertainties of the combination, respectively.

The ratios of branching fractions for each previously observed mode are

$$\frac{\mathcal{B}\left(B_s^0 \to K_{\mathrm{S}}^0 \pi^+ \pi^-\right)}{\mathcal{B}\left(B^0 \to K_{\mathrm{S}}^0 \pi^+ \pi^-\right)} = 0.26 \pm 0.04(\mathrm{stat.}) \pm 0.02(\mathrm{syst.}) \pm 0.01(\mathrm{f_s/f_d}),$$

$$\frac{\mathcal{B}\left(B^0 \to K_{\mathrm{S}}^0 K^\pm \pi^\mp\right)}{\mathcal{B}\left(B^0 \to K_{\mathrm{S}}^0 \pi^+ \pi^-\right)} = 0.17 \pm 0.02(\mathrm{stat.}) \pm 0.00(\mathrm{syst.}),$$

$$\frac{\mathcal{B}\left(B_s^0 \to K_{\mathrm{S}}^0 K^\pm \pi^\mp\right)}{\mathcal{B}\left(B^0 \to K_{\mathrm{S}}^0 \pi^+ \pi^-\right)} = 1.84 \pm 0.07(\mathrm{stat.}) \pm 0.02(\mathrm{syst.}) \pm 0.04(\mathrm{f_s/f_d}),$$

$$\frac{\mathcal{B}\left(B^0 \to K_{\mathrm{S}}^0 K^+ K^-\right)}{\mathcal{B}\left(B^0 \to K_{\mathrm{S}}^0 \pi^+ \pi^-\right)} = 0.59 \pm 0.02(\mathrm{stat.}) \pm 0.01(\mathrm{syst.}),$$

$$\tag{4.35}$$

4.8.3 $B_s^0 \to K_{\mathrm{S}}^0 K^+ K^-$ observation significance

The significance of the observation of the $B_s^0 \to K_{\mathrm{S}}^0 K^+ K^-$ mode is derived from a likelihood scan of each $B_s^0 \to K_{\mathrm{S}}^0 K^+ K^-$ yield in the loose BDT optimization. The distribution of likelihood is fitted using a bifurcated Gaussian, smeared by a Gaussian to account for systematic uncertainties. Figure 4.18 shows these likelihood scans for each data-taking period and K_{S}^0 reconstruction mode.

In order to estimate the significance in each category, we evaluate the difference between the log-likelihood of the nominal fit and that of a fit where the branching fraction is set to 0 (this difference is referred to as "likelihood ratio" and can be directly read from the likelihood scans of Fig. 4.18).

Table 4.23 shows the significances derived from likelihood ratios in each of the corresponding scans. These significances are then summed in quadrature to obtain the global significance of the $B_s^0 \to K_{\mathrm{S}}^0 K^+ K^-$ observation.[4]

---

[4]Statistical correlations would make this approach invalid. They are however evaluated as below the percent level, and are thus ignored.
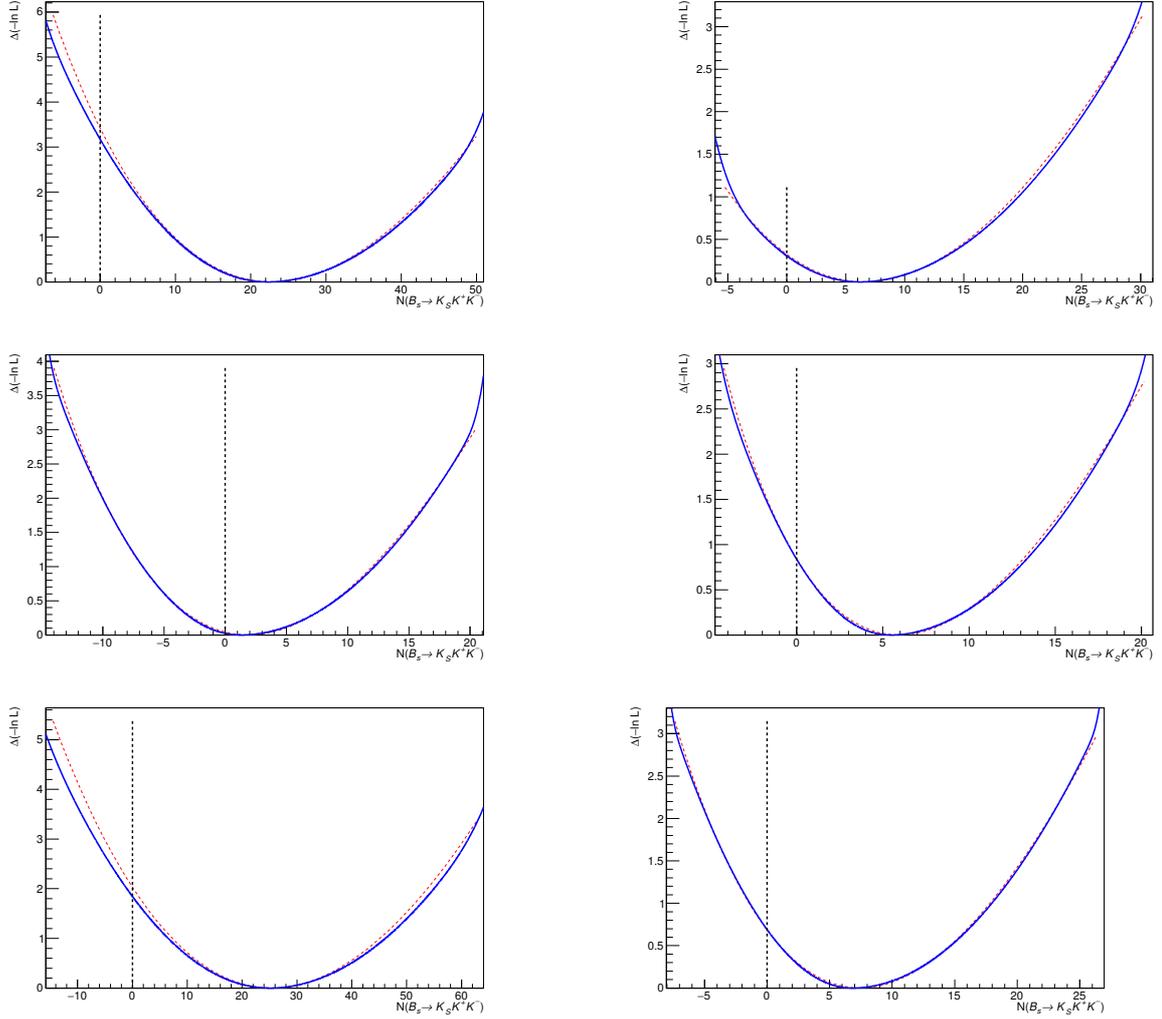
Figure 4.18 – Likelihood scans of $B_s^0 \to K_S^0 K^+ K^-$ yields. The blue line indicates the total likelihood (including systematics), whereas the red, dotted line is statistical only. The dashed, vertical line indicates the $N(B_s^0 \to K_S^0 K^+ K^-) = 0$ hypothesis. Left: Down-Down K_S^0 reconstruction mode. Right: Long-Long K_S^0 reconstruction mode. Different data-taking periods are shown on top (2011), middle (2012 pre-June), and bottom (2012 post-June).

4.8.4   Comparison with previous measurements

All reported branching fractions have already been measured with the $1\,\mathrm{fb}^{-1}$ LHCb dataset from 2011, and the $B^0 \to K_S^0 K^+ K^-$ branching fraction has been accurately measured by B factories. We perform a naive comparison with previous measurements neglecting correlations between datasets for the previous LHCb measurement.[5]

---

[5]These correlations need a specific treatment as the stripping, trigger, and selection are different in the two analyses.

Table 4.23 – Significance of $B_s^0 \to K_S^0 K^+ K^-$ yields for each data category, including systematics. Global significance is obtained by summing individual significances in quadrature.

| | 2011 | | 2012 pre-June | | 2012 post-June | |
	DD	LL	DD	LL	DD	LL
Significance	2.6σ	0.7σ	0.2σ	1.2σ	2.0σ	1.2σ
Global						$\mathbf{3.7\sigma}$
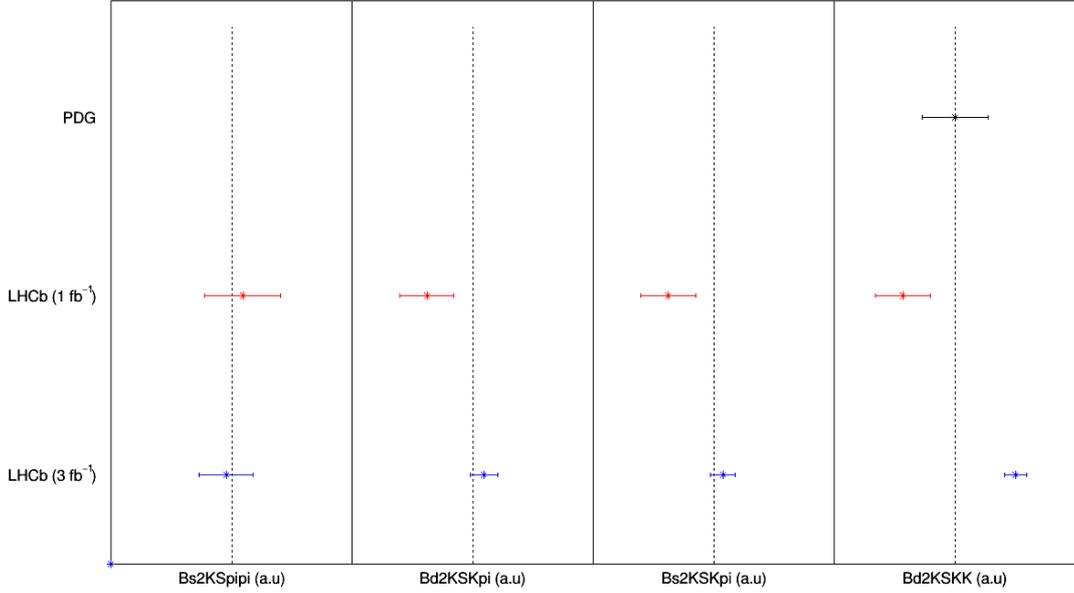
Figure 4.19 – Measured branching fractions relative to that of $B^0 \to K_S^0 \pi^+ \pi^-$ for each previously observed $B_{d,s}^0 \to K_S^0 h^+ h'^-$ mode, in arbitrary units (a.u). The "PDG" measurement is computed neglecting correlations between $B^0 \to K_S^0 K^+ K^-$ and $B^0 \to K_S^0 \pi^+ \pi^-$ measurements. The central value is set to the PDG result when existing, otherwise it is set to the weighted average of the LHCb measurements.

Figure 4.19 shows the comparison between available measurements for signal modes, except for $B_s^0 \to K_S^0 K^+ K^-$. The branching fractions are represented in arbitrary units, and only the size of uncertainties and distance between central values are meaningful. The agreement between the two LHCb results is satisfactory for all modes, with the exception of the $B^0 \to K_S^0 K^+ K^-$ mode. Table 4.24 shows the significances of the difference between the two LHCb results, assuming a correlation of 30% between the two datasets.

Table 4.24 – Significance of the difference between the LHCb measurement with 2011 data only and the current measurement.

	$\dfrac{\mathcal{B}\left(B_s^0 \to K_S^0 \pi^+ \pi^-\right)}{\mathcal{B}\left(B^0 \to K_S^0 \pi^+ \pi^-\right)}$	$\dfrac{\mathcal{B}\left(B^0 \to K_S^0 K^\pm \pi^\mp\right)}{\mathcal{B}\left(B^0 \to K_S^0 \pi^+ \pi^-\right)}$	$\dfrac{\mathcal{B}\left(B_s^0 \to K_S^0 K^\pm \pi^\mp\right)}{\mathcal{B}\left(B^0 \to K_S^0 \pi^+ \pi^-\right)}$	$\dfrac{\mathcal{B}\left(B^0 \to K_S^0 K^+ K^-\right)}{\mathcal{B}\left(B^0 \to K_S^0 \pi^+ \pi^-\right)}$
Significance (σ)	0.4	2.2	2.1	4.3

4.9   Conclusion

We have observed the $B_s^0 \to K_S^0 K^+ K^-$ decay using the $3\,\mathrm{fb}^{-1}$ dataset from Run I with a significance of $3.7\,\sigma$. Once the efficiency of this signal decay is computed, we will report a branching fraction measurement relative to that of $B^0 \to K_S^0 \pi^+ \pi^-$.

The measurements of all previously observed $B_{d,s}^0 \to K_S^0 h^+ h'^-$ modes have been updated, and is in good agreement with previous measurements, with the exception of $B^0 \to K_S^0 K^+ K^-$. Indeed, the measurement of the $B^0 \to K_S^0 K^+ K^-$ branching fraction with the reference B-factory measurement is good, but it is $4.3\,\sigma$ away from the measurement performed with LHCb 2011 data.

Chapter 5

Dalitz-plot analysis of $B^0 \to K^0_S K^+ K^-$

In this chapter, I describe the Dalitz-plot analysis of $B^0 \to K^0_S K^+ K^-$ decays. In Sec. 5.1, I present the strategy of this analysis. I then describe the reoptimization of the selection criterion on the BDT output variable in Sec. 5.2 and the event species present in the dataset in Sec. 5.3. The study of background sources and their distributions over the Dalitz plot is presented in Sec. 5.4, and the modelling of efficiency variation over the Dalitz plane is detailed in Sec. 5.5. I then present the data fit model in Sec. 5.6, and the fit validation procedure in Sec. 5.7. The evaluation of systematic uncertainties on the isobar parameters is discussed in Sec. 5.8, and the results of the fit to data are presented in Sec. 5.9.

5.1 Analysis context and strategy

The BaBar and Belle experiments have performed a time-dependent flavour-tagged amplitude analysis of $B^0 \to K^0_S K^+ K^-$ decays, and measured the angle β_{eff} in this mode [40,67]. These measurements are consistent with the value of β extracted from $b \to c\bar{c}s$ transitions. A particularity of the $B^0 \to K^0_S K^+ K^-$ mode is that the amplitude is dominated by a nonresonant component that is not clearly understood. The analysis of this mode in the LHCb environment will provide another insight into this nonresonant component. Additionally, a wide resonant structure has been seen by both BaBar and Belle in the $K^+ K^-$ spectrum. While Belle modelled it by the $f_0(1500)$, the BaBar experiment used a combination of the $f_0(1500)$, the $f_0(1710)$, and the $f'_2(1525)$. In the following, we take as a reference the BaBar result [40], which includes the resonances shown in Table 5.1 along with their lineshapes.

We aim at measuring the amplitude of $B^0 \to K^0_S K^+ K^-$ decays over the Dalitz plane using the isobar model described in Sec. 1.3.4. We do not consider CP violation in the model for this iteration of the analysis. Indeed, performing a CP-sensitive analysis requires either flavour tagging or the presence of flavour-specific structures such as $K^{*\pm}$ resonances. The second option is not relevant to this analysis as the baseline model only includes $K^+ K^-$ resonances and no $K^0_S K^+$ or $K^0_S K^-$ contributions. The size of the current

Table 5.1 – List of the resonances composing the BaBar result [40] and of their lineshapes. "Relativistic BW" stands for Relativistic Breit-Wigner.

Resonance	Lineshape
ϕ^0	Relativistic BW
$f_0(980)$	Flatté
$f_0(1500)$	Relativistic BW
$f_2'(1525)$	Relativistic BW
$f_0(1710)$	Relativistic BW
χ_{c0}	Relativistic BW
NR(S-wave)	Second-order polynomial
NR(P-wave)	Second-order polynomial

data sample does not allow the use of flavour tagging. As the dataset is untagged, the B^0 and $\overline{B}^0$ amplitudes are added incoherently, and the decay rate as a function of the Dalitz plot is

$$I(m^2_{K^0_S K^+}, m^2_{K^0_S K^-}) = \left| \mathcal{A}(m^2_{K^0_S K^+}, m^2_{K^0_S K^-}) \right|^2 + \left| \overline{\mathcal{A}}(m^2_{K^0_S K^+}, m^2_{K^0_S K^-}) \right|^2, \qquad (5.1)$$

where $\mathcal{A}$ and $\overline{\mathcal{A}}$ are the decay amplitudes of the B^0 and the $\overline{B}^0$, respectively. Replacing these amplitudes by their expressions in the isobar model, we get

$$I = \left| \left(\sum_j c_j e^{i\phi_j} F_j(m^2_{K^0_S K^+}, m^2_{K^0_S K^-}) \right) \right|^2 + \left| \left(\sum_j \overline{c}_j e^{i\overline{\phi}_j} \overline{F}_j(m^2_{K^0_S K^+}, m^2_{K^0_S K^-}) \right) \right|^2. \quad (5.2)$$

The lineshapes F_j and $\overline{F}_j$ are related by

$$F_j(m^2_{K^0_S K^+}, m^2_{K^0_S K^-}) = \overline{F}_j(m^2_{K^0_S K^-}, m^2_{K^0_S K^+}). \qquad (5.3)$$

The exchange operator between K^+ and K^- has a signature

$$\eta = (-1)^{L_{K^+ K^-}}, \qquad (5.4)$$

where $L_{K^+ K^-}$ is the orbital angular momentum between K^+ and K^-. In the decay of a pseudoscalar particle to three pseudoscalars, and in the case of a $K^+ K^-$ resonance j, this signature is

$$\eta_j = (-1)^{S_j}, \qquad (5.5)$$

where S_j is the spin of the resonance. This means that for a $K^+ K^-$ resonance j,

$$F_j(m^2_{K^0_S K^+}, m^2_{K^0_S K^-}) = (-1)^{S_j} \overline{F}_j(m^2_{K^0_S K^+}, m^2_{K^0_S K^-}). \qquad (5.6)$$

We define

$$\delta_{jk} = \arg(F_j F_k), \tag{5.7}$$

$$C_{jk} = \frac{c_j c_k}{\overline{c}_j \overline{c}_k}, \tag{5.8}$$

$$\beta_{jk} = \overline{\phi}_{jk} - \phi_{jk}, \tag{5.9}$$

$$\eta_{jk} = (-1)^{S_j + S_k}. \tag{5.10}$$

$$\tag{5.11}$$

The parameter δ_{jk} is implicitly a function of the Dalitz plot. The C_{jk} and β_{jk} terms contain information on the direct and indirect CP violation, respectively.[1] Developing the sums and looking only at the interference term between a resonance j and a resonance k, we obtain

$$I_{jk} = 2c_j c_k F_j F_k \left[(1 + C_{jk}\eta_{jk}) \cos(\phi_{jk} + \delta_{jk}) \cos(\beta_{jk}) - C_{jk}\eta_{jk} \sin(\phi_{jk} + \delta_{jk}) \sin(\beta_{jk}) \right] \tag{5.12}$$

where $F_{j,k}$ is implicitly function of $m^2_{K^0_S K^+}$ and $m^2_{K^0_S K^-}$.

In the case of the current analysis, we do not expect sufficient statistical power to measure the C_{jk} and β_{jk} parameters. Considering the case where no CP violation occurs and detection asymmetry is negligible ($C_{jk} = 1$ and $\beta_{jk} = 0$), we observe that the interference term is

$$I_{jk} \propto (1 + \eta_{jk}) \cos(\phi_{jk} + \delta_{jk}). \tag{5.13}$$

This means that in our particular case of an untagged analysis of $B^0 \to K^0_S K^+ K^-$, considering only $K^+ K^-$ resonances, we are not sensitive to the relative phase between even and odd partial waves (ϕ_{jk}). The residual sensitivity originates from CP violation, possibly, which allows in principle to measure these effects even in an untagged, time-independent analysis. A similar calculation has been performed in [68, 69]. Section 5.6 discusses the adaptations of the baseline model that we implement to address this issue.

The amplitude analysis of $B^0 \to K^0_S K^+ K^-$ decays uses the following inputs from the branching fraction measurement:

- the same stripping, trigger requirements, preselection, and BDT training;

- tight PID selection criteria are applied;

- yields of signal and background species are extracted using the same B-meson candidate invariant-mass fit model;

- the distribution of cross-feed events over the Dalitz plane is estimated using the $_s\mathcal{P}lot$ method described in Sec. 4.2.3.

---

[1]In this expression, the possible production or detection asymmetry between B^0 and $\overline{B}^0$ is absorbed in a different magnitude convention for c_k and $\overline{c}_k$. This effect has to be taken into account before any statement is issued from a measurement of C_{jk}.

We reoptimize the BDT selection criteria in order to get better uncertainties on the parameters of the isobar model. We also estimate the variations of the efficiency across the Dalitz plane using a similar method to that of the branching fraction measurement.

Using the `Laura++` package [70], we perform an extended maximum likelihood fit to data events selected in a window of $\pm 2.5\sigma$ around the B^0 invariant mass. In order to respect the Dalitz-plot boundaries, both the B-meson candidate mass and that of the K_s^0 are constrained to their nominal values, and momenta of the daughters are refitted taking these constraints into account.

Due to the large number of parameters in the fit model and the small sensitivity to some of them, each fit to data is performed 1000 times with randomized initial parameters in order to find the best minimum. Multiple solutions could appear in an amplitude analysis due to many different reasons. In particular, this could happen due to interference between two broad scalar resonances. In that case solutions typically appear in pairs: one with larger fit fractions of the two resonances and destructive interference between them, and another with smaller fit fractions and constructive interference. We consider all solutions within 4.5 NLL units from the best minimum, and study the associated fit fractions.[2]

We consider variations of the fit model by adding and removing resonances, or changing their distribution across the Dalitz plane. We compare the agreement of each model with the data using the minimum negative log-likelihood and goodness-of-fit estimators. Details of this procedure are given in Sec. 5.6.

5.2 Reoptimization of the BDT selection

5.2.1 Strategy of the reoptimization

The BDT method described in Sec. 4.1 produces an output variable on which we apply a selection to reject combinatorial background events. Our goal is to use an optimized BDT selection criterion that yields the smallest uncertainties on the isobar parameters extracted from the fit to the Dalitz plane. While there exist widely used optimization methods adapted to the measurement of branching fractions, this is not the case for Dalitz-plot analyses. In this study, we first perform simplified Dalitz plot fits to several datasets obtained with different BDT selections and compare the uncertainties on the isobar parameters. We then attempt to find a simple figure of merit that yields similar conclusions in order to facilitate exporting the the results of this study to future Dalitz-plot analyses of this mode.

In the case of the measurement of the branching fraction of an already-observed mode, the figure of merit

$$FoM(N_\mathrm{S}, N_\mathrm{B}) = \frac{N_\mathrm{S}}{\sqrt{N_\mathrm{S} + N_\mathrm{B}}}, \tag{5.14}$$

---

[2]4.5 negative log-likelihood units correspond to three Gaussian standard deviations.

where N_S (N_B) is the number of signal (background) events, is frequently adopted to maximise the significance of the measurement. In the case of an unobserved mode, the Punzi figure of merit

$$FoM(N_S, N_B) = \frac{\epsilon_{\mathrm{sig}}}{\sqrt{\frac{a}{2} + N_B}} \tag{5.15}$$

is adopted, where ϵ_{sig} is the estimated efficiency of the signal, and a is a parameter to determine by the analyst. Although in the case of a Dalitz-plot analysis, no such standard solution exists, the figure of merit

$$FoM(N_S, N_B) = \frac{N_S^2}{(N_S + N_B)^{3/2}} \tag{5.16}$$

is sometimes used.[3]

In order to optimize the cut value on the BDT output variable for the Dalitz-plot analysis, we define a series of lower cuts, λ_{WP}, on the BDT variable λ (*i.e.* cutting out events with $\lambda \leq \lambda_{\mathrm{WP}}$). We refer to the different values of λ_{WP} as "working points". They correspond to values in between those from the loose and the tight optimizations (λ_{loose} and λ_{tight}, respectively)

$$\lambda^{\mathrm{WP}} = \lambda^{\mathrm{loose}} + \alpha^{\mathrm{WP}}(\lambda^{\mathrm{tight}} - \lambda^{\mathrm{loose}}). \tag{5.17}$$

The parameter α^{WP}, which has a different value for each working point, runs between 0 and 1.[4] We perform separate invariant-mass fits on the samples corresponding to each working point, and extract the number of signal, combinatorial background, and cross-feed events as described in Sec. 5.3. We then use these yields and isobar parameters extracted in the BaBar study of the $B^0 \to K_S^0 K^+ K^-$ mode [40] as a baseline model to generate pseudo-data distributions over the Dalitz plot for each working point. We finally perform a simplified amplitude analysis on the pseudo-data corresponding to each working point using the BaBar results as a signal model, with the following guidelines:

- the same Dalitz-plot distribution of combinatorial background is used for all working points, as a relaxed BDT selection criterion (described in Sec. 5.4) is applied to samples from which we extract this distribution;

- Gaussian constraints in the invariant-mass fit model used to extract the yields of the partially reconstructed backgrounds and cross-feeds are re-evaluated for each working point, as the corresponding efficiencies vary;

---

[3]The optimum value obtained by this figure of merit typically lies between those from $N_S/\sqrt{N_S + N_B}$ and Punzi-type figures of merit.

