Time-dependent Dalitz-Plot Analysis of the Charmless Decay $B^0 \rightarrow K_S^0 \pi^+ \pi^-$ at *BABAR*

Pablo del Amo Sánchez

Thesis submitted for the degree of Doctor of Philosophy

Particle Physics Group,
School of Physics and Astronomy,
University of Birmingham.

October 2007.

Abstract

A time-dependent Dalitz-plot analysis of the decay channel $B^0 \to K^0_S \pi^+\pi^-$ has been performed on a data sample of 383 ± 3 millions of $B\overline{B}$ pairs recorded by the *BABAR* detector at the PEP-II B Factory. Branching fractions, direct and time-dependent asymmetries of different resonant modes have been measured, as well as the relative phases between them, among which that between $B^0 \to K^{*+}(892)\pi^-$ and $\overline{B}^0 \to K^{*-}(892)\pi^+$ is relevant for the extraction of the unitarity triangle angle γ. A summary of the results follows, the branching fractions quoted being averaged over *CP* conjugate states:

- $\mathcal{B}(B^0 \to K^0\pi^+\pi^-) = (45.8 \pm 1.6) \times 10^{-6}$
- $\mathcal{B}(K^{*\pm}(892)\pi^\mp; K^{*\pm}(892) \to K^0\pi^\pm) = (5.1 \pm 0.5) \times 10^{-6}$
- $\mathcal{B}(K_0^{*\pm}(1430)\pi^\mp; K_0^{*\pm}(1430) \to K^0\pi^\pm) = (23.4 \pm 1.4) \times 10^{-6}$
- $\mathcal{B}(\rho^0(770)K^0; \rho^0(770) \to \pi^+\pi^-) = (4.4 \pm 0.6) \times 10^{-6}$
- $\mathcal{B}(f_0(980)K^0; f_0(980) \to \pi^+\pi^-) = (7.4 \pm 0.7) \times 10^{-6}$
- $\mathcal{B}(K^0\pi^+\pi^-$ non resonant$) = (3.1 \pm 0.8) \times 10^{-6}$
- $\mathcal{S}(\rho^0(770)K^0) = -0.06 \pm 0.37$
- $\mathcal{S}(f_0(980)K^0) = -0.77 \pm 0.21$
- $\varphi(\overline{B}^0 \to K^{*-}(892)\pi^+,\ B^0 \to K^{*+}(892)\pi^-) = (-6 \pm 31)^\circ$

- All $\mathcal{A}_{CP}$ for the resonant modes quoted above are found to be consistent with zero.

A mi familia, por todo el apoyo que me habéis dado,
y en especial a mi padre,
por incitar la curiosidad que me ha traído hasta aquí.

Acknowledgements

I would like to thank my supervisor, Chris Hawkes, for his advice and encouragement, and the rest of the members of the Birmingham and Warwick *BABAR* groups (Tom Latham, Tim Gershon, Jelena Ilić, Alan Watson, John Back, Nitesh Soni and Gagan Mohanty). Professionally I have learnt a lot from them, and personally, they have always made me feel integrated and part of the group.

My thanks are also extended to those at the Particle Physics group, the University and otherwise that have helped make my stay in the UK enjoyable.

I would also like to thank the Particle Physics group of the University of Birmingham, the School of Physics and STFC (formerly PPARC) for providing the funding for my Ph.D.

I am grateful for the use of the computing facilities of the University, the e-Science and the BlueBear clusters, that have helped in carrying out the work described in this thesis, and to Lawrie Lowe for his assistance in using them. The main bulk of the computing needed for this work, however, has been carried out in the facilities made available by the *BABAR* Collaboration at SLAC and RAL.

I also want to say how much I appreciate my supervisor, Alan Watson and the examiners (Glen Cowan and Nigel Watson) being so flexible, ready to read a thesis with such short notice to suit the tight schedule imposed by my other commitments.

Finally, thanks to Daniel and Jelena for being the only ones in a 100 miles radius to believe that I would finish on time.

Declaration

I declare that no part of the work in this thesis has been previously presented to this, or any other university as part of the requirements of a higher degree.

The data on which the analysis has been performed were recorded by the *BABAR* Collaboration with the *BABAR* detector. The author has contributed to the collection and treatment of the data by taking detector shifts and producing K_S^0 efficiency correction tables available to all members of the Collaboration.

Two parallel though independent Dalitz analyses of the $B^0 \to K_S^0\pi^+\pi^-$ mode have been carried out within the Collaboration. The author has contributed extensively to one of those efforts, led by Thomas Latham and in which Tim Gershon, Jelena Ilić, Nitesh Soni and Gagan Mohanty also worked. The event reconstruction uses software developed by the Collaboration, and the selection stages employed the packages `QnBUser` and `CharmlessFitter`, written respectively by Fergus Wilson and Thomas Latham. The software used to perform the Dalitz-plot fit, `Laura++`, was first developed by Paul Harrison and John Back. Thomas Latham, with the collaboration of the author, extended it to the time-dependent case. In particular, the author derived and implemented the formulae for the analytic treatment of the Δt resolution model (Section 4.4.4), included the possibility of using the $m_{\rm ES}$ and ΔE variables in the fit (Section 5.3), developed and implemented Tim Gershon's ideas on the treatment of Self Cross Feed (Section 5.6) and finalized the handling of backgrounds in the fit (Section 5.8). The author also tested the fit extensively in a variety of different scenarios.

Tom Latham, Gagan Mohanty and Nitesh Soni produced the event selection (Section 5.4). The last two also worked out the optimal combination of the different background discriminating variables, coming up with the current configuration of the Fisher discriminant (Section 4.5.3).

The studies of the dependencies of the background discriminating variables on the Dalitz-plot position, their implementation and the extraction of a number of PDF shapes were made by Jelena Ilić, who also did the $B\overline{B}$ background identification. The author put together the final background model in close collaboration with her (Sections 5.7, 5.8).

Finally, the work described in Chapter 6 was carried out solely by the author.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this Chapter we briefly state the physics interest of the analysis that is the subject of this thesis, the charmless 3-body decay