[4]The tight optimization of the BDT is then strictly equal to the working point defined using $\alpha^{\mathrm{WP}} = 1$.

Table 5.2 – Values of the parameter α^{WP} for each working point, along with the number of signal events and the proportion of signal events in the signal region. We considered all together data-taking periods and K_{S}^0 reconstruction modes.

	WP0	WP1	WP2	WP3	WP4	WP5	WP6	WP7	WP8	WP9
α^{WP}	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.85	0.95
N_S	1474.2	1438.8	1377.8	1323.2	1271.0	1212.4	1152.5	1086.3	986.1	905.0
$N_S/N_{\mathrm{tot}}(\%)$	89.2	90.4	91.2	92.1	93.0	94.0	94.6	95.2	96.0	96.0

- as we apply tight PID selection criteria for all working points, we use tight PID efficiency maps;

- selection efficiency maps are estimated using a linear interpolation between the loose and the tight selection efficiency maps, using the parameter α^{WP};[5]

- only the systematic uncertainties related to biases, efficiency modelling, backgrounds distributions modelling, and the knowledge of the yields of the different event species are estimated;

- when varying the yields within their uncertainties in order to evaluate the corresponding systematic uncertainty, only their statistical uncertainties, which are dominant, are taken into account.

From this simplified analysis, we extract the total uncertainties on all the isobar parameters for each working point.

5.2.2 Results of the reoptimization

Table 5.2 shows the values of α^{WP} used in the study and the names of the associated working points. The same table details the corresponding proportion of signal events and the number of signal events, both extracted from an invariant-mass fit to data. To obtain these numbers, we considered all together the different data-taking periods and K_{S}^0 reconstruction modes. As the working points get closer to the tight optimization, the purity of the samples stabilize while the numbers of signal events continue to diminish.

Figure 5.1 shows the total relative uncertainties on the fit fractions of each resonance, with respect to values obtained for "WP0". The main conclusion of this study is that it is difficult to point a clear overall optimum, as variations are observed between resonances and as, with the current dataset, uncertainties are large. For instance, we notice that optima of nonresonant contributions tend to correspond to tighter BDT selection criteria than those of resonant contributions. Indeed, nonresonant contributions are competing with combinatorial background over large parts of the Dalitz plane.

---

[5]The absolute value of the efficiency of the BDT selection is not linear between the loose and the tight selection cut values, but only the variations of the efficiency across the Dalitz plane are relevant here.
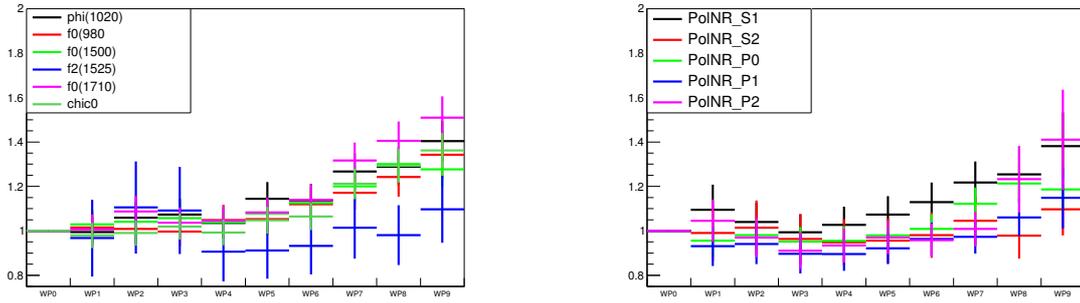
Figure 5.1 – Total relative uncertainties on fit fractions of different isobars (see the legend in each graph) for each of the working points. Uncertainties are scaled with respect to those of "WP0". Left: Resonant contributions. Right: nonresonant contributions; where S and P correspond to S and P-waves, respectively, followed by the degree of the polynomial term attached to that contribution. The constant term of the S-wave ("PolNR_S0") is the fixed reference amplitude and thus it is not shown here.

Table 5.3 – Values of the BDT selection cuts that are chosen for the different data-taking periods and K_S^0 reconstruction modes.

| | 2011 | | 2012 pre-June | | 2012 post-June | |
	DD	LL	DD	LL	DD	LL
Selection cut value	-0.025	-0.104	0.081	0.01	-0.027	-0.055

Figure 5.2 shows the values of four different figures of merit evaluated for each working point, relative to the values obtained for "WP0". As expected, the optimum values for the $N_S/\sqrt{N_S + N_B}$ and Punzi-type figures of merit are close to "WP0" and "WP9", respectively.[6] The figure of merit $N_S^2/(N_S + N_B)^{3/2}$ is maximal between "WP1" and "WP3", thus pointing to a selection close to the loose BDT optimization. The profile of this figure of merit and the corresponding optimal selection are the most similar to the profile of uncertainties shown in Fig. 5.1. The figure of merit that we chose for this analysis is therefore that defined in Eq. 5.16. Table 5.3 summarizes the corresponding BDT selection cut values for the different data-taking periods and K_S^0 reconstruction modes.

_____

[6]The loose and tight optimizations of the BDT have been obtained using an approximate invariant-mass fit model and in each spectrum separately. It explains the fact that the optimum for the Punzi figure of merit is not "WP9".
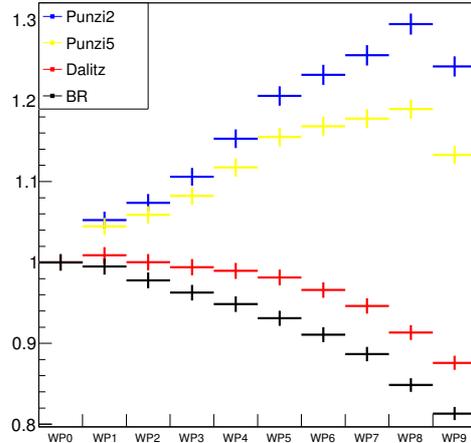
Figure 5.2 – Values of different figures of merit evaluated in each of the working points. The "BR" label (black points) refers to the $N_S/\sqrt{N_S + N_B}$ figure of merit, whereas "Dalitz" (red points) refers to $N_S^2/(N_S + N_B)^{3/2}$. The "Punzi2" (blue points) and "Punzi5" (yellow points) labels refer the to Punzi figures of merit calculated with $a = 2$ and $a = 5$, respectively.

5.3 Yields of the signal and background species

Data events are selected in a invariant-mass window around the B^0 signal peak. As we do not perform an extended fit, yields of signal and background species are estimated from the invariant-mass fit. The estimated number of events N_i^{window} for an event species i in a mass window defined between m_{min} and m_{max} is

$$N_i^{\text{window}} = N_i \frac{\int_{m_{\text{min}}}^{m_{\text{max}}} f_i(m)dm}{\int_{5150}^{5800} f_i(m)dm}, \tag{5.18}$$

where N_i is the number of events from the event species i extracted from the mass fit in the whole mass range, defined as the 5150–5800 MeV/c^2 interval, and $f_i(m)$ is the distribution of the event species i. As mentioned in Sec. 5.1, we define signal mass windows following

$$m_{\text{min}} = \mu - 2.5\sigma, m_{\text{max}} = \mu + 2.5\sigma, \tag{5.19}$$

where μ and σ are the values of the corresponding parameters in the double Crystal-Ball distribution that describes the $B^0 \to K_S^0 K^+ K^-$ signal.

Table 5.4 shows the estimated number of events for each event species, K_S^0 reconstruction mode, and data-taking period. In the following, we neglect events from the partially

Table 5.4 – Number of events for each event species for the different data-taking periods and K_S^0 reconstruction mode. Purity is defined as the proportion of signal events in the sample. "B^0 CF" refers to the cross-feed from $B^0 \to K_S^0 K^\pm \pi^\mp$ events, whereas "B_s^0 CF" refers to the cross-feed from $B_s^0 \to K_S^0 K^\pm \pi^\mp$ events. Similarly, "PR1" and "PR2" refer to charmed and charmless partially reconstructed backgrounds, respectively.

Year	K_S^0 rec. mode	Signal	Comb.	B^0 CF	B_s^0 CF	PR1	PR2	Purity(%)
2011	DD	271.5	39.2	6.1	0.0	0.0	0.0	85.7
2011	LL	210.5	39.5	3.8	0.0	0.0	0.0	82.9
2012a	DD	176.1	29.9	5.2	0.0	0.0	0.0	83.4
2012a	LL	114.9	10.0	2.2	0.0	0.0	0.0	90.3
2012b	DD	649.9	88.8	11.2	0.1	0.0	0.0	86.7
2012b	LL	330.1	23.1	4.9	0.0	0.0	0.0	92.2

reconstructed background. Due to the blinding of the $B_s^0 \to K_S^0 K^+ K^-$ yield in the mass fit, we can only estimate the number of events in the mass window through an educated guess. Postulating that

$$N(B_s^0)(\text{Year})(K_S^0) = \frac{N(B^0)(\text{Year})(K_S^0)}{\lambda^2}, \tag{5.20}$$

where λ is the sine of the Cabibbo angle, we find fewer than two events of $B_s^0 \to K_S^0 K^+ K^-$ in the signal window for all data-taking periods and K_S^0 reconstruction modes combined. We thus ignore this contribution in the following.

5.4 Background distributions

5.4.1 Combinatorial background modelling

The nature of combinatorial background in this analysis is discussed in Sec. 4.3.5. We expect that events with a B candidate mass larger than $5550\,\text{MeV}/c^2$ originate only from combinatorial background,[7] thus we use the distribution over the Dalitz plot of events from this sideband to model the combinatorial background distribution.

The small number of upper-mass sideband events limits the understanding of the combinatorial background. In order to estimate more accurately the distribution of these events over the Dalitz plot, we relax the BDT selection. We check the dependency of this distribution with respect to the BDT output value by splitting the dataset in the upper-mass sideband in ten samples with roughly the same number of events in different

---

[7]This threshold is larger than that used for the BDT training as in the $K_S^0 K^+ K^-$ mode cross-feeds from $B_s^0 \to K_S^0 K^\pm \pi^\mp$ are not negligible at $5450\,\text{MeV}/c^2$.

Table 5.5 – Values of the BDT output variable used to split the 2011, Down-Down dataset, and number of events in each dataset.

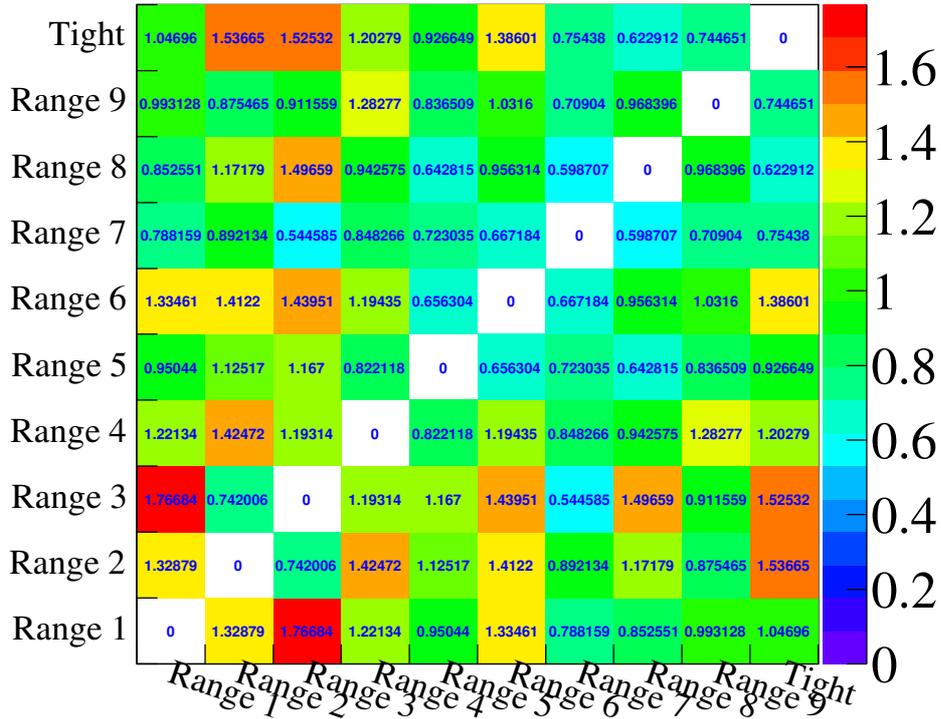| Minimum BDT value	-0.12	-0.10	-0.09	-0.06	-0.04	-0.01	0.02	0.06	0.12	0.22
Number of events	35	28	32	29	30	37	24	32	32	29

Figure 5.3 – Calculated χ^2/ndf between distributions of upper-mass sideband events from 2011, Down-Down data samples. Each distribution corresponds to an interval of BDT output variable.

BDT intervals. Table 5.5 shows the BDT selection cut values used to split the dataset and Fig. 5.3 shows the calculated χ^2/ndf between pairs of these distributions for 2011, Down-Down events. As there is no clear trend, we relax the BDT selection to that corresponding to the loose optimization of the BDT.

Figure 5.4 shows the distribution of combinatorial background events over the Dalitz plane for all K_s^0 reconstruction modes and data-taking periods, using the relaxed BDT

cut value.

5.4.2 Cross-feeds modelling

The sWeighting procedure that is described in Sec. 4.7 can be used to estimate the distribution of cross-feed events across the Dalitz plane as well. Figure 5.5 shows the distribution across the Dalitz plot of the two cross-feed species that are present in the mass window, along with the uncertainties on these distributions for 2011, Down-Down events.

5.5 Efficiency variations across the Dalitz plot

As discussed in Sec. 4, the efficiency of signal events is not flat across the Dalitz plane, for instance because of the limited geometrical acceptance of the LHCb detector. This non-uniformity has to be taken into account in the fit to data, as it results in a distortion of the observed distribution of signal events over the Dalitz plane. We use a similar approach as in Sec. 4.1, breaking down the total efficiency into three multiplicative contributions:

- ϵ^{gen} is the efficiency of the generator-level cuts applied to the Monte-Carlo samples that are used to evaluate the rest of the efficiencies;

- $\epsilon^{\text{sel|gen}}$ is the efficiency of the trigger, reconstruction, stripping, and selection methods. It is evaluated using signal Monte-Carlo samples, corrected for discrepancies between simulation and data. Vetoes on charmed contributions are removed if they do not include a mis-ID hypothesis, as they are taken into account in the Dalitz-plot fit;

- $\epsilon^{\text{PID|(sel|gen)}}$ is the efficiency of the PID requirements. It is evaluated by a data-driven approach on signal Monte-Carlo samples, using the `PIDcalib` package.

These three contributions are then multiplied together to get the total efficiency

$$\epsilon^{\text{tot}} = \epsilon^{\text{gen}}\epsilon^{\text{sel|gen}}\epsilon^{\text{PID|(sel|gen)}} \tag{5.21}$$

In the following section, I detail the calculation of these efficiencies and their uncertainties. Firstly, I present the methods used in the evaluation of uncertainties, then the extraction of generator-level efficiencies, of the selection efficiency along with all the relevant corrections of data/MC discrepancies, and finally of PID efficiencies. As there is no source of stiff variation of efficiency across the Dalitz plane[8], we smooth efficiency histograms using a two-dimensional cubic interpolation ("spline").

---

[8]Vetoes on charmed resonances constitute such a source but are considered independently.

5.5.1 Uncertainty estimation procedure

We evaluate the asymmetric uncertainty on the efficiency in each bin of the Dalitz plot independently. In the simple case where the efficiency is

$$\epsilon = \frac{N_{\text{passed}}}{N_{\text{total}}}, \tag{5.22}$$

the uncertainty on the efficiency ϵ can be determined using Clopper-Pearson intervals, implemented in the TEfficiency package [71]. These intervals provide a coverage of the uncertainty that is always conservative, thus suitable for the evaluation of systematic uncertainties. However, this technique does not extend to weighted events, especially if the uncertainty on the weights has to be taken into account. In that case, we evaluate the uncertainty as follows:

- create 500 new samples using the bootstrapping method. This method creates a new sample out of an original one by randomly resampling the events. Event weights are randomized within their uncertainties;

- evaluate the efficiency histogram for each sample;

- in each bin of the efficiency histograms, fit the distribution of efficiencies in the 500 samples using a bifurcated Gaussian. The upper and lower uncertainties are assigned to the right and left width parameters values, respectively.

All sources of uncertainties are considered as uncorrelated, and are thus summed in quadrature to estimate the total uncertainty for a given contribution. Likewise, the total uncertainty on the efficiency is calculated by propagating the uncertainties on ϵ^{gen}, $\epsilon^{\text{sel|gen}}$, and $\epsilon^{\text{PID|(sel|gen)}}$ assuming no correlation between them.

5.5.2 Acceptance of the generator-level cut

In order to save CPU, we apply in this analysis a generator-level cut requiring that both the K^+ and the K^- are generated inside of the detector acceptance. This acceptance is modelled as the $\theta \in [0.01\,\text{rad}, 0.4\,\text{rad}]$ interval, θ being the angle formed between a track and the z axis. Section 2.5 presented the principle of generator-level cuts.

We generate a sample of 50,000 $B^0 \to K_{\text{s}}^0 K^+ K^-$ events with no generator-level cut applied and a flat distribution over the square Dalitz plane. The K_{s}^0 reconstruction mode is not relevant here, nor is the difference between pre-June and post-June trigger configurations in 2012. The generator-level efficiency in each bin of the Dalitz plot is the ratio of the number of events that pass the cut

$$\theta_{K^+,K^-} \in [0.01\,\text{rad}, 0.4\,\text{rad}]. \tag{5.23}$$

Figure 5.6 shows the results of this procedure for both 2011 and 2012 conditions, along with the generated distributions. The non-uniformity of the generated distribution across the Dalitz plane is a well-known feature of the `flatSqDalitz` generation, and does not impact the extraction of efficiencies. Figure 5.7 shows the asymmetric uncertainties on the generator-level efficiencies, calculated using Clopper-Pearson intervals.

5.5.3 Selection efficiency

The selection efficiency is determined in each bin of the Dalitz plot following

$$\epsilon^{\text{sel}|\text{gen}} = \frac{N_{\text{sel}}}{N_{\text{gen}}}, \tag{5.24}$$

where N_{sel} is the number of events that pass the stripping, trigger, preselection, and BDT selection. Correction factors are applied to account for data/MC differences in the tracking and trigger efficiencies.

Re-weighting of the MC

Tracking efficiency in LHCb depends, among other variables, on the momentum p and the pseudorapidity η of each particle, along with the number of tracks in the event. These variables are not exactly modelled by the Monte-Carlo simulation. We use data events in the signal mass window as a reference for the distribution of these variables in data, and reweight Monte-Carlo samples to match these distributions. The uncertainty on these weights is estimated using the bootstrap method described in Sec. 5.5.1.

Tracking efficiency correction

The LHCb experiment disposes of reference tables to correct for data/MC discrepancies in tracking efficiency of `Long` tracks. These tables are produced using decays that have both a large production rate and large branching fractions. Weights depending on a two-dimensional binning of p and η are calculated using these tables.

Figure 5.8 shows the tracking-efficiency corrections that we apply depending on the Dalitz-plot coordinates. These corrections are close to unity, but show a dependency on the Dalitz plot.

`L0Hadron` trigger efficiency correction

As described in Sec. 2.4, the trigger system in LHCb consists of three steps: `L0`, `Hlt1`, and `Hlt2`. The efficiency of the `Hlt1` and `Hlt2` is well modelled in the simulation, but there are significant differences in the `L0Hadron_TOS` line efficiencies that we use in this analysis.[9] Indeed, this line is fired up each time there is a large enough deposit of transverse energy E_{T} in one cluster of the hadronic calorimeter, and modelling the response of

---

[9]There are also data/MC differences in the `L0Electron_TOS` line, but we do not use it.

this calorimeter requires an excellent understanding of hadronic showers in thick materials and of the aging of these materials.

In order to correct the efficiencies for the data/MC differences in `L0Hadron_TOS` efficiency, we split the sample between trigger-on-signal (TOS) and trigger-independent-of-signal (TIS) samples. We require that events in the TIS sample do not fire `L0Hadron_TOS` in order to build exclusive samples.

We dispose of tables where the efficiency of the `L0Hadron_TOS` trigger as a function of the energy of a given cluster formed in the calorimeter. Clusters can be formed by one or several particles, and also partially overlap. Each cluster i having a probability $p(E_{T,i})$ of firing the trigger depending on its transverse energy $E_{T,i}$, the total data-driven efficiency estimation of the `L0Hadron_TOS` trigger is

$$\epsilon_{\text{data}}^{\text{TOS}} = 1 - \prod_i (1 - p(E_{T,i})), \tag{5.25}$$

whereas the efficiency of the `L0Global_TIS`&!`L0Hadron_TOS` trigger is

$$\epsilon_{\text{data}}^{\text{TIS\&!TOS}} = \prod_i (1 - p(E_{T,i})). \tag{5.26}$$

We estimate in each bin of the Dalitz plot the correction factor on `L0Hadron` trigger efficiency for TIS(TOS) events as

$$k_{\text{L0}}^{\text{TIS(TOS)}} = \frac{\epsilon_{\text{data}}^{\text{TIS(TOS)}}}{\epsilon_{\text{MC}}^{\text{TIS(TOS)}}}, \tag{5.27}$$

where $\epsilon_{\text{MC}}^{\text{TIS(TOS)}}$ is the `L0Hadron` trigger efficiency calculated on Monte-Carlo samples as the proportion of events that do not pass the `L0Hadron_TOS` trigger.

We consider possible uncertainties originating from limited statistics and from tables values. Both are estimated using the bootstrap method described in Sec. 5.5.1, and are summed in quadrature. Figure 5.9 shows calculated corrections for Down-Down simulated events in the TOS trigger category, where all data-taking periods are considered together, along with upper and lower uncertainties on these corrections. The correction factor varies significantly across the Dalitz plot, thus stressing the importance of applying this correction.

Total selection efficiency

The total selection efficiency is calculated by summing the TOS and TIS contributions following

Table 5.6 – Binning on the momentum and transverse momentum of each particle used in the `PIDcalib` method.

Variable	Binning
p	3000–9300–15600–18515–28325–40097–59717–100000
η	1.5–2.4975–2.7075–3.0575–3.3725–3.7225–4.0025–5.

$$\epsilon^{\text{sel}|\text{gen}} = \frac{f_{\text{data}}^{\text{TOS}}}{f_{\text{MC}}^{\text{TOS}}} k_{\text{L0}}^{\text{TOS}} \epsilon^{\text{sel}|\text{gen},\text{TOS}} + \frac{1 - f_{\text{data}}^{\text{TOS}}}{1 - f_{\text{MC}}^{\text{TOS}}} k_{\text{L0}}^{\text{TIS\&!TOS}} \epsilon^{\text{sel}|\text{gen},\text{TIS\&!TOS}} \qquad (5.28)$$

where $f_{\text{data(MC)}}^{\text{TOS(TIS\&!TOS)}}$ is the fraction of events in data (Monte-Carlo) for which `L0Hadron_TOS` is (not) fired, and $\epsilon^{\text{sel}|\text{gen},\text{TOS(TIS\&!TOS)}}$ is the selection efficiency calculated using Eq. 5.24 on the subset of Monte-Carlo samples in which `L0Hadron_TOS` is (not) fired.

Figure 5.10 shows the selection efficiency for 2011, Down-Down events, along with the uncertainties on efficiencies, in the TOS and TIS&!TOS trigger categories. Large structures that can be seen across the Dalitz plane are expected to originate from stripping and trigger, while the BDT selection method itself was designed not to bias the Dalitz plot.

5.5.4 PID efficiency

As described in Sec. 2.5.3, a realistic estimate of the efficiency of a PID requirement on Monte-Carlo samples can be estimated by the `PIDcalib` package. This technique attributes a weight to each event that estimates the expected efficiency of the PID requirement on this event. In each bin A of the Dalitz plot, the efficiency is then

$$\epsilon^{\text{PID}|(\text{sel}|\text{gen})}(A) = \frac{\sum_{e \in A} w_e}{N_{e \in A}}, \qquad (5.29)$$

where w_e is the weight attributed by the `PIDcalib` method to event e.

The `PIDcalib` method takes into account the dependency of PID efficiency on the momentum p and pseudorapidity η of each particle, and on the number of tracks in the overall event. We decide to integrate out the dependency on the number of tracks, and consider the binning described in Table 5.6.

Figure 5.11 shows the PID efficiency of events of the 2011 data-taking period, in the Down-Down K_{S}^0 reconstruction mode, along with the uncertainties on these efficiencies.

5.5.5 Total efficiencies

We combine all the efficiencies previously calculated using Eq. 5.21, and propagate the uncertainties accordingly. Figure 5.12 shows the results of the efficiency calculation for 2011 events reconstructed in the Down-Down mode.

The principal source of uncertainties on the total efficiency is the selection efficiency, and especially the correction on the `L0Hadron` efficiency.
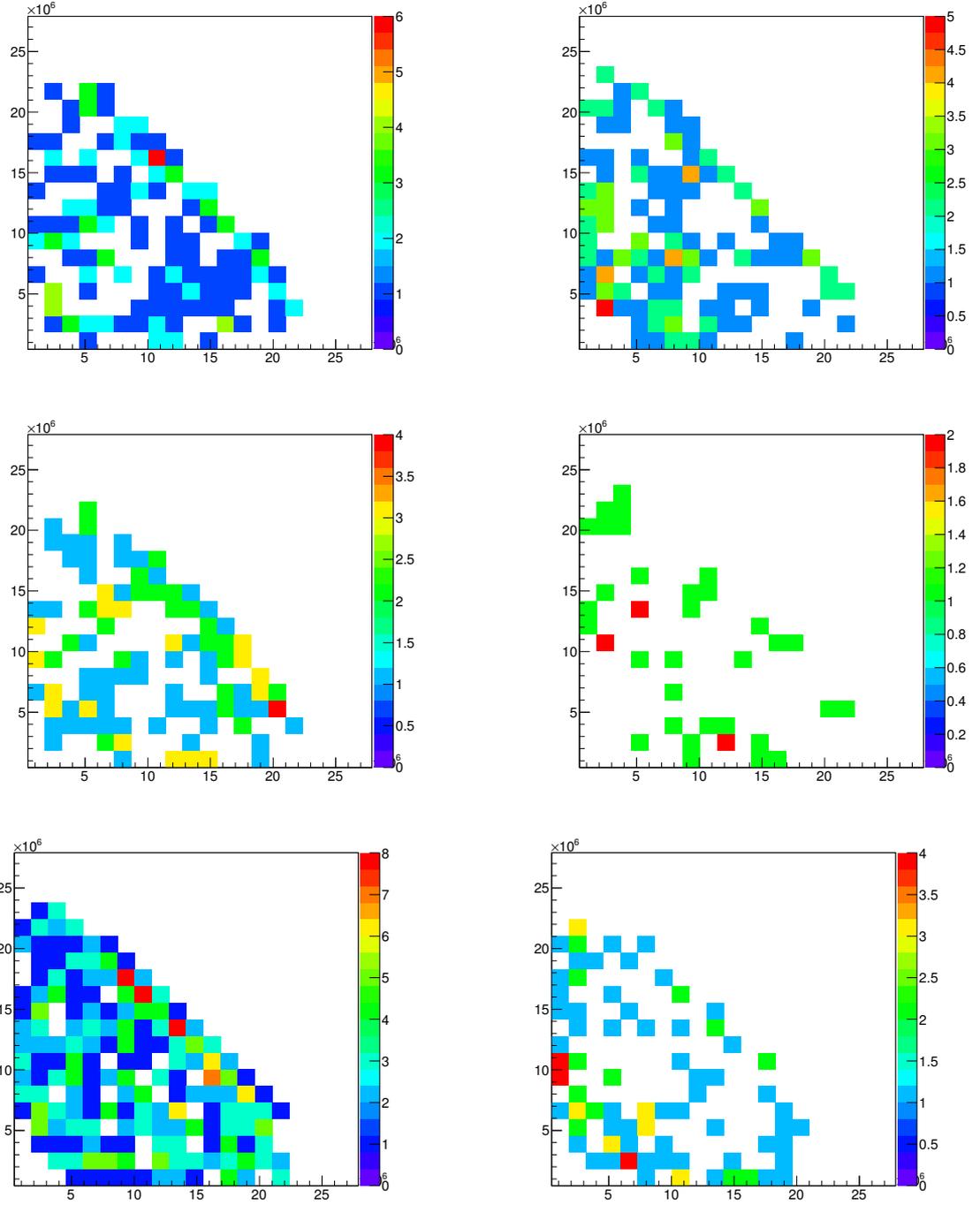
Figure 5.4 – Distributions of combinatorial background events over the Dalitz plane. Events from Down-Down (Long-Long) K_S^0 reconstruction mode are shown on the left (right). Top: 2011 events. Middle: 2012 pre-June events. Bottom: 2012 post-June events.
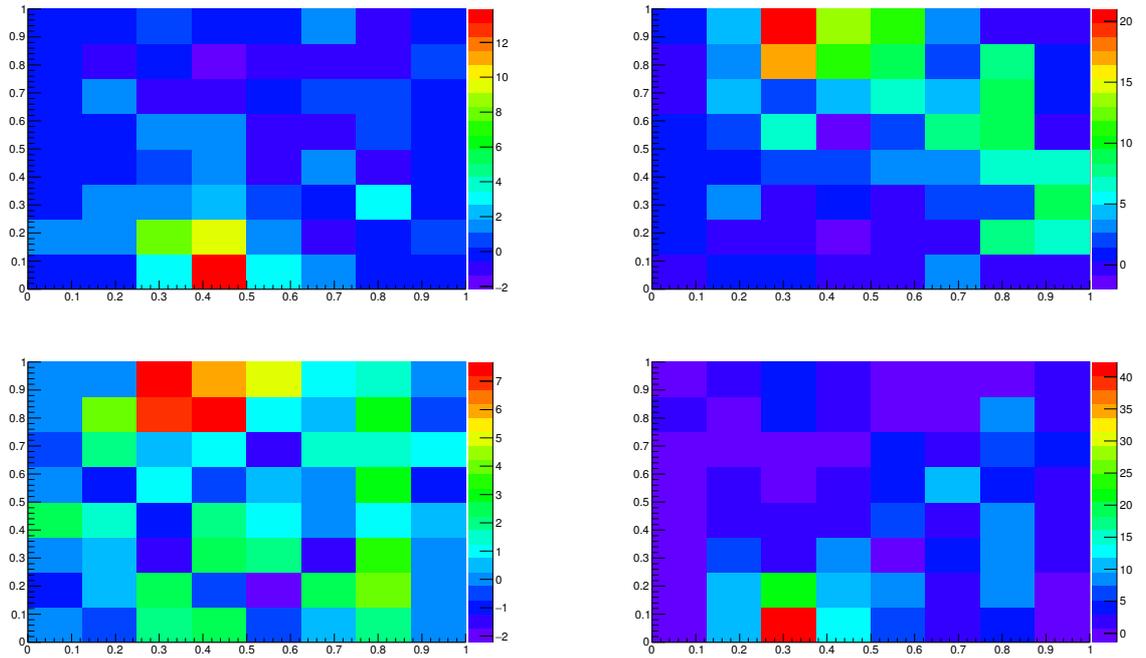
Figure 5.5 – Distribution of cross-feed events coming from $B^0 \to K_S^0 K^+ \pi^-$ (top,left), $B^0 \to K_S^0 K^+ \pi^-$ (top,right), $B^0 \to K_S^0 K^- \pi^+$ (bottom, left), and $B_s^0 \to K_S^0 K^- \pi^+$ (bottom, right), for 2012 post-June, Down-Down K_S^0 reconstruction mode.
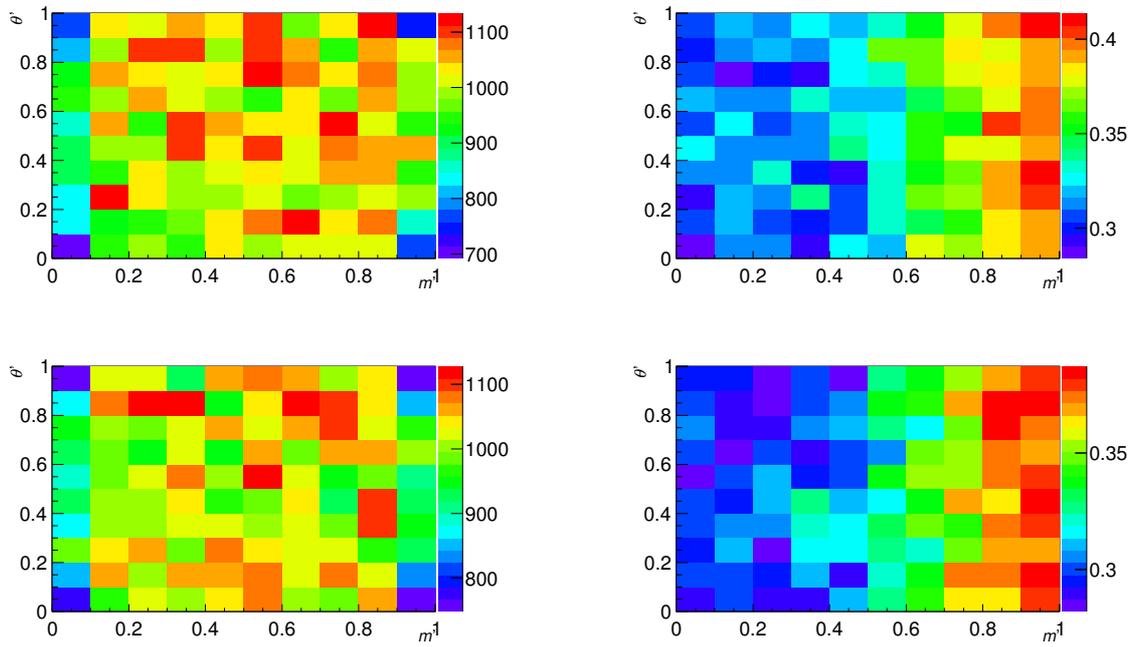
Figure 5.6 – Left: distribution of generated events over the Dalitz plane. Right: calculated efficiency of the generator-level cut. Top: 2011 conditions. Bottom: 2012 conditions.
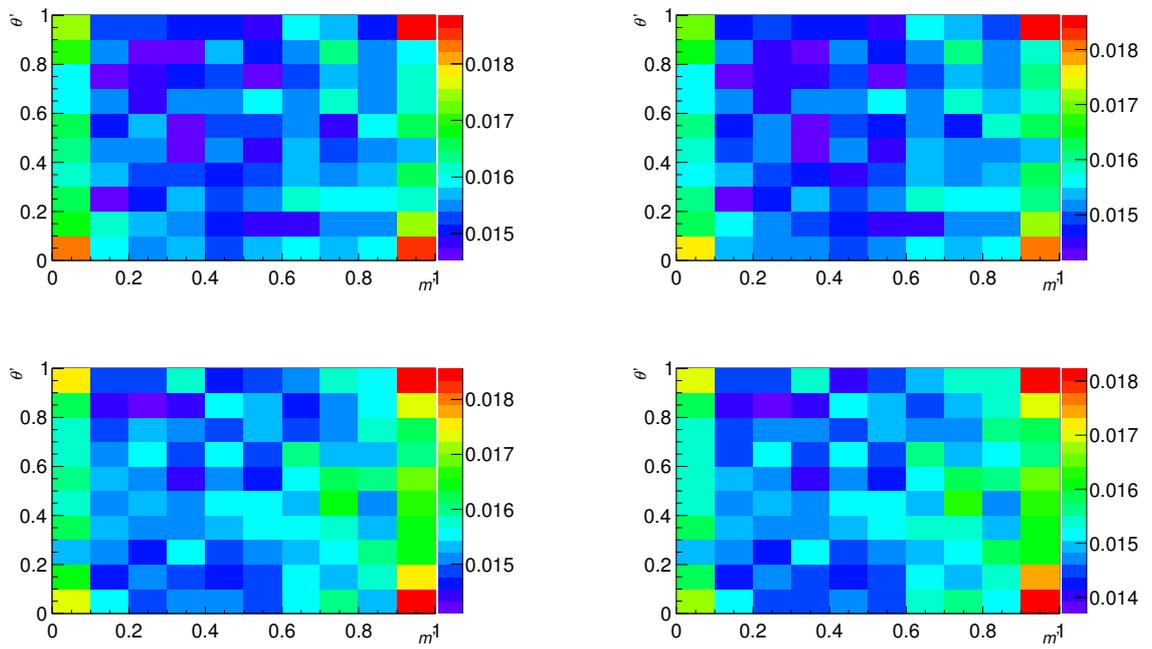
Figure 5.7 – Left: upper uncertainties on the generator-level efficiency. Right: lower uncertainties on the generator-level efficiency. Top: 2011 conditions. Bottom: 2012 conditions.
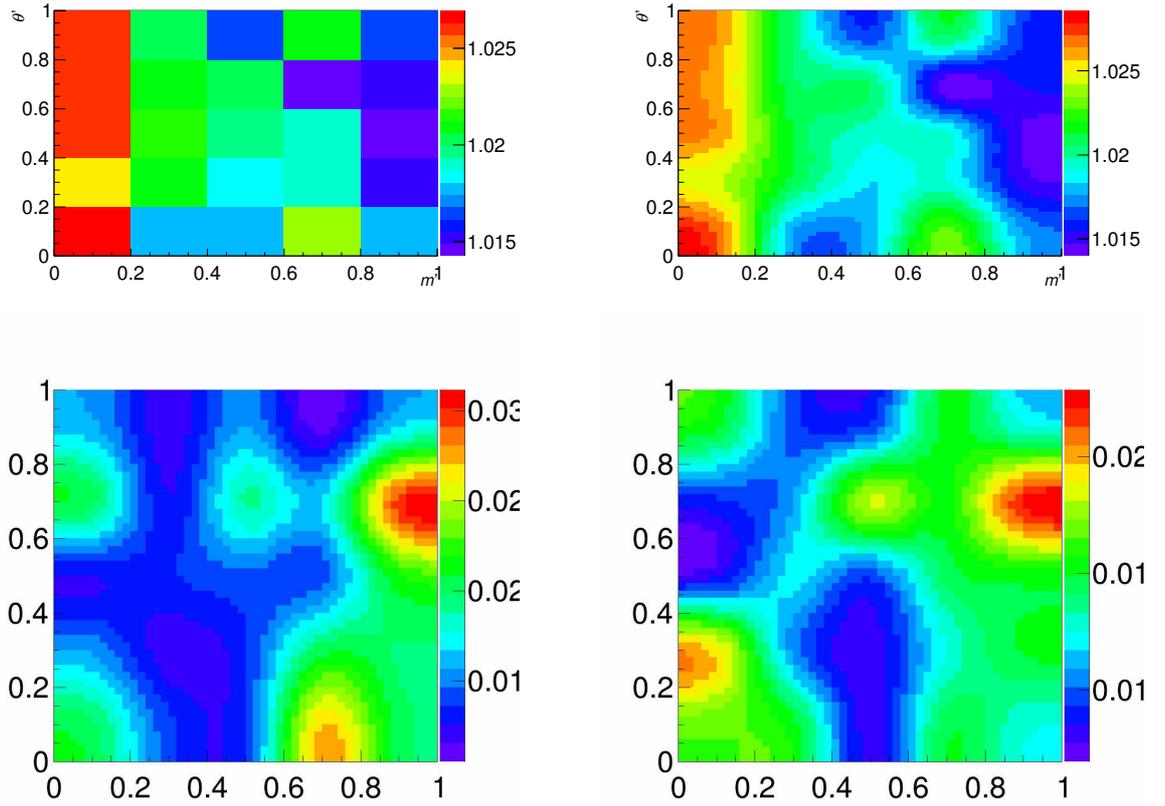
Figure 5.8 – Tracking efficiency corrections applied to 2011, Down-Down Monte-Carlo samples as a function of the square Dalitz-plot. Top, left: raw histograms. Top, right: splined histograms. Bottom: splined upper (lower) uncertainties are shown on the left (right).
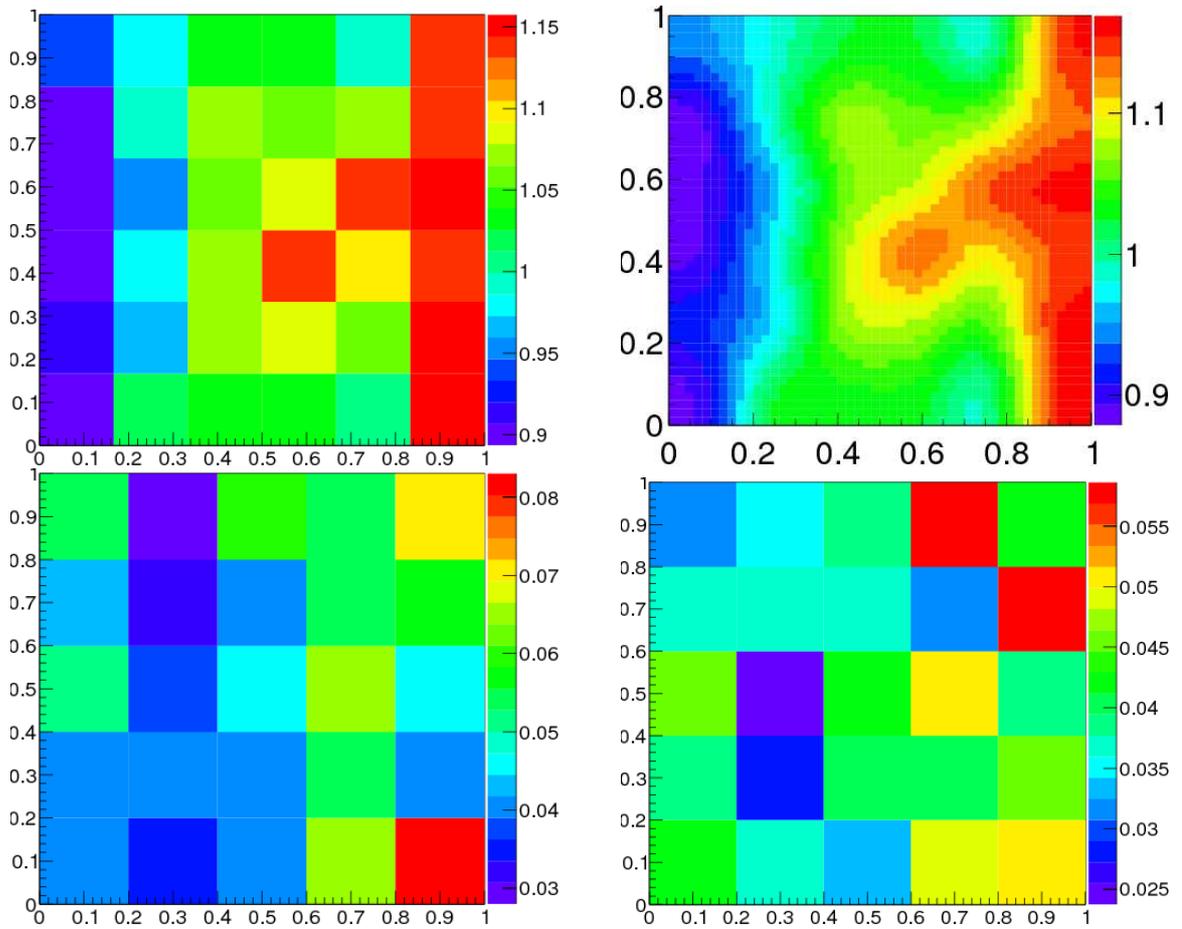
Figure 5.9 – Top, left: correction factors for TOS, Down-Down events. Top, right: splined correction factors. Bottom: upper (lower) uncertainties on correction factors are shown on the left (right).
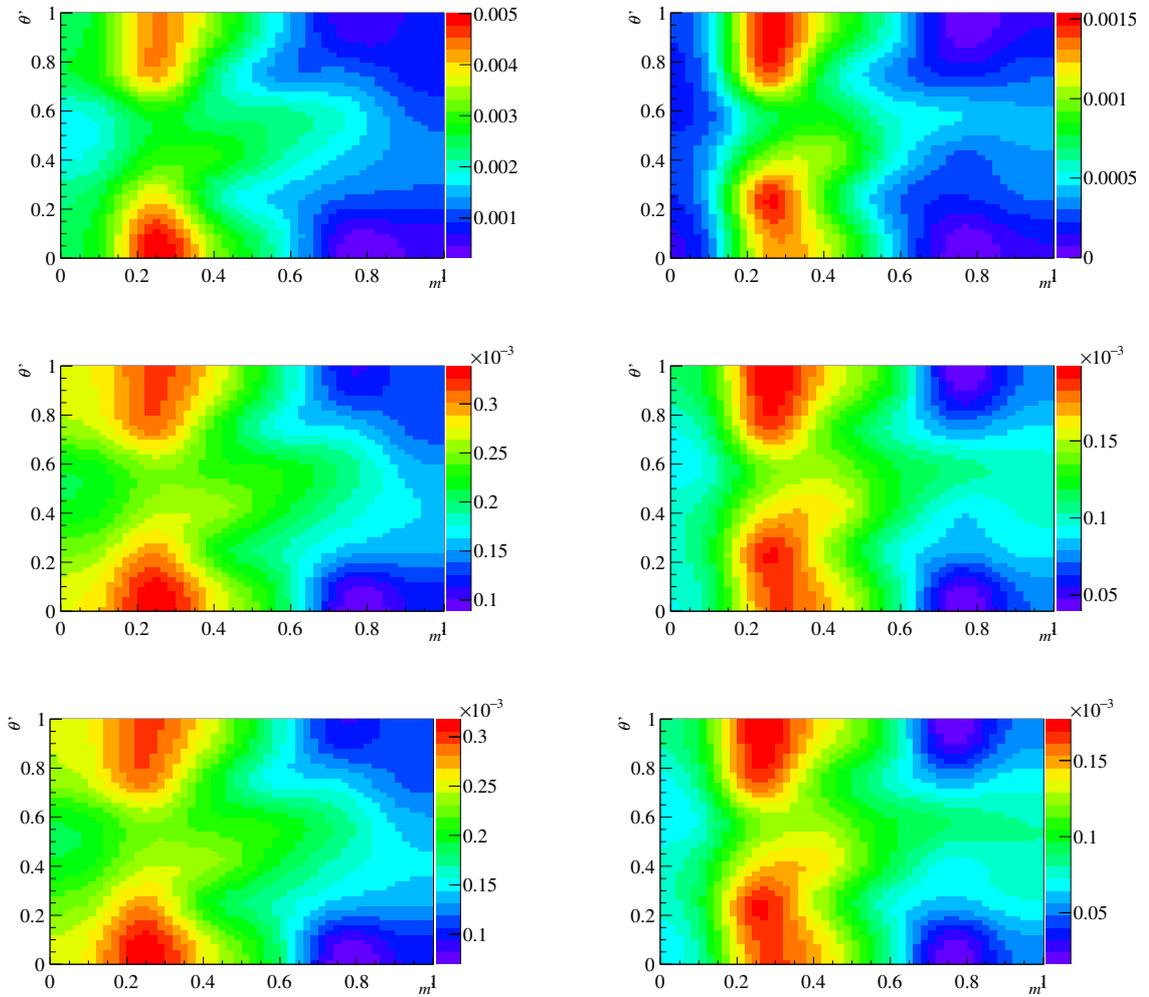
Figure 5.10 – Top: total selection efficiency for 2011, Down-Down events, in the TOS (left) and TIS&!TOS (right) trigger category. Middle: upper uncertainty on the selection efficiency for 2011, Down-Down events, in the TOS (left) and TIS&!TOS (right) trigger category. Bottom: lower uncertainty on the selection efficiency for 2011, Down-Down events, in the TOS (left) and TIS&!TOS (right) trigger category.

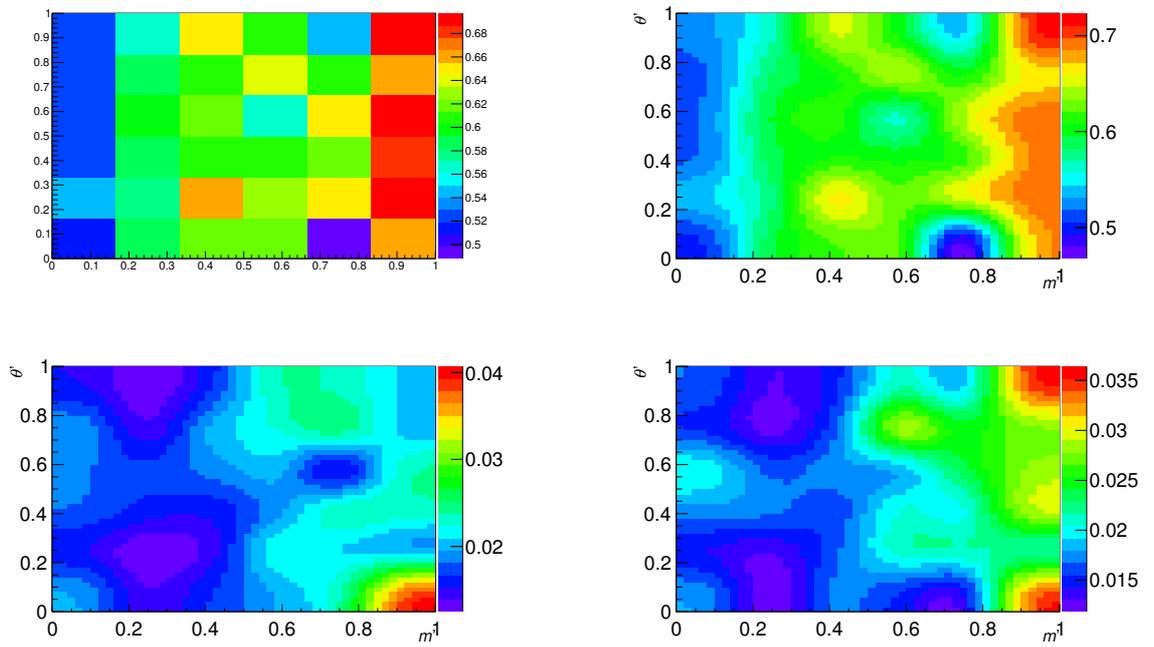Figure 5.11 – Top: distributions over the Dalitz plot of the PID efficiency as extracted from Monte-Carlo samples (left), and splined (right). Bottom left (right): upper (lower) uncertainties on the PID efficiency.
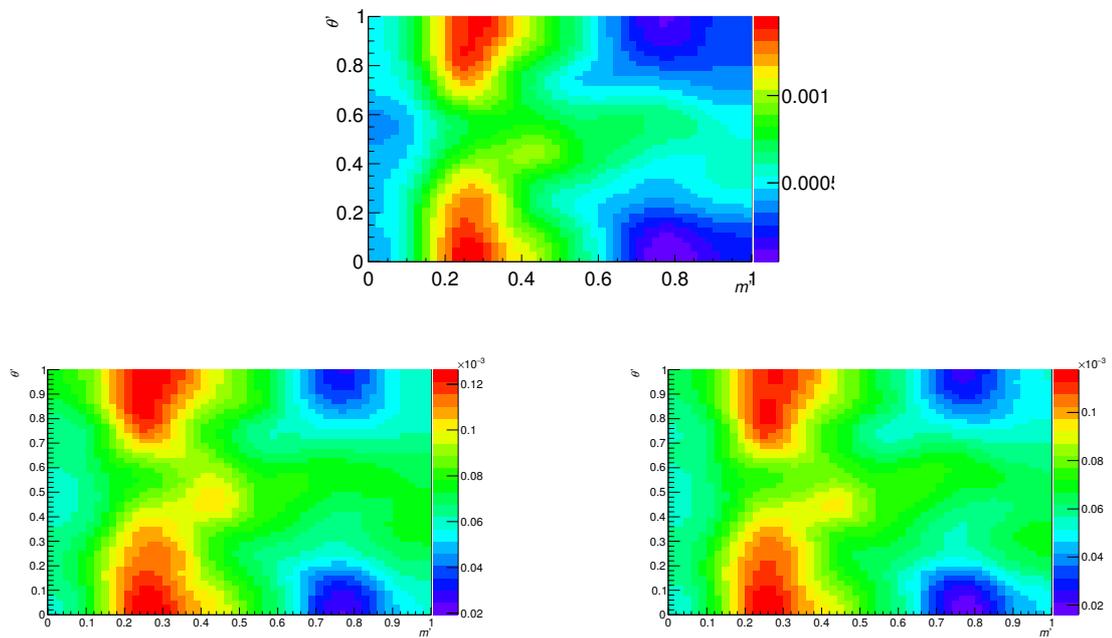
Figure 5.12 – Top: variations of the total signal efficiency over the Dalitz plot, for 2011, Down-Down events. Bottom: upper (lower) uncertainties on the total signal efficiency are shown on the left (right).

5.6 Data-fit model

The main goal of this analysis is to provide a starting point for future analyses of the $B^0 \to K_S^0 K^+ K^-$ mode in LHCb, as well as bringing additional information that can be combined with the latest result from BaBar. The baseline model of our analysis is similar to that from BaBar, and we present it in Sec. 5.6.1. We aim at improving its description of the data by adding resonances, removing resonances, or changing their parameterization. We compare the quality of different models by means of goodness-of-fit methods, presented in App. B. Section 5.6.2 presents the results of this fit to data.

5.6.1 Baseline model

Table 5.7 details the intermediate resonances composing the baseline fit model, and their parameterizations.

The nonresonant (NR) component of the amplitude is described as the sum of a S-wave and a P-wave, both modelled as a second-degree polynomial in the parameter Ω, defined as

$$\Omega = \frac{1}{2}(m_B + \frac{1}{3}(m_{K^+} + m_{K^-} + m_{K_S^0})). \tag{5.30}$$

Fits to data using different models and sets of fixed parameter consistently result in a small ($< 0.1\%$) fit fraction for the first-degree term of the P-wave. In order to improve the stability of the fit, we remove this component. The nonresonant amplitude is then the sum of five terms, PolNR(S0), PolNR(S1), PolNR(S2), PolNR(P0), and PolNR(P2), where the "S" and "P" letters stand for S-wave and P-wave, respectively. The indexes 0, 1, and 2 designate the degree of the polynomial term in Ω.

Table 5.7 – Modelling of the resonances used in the model. Masses and widths are given in MeV/c^2.

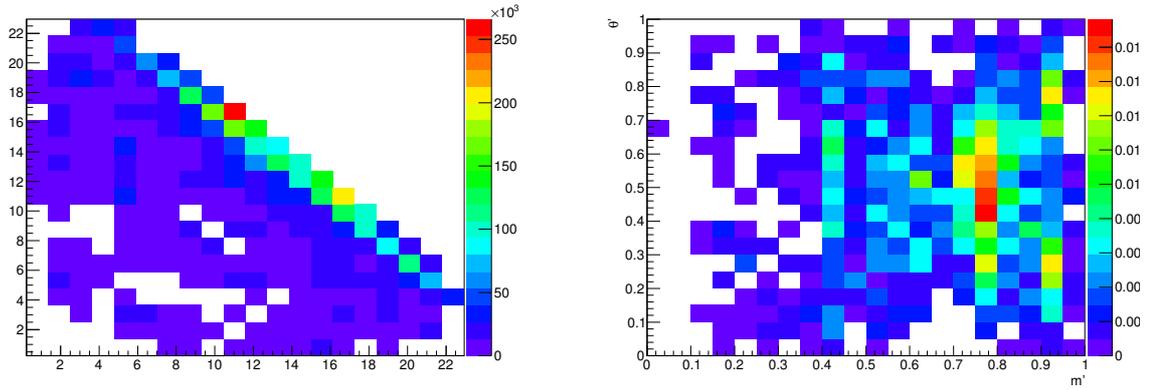| Resonance	Shape parameters	Lineshape
ϕ^0	m = 1019.455 ± 0.020, Γ = 4.26 ± 0.04	Rel. BW
$f_0(980)$	m = 965 ± 10, g_π = (0.165 ± 0.018) GeV^2/c^4, g_K/g_π = 4.21 ± 0.33	Flatté
$f_0(1500)$	m = 1505 ± 6, Γ = 109 ± 7	Rel. BW
$f_0(1710)$	m = 1720 ± 6, Γ = 135 ± 8	Rel. BW
$f_2'(1525)$	m = 1525 ± 5, Γ = 73^{+6}_{-5}	Rel. BW
χ_{c0}	m = 3414.75 ± 0.31, Γ = 10.3 ± 0.6	Rel. BW

Figure 5.13 – Distribution of data events from all data-taking periods and K_S^0 reconstruction modes over the Dalitz plot (left) and the square Dalitz plot (right).
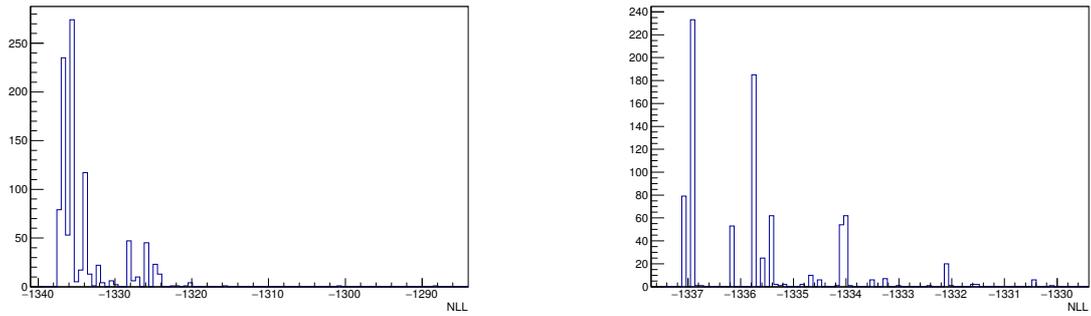
Figure 5.14 – (left) Distribution of negative log-likelihood for all converged fits on data. (right) Distribution of negative log-likelihood for converged fits close to the best minimum.

5.6.2 Fit results

Figure 5.13 shows the distribution of data events over the Dalitz plot and the square Dalitz plot, combining datasets from all K_S^0 reconstruction modes and all data-taking periods.

We perform 1000 fits to data using randomized initial values for all parameters, and show the obtained likelihood values in Fig. 5.14. A clear best minimum is present, and there are 15 secondary minima within a 3σ interval. Table 5.8 shows the isobar parameters and fit fractions for the best minimum, along with the statistical uncertainties. The sum of fit fractions is different from 100%, indicating significant interference between resonances, as expected. Table 5.9 shows the interference fit fractions between all the intermediate states.

Figure 5.15 shows the fit fractions of all the intermediate states for the best minimum

134

Table 5.8 – Isobar parameters and fit fractions for the best minimum, along with uncertainties as estimated by `MIGRAD`.

Resonance	Fit fraction (%)	c	ϕ
ϕ^0 (1020)	14.02	0.70±0.14	0.35 (fixed)
$f_0(980)$	28.34	0.99±0.19	1.92±0.62
$f_0(1500)$	4.50	0.40±0.09	-1.77±0.43
$f_2'(1525)$	4.28	0.39±0.10	0.14±0.39
$f_0(1710)$	1.70	0.24±0.07	-0.19±0.39
χ_{c0}	2.87	0.32±0.07	-1.29±0.30
PolNR(S0)	20.16	0.84 (fixed)	0.00 (fixed)
PolNR(S1)	9.38	0.57±0.11	-3.30±0.30
PolNR(S2)	2.97	0.32±0.13	3.84±0.37
PolNR(P0)	23.86	0.91±0.21	1.13 (fixed)
PolNR(P2)	8.41	0.54±0.15	-2.29 (fixed)
Sum	120.48		

Table 5.9 – Measured interference fit fractions corresponding to the best minimum. The A0–10 indexes correspond to, in order, ϕ^0, $f_0(980)$, $f_0(1500)$, $f_2'(1525)$, $f_0(1710)$, χ_{c0}, PolNR_S0, PolNR_S1, PolNR_S2, PolNR_P0, and PolNR_P2.

	A0	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
A0	14.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-7.24	5.69
A1		28.34	1.17	0.00	-4.56	0.54	-1.80	-3.57	9.99	0.00	0.00
A2			4.50	0.00	2.47	-0.02	3.92	3.35	-1.24	0.00	0.00
A3				4.28	-0.00	0.00	0.00	0.00	0.00	0.00	0.00
A4					1.70	-0.11	-1.17	0.56	-1.01	0.00	0.00
A5						2.87	1.25	-0.13	-0.14	0.00	0.00
A6							20.16	-1.71	-8.31	0.00	0.00
A7								9.38	0.59	0.00	0.00
A8									2.97	0.00	0.00
A9										23.86	-19.02
A10											8.41

and the 15 closest secondary minima. The fit fractions of the ϕ^0 corresponding to different minima are similar, whereas mirror solutions for the χ_{c0} and the $f_0(1710)$ are clearly distinguishable. The fit fractions of broad scalars such as the $f_0(980)$ and the nonresonant S-wave strongly vary between the solutions. We thus do not extract a Q2B branching fraction for these modes.
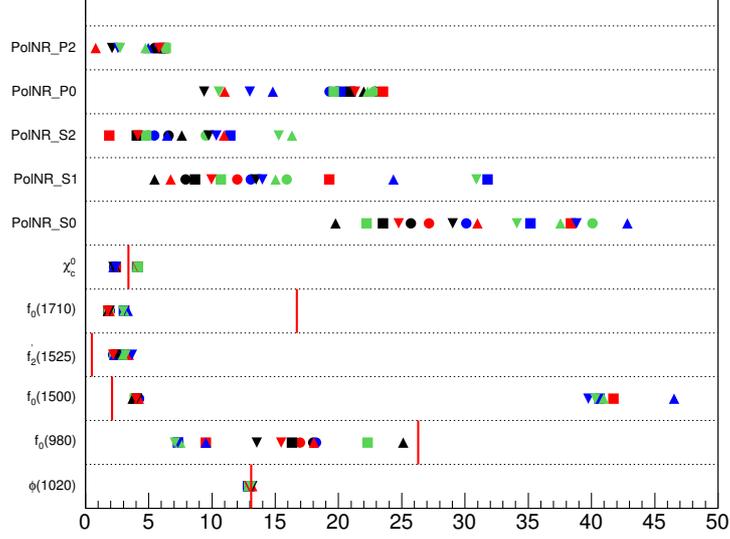
Figure 5.15 – Fit fractions (in %) of each intermediate state for the best minimum and the 15 closest secondary minima. Each solution corresponds to an unique marker (colour and form). The best minimum is marked by a black circle. The fit fraction measured by BaBar is indicated by the red, vertical line for each resonance.

Figures 5.16 and 5.17 show projections of the fit results on different Dalitz-plot variables and the pulls distribution over the Dalitz plane, all the K_S^0 reconstruction modes and data-taking periods taken together. We notice several localized discrepancies, but an overall satisfactory agreement, especially for m′ and θ′.

We also calculate the angular moments, defined as

$$< P_l(\cos(\theta_{K^+K^-})) > = \int_{-1}^{1} d\Gamma P_l(\cos(\theta_{K^+K^-}) d\cos(\theta_{K^+K^-})), \tag{5.31}$$

where P_l is the lth Legendre polynomial, Γ is the differential decay rate, and $\theta_{K^+K^-}$ is the helicity angle between K^+ and K_S^0. They constitute an alternative representation to the ordinary DP projection, and provide more information as they probe the angular distribution of data. Indeed, considering that there is no partial-wave amplitudes of spin 3 or higher, the amplitude writes as

$$\begin{aligned}\mathcal{A}(m_{K^+K^-}, \theta_{K^+K^-}) = &\mathcal{A}_S(m_{K^+K^-}, \theta_{K^+K^-}) P_0(\cos(\theta_{K^+K^-})) \\ &+ \mathcal{A}_P(m_{K^+K^-}, \theta_{K^+K^-}) e^{i\phi_P(m_{K^+K^-})} P_1(\cos(\theta_{K^+K^-})) \\ &+ \mathcal{A}_D(m_{K^+K^-}, \theta_{K^+K^-}) e^{i\phi_D(m_{K^+K^-})} P_2(\cos(\theta_{K^+K^-})), \tag{5.32}\end{aligned}$$
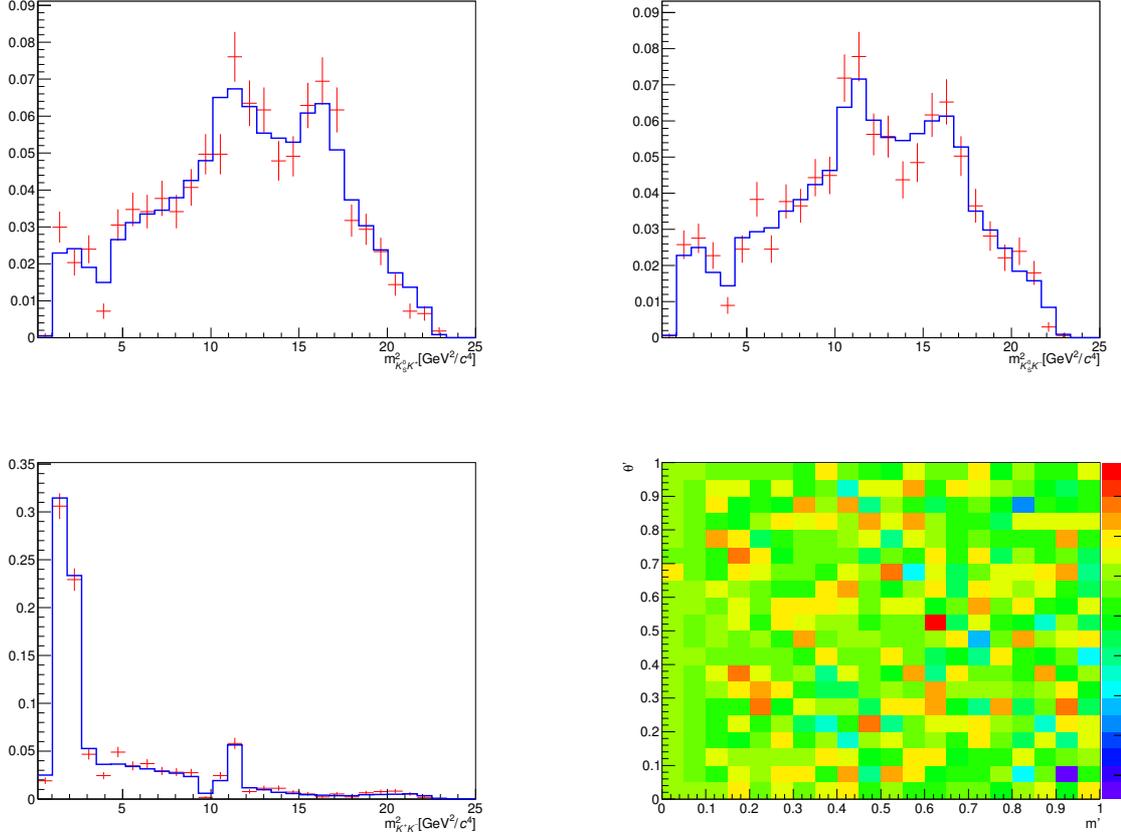
Figure 5.16 – Distributions of $m^2_{K^0_S K^+}$ (top, left), $m^2_{K^0_S K^-}$ (top, right), and $m^2_{K^+ K^-}$ (bottom, left) in data events (red dots) and in a Monte-Carlo sampled from the fit results (blue dots). Bottom, right: pull distribution between data events and the fit result.

where $\mathcal{A}_{S,P,D}$ and $\phi_{P,D}$ are real-valued functions of $m_{K^+ K^-}$ (ϕ_S is absorbed in the definition of the phases). Using the orthogonality of Legendre polynomials

$$\int_{-1}^{1} P_i(\cos\theta_{K^+K^-})P_j(\cos\theta_{K^+K^-})d\cos\theta_{K^+K^-} = \frac{2}{2l+1}\delta_{ij}, \tag{5.33}$$

Figure 5.17 – Distribution of m′ (left) and θ′ (right) in data events (red dots) and in a Monte-Carlo sampled from the fit results (blue dots).

we can relate the average of angular moments to the amplitude[10]

$$
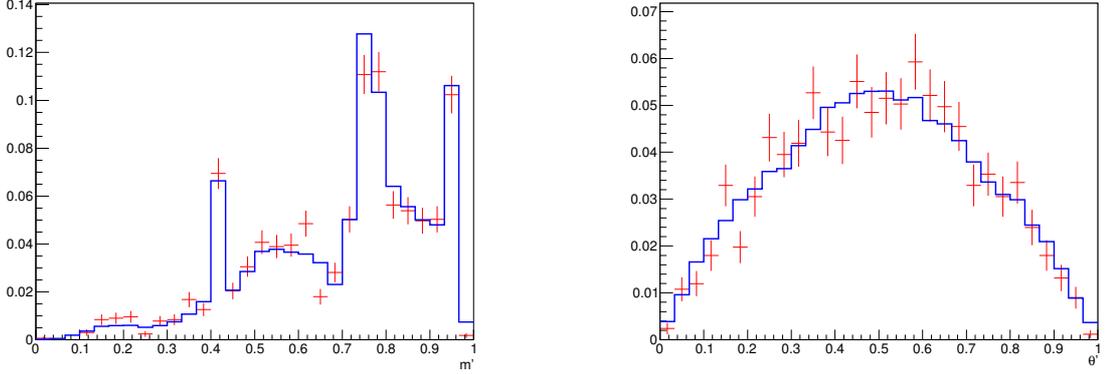< P_0 > = \frac{|\mathcal{A}_S|^2 + |\mathcal{A}_P|^2 + |\mathcal{A}_D|^2}{\sqrt{2}};
$$

$$
< P_1 > = \sqrt{2}\mathcal{A}_S\mathcal{A}_P \cos(\phi_P) + \frac{2\sqrt{10}}{5}\mathcal{A}_\mathcal{P}\mathcal{A}_\mathcal{D} \cos(\phi_P - \phi_D);
$$

$$
< P_2 > = \sqrt{\frac{2}{5}}\mathcal{A}_P^2 + \frac{\sqrt{10}}{7}\mathcal{A}_D^2 + \sqrt{2}\mathcal{A}_S\mathcal{A}_D \cos \phi_D;
$$

$$
< P_3 > = \frac{3}{5}\sqrt{\frac{30}{7}}\mathcal{A}_P\mathcal{A}_D \cos(\phi_P - \phi_D);
$$

$$
< P_4 > = \frac{\sqrt{18}}{7}\mathcal{A}_D^2. \tag{5.34}
$$

The analysis is not flavour-tagged, and as a result we observe the sum of B^0 and $\overline{B}^0$ contributions. Assuming no CP violation and as the model only includes K^+K^- resonances, the partial-wave amplitudes fulfill

$$
\mathcal{A}_S = \overline{\mathcal{A}_S}
$$
$$
\mathcal{A}_P = -\overline{\mathcal{A}_P}
$$
$$
\mathcal{A}_D = \overline{\mathcal{A}_D}, \tag{5.35}
$$

where $\overline{\mathcal{A}}$ refers to the $\overline{B}^0$ amplitude. As a result, terms that are a product between odd and even waves cancel out. The $< P_1 >$ and $< P_3 >$ terms only contain such terms, and

---

[10]Dependencies on $m_{K^+K^-}$ have been dropped for clarity purposes.

138

Figure 5.18 – Distributions of angular moments as a function of $m_{K^+K^-}$ in data (red points) and in a Monte-Carlo sampled from the fit results (blue line) for $l = 0, 1, 2, 3$ (top, left; top, right; bottom, left; bottom, right).

we thus expect a constant distribution compatible with 0 in data.[11]

Figures 5.18 and 5.19 show the projection of data events and of the fit result on these variables for $l < 4$ as a function of $m_{K^+K^-}$ in the entire mass range and in the $m_{K^+K^-} < 2\,\text{GeV}/c^2$ interval, respectively. The agreement between data and the model is satisfying.

5.7 Fit validation

The stability of the fit model is ensured by means of toy studies comparable to those described in Sec. 4.5. This procedure also allows to estimate the biases and the statistical

---

[11]This is equivalent to the fact that we are not sensitive to relative phases between even and odd waves, as $< P_1 >$ and $< P_3 >$ can be interpreted as the mean effect of the interference between even and odd partial waves.
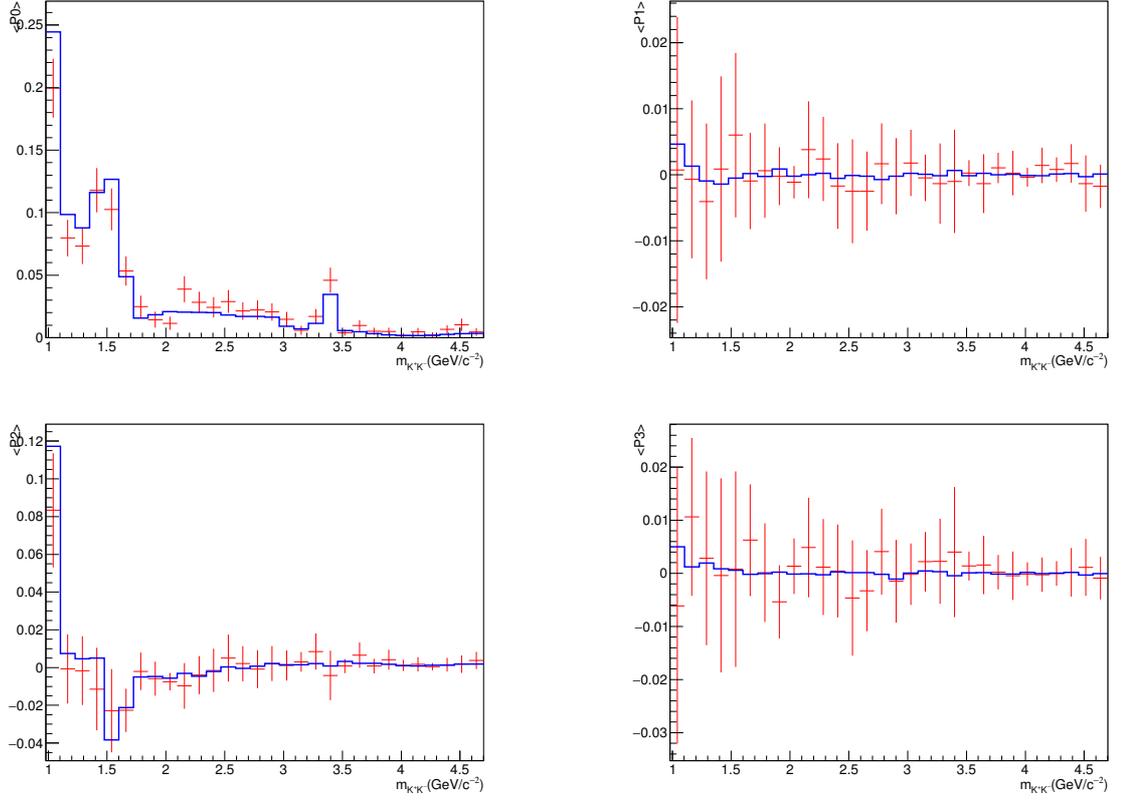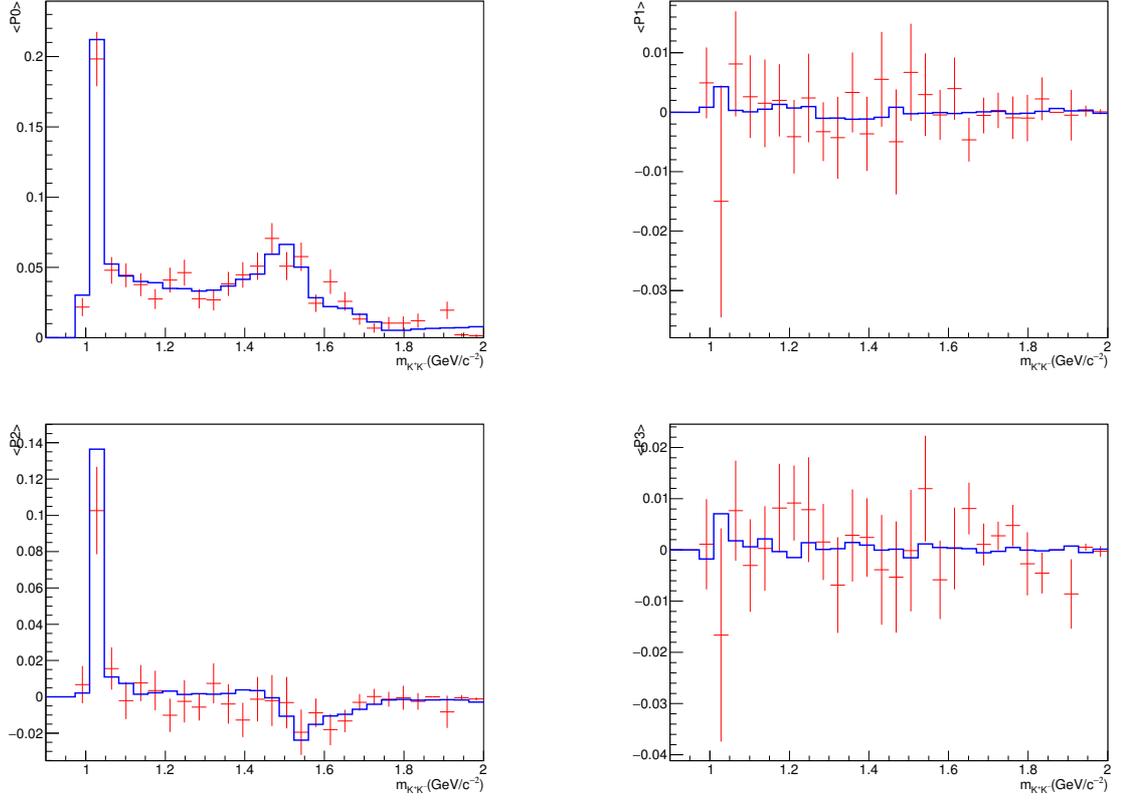
Figure 5.19 – Distributions of angular moments as a function of $m_{K^+K^-}$ in data (red points) and in a Monte-Carlo sampled from the fit results (blue line) for $l = 0, 1, 2, 3$ (top, left; top, right; bottom, left; bottom, right), in the $m_{K^+K^-} < 2\,\mathrm{GeV}/c^2$ interval.

uncertainties on fit fractions. Table 5.10 summarizes the isobar parameters for which a significant bias is observed.

Fit fractions are non-linear combinations of isobar parameters (see Eq. 1.56), and the estimation of their statistical uncertainty from the fit is difficult. We estimate the asymmetric uncertainty on fit fractions using the toy studies previously mentioned, by fitting the resulting distribution of residuals with a bifurcated Gaussian. We check that the interval defined as such contains 68% of toys, and thus that the uncertainties are correctly covered. Figure 5.20 show some of these residual distributions, and Table 5.11 shows the resulting uncertainties, along with the measured biases.

Additionally, we perform a likelihood scan of each fit parameter in order to check their consistency with the uncertainties on isobar parameters. Figure 5.21 shows the result of this procedure on the parameters of the $f_0(980)$ contribution, along with the uncertainty provided by the nominal fit. Table 5.12 details the uncertainties as obtained

Table 5.10 – Summary of significant biases of isobar parameters, along with the values obtained for the best minimum.

Parameter	Value	Bias
$c(\phi^0)$	0.6	-0.1
$c(f_0(980))$	0.7	-0.067
$\phi(f_0(980))$	2.4	-0.95
$c(f_0(1500))$	0.34	-0.075
$\phi(f_0(1500))$	-1.4	-0.49
$c(f_2'(1525))$	0.27	-0.065
$\phi(f_2'(1525))$	0.28	-0.55
$c(f_0(1710))$	0.23	-0.037
$\phi(f_0(1710))$	-0.0071	-0.28
$c(\chi_{c0})$	0.25	-0.051
$\phi(\chi_{c0})$	-1.3	-0.052
$c(\text{PolNR(S1)})$	0.47	-0.0049
$\phi(\text{PolNR(S1)})$	-3.1	-0.62
$c(\text{PolNR(S2)})$	0.42	-0.0037
$\phi(\text{PolNR(S2)})$	3.9	0.34
$c(\text{PolNR(P0)})$	0.79	-0.18
$c(\text{PolNR(P2)})$	0.39	-0.073
$\phi(\text{PolNR(P2)})$	-2.3	0.26

Table 5.11 – Statistical uncertainties and biases of fit fractions estimated by fitting the distribution of residuals of toy experiments with a bifurcated Gaussian.

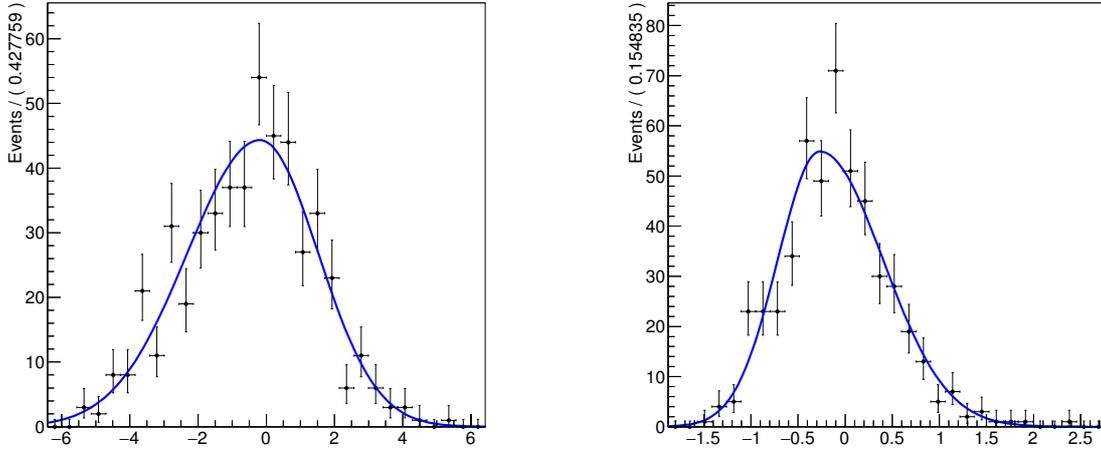| Parameter	Value (%)	Bias (%)	Lower unc. (%)	Upper unc. (%)
$FF(\phi^0)$	13	-0.32	1.4	2
$FF(f_0(980))$	18	-9.6	3.6	21
$FF(f_0(1500))$	4.2	-0.97	0.96	2.3
$FF(f_2'(1525))$	2.7	-0.52	0.64	1.8
$FF(f_0(1710))$	1.9	-0.65	0.46	2.1
$FF(\chi_{c0})$	2.3	-0.31	0.28	0.79
$FF(\text{PolNR(S0)})$	26	2.3	5.7	16
$FF(\text{PolNR(S1)})$	7.9	0.17	2.2	9
$FF(\text{PolNR(S2)})$	6.6	-2.6	1.8	11
$FF(\text{PolNR(P0)})$	23	-2.6	5.1	4.7
$FF(\text{PolNR(P2)})$	5.5	-0.63	3.6	3.7

Figure 5.20 – Distributions of residuals of the ϕ^0 and χ_{c0} fit fractions (left and right, respectively).

from `MINOS` and the uncertainties from the likelihood scans for each isobar parameter.[12] We notice that uncertainties from the likelihood scans are systematically larger than the uncertainties from `MINOS`. The values are however close, and the significant asymmetry of the uncertainties for certain parameters justifies the use of `MINOS` uncertainties in the following rather than those from `MIGRAD`.

5.8 Evaluation of systematic uncertainties

5.8.1 Fit-bias estimation

As described in Sec. 5.7, the model is validated using pseudo-experiments. For each parameter, we extract the average bias and, in case it is significant, assign the systematic uncertainty

$$\Delta_X = \mid \frac{\delta_X}{2} \mid, \tag{5.36}$$

where δ_X is the bias measured on the parameter X.

5.8.2 General method to evaluate systematic uncertainties

The method to extract systematic uncertainties is similar to that exposed in Sec 4.6. We divide systematic uncertainties in two categories:

---

[12]The 1σ interval around a minimum can be defined by the closest values for which the NLL is 0.5 larger than at that minimum.
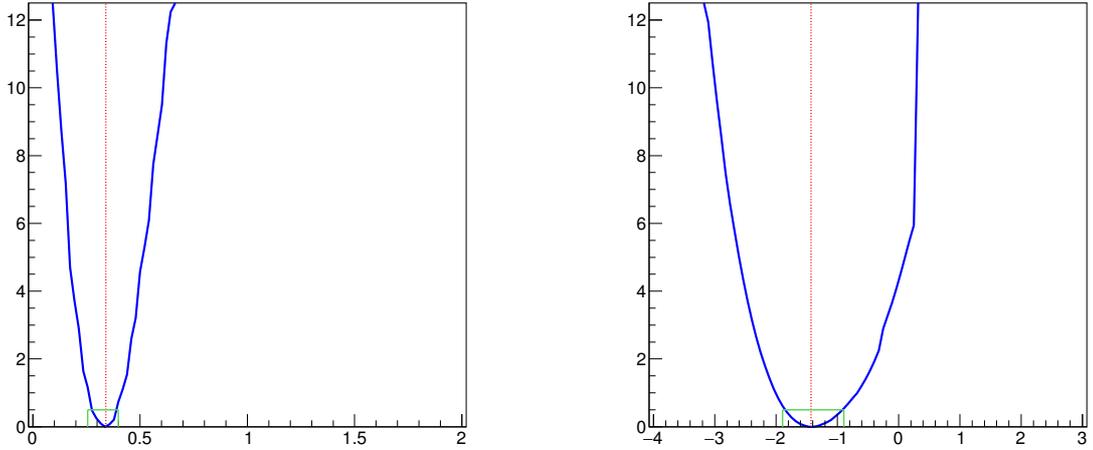
Figure 5.21 – Likelihood scans of the magnitude (left) and the phase (right) of the $f_0(1500)$, shifted by the minimum value and in the [0,12.5] interval (corresponding to a 5σ distance from the minimum). The vertical, dotted red line indicates the minimum found by the nominal fit. The values for which the NLL crosses the $\Delta(\text{NLL}) = 0.5$ line (1σ) are indicated by the green line.

- systematic uncertainties related to assumptions on fixed parameters, such as efficiencies, resonance parameters, and yields;

- systematic uncertainties related to the choice of the model used to fit the data.

The first kind of uncertainties is estimated by varying the fixed parameters within their uncertainties, taking correlations into account whenever possible, and fitting the model to data with the randomized parameter. The uncertainty Δ_X on a parameter X is then

$$\Delta_X = \sqrt{\sigma_X^2 + (\frac{\mu_X}{2})^2}, \tag{5.37}$$

where μ_X and σ_X are mean and the rms of the distribution of residuals calculated between the nominal fit and the fits using randomized parameters.

The second kind of uncertainties is estimated by fitting the concurrent model to data, and using the result of that fit to generate pseudo-experiments. These pseudo-experiments are then fitted using both the nominal and the concurrent model. The systematic uncertainty on a parameter X is then calculated using Eq. 5.37, where μ_X and σ_X are the mean and the rms of the distribution of residuals calculated between the fits using the nominal model and those using the concurrent model.

Table 5.12 – Isobar parameters along with uncertainties extracted from `MIGRAD`, `MINOS`, or the likelihood scans.

Parameter	Value	Stat. (fit)	Stat. (scan)	Stat. (`MINOS`)
$c(\phi^0)$	0.6	± 0.072	$\pm^{0.088}_{0.096}$	$\pm^{0.062}_{0.048}$
$\phi(\phi^0)$	0.35	fixed	fixed	fixed
$c(f_0(980))$	0.7	± 0.16	$\pm^{0.17}_{0.22}$	$\pm^{0.064}_{0.062}$
$\phi(f_0(980))$	2.4	± 0.79	$\pm^{0.65}_{0.71}$	$\pm^{0.41}_{0.63}$
$c(f_0(1500))$	0.34	± 0.054	$\pm^{0.058}_{0.085}$	$\pm^{0.043}_{0.038}$
$\phi(f_0(1500))$	-1.4	± 0.45	$\pm^{0.54}_{0.46}$	$\pm^{0.37}_{0.3}$
$c(f_2'(1525))$	0.27	± 0.056	$\pm^{0.063}_{0.06}$	$\pm^{0.058}_{0.046}$
$\phi(f_2'(1525))$	0.28	± 0.43	$\pm^{0.47}_{0.45}$	$\pm^{0.37}_{0.29}$
$c(f_0(1710))$	0.23	± 0.044	$\pm^{0.048}_{0.054}$	$\pm^{0.034}_{0.033}$
$\phi(f_0(1710))$	-0.0071	± 0.4	$\pm^{0.47}_{0.39}$	$\pm^{0.34}_{0.24}$
$c(\chi_{c0})$	0.25	± 0.038	$\pm^{0.044}_{0.038}$	$\pm^{0.035}_{0.027}$
$c(\mathrm{PolNR(S0)})$	0.84	fixed	fixed	fixed
$\phi(\mathrm{PolNR(S0)})$	0	fixed	fixed	fixed
$c(\mathrm{PolNR(S1)})$	0.47	± 0.089	$\pm^{0.096}_{0.088}$	$\pm^{0.098}_{0.14}$
$\phi(\mathrm{PolNR(S1)})$	-3.1	± 0.3	$\pm^{-0.034}_{0.39}$	$\pm^{0.24}_{0.16}$
$c(\mathrm{PolNR(S2)})$	0.42	± 0.11	$\pm^{0.12}_{0.13}$	$\pm^{0.18}_{0.23}$
$\phi(\mathrm{PolNR(S2)})$	3.9	± 0.21	$\pm^{-0.29}_{0.20}$	$\pm^{0.22}_{0.6}$
$c(\mathrm{PolNR(P0)})$	0.79	± 0.12	$\pm^{0.12}_{0.13}$	$\pm^{0.11}_{0.088}$
$\phi(\mathrm{PolNR(P0)})$	1.1	fixed	fixed	fixed
$c(\mathrm{PolNR(P2)})$	0.39	± 0.12	$\pm^{0.11}_{0.13}$	$\pm^{0.065}_{0.058}$
$\phi(\mathrm{PolNR(P2)})$	-2.3	fixed	fixed	fixed

5.8.3 Efficiencies

The method to evaluate efficiencies across the Dalitz plot along with the uncertainties on the efficiency values has been discussed in Sec. 5.5. We fit the model to data using alternative efficiency maps obtained by varying the nominal efficiency maps within their uncertainties. We neglect correlations between data-taking periods and neighbouring bins in the histograms, and show the results in Table 5.13.

5.8.4 Signal and background yields estimations

As explained in Sec. 5.3, signal and background yields are fixed in the fit to data, and taken from the invariant-mass fit. In order to take into account all possible correlations, we use the full covariance matrix from the invariant-mass fit to obtain a set of alternative yields and shape parameters. The integrals of the contributions in the signal mass window are recalculated.

Table 5.14 shows the systematic uncertainties related to the estimation of signal and

Table 5.13 – Systematic uncertainties on isobar parameters and fit fractions related to the estimation of efficiencies.

Resonance	Fit fraction (%)	c	ϕ
ϕ^0	0.17	0.03	fixed
$f_0(980)$	4.5	0.1	0.68
$f_0(1500)$	0.18	0.017	0.25
$f_2'(1525)$	0.17	0.017	0.25
$f_0(1710)$	0.12	0.012	0.21
χ_{c0}	0.052	0.013	0.14
PolNR(S0)	3	fixed	fixed
PolNR(S1)	1.1	0.024	0.12
PolNR(S2)	1.6	0.035	0.16
PolNR(P0)	1	0.047	fixed
PolNR(P2)	1.1	0.049	fixed

Table 5.14 – Systematic uncertainties on isobar parameters and fit fractions related to the estimations of the yields of event species.

Resonance	Fit fraction (%)	c	ϕ
ϕ^0	0.044	0.003	fixed
$f_0(980)$	0.23	0.0026	0.026
$f_0(1500)$	0.012	0.0019	0.012
$f_2'(1525)$	0.032	0.0032	0.0089
$f_0(1710)$	0.0073	0.0016	0.01
χ_{c0}	0.0083	0.0015	0.0045
PolNR(S0)	0.34	fixed	fixed
PolNR(S1)	0.072	0.0025	0.008
PolNR(S2)	0.054	0.0039	0.0019
PolNR(P0)	0.42	0.013	fixed
PolNR(P2)	0.11	0.0025	fixed

background yields on isobar parameters and fit fractions.

5.8.5 Background shapes

The distributions of background events over the Dalitz plot are considered separately for combinatorial and cross-feed contributions. We perform 200 fits to data, varying the histograms representing these distributions within their uncertainties. Neighbouring bins are considered as uncorrelated, as are histograms for different data-taking periods and K_S^0 reconstruction modes.

Table 5.15 – Systematic uncertainties on isobar parameters and fit fractions related to the modelling of backgrounds.

Resonance	Fit fraction (%)	c	ϕ
ϕ^0	0.028	0.002	fixed
$f_0(980)$	0.26	0.0051	0.024
$f_0(1500)$	0.015	0.0012	0.012
$f_2'(1525)$	0.035	0.0021	0.011
$f_0(1710)$	0.017	0.0012	0.0098
χ_{c0}	0.01	0.0011	0.006
PolNR(S0)	0.18	fixed	fixed
PolNR(S1)	0.11	0.0036	0.01
PolNR(S2)	0.13	0.0039	0.0066
PolNR(P0)	0.13	0.0043	fixed
PolNR(P2)	0.16	0.0054	fixed

Table 5.15 shows the systematic uncertainties on isobar parameters and fit fractions related to the shape of the combinatorial background shape and the cross-feeds.

5.8.6 Total experimental systematic uncertainties

The previous systematic uncertainties related to the imperfect knowledge of the event yield species and experimental setup are summed in quadrature and reported independently from other sources of systematic uncertainties in the final result. Table 5.16 shows the experimental systematic uncertainties on the isobar parameters and fit fractions for the preferred solution.

5.8.7 Resonance shape parameters

Resonance parameters such as masses and widths are rather well-known inputs from other experiments. We vary the mass and the width of each resonant component of the Dalitz-plot model, neglecting any correlation, and show the results in Table 5.17.

We also consider a systematic uncertainty related to Blatt-Weisskopf barrier factors, described in Sec. 1.3.5. When varying the values of these factors, we assume that it remains the same for all resonances. We vary independently the barrier factors of the mother particle and the resonances, and add the resulting systematic uncertainties in quadrature. Table 5.18 shows this systematic uncertainty.

The overall systematic uncertainty related to resonance shape parameters on a given quantity is the sum in quadrature of all the previously mentioned systematic uncertainties.

Table 5.16 – Summary of systematic uncertainties on fit fractions arising from the knowledge of event species and experimental setup.

Parameter	Efficiency	Shape cf.	Shape comb.	Yield ratios	Total
FF(ϕ^0)	0.17	TBD	0.028	0.044	0.18
FF($f_0(980)$)	4.5	TBD	0.26	0.23	4.5
FF($f_0(1500)$)	0.18	TBD	0.015	0.012	0.18
FF($f_2'(1525)$)	0.17	TBD	0.035	0.032	0.18
FF($f_0(1710)$)	0.12	TBD	0.017	0.0073	0.12
FF(χ_{c0})	0.052	TBD	0.01	0.0083	0.053
FF(PolNR(S0))	3	TBD	0.18	0.34	3
FF(PolNR(S1))	1.1	TBD	0.11	0.072	1.1
FF(PolNR(S2))	1.6	TBD	0.13	0.054	1.6
FF(PolNR(P0))	1	TBD	0.13	0.42	1.1
FF(PolNR(P2))	1.1	TBD	0.16	0.11	1.1

Table 5.17 – Systematic uncertainties on isobar parameters and fit fractions originating from fixed parameters in the lineshapes.

Resonance	Fit fraction (%)	c	ϕ
ϕ^0	0.065	0.019	fixed
$f_0(980)$	4.1	0.098	0.19
$f_0(1500)$	0.3	0.018	0.14
$f_2'(1525)$	0.13	0.013	0.073
$f_0(1710)$	0.21	0.012	0.086
χ_{c0}	0.042	0.0091	0.064
PolNR(S0)	1.7	fixed	fixed
PolNR(S1)	0.8	0.017	0.08
PolNR(S2)	1	0.031	0.11
PolNR(P0)	0.38	0.029	fixed
PolNR(P2)	0.19	0.015	fixed

5.8.8 Fixed isobar parameters

As described in Sec. 5.6, some isobar parameters are fixed in the fit to data as the sensitivity of an untagged analysis to these parameters is limited.[13] A systematic uncertainty related to the fixed parameters is assigned by varying their values within their uncertainties provided by the BaBar experiment. We neglect correlations between these uncertainties. Table 5.19 shows the resulting systematic uncertainties.

---

[13]The isobar parameters related to the constant term of the nonresonant S-wave are not included in this, as they set the reference for both the phases and the magnitudes.

Table 5.18 – Systematic uncertainties on isobar parameters and fit fractions originating from Blatt-Weisskopf barrier factors.

Resonance	FF	c	ϕ
ϕ^0	1.1	0.07	fixed
$f_0(980)$	1.1	0.062	0.092
$f_0(1500)$	0.092	0.034	0.16
$f_2'(1525)$	0.99	0.047	0.16
$f_0(1710)$	0.096	0.029	0.088
χ_{c0}	0.095	0.031	0.019
PolNR(S0)	5.5	fixed	fixed
PolNR(S1)	0.93	0.059	0.025
PolNR(S2)	1.3	0.0098	0.065
PolNR(P0)	5.5	0.12	fixed
PolNR(P2)	1.6	0.093	fixed

Table 5.19 – Systematic uncertainties on isobar parameters and fit fractions originating from fixed isobar parameters aside from the parameters of the constant term of the nonresonant S-wave.

Resonance	Fit fraction (%)	c	ϕ
ϕ^0	0.35	0.073	fixed
$f_0(980)$	6.5	0.07	0.85
$f_0(1500)$	0.13	0.056	0.4
$f_2'(1525)$	0.49	0.039	0.38
$f_0(1710)$	0.079	0.035	0.33
χ_{c0}	0.1	0.047	0.099
PolNR(S0)	11	fixed	fixed
PolNR(S1)	0.82	0.053	0.3
PolNR(S2)	3	0.043	0.064
PolNR(P0)	6	0.19	fixed
PolNR(P2)	2.4	0.13	fixed

5.8.9   Model uncertainties

We consider the possible presence of the following additional resonances: $f_0(1370)$, $f_2(1270)$, $f_2(2010)$, $f_2(2300)$, and $\phi(1680)$. No significant contribution from any of these resonances is found, and they are then only included in the model to evaluate systematic uncertainties.

In order to estimate a systematic uncertainty related to an alternative model, we generate 200 toy experiments using the fit of this model to data. We then fit each of these toys with the baseline model and the alternative model, and fit the distribution of the

Table 5.20 – Systematic uncertainties on fit fractions originating from the addition of a resonance in the model. They are then summed in quadrature in order to yield the total systematic uncertainty related to the model.

	$f_0(1370)$	$f_2(1270)$	$f_2(2010)$	$f_2(2300)$	$\phi 1370$	Sum
ϕ^0	0.58	0.89	0.67	0.00075	0.64	1.4
$f_0(980)$	13.	8.5	3.4	0.023	3.9	17.
$f_0(1500)$	2.2	0.64	0.39	0.0012	0.48	2.3
$f_2'(1525)$	0.59	0.73	0.52	0.00079	1.7	2.0
$f_0(1710)$	1.2	0.58	0.78	0.0011	0.69	1.7
χ_{c0}	0.13	0.22	0.17	0.00039	0.16	0.34
PolNR(S0)	14.	12.	8.1	0.018	5.9	21.
PolNR(S1)	8.5	11.	3.8	0.012	1.9	14.
PolNR(S2)	7.0	5.4	2.9	0.010	2.5	9.6
PolNR(P0)	3.7	4.6	4.8	0.0048	4.4	8.8
PolNR(P2)	1.9	3.6	2.2	0.0042	1.4	4.8

differences of the fit fractions between the two fits with a Gaussian. The corresponding systematic uncertainty is then

$$\Delta_X = \sqrt{\left(\frac{\mu}{2}\right)^2 + \sigma^2}, \tag{5.38}$$

where μ and σ are the mean and the width of the Gaussian.

Table 5.20 summarizes the model uncertainties on each fit fraction related to the addition of one of these resonances in the model. As expected, this is the largest contribution to the systematic uncertainties.

The addition of the $a_0^0(980)$ or the $a_0^\pm(980)$ causes the fit to converge to a solution that is rather far away from the global minimum. We thus do not assign a systematic uncertainty for it.

5.9 Conclusion

We have performed a preliminary Dalitz-plot analysis of the $B^0 \to K_S^0 K^+ K^-$ decay mode in LHCb, using $3\,\mathrm{fb}^{-1}$ of data from Run I, taking the result from BaBar [40] as a reference model. The distribution of events over the Dalitz plot is overall well modelled, as shown by Figs. 5.16 and 5.17. Table 5.21 shows the fit fractions of the different resonances along with their statistical and systematic uncertainties.

Table 5.21 – Fit fractions of each resonance, along with their statistical and systematic uncertainties. Each value is in %. The quoted value is corrected for the bias. The "Total" column refers to the total systematic uncertainty; "Exp." refers to the experimental systematic uncertainty, "FixParams" to the systematics uncertainty related to fixed isobar parameters, "Res. par." to the systematic uncertainty related to fixed resonance parameters, and "Add. res." refers to the systematic uncertainty related to alternative models.

	Value	Stat.	Bias	Exp.	FixParams	Res. par.	Add. res.	Total
FF(ϕ^0)	13.	$\pm^{1.7}_{2.1}$	0.090	0.18	0.35	1.1	1.4	1.8
FF($f_0(980)$)	13.	$\pm^{23.}_{2.2}$	5.5	4.5	6.5	4.3	17.	20.
FF($f_0(1500)$)	3.7	$\pm^{2.3}_{1.0}$	0.53	0.18	0.13	0.32	2.3	2.4
FF($f'_2(1525)$)	2.9	$\pm^{1.3}_{1.3}$	0.15	0.18	0.49	1.0	2.0	2.3
FF($f_0(1710)$)	2.2	$\pm^{1.8}_{1.8}$	0.28	0.12	0.079	0.23	1.7	1.7
FF(χ_{c0})	2.2	$\pm^{0.67}_{0.45}$	0.13	0.053	0.10	0.10	0.34	0.40
FF(PolNR(S0))	25.	$\pm^{18.}_{5.6}$	0.99	3.0	11.	5.8	21.	25.
FF(PolNR(S1))	7.3	$\pm^{9.0}_{1.8}$	0.63	1.1	0.82	1.2	14.	15.
FF(PolNR(S2))	4.5	$\pm^{12.}_{1.5}$	2.1	1.6	3.0	1.7	9.6	11.
FF(PolNR(P0))	21.	$\pm^{5.9}_{4.8}$	2.2	1.1	6.0	5.5	8.8	12.
FF(PolNR(P2))	4.4	$\pm^{5.1}_{20.}$	1.0	1.1	2.4	1.7	4.8	5.9

The fit model has many (16) solutions within 4.5 NLL units from the best minimum, corresponding to different interference patterns and fit fractions. Resolving these solutions, which could be made possible with a larger dataset, would help to make a final interpretation of the result. However, the fit fraction of the ϕ^0 does not depend on a specific minimum, and its value is compatible with the result from the B factories. We thus determine the Q2B branching fraction of this mode

$$\mathcal{B}(B^0 \to K^0_S(\phi^0 \to K^+K^-) = \left(1.63 \pm^{0.2}_{0.3} (\text{stat}) \pm 0.2(\text{syst}) \pm 0.2(\text{BF})\right) \times 10^{-6}, \quad (5.39)$$

where the uncertainties are statistical, systematic, and due to the uncertainty on $\mathcal{B}\,(B^0 \to K^0_S K^+K^-)$, respectively. This branching fraction is compatible with the PDG value

$$\mathcal{B}_{\text{PDG}}(B^0 \to K^0\phi^0) = (7.3 \pm 0.7) \times 10^{-6}. \quad (5.40)$$

We also extract branching fractions for the $B^0 \to K^0_S \chi_{c0}(\to K^+K^-)$, $B^0 \to K^0_S f'_2(1525)(\to K^+K^-)$, and $B^0 \to K^0_S f_0(1710)(\to K^+K^-)$ modes

$$\mathcal{B}(B^0 \to K^0_S \chi_{c0}(\to K^+K^-)) = \left(0.28 \pm^{0.08}_{0.06} (\text{stat}) \pm 0.05(\text{syst}) \pm 0.04(\text{BF})\right) \times 10^{-6},$$
$$\mathcal{B}(B^0 \to K^0_S f'_2(1525)(\to K^+K^-)) = (0.36 \pm 0.16(\text{stat}) \pm 0.29(\text{syst}) \pm 0.05(\text{BF})) \times 10^{-6}.$$
$$\mathcal{B}(B^0 \to K^0_S f_0(1710)(\to K^+K^-)) = (0.27 \pm 0.22(\text{stat}) \pm 0.21(\text{syst}) \pm 0.03(\text{BF})) \times 10^{-6}.$$
$$(5.41)$$

The branching fraction of $B^0 \to K^0_s f'_2(1525)(\to K^+ K^-)$ is in good agreement with PDG values. There is no such reference for $\mathcal{B}\,(B^0 \to K^0_s \chi_{c0}(\to K^+ K^-))$, but the value is consistent with both BaBar and Belle measurements. Finally, the branching fraction of $f_0(1710)$ differs significantly from the $(2.2 \pm 0.45) \times 10^{-6}$ value of the PDG.

Conclusion

The Run I dataset from the LHCb experiment, corresponding to $3\,\mathrm{fb}^{-1}$ recorded at centre-of-mass energies of 7 and 8 TeV, has been analyzed in order to search for the $B_s^0 \to K_\mathrm{S}^0 K^+ K^-$ decay and update LHCb measurement of the branching fraction of other $B_{d,s}^0 \to K_\mathrm{S}^0 h^+ h^{'-}$ modes. Furthermore, an untagged, time-independent Dalitz-plot analysis of $B^0 \to K_\mathrm{S}^0 K^+ K^-$ is performed in order to extract Q2B branching fractions.

The $B_s^0 \to K_\mathrm{S}^0 K^+ K^-$ mode is observed for the first time with a global significance of $3.7\,\sigma$. The results obtained of the other $B^0 \to K_\mathrm{S}^0 h^\pm h^{'\mp}$ relative branching fractions are

$$\frac{\mathcal{B}\left(B_s^0 \to K_\mathrm{S}^0 \pi^+ \pi^-\right)}{\mathcal{B}\left(B^0 \to K_\mathrm{S}^0 \pi^+ \pi^-\right)} = 0.26 \pm 0.04(\mathrm{stat.}) \pm 0.02(\mathrm{syst.}) \pm 0.01(\mathrm{f_s/f_d}),$$

$$\frac{\mathcal{B}\left(B^0 \to K_\mathrm{S}^0 K^\pm \pi^\mp\right)}{\mathcal{B}\left(B^0 \to K_\mathrm{S}^0 \pi^+ \pi^-\right)} = 0.17 \pm 0.02(\mathrm{stat.}) \pm 0.00(\mathrm{syst.}),$$

$$\frac{\mathcal{B}\left(B_s^0 \to K_\mathrm{S}^0 K^\pm \pi^\mp\right)}{\mathcal{B}\left(B^0 \to K_\mathrm{S}^0 \pi^+ \pi^-\right)} = 1.84 \pm 0.07(\mathrm{stat.}) \pm 0.02(\mathrm{syst.}) \pm 0.04(\mathrm{f_s/f_d}),$$

$$\frac{\mathcal{B}\left(B^0 \to K_\mathrm{S}^0 K^+ K^-\right)}{\mathcal{B}\left(B^0 \to K_\mathrm{S}^0 \pi^+ \pi^-\right)} = 0.59 \pm 0.02(\mathrm{stat.}) \pm 0.01(\mathrm{syst.}),$$

$$\tag{5.42}$$

They are compatible with results obtained by B-factories, and coherent with previous measurements from LHCb, at the exception of $B^0 \to K_\mathrm{S}^0 K^+ K^-$. These results are used as a baseline to extract signal yields and background distributions for the amplitude analyses of $B^0 \to K_\mathrm{S}^0 \pi^+ \pi^-$, $B_s^0 \to K_\mathrm{S}^0 K^\pm \pi^\mp$, and $B^0 \to K_\mathrm{S}^0 K^+ K^-$.

The first flavour-untagged, time-independent Dalitz-plot analysis of $B^0 \to K_\mathrm{S}^0 K^+ K^-$ in LHCb is performed with a reoptimized BDT selection. The amplitude is modelled as the sum of a ϕ^0, $f_0(980)$, $f_0(1500)$, $f_0(1710)$, $f_2'(1525)$, χ_{c0}, and a nonresonant component. This nonresonant amplitude is described similarly to that from the latest BaBar analysis, as a sum of an S-wave and a P-wave, both modelled as second-degree polynomials. Fit fractions are extracted, and quasi-two-body branching fractions are measured for the Q2B modes $B^0 \to K_\mathrm{S}^0 \phi^0$, $B^0 \to K_\mathrm{S}^0 f_2'(1525)$, $B^0 \to K_\mathrm{S}^0 \chi_{c0}$, and $B^0 \to K_\mathrm{S}^0 f_0(1710)$. These branching fractions are compatible with results from B factories, with the exception of the latter.

The study of $B_{d,s}^0 \to K_\mathrm{S}^0 h^+ h^{'-}$ decays will benefit from a growing dataset in the next few years; amplitude analyses of the suppressed modes such as $B_s^0 \to K_\mathrm{S}^0 \pi^+ \pi^-$ could

become possible for the first time. Amplitude analyses of the favoured modes such as $B^0 \rightarrow K^0_S K^+ K^-$ will enter a new phase, with the addition of a tagging information that will allow to disentangle B and $\overline{B}$ contributions. Indeed, this information is necessary not only to perform CP violation measurements, but also to measure the relative phases of even and odd partial waves.

Appendix A

Correcting sWeights in the presence of fixed yields

Subtracting background from the signal in physics analyses can be performed in several ways. We described in Sec. 4.2.3 the $_s\mathcal{P}lots$ procedure [65], which is an efficient and widespread method to do so; it is implemented in the RooStats package [72]. In this appendix, we focus on the effect of fixed yields on the results, which is documented in Ref. [73], Annexes B.1 and B.2.

Section A.1 shortly reminds the $_s\mathcal{P}lots$ subtraction procedure and its modifications in the presence of fixed yields that was presented in Sec. 4.2.3. Section A.2 then reviews the implementation of $_s\mathcal{P}lots$ in RooStats and the issues it may introduce. Finally, an alternative constructor to the RooStats::SPlot class that would solve these issues is proposed in Sec A.3, along with some tests.

A.1 $_s\mathcal{P}lots$ with fixed yields

We consider a model with N_S event species; the yield of a species k is noted N_k and its normalized PDF f_k. The $_s\mathcal{P}lots$ procedure allows to use the information from a fit performed on a discriminating variable X to extract the distributions of the control variable Y for the different species.

A main ingredient of the $_s\mathcal{P}lots$ calculation is the covariance matrix V of the fit, which can be taken from the output of a fit routine (*e.g.* TMinuit [74]). Alternatively, its inverse can be directly computed using:

$$V_{ij}^{-1} = \sum_{e=1}^{N} \frac{f_i(e)f_j(e)}{(\sum_{k=1}^{N_S} N_k f_k(e))^2}, \tag{A.1}$$

where the sum is running over N events, and $f_i(e)$ designates the value of the PDF f_i for

the event e. We can then use the covariance matrix to compute, for each event species n, the per-event sWeight $sP_n(e)$, using:

$$sP_n(e) = \frac{\sum\limits_{j=1}^{N_S} V_{nj} f_j(X_e)}{\sum\limits_{k=1}^{N_S} N_k f_k(X_e)}. \tag{A.2}$$

The distribution of the event species n on the control variable Y is then estimated by:

$$N_n.sM_n(Y).\delta Y = \sum_{e \subset [Y-\delta Y, Y+\delta Y]} sP_n(e). \tag{A.3}$$

We now introduce another event species in the model, with a fixed yield N_0 and a normalized PDF f_0. The covariance matrix changes, and its inverse becomes:

$$V_{ij}^{-1} = \sum_{e=1}^{N} \frac{f_i(e)f_j(e)}{(\sum\limits_{k=1}^{N_S} N_k f_k(e) + N_0 f_0(e))^2}, \tag{A.4}$$

whereas the per-event sWeight becomes:

$$sP_n(e) = \frac{\sum\limits_{j=1}^{N_S} V_{nj} f_j(X_e)}{\sum\limits_{k=1}^{N_S} N_k f_k(X_e) + N_0 f_0(X_e)} = \frac{\sum\limits_{j=1}^{N_S} V_{nj} f_j(X_e)}{\sum\limits_{k=1}^{N_S} N_k f_k(X_e)} \frac{\sum\limits_{k=1}^{N_S} N_k f_k(X_e)}{\sum\limits_{k=1}^{N_S} N_k f_k(X_e) + N_0 f_0(X_e)}. \tag{A.5}$$

This expression differs from Eq. A.2 only by an event-by-event factor that depends on the yields N_k and the PDFs f_k.

In the case where the distribution of the control variable for the species with fixed yield, $M_0(Y)$, is known, the distribution of the control variable Y for the species n is:

$$N_n.sM_n(Y).\delta Y = \sum_{e \subset [Y-\delta Y, Y+\delta Y]} sP_n(e) + c_n.M_0(Y), \tag{A.6}$$

where

$$c_n = N_n - \sum_{j=1}^{N_S} V_{nj} \tag{A.7}$$

is a coefficient depending uniquely on the considered event species. It quantifies the impact of the species with fixed yields on $sM_n(Y)$, and vanishes only if $N_0 = 0$.

One of the issues of extracting sWeights using a tool that treats the varied and fixed yields on an equal footing is related to the calculation of the covariance matrix. To illustrate this, we consider a model including three species: a signal S, a combinatorial background C, and another background B with a fixed yield. Using Eq. A.4, we calculate the inverse of the covariance matrix, and obtain:

$$V^{-1} = \begin{pmatrix} (V^{-1})_{SS} & (V^{-1})_{SC} & (V^{-1})_{SB} \\ (V^{-1})_{CS} & (V^{-1})_{CC} & (V^{-1})_{CB} \\ (V^{-1})_{BS} & (V^{-1})_{BC} & (V^{-1})_{BB} \end{pmatrix}, \tag{A.8}$$

which incorrectly includes terms related to the species with fixed yields. As the terms $(V^{-1})_{SB}$ and $(V^{-1})_{CB}$ do not vanish *a priori*, the inverse of this matrix has no clear link with the correct covariance matrix. Also, we notice that in the case where $f_B = f_C$ or $f_B = f_S$, the matrix is no longer invertible (it has two identical columns and lines), whereas sWeights should still be calculable. This shows that there is something fundamentally flawed with this approach.

A.2 `RooStats` implementation of the $_s\mathcal{P}lots$ method

The `RooStats::SPlot` method, used to calculate sWeights, computes the inverse of the covariance matrix using Eq. A.1 with a list of yields that the user provides as an argument. The covariance matrix itsef is then obtained from its inverse. As shown in Sec. A.1, in the case where there are some fixed yields in the arguments, this results in an incorrect covariance matrix.

However, building the `RooStats::SPlot` object using only the varied yields is also incorrect, as it would result in using Eq. A.1 and Eq. A.2 rather than Eq. A.4 and Eq. A.5. Correcting the sWeights event-by-event using Eq. A.5 is not possible either, as the covariance matrix is not correctly calculated.

A.3 Proposed method and test

It is clear from Sec. A.1 that it is necessary to differenciate the fixed yields from the others in the `RooStats::SPlot` object. To address this requirement, we propose an alternative constructor to the `RooStats::SPlot` object, shown in Fig. A.1 along with the original

constructor. We stress that this alternative constructor does not perform the correction shown in Eq. A.6. However, it calculates and stores the c_n coefficients in a new attribute of the class, as shown in Fig. A.2 and Fig. A.3. We also add a test to the original constructor, to check if all the arguments are indeed varying yields.

We test this alternative constructor and the c_n extraction tool using a toy model containing 3 event species: a signal S, a combinatorial background C, and a peaking background B with a fixed yield. The PDFs of these species on the discriminating variable X and on the control variable Y are taken as:

- Signal (S): Gaussian ($\mu = 0$, $\sigma = 0.1$) for X; Gaussian ($\mu = 0$, $\sigma = 0.05$) for Y.

- Combinatorial background (C): Constant for X; Gaussian ($\mu = $ -0.5, $\sigma = 0.05$) for Y.

- Fixed background. (B): Gaussian ($\mu = 0$, $\sigma = 0.05$) for X; Gaussian ($\mu = 0.5$, $\sigma = 0.05$) for Y.

Both variables X and Y are defined in the interval [-1,1]. This model is chosen in order to ensure a sizable nuisance of the species with fixed yield on the discriminating variable, whereas its impact on the control variable is easy to spot. We generate a sample of 16,000 events, including 5,000 signal events, 1,000 fixed background events, and 10,000 combinatorial background events. We then assume a wrong hypothesis on the yield of the fixed background ($N_B = 1200$), in order to simulate the general case where the value of the fixed yield is not precisely known. We show the result of the one-dimensional fit on the variable X using this model, along with the projection of this fit on the variable Y, in Fig. A.4.

We consider two approaches, A and B: the former is the calculation of the sWeights with the original `RooStats::SPlot` method, providing only the list of varying yields to the `RooStats::SPlot` constructor; the latter is the proposed approach, where we use the alternative `RooStats::SPlot` constructor providing the list of all yields and the list of fixed yields.

For each of these two methods, Fig. A.5 shows the samples with signal and background weights applied, both before and after the correction of Eq. A.6. The results are satisfactory for approach B after the c_n correction, for both the signal and the combinatorial background distributions. On the contrary, approach A provides an acceptable description of the signal shape, but shows large discrepancies for the combinatorial background.

For each event species, Table. A.1 also shows the sum of sWeights and of its associated c_n, compared to the fitted yield. According to Ref. [73], the sum has to be compatiable with the yield, which is clearly not the case for approach A.

Species	$\sum_e sP_A(e)+c_A$	$\sum_e sP_B(e)+c_B$	Fitted yield
Signal	-7336.15	4975.27	4975.35
Comb.	16361.5	10007.5	10007.5
Bkg	-	-	1200.

Table A.1 – Summary of yields and sum of sWeights for the two approaches. As shown in Ref. [73], in both approaches the sum of all yields extracted from the fit is not equal to the total number of events.

Changes in the implementation of the usual constructor:

```
SPlot::SPlot(const char* name, const char* title,
             RooDataSet& data, RooAbsPdf* pdf,
             const RooArgList &yieldsList, const RooArgSet &projDeps,
             bool includeWeights, bool cloneData, const char* newName):
  TNamed(name, title)
{
  /* Original body of the SPlot constructor.*/
  // Add check that yieldsList contains only varying yields
  iter = yieldsList.createIterator() ;
  RooRealVar *var;
  while((var=(RooRealVar*)iter->Next())) {
    if (var->isConstant()) {
      //Throw exception and error message.
    }}
  delete iter ;
  //Call method to build sWeights
  this->AddSWeight(pdf,yieldsList,projDeps,includeWeights);
}
```

Implementation of the alternative constructor:

```
SPlot::SPlot(const char* name, const char* title,
             RooDataSet& data, RooAbsPdf* pdf,
             const RooArgList &allYieldsList,const RooArgList &fixedYields,
             const RooArgSet &projDeps,
             bool includeWeights, bool cloneData, const char* newName):
  TNamed(name, title)
{
  /* Original body of the SPlot constructor.*/
  // Add check that yieldsList contains only varying yields
  //Check that fixed yields are in the allYields arguments
  iter = fixedYields.createIterator();
  while((arg=(RooAbsArg*)iter->Next()))
    if (!(allYieldsList.contains(*arg)))
      {
        //Throw exception and error message.
      }
  //Call new method to build sWeights, with fixed yields
  this->AddSWeight(pdf,allYieldsList,fixedYields,projDeps,includeWeights);
}
```

Figure A.1 – Snippets of code showing the original (top) and the alternative (bottom) constructors for the `RooStats::SPlot` class. Providing an empty RooArgSet as the fixedYieldsList argument of the alternative constructor yields the same results as calling the original constructor with the same arguments.

```cpp
void SPlot::AddSWeight( RooAbsPdf* pdf, const RooArgList &allYieldsList,
                        const RooArgList &fixedYields,
                        const RooArgSet &projDeps, bool includeWeights)
{
  /*Usual body of the method: stores the constant parameters (other than
    yields) values */
  //Stores which indexes of the allYieldsList are the variable ones.
  TIterator *it = allYieldsList.createIterator();
  RooAbsArg* arg;
  unsigned int iArg(0);
  std::vector<unsigned int> varIndexes;
  while ((arg = (RooAbsArg*) it->Next()) != NULL){
      if (!(fixedYields.find(arg->GetName())))
        varIndexes.push_back(iArg);
      iArg++;}
  //We now have two indexes over which we iterate
  Int_t nAllSpec = allYieldsList.getSize();
  Int_t nVarSpec = allYieldsList.getSize() - fixedYields.getSize();

  /*Usual body of the method:
    -stores the initial yield parameters
    -calculate the value of the component pdf for each event and
    species.*/
  //Inverse of the covariance matrix
  TMatrixD covInv(nVarSpec, nVarSpec);
  /*Initialisation*/
  for (Int_t ievt = 0; ievt < numevents; ++ievt){
      fSData->get(ievt) ;
      // Sum for the denominator
      Double_t dsum(0);
      for(Int_t k = 0; k < nAllSpec; ++k)
        dsum += pdfvalues[ievt][k] * yieldvalues[k] ;

      for(Int_t n=0; n<nVarSpec; ++n)
        for(Int_t j=0; j<nVarSpec; ++j)
            if(includeWeights == kTRUE)
              covInv(n,j) += fSData->weight()
                *pdfvalues[ievt][varIndexes[n]]
                *pdfvalues[ievt][varIndexes[j]]/(dsum*dsum) ;
            else
              covInv(n,j) +=
                pdfvalues[ievt][varIndexes[n]]
                *pdfvalues[ievt][varIndexes[j]]/(dsum*dsum) ;}
  // Invert to get the covariance matrix
  TMatrixD covMatrix(TMatrixD::kInverted,covInv);
```

Figure A.2 – First part of the implementation of the new constructor.

```cpp
// Create and label the variables
// used to store the SWeights
for(Int_t k=0; k<nVarSpec; ++k){
  /*Usual variables*/

  //c_n coefficients
  wname = std::string(yieldvars[varIndexes[k]]->GetName())+"_c";
  var = new RooRealVar(wname.c_str(),wname.c_str(),0);
  double cVal = yieldvalues[varIndexes[k]];
  for (Int_t n = 0 ; n<nVarSpec; ++n)
    cVal -= covMatrix[k][n];
  var->setVal(cVal);
  fSWeightCoefs.add(*var); //new attribute of the class.
}
// Create and fill a RooDataSet with the SWeights
RooDataSet* sWeightData = new RooDataSet("dataset", "", sweightset);
for(Int_t ievt = 0; ievt < numevents; ++ievt){
    fSData->get(ievt) ;
    // sum for denominator
    Double_t dsum(0);
    for(Int_t k = 0; k < nAllSpec; ++k)
      dsum += pdfvalues[ievt][k] * yieldvalues[k] ;
    // covariance weighted pdf for each specie
    for(Int_t n=0; n<nVarSpec; ++n){
      Double_t nsum(0) ;
      for(Int_t j=0; j<nVarSpec; ++j)
        nsum += covMatrix(n,j) * pdfvalues[ievt][varIndexes[j]] ;
      if(includeWeights == kTRUE)
        sweightvec[n]->setVal(fSData->weight() * nsum/dsum) ;
      else
        sweightvec[n]->setVal( nsum/dsum) ;
      /* Fill the dataset with sweightvec*/
    }}
// Add the SWeights to the original data set
fSData->merge(sWeightData);
/*Reinitialise all parameters and yields*/
return;
}
```

Figure A.3 – Second part of the implementation of the new constructor.
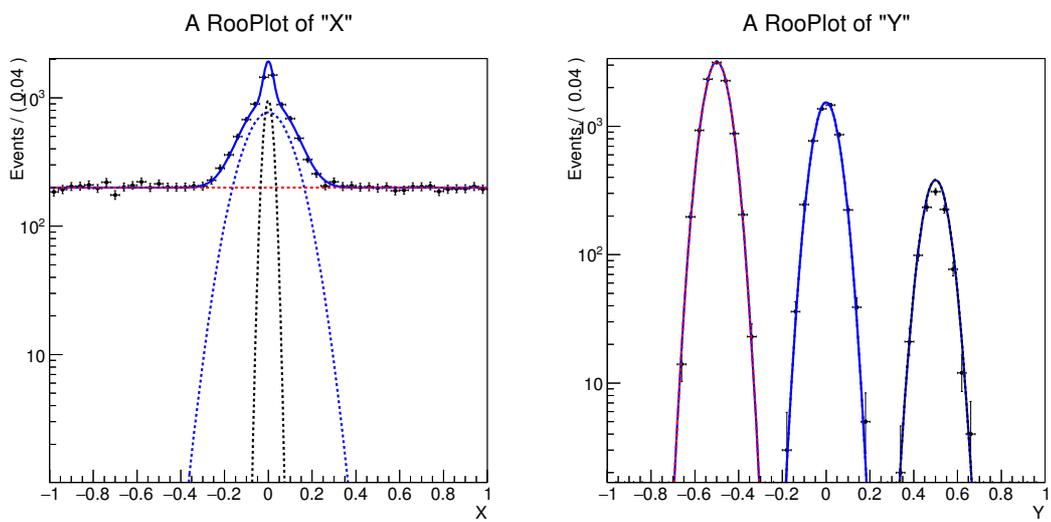
Figure A.4 – Result of the fit on X, projected on the X dimension (left) and the Y dimension (right). The signal is displayed in blue, the combinatorial background in red, and the peaking background in black. As expected, the projection of the result on the Y dimension shows that we overestimated (on purpose) the number of background events.
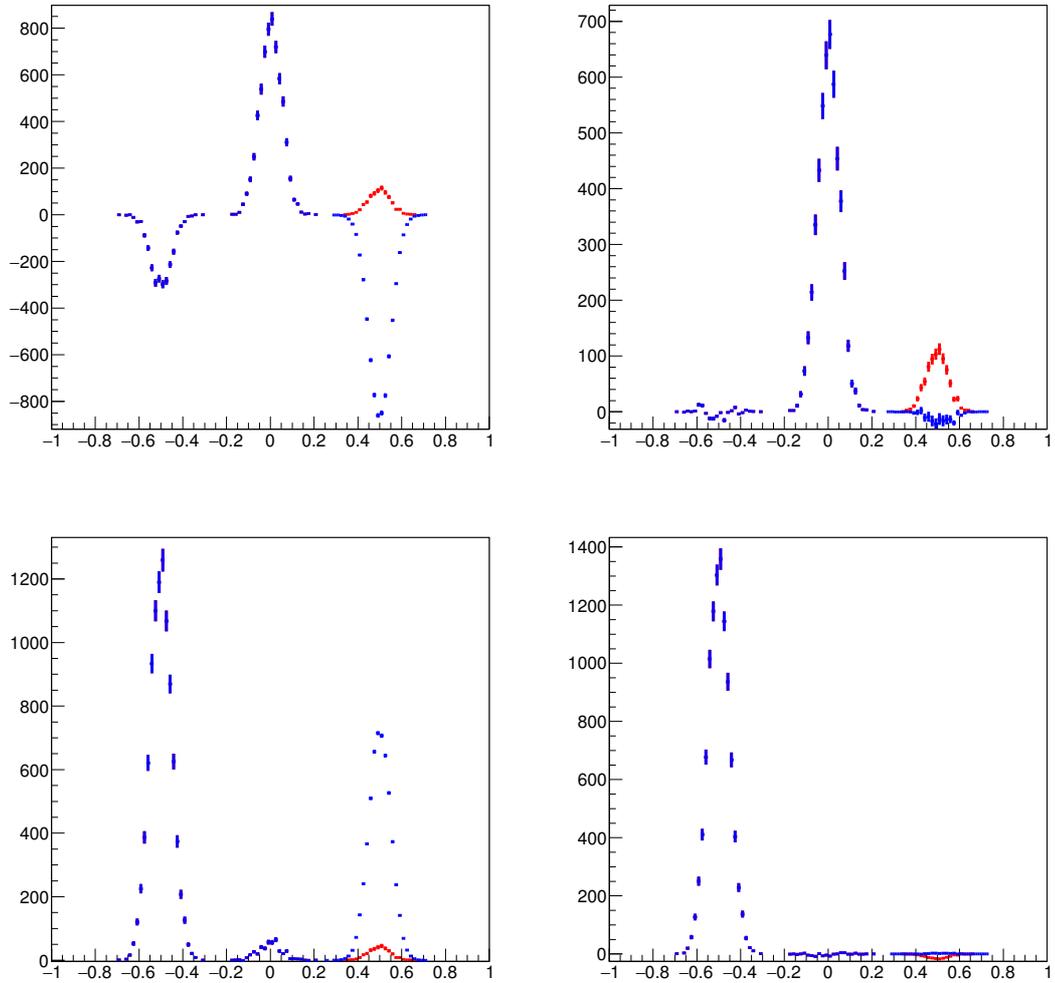
Figure A.5 – Distribution of Y in data samples sWeighted according to signal (combinatorial) weights. Results of the approach A(B) are shown on the left (right). The red points correspond to the distribution of the sum of sWeights, whereas the blue points represent the final distribution (after correction using Eq. A.6).

A.4 Conclusion

Models that include an event species with fixed yields require a specific treatment when using the $_s\mathcal{P}lots$ method, that the `RooStats::SPlot` implementation does not provide. In this note, we propose a straightforward modification of this class that allows to extract the correct $_s\mathcal{P}lots$ in the case of fixed yields. We also implement some precautions in the code to help analysts avoiding the use of the wrong method. These modifications do not remove the need to correct the distributions using Eq. A.6, but allow to calculate the c_n coefficients inside the `RooStats::SPlot` object in a way that is coherent with the sWeights extraction.

We tested this additional code, both in terms of compatibility with the former implementation (not shown here), and in terms of expected results. The results are satisfactory, and show a clear improvement compared to the original approach, especially in terms of normalization properties.

Appendix B

Goodness-of-fit criteria

Estimating the relative quality of two fits is usually done by a χ^2 calculation or a likelihood estimation. However, while these methods yield a good indicator of the quality of a fit, they tend not to distinguish a localized discrepancy between two fits, especially when the fit is performed on more than one dimension. The addition or removal of a resonance in a Dalitz-plot model constitutes such a localized difference, and thus we use additional tools to assess the relative quality of the fit of two models on data. These additional tools provide us with so-called goodness-of-fit criteria, defined on an event-by-event basis. In the following, we describe two of these methods, presented in [75].

These methods define for two given samples a statistic T that takes a value of t when comparing data and a Monte-Carlo sample generated using a model fitted on data. Given the expected distribution g of this statistic in the case where the model is the parent PDF of data, we define the p-value of a fit

$$p = \int\limits_{t}^{+\infty} g(T)dT. \tag{B.1}$$

This p-value is what is used to compare two different fits to data.

Mixed-sample estimation

The idea of mixed-sample estimation is that the mixing between two samples is maximal if and only if f and g have been generated by the same underlying PDF. Figure B.1 shows the example of two samples generated using the same PDF and different PDFs to illustrate that idea.

A mixing indicator I on two points x and y can be defined as

$$I(x,y) = \begin{cases} 1 \text{ if x and y belong to the same sample,} \\ 0 \text{ otherwise.} \end{cases} \tag{B.2}$$
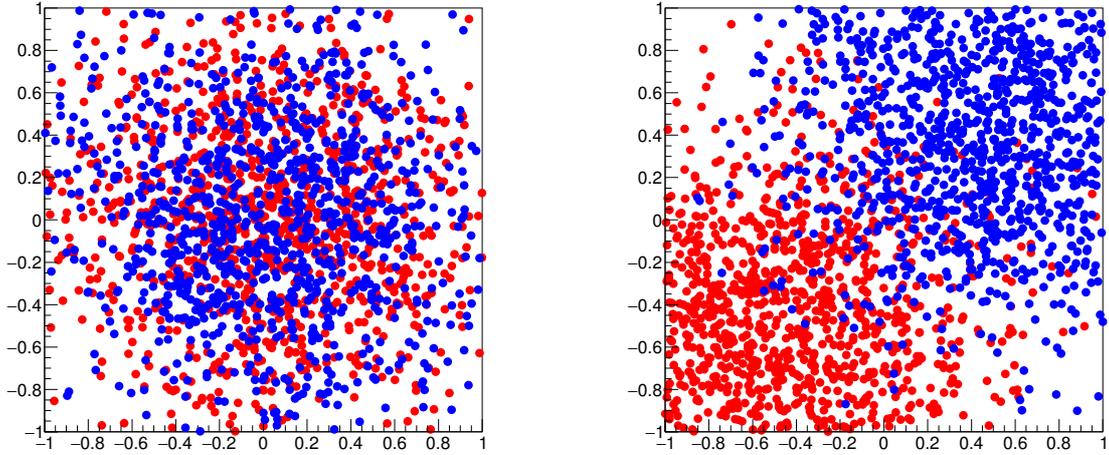
Figure B.1 – Example of two samples (red and blue points) generated with the same underlying PDF (left), and different PDFs (right).

The mixing statistic T(A,B) for two samples A and B, containing n_a and n_b points, is defined as

$$T = \frac{1}{n_k(n_a + n_b)} \sum_{i=1}^{n_a+n_b} \sum_{k=1}^{n_k} I(i,k),$$

(B.3)

where n_k is an arbitrary number of neighbours of a point that are considered. A large value of this parameter reduces statistical fluctuations, but it also reduces the resolution of the method as further points are considered.

This statistic is maximal for minimally mixed samples, and can be used to assess the quality of the fit of a model to data. Indeed, if the fitted model is used to generate a Monte-Carlo sample, this statistic calculated using data and the simulated sample yields an indication on the quality of the fit.

In the case where the two underlying PDFs are the same, this statistic converges to a Gaussian distribution of mean

$$\mu_T = \frac{n_a(n_a - 1) + n_b(n_b - 1)}{(n_a + n_b)(n_a + n_b - 1)}.$$

(B.4)

The width of this distribution depends on the PDF and on the choice of n_k, but has a limiting value

$$\lim_{n,n_k \to \infty} \sigma_T^2 = \frac{1}{nn_k}\left(\frac{n_a n_b}{n^2} + 4\frac{n_a^4 n_b^4}{n^4}\right), \tag{B.5}$$

where $n = n_a + n_b$. We assume that this limiting value is reached in our conditions.[1] The distribution g used to calculate the p-value in Eq. B.1 is then the limit Gaussian distribution.

Point-to-point dissimilarity methods

In the case where the real parent distribution $f(x)$ of data is known,

$$d = \frac{1}{2}\int (f(x) - f_0(x))^2 dx \tag{B.6}$$

is a distance between this parent distribution and the fitted model $f_0(x)$. In practical cases, the real parent distribution $f(x)$ of data is unknown, but Eq B.6 can be generalized by introducing a correlation function $\psi(x, x')$, and the function

$$T = \frac{1}{2}\iint (f(x) - f_0(x))(f(x') - f_0(x'))\psi(x, x')dxdx'. \tag{B.7}$$

is defined. This generalization allows to calculate a similar quantity for samples rather than distributions. Indeed, developing the expression and replacing integrals by sums we define for two samples A and B the statistic

$$T = \frac{1}{n_A(n_A - 1)}\sum_{x \in A}\sum_{x' \in A} \psi(x, x')dxdx' + \frac{1}{n_B(n_B - 1)}\sum_{x \in B}\sum_{x' \in B} \psi(x, x')dxdx'$$
$$- \frac{1}{n_A n_B}\sum_{x \in A}\sum_{x' \in B} \psi(x, x')dxdx'. \tag{B.8}$$

The first and second terms of this expression can be considered as statistical uncertainties on data and Monte-Carlo samples, the third term being a correlation term between the two. Monte-Carlo samples are often generated with large amount of points, so the second term can be neglected as it is computationally heavy.

Two forms of the correlating function $\psi(x, x')$ have been studied in [75]:

- $\psi_{\text{Gaus}} = \exp\left(\frac{(x-x')^2}{2\sigma(x)\sigma(x')}\right)$

---

[1]Indeed, the convergence to this limit is demonstrated as being very fast in [75].

- $\psi_{\mathrm{Log}} = \log(|x - x'| + \epsilon)$

The parameter ϵ is an arbitrary number that keeps the ψ_{Log} from exhibiting a pole when x and x' are too close. The widths $\sigma(x)$ and $\sigma(x')$ in the Gaussian correlating function are shown in [75] to be optimal in terms of discriminating power when they are proportional to the inverse of the PDF:

$$\sigma(x) = \frac{\bar{\sigma}}{f(x)}. \tag{B.9}$$

The preferred range for the parameter $\bar{\sigma}$ is $[\bar{\Gamma}, 2\bar{\Gamma}]$ where

$$\bar{\Gamma} = \frac{\sum \mathrm{FF_i}\Gamma_i}{\sum \mathrm{FF_i}}, \tag{B.10}$$

the index i running over all resonances in the model, and Γ_i and $\mathrm{FF_i}$ designate their widths and fit fractions, respectively.

Appendix C

Fast MC method for background studies - other channels

We show in this section the results of smearing various partially reconstructed backgrounds using resolution functions extracted from a $B^0 \to (K^{*0} \to K^0_S \pi^0)(\rho^0 \to \pi^+\pi^-)$ sample.

C.1 $\quad B^0 \to K^0_S (\eta \to \pi^+\pi^-\pi^0)$

This channel belongs to the $K^0_S \pi^+\pi^-$ spectrum, but its kinematics are different from the ones of $B^0 \to K^{*0}\rho^0$.

Figure C.1 shows the distribution of the $m_{K^0_S h^+ h'}$ near the threshold for reconstructed and fast MC events, as well as the resolution distributions of $m_{K^0_S h h'}$. The distributions agree quite well on $m_{K^0_S h h'}$, with a well-reproduced behaviour at the threshold. As before, the distribution of $\Delta_{m_{K^0_S h h'}}$ is slightly narrower in the fast MC case. Fig. C.2 shows the distributions of events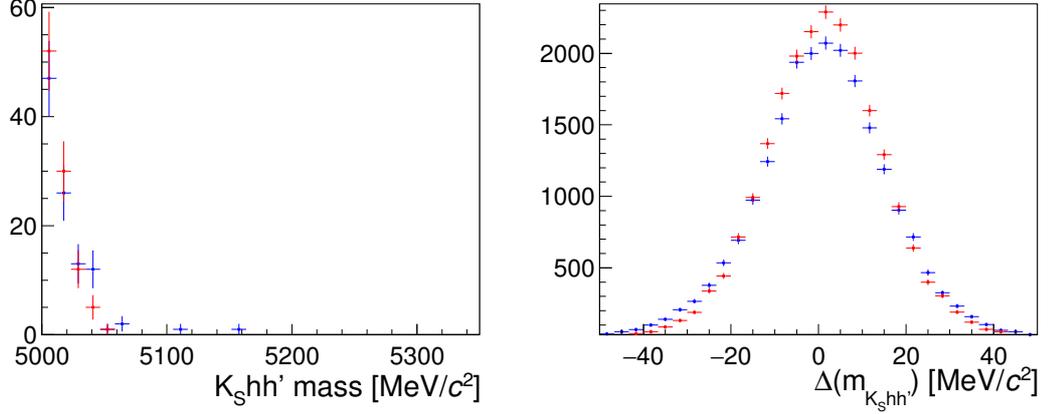 over the Dalitz plane for reconstructed and fast MC events, along with the distribution of the pulls between these two distributions. The distributions are similar, and the pulls are rather small and show no structure. Overall, the result we obtain on this channel are satisfactory.

C.2 $\quad B^0 \to K^0_S \pi^+\pi^-\gamma$

This channel also belongs to the $K^0_S \pi^+\pi^-$ spectrum. As the missed particle in this channel is massless, its $m_{K^0_S h h'}$ distribution goes under the signal peak, which makes it especially dangerous for our analysis.

Figure C.3 shows the distribution of the $m_{K^0_S h h'}$ near the threshold for reconstructed and fast MC events, as well as the resolution distributions of $m_{K^0_S h h'}$. The distributions agree quite well on $m_{K^0_S h h'}$, with a well-reproduced behaviour at the threshold. As before, the distribution of $\Delta_{m_{K^0_S h h'}}$ is slightly narrower in the fast MC case. Fig. C.4 shows the distributions of events over the Dalitz plane for reconstructed and fast MC events,

Figure C.1 – Left: Distributions of $m_{K_S^0 h^+ h'}$ for reconstructed (blue) and fast MC (red) $B^0 \to K_S^0 \eta$ events. Right: Resolutions of $m_{K_S^0 h^+ h'}$ for reconstructed (blue) and fast MC (red) $B^0 \to K_S^0 \eta$ events.

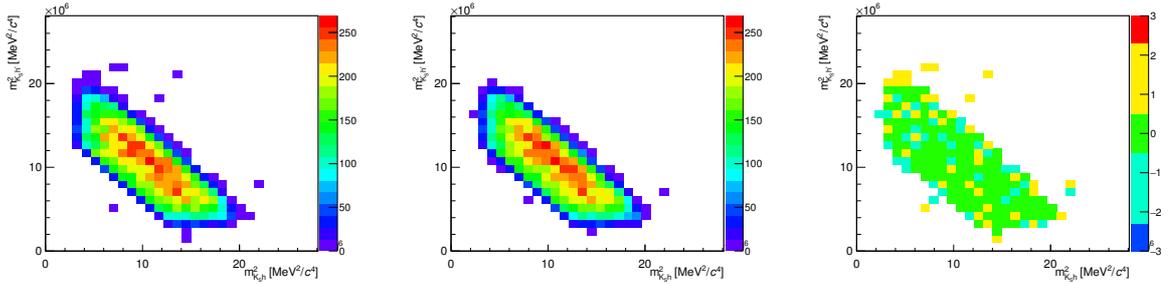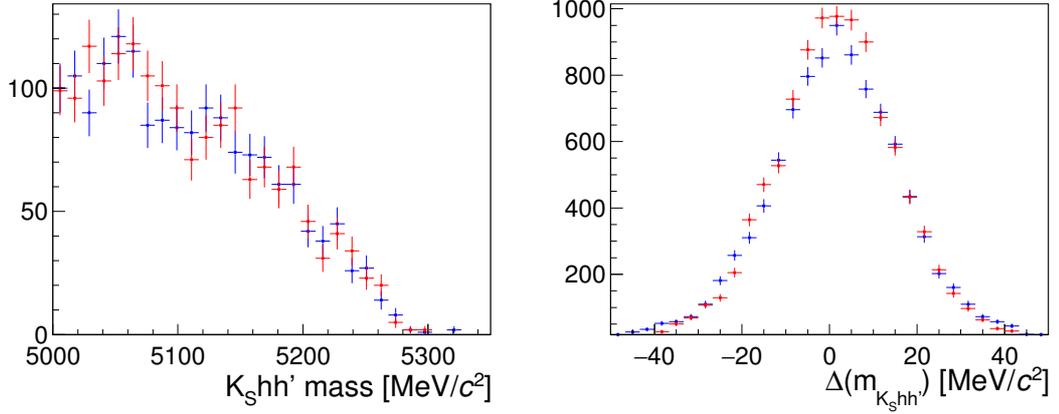
Figure C.2 – Distributions of $m_{K_S^0 h}^2$ versus $m_{K_S^0 h'}^2$ for the $B^0 \to K_S^0 \eta$ mode, with DD reconstruction of the K_S^0. Left: Reconstructed events. Middle: Fast MC events. Right: Pulls between the two distributions.

along with the distribution of the pulls between these two distributions. The distributions are similar, and the pulls are rather small and show no structure. Overall, the result we obtain on this channel are satisfactory.

C.3 $B^0 \to (K^{*0} \to K_S^0 \pi^+) \pi^+ \pi^-$

In this channel, we are missing a charged particle (π^+). We expect the $m_{K_S^0 h^+ h'}$ distribution to be similar to that in $B^0 \to K^{*0} \rho^0$, but the Dalitz plane distributions of the two modes are expected to differ.

Figure C.5 shows the distribution of the $m_{K_S^0 h^+ h'}$ near the threshold for reconstructed

Figure C.3 – Left: Distributions of $m_{K_S^0 h^+ h'}$ for reconstructed (blue) and fast MC (red) $B^0 \to K_S^0 \pi^+ \pi^- \gamma$ events. Right: Resolutions of $m_{K_S^0 h^+ h'}$ for reconstructed (blue) and fast MC (red) $B^0 \to K_S^0 \pi^+ \pi^- \gamma$ events.
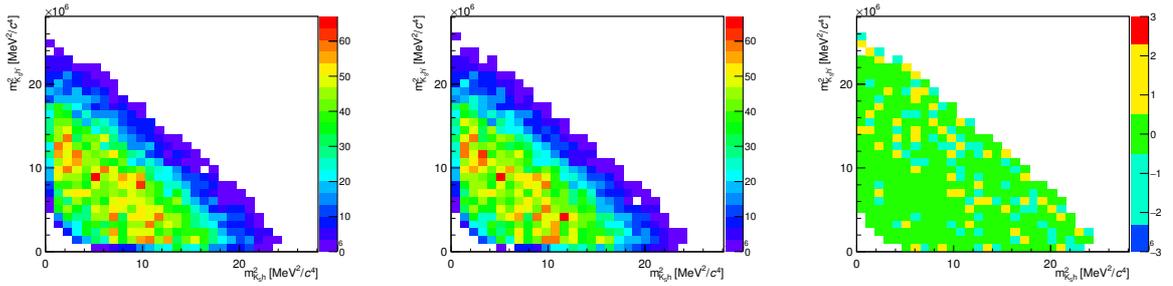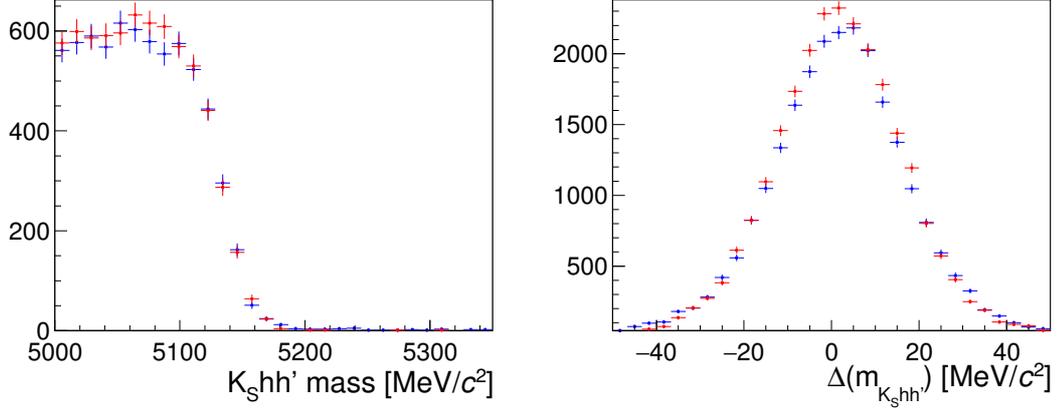

Figure C.4 – Distributions of $m_{K_S^0 h}^2$ versus $m_{K_S^0 h'}^2$ for the $B^0 \to K_S^0 \pi^+ \pi^- \gamma$ mode, with DD reconstruction of the K_S^0. Left: Reconstructed events. Middle: Fast MC events. Right: Pulls between the two distributions.

and fast MC events, as well as the resolution distributions of $m_{K_S^0 h^+ h'}$. The distributions agree quite well on $m_{K_S^0 h^+ h'}$, with a well-reproduced behaviour at the threshold. As before, the distribution of $\Delta_{m_{K_S^0 h^+ h'}}$ is slightly narrower in the fast MC case. Fig. C.6 shows the distributions of events over the Dalitz plane for reconstructed and fast MC events, along with the distribution of the pulls between these two distributions. The distributions are similar, and the pulls are rather small and show no structure. Overall, the result we obtain on this channel are satisfactory.

v

Figure C.5 – Left: Distributions of $m_{K^0_S h^+ h'}$ for reconstructed (blue) and fast MC (red) $B^+ \to K^{*+}\pi^+\pi^-$ events. Right: Resolutions of $m_{K^0_S h^+ h'}$ for reconstructed (blue) and fast MC (red) $B^+ \to K^{*+}\pi^+\pi^-$ events.
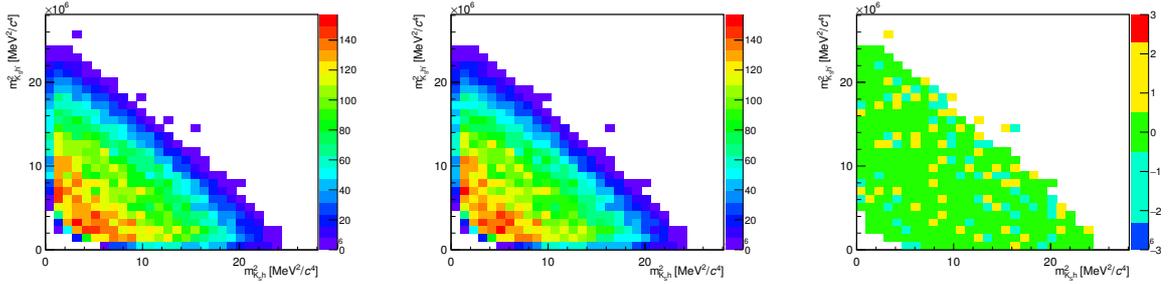

Figure C.6 – Distributions of $m^2_{K^0_S h}$ versus $m^2_{K^0_S h'}$ for the $B^+ \to K^{*+}\pi^+\pi^-$ mode, with DD reconstruction of the K^0_S. Left: Reconstructed events. Middle: Fast MC events. Right: Pulls between the two distributions.

C.4 $B^0 \to (K^{*0} \to K^0_S \pi^0)(\phi \to K^+ K^-)$

This mode does not belong the $K^0_S \pi^+\pi^-$ spectrum. However, it is still interesting to test the limits of the method with this mode, similar to $B^0 \to K^{*0}\rho^0$ but with a ϕ instead of the ρ^0.

Figure C.7 shows the distribution of the $m_{K^0_S h^+ h'}$ near the threshold for reconstructed and fast MC events, as well as the resolution distributions of $m_{K^0_S h^+ h'}$. The distributions agree quite well on $m_{K^0_S h^+ h'}$, with a well-reproduced behaviour at the threshold. As before, the distribution of $\Delta_{m_{K^0_S h^+ h'}}$ is slightly narrower in the fast MC case.   Fig. C.8 shows the distributions of events over the Dalitz plane for reconstructed and fast MC events, along with the distribution of the pulls between these two distributions. The distributions

are similar, and the pulls are rather small and show no structure. Overall, the result we obtain on this channel are satisfactory.
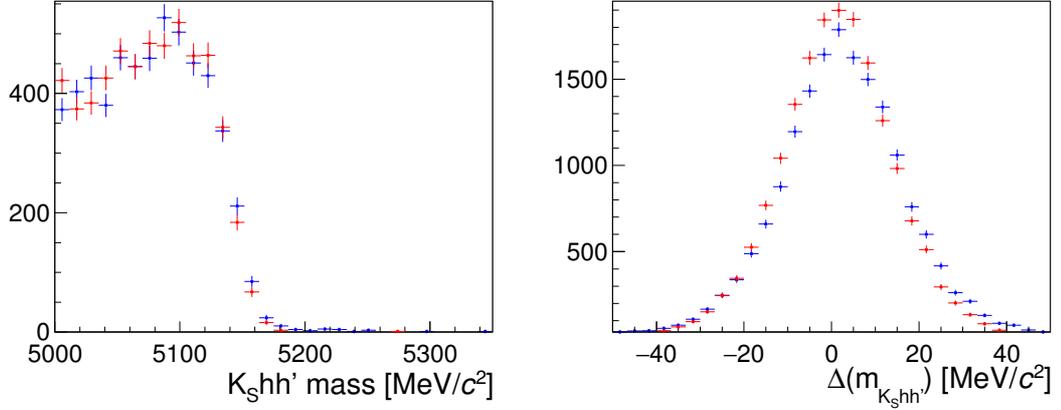

Figure C.7 – Left: Distributions of $m_{K_S^0 h^+ h'}$ for reconstructed (blue) and fast MC (red) $B^0 \to K^{*0}\phi$ events. Right: Resolutions of $m_{K_S^0 h^+ h'}$ for reconstructed (blue) and fast MC (red) $B^0 \to K^{*0}\phi$ events.

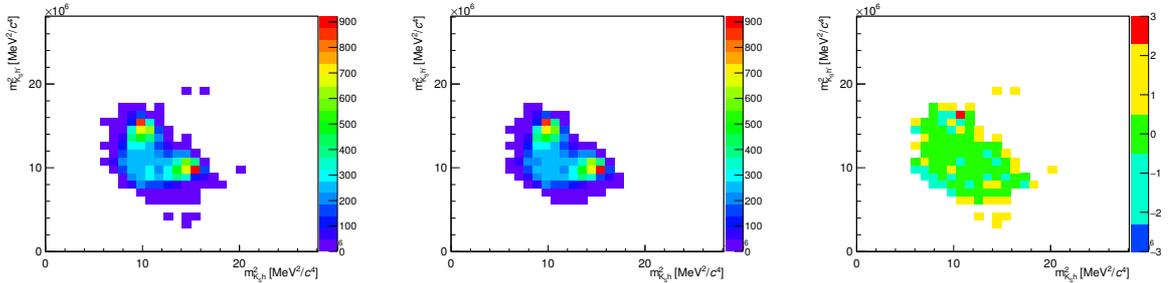Figure C.8 – Distributions of $m^2_{K_S^0 h}$ versus $m^2_{K_S^0 h'}$ for the $B^0 \to K^{*0}\phi$ mode, with DD reconstruction of the K_S^0. Left: Reconstructed events. Middle: Fast MC events. Right: Pulls between the two distributions.

C.5 $B^+ \to (K^{*0} \to K_S^0 \pi^+)(\phi \to K^+ K^-)$

This mode does not belong the $K_S^0 \pi^+ \pi^-$ spectrum. However, it is still interesting to test the limits of the method with this mode, similar to $B^0 \to K^{*0}\rho^0$ but with a ϕ instead of the ρ^0, and missing a π^+.

Figure C.9 shows the distribution of the $m_{K_S^0 h^+ h'}$ near the threshold for reconstructed and fast MC events, as well as the resolution distributions of $m_{K_S^0 h^+ h'}$. The distributions agree quite well on $m_{K_S^0 h^+ h'}$, with a well-reproduced behaviour at the threshold. As before, the distribution of $\Delta_{m_{K_S^0 h^+ h'}}$ is slightly narrower in the fast MC case. Fig. C.10 shows the distributions of events over the Dalitz plane for reconstructed and fast MC events, along with the distribution of the pulls between these two distributions. The distributions are similar, and the pulls are rather small and show no structure. Overall, the result we obtain on this channel are satisfactory.
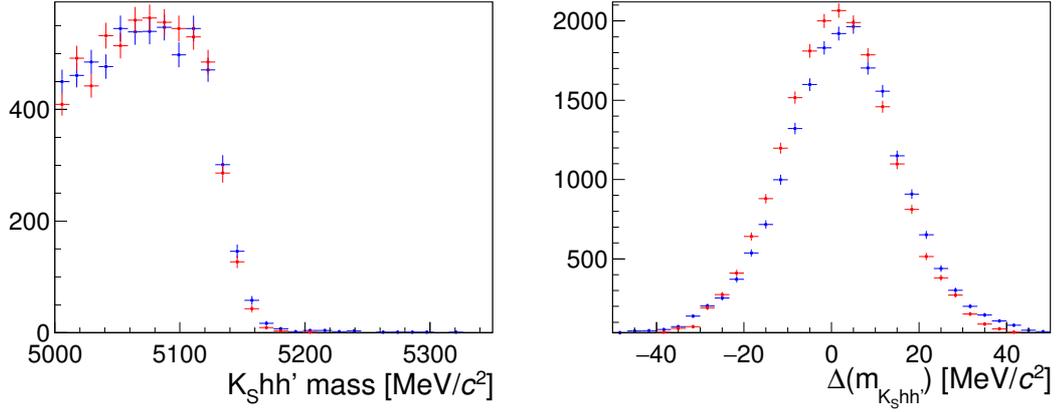

Figure C.9 – Left: Distributions of $m_{K_S^0 h^+ h'}$ for reconstructed (blue) and fast MC (red) $B^+ \to K^{*+}\phi$ events. Right: Resolutions of $m_{K_S^0 h^+ h'}$ for reconstructed (blue) and fast MC (red) $B^+ \to K^{*+}\phi$ events.
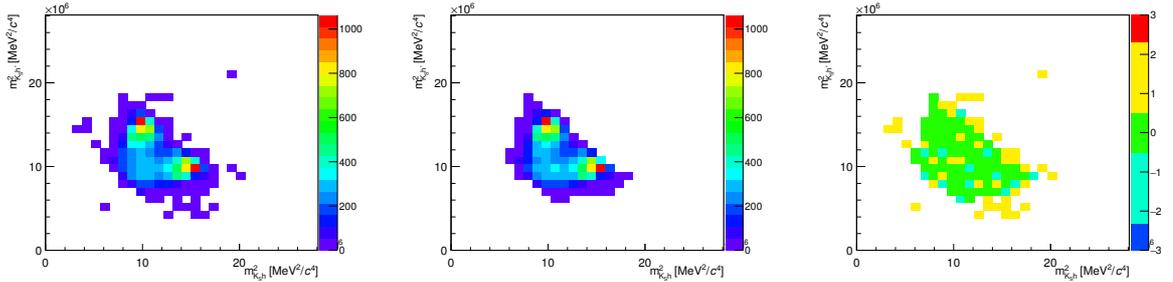

Figure C.10 – Distributions of $m_{K_S^0 h}^2$ versus $m_{K_S^0 h'}^2$ for the $B^+ \to K^{*+}\phi$ mode, with DD reconstruction of the K_S^0. Left: Reconstructed events. Middle: Fast MC events. Right: Pulls between the two distributions.

C.6 $\quad B_s^0 \to (K^{*0} \to K_S^0 \pi^0)(\phi \to K^+ K^-)$

This mode does not belong the $K_S^0 \pi^+ \pi^-$ spectrum. However, it is still interesting to test the limits of the method with this mode, similar to $B^0 \to K^{*0} \rho^0$ but with a ϕ instead of the ρ^0.

Figure C.11 shows the distribution of the $m_{K_S^0 h^+ h'}$ near the threshold for reconstructed and fast MC events, as well as the resolution distributions of $m_{K_S^0 h^+ h'}$. The distributions agree quite well on $m_{K_S^0 h^+ h'}$, with a well-reproduced behaviour at the threshold. As before, the distribution of $\Delta_{m_{K_S^0 h^+ h'}}$ is slightly narrower in the fast MC case. Fig. C.12 shows the distributions of events over the Dalitz plane for reconstructed and fast MC events, along with the distribution of the pulls between these two distributions. The distributions are similar, and the pulls are rather small and show no structure. Overall, the result we obtain on this channel are satisfactory.
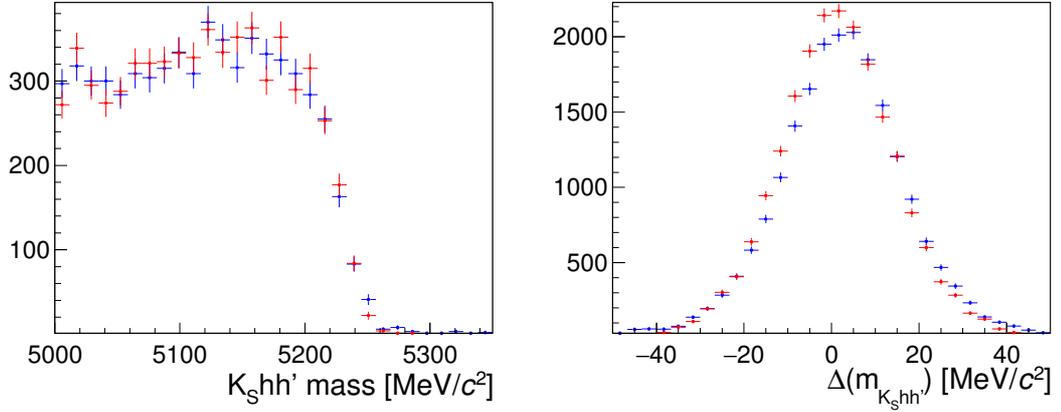

Figure C.11 – Left: Distributions of $m_{K_S^0 h^+ h'}$ for reconstructed (blue) and fast MC (red) $B_s^0 \to K^{*0}\phi$ events. Right: Resolutions of $m_{K_S^0 h^+ h'}$ for reconstructed (blue) and fast MC (red) $B_s^0 \to K^{*0}\phi$ events.
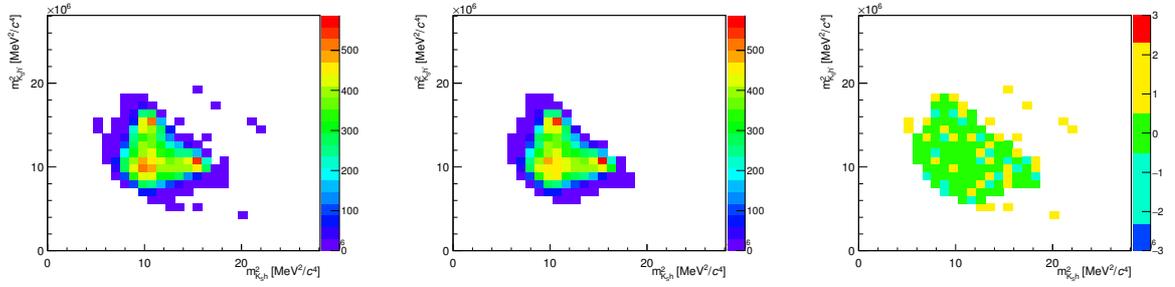
Figure C.12 – Distributions of $m^2_{K^0_S h}$ versus $m^2_{K^0_S h'}$ for the $B^0_s \to K^{*0}\phi$ mode, with DD reconstruction of the K^0_S. Left: Reconstructed events. Middle: Fast MC events. Right: Pulls between the two distributions.

Bibliography

[1] S. L. Glashow, *Partial symmetries of weak interactions*, Nucl. Phys. **22** (1961) 579.

[2] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[3] A. Salam, *Weak and electromagnetic interactions*, Conf. Proc. **C680519** (1968) 367.

[4] Particle Data Group, K. A. Olive *et al.*, *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001.

[5] G. F. Giudice, *Naturalness after LHC8*, PoS **EPS-HEP2013** (2013) 163, arXiv:1307.7879.

[6] A. Riotto, *Theories of baryogenesis*, in *High energy physics and cosmology. Proceedings, Summer School, Trieste, Italy, June 29-July 17, 1998*, pp. 326–436, 1998. arXiv:hep-ph/9807454.

[7] A. D. Sakharov, *Violation of CP invariance, C asymmetry, and baryon asymmetry of the Universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.

[8] F. R. Klinkhamer and N. S. Manton, *A Saddle Point Solution in the Weinberg-Salam Theory*, Phys. Rev. **D30** (1984) 2212.

[9] 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, arXiv:1207.7214.

[10] 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, arXiv:1207.7235.

[11] T. D. Lee and C.-N. Yang, *Question of parity conservation in weak interactions*, Phys. Rev. **104** (1956) 254.

[12] C. S. Wu *et al.*, *Experimental test of parity conservation in beta decay*, Phys. Rev. **105** (1957) 1413.

[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.

[14] L. D. Landau, *On the conservation laws for weak interactions*, Nucl. Phys. **3** (1957) 127.

[15] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π decay of the K_2^0 Meson*, Phys. Rev. Lett. **13** (1964) 138.

[16] BaBar collaboration, B. Aubert *et al.*, *Observation of CP violation in the B^0 meson system*, Phys. Rev. Lett. **87** (2001) 091801, `arXiv:hep-ex/0107013`.

[17] CPLEAR collaboration, A. Angelopoulos *et al.*, *First direct observation of time reversal noninvariance in the neutral kaon system*, Phys. Lett. **B444** (1998) 43.

[18] P. G. Harris *et al.*, *New experimental limit on the electric dipole moment of the neutron*, Phys. Rev. Lett. **82** (1999) 904.

[19] Heavy Flavor Averaging Group (HFAG), Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2014*, `arXiv:1412.7515`.

[20] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

[21] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak interactions with lepton-hadron symmetry*, Phys. Rev. **D2** (1970) 1285.

[22] M. Kobayashi and T. Maskawa, *CP violation in the renormalizable theory of the weak interaction*, Prog. Theor. Phys. **49** (1973) 652.

[23] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[24] A. J. Buras, M. E. Lautenbacher, and G. Ostermaier, *Waiting for the top quark mass, $K^+ \rightarrow \pi^+\nu\overline{\nu}$, $B_{(s)}^0$-$\overline{B}_{(s)}^0$ mixing and CP asymmetries in B decays*, Phys. Rev. **D50** (1994) 3433, `arXiv:hep-ph/9403384`.

[25] J. Charles *et al.*, *Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$ New Physics*, Phys. Rev. **D91** (2015), no. 7 073007, `arXiv:1501.05013`.

[26] KTeV, E. Abouzaid *et al.*, *Precise Measurements of Direct CP Violation, CPT Symmetry, and Other Parameters in the Neutral Kaon System*, Phys. Rev. **D83** (2011) 092001, `arXiv:1011.0127`.

[27] UA1, C. Albajar *et al.*, *Search for B^0-$\overline{B}^0$ oscillations at the CERN proton-antiproton Collider. 2.*, Phys. Lett. **B186** (1987) 247, [Erratum: Phys. Lett.B197,565(1987)].

[28] ARGUS, H. Albrecht *et al.*, *Observation of B^0-$\overline{B}^0$ Mixing*, Phys. Lett. **B192** (1987) 245, [,51(1987)].

[29] CDF, A. Abulencia *et al.*, *Observation of B_s^0-$\bar{B}_s^0$ Oscillations*, Phys. Rev. Lett. **97** (2006) 242003, `arXiv:hep-ex/0609040`.

[30] LHCb, R. Aaij *et al.*, *Precision measurement of the B_s^0-$\bar{B}_s^0$ oscillation frequency with the decay $B_s^0 \to D_s^- \pi^+$*, New J. Phys. **15** (2013) 053021, `arXiv:1304.4741`.

[31] BaBar, 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`.

[32] I. Adachi *et al.*, *Precise measurement of the CP violation parameter $sin\phi_1$ in $B^0 \to (c\bar{c})K^0$ decays*, Phys. Rev. Lett. **108** (2012) 171802, `arXiv:1201.4643`.

[33] LHCb, R. Aaij *et al.*, *Measurement of the time-dependent CP asymmetries in $B_s^0 \to J/\psi K_S^0$*, JHEP **06** (2015) 131, `arXiv:1503.07055`.

[34] R. H. Dalitz, *On the analysis of tau-meson data and the nature of the tau-meson*, Phil. Mag. **44** (1953) 1068.

[35] BaBar collaboration, B. Aubert *et al.*, *An amplitude analysis of the decay $B^\pm \to \pi^\pm \pi^\pm \pi^\mp$*, Phys. Rev. **D72** (2005) 052002, `arXiv:hep-ex/0507025`.

[36] BaBar, 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** (2013), no. 1 012003, `arXiv:1304.3503`.

[37] G. N. Fleming, *Recoupling effects in the isobar model. 1. General formalism for three-pion scattering*, Phys. Rev. **135** (1964) B551.

[38] D. Morgan, *Phenomenological analysis of $I = 1/2$ single-pion production processes in the energy range 500 to 700 MeV*, Phys. Rev. **166** (1968) 1731.

[39] D. Herndon, P. Soding, and R. J. Cashmore, *A generalised isobar model formalism*, Phys. Rev. **D11** (1975) 3165.

[40] BaBar, J. P. Lees *et al.*, *Study of CP violation in Dalitz-plot analyses of $B^0 \to K^+K^-K_{(S)}^0$, $B^+ \to K^+K^-K^+$, and $B^+ \to K_{(S)}^0 K_{(S)}^0 K^+$*, Phys. Rev. **D85** (2012) 112010, `arXiv:1201.5897`.

[41] J. M. Blatt and V. F. Weisskopf, *Theoretical nuclear physics*, John Wiley & Sons, New York, 1952.

[42] C. Zemach, *Determination of the spins and parities of resonances*, Phys. Rev. **140** (1965) B109.

[43] 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** (1968) 244.

[44] S. M. Flatte, *Coupled - Channel analysis of the π-η and K-$\overline{K}$ systems near K-$\overline{K}$ Threshold*, Phys. Lett. **B63** (1976) 224.

[45] I. J. R. Aitchison, *K-Matrix formalism for overlapping resonances*, Nucl. Phys. **A189** (1972) 417.

[46] H.-Y. Cheng, *Theoretical Overview of Hadronic Three-body B Decays*, arXiv:0806.2895.

[47] H.-Y. Cheng, *Charmless 3-body B Decays: Resonant and Nonresonant contributions*, Int. J. Mod. Phys. **A23** (2008) 3229, arXiv:0711.1416.

[48] BaBar, 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.

[49] Belle, J. Dalseno *et al.*, *Time-dependent Dalitz Plot measurement of CP parameters in $B^0 \to K^0_{(S)}\pi^+\pi^-$ decays*, Phys. Rev. **D79** (2009) 072004, arXiv:0811.3665.

[50] Belle, Y. Nakahama *et al.*, *Measurement of CP violating asymmetries in $B^0 \to K^0_S K^+ K^-$ decays with a time-dependent Dalitz approach*, Phys. Rev. **D82** (2010) 073011, arXiv:1007.3848.

[51] BaBar collaboration, P. del Amo Sanchez *et al.*, *Observation of the Rare Decay $B^0 \to K^0_S K^\pm \pi^\mp$*, Phys. Rev. **D82** (2010) 031101, arXiv:1003.0640.

[52] LHCb collaboration, R. Aaij *et al.*, *Study of $B^0_{(s)} \to K^0_S h^+ h'^-$ decays with first observation of $B^0_s \to K^0_S K^\pm \pi^\mp$ and $B^0_s \to K^0_S \pi^+\pi^-$*, JHEP **10** (2013) 143, arXiv:1307.7648.

[53] ATLAS, G. Aad *et al.*, *Search for the decay $B^0_s \to \mu\mu$ with the ATLAS detector*, Phys. Lett. **B713** (2012) 387, arXiv:1204.0735.

[54] CMS, A. Heister, *Search for Bs,B0->mumu decays at CMS*, PoS **Beauty2014** (2015) 048.

[55] LHCb collaboration, *LHCb magnet: Technical Design Report*, CERN-LHCC-2000-007. LHCb-TDR-001.

[56] M. Vesterinen, *Considerations on the LHCb dipole magnet polarity reversal*, Tech. Rep. LHCb-PUB-2014-006. CERN-LHCb-PUB-2014-006, CERN, Geneva, Apr, 2014.

[57] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) 09007, arXiv:1405.7808.

[58] LHCb collaboration, J. Alves, A.Ãugusto *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[59] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.

[60] LHCb collaboration, *LHCb RICH: Technical Design Report*, CERN-LHCC-2000-037. LHCb-TDR-003.

[61] V. Gligorov, C. Thomas, and M. Williams, *The HLT inclusive B triggers*, .

[62] LHCb collaboration, *Performance of flavor tagging algorithms optimised for the analysis of $B_s^0 \to J/\psi\phi$*, LHCb-CONF-2012-026.

[63] LHCb collaboration, *Branching fraction measurements of $B_{d,s}^0$ decays to $K_S^0 hh'$ final states, including first observation of $B_s^0 \to K_S K\pi$*, LHCb-CONF-2012-023.

[64] G. Punzi, *Sensitivity of searches for new signals and its optimization*, in *Statistical Problems in Particle Physics, Astrophysics, and Cosmology* (L. Lyons, R. Mount, and R. Reitmeyer, eds.), p. 79, 2003. arXiv:physics/0308063.

[65] M. Pivk and F. R. Le Diberder, *sPlot: a statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

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

[67] Belle, A. Garmash *et al.*, *Study of B meson decays to three body charmless hadronic final states*, Phys. Rev. **D69** (2004) 012001, arXiv:hep-ex/0307082.

[68] S. Gardner, *Direct CP violation in untagged B meson decays*, Phys. Lett. **B553** (2003) 261, arXiv:hep-ph/0203152.

[69] S. Gardner and J. Tandean, *Observing direct CP violation in untagged B meson decays*, Phys. Rev. **D69** (2004) 034011, arXiv:hep-ph/0308228.

[70] Laura++ *Dalitz plot fitting package*, University of Warwick, http://laura.hepforge.org/.

[71] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. **A389** (1997) 81.

[72] *RooStats package webpage*, https://root.cern.ch/root/html/ROOFIT_ROOSTATS_Index.htm

[73] M. Pivk and F. R. Le Diberder, *sPlot: a statistical tool to unfold data distributions*, .

[74] *TMinuit webpage*, https://root.cern.ch/doc/v606/classTMinuit.html.

[75] M. Williams, *How good are your fits? Unbinned multivariate goodness-of-fit tests in high energy physics*, JINST **5** (2010) P09004, arXiv:1006.3019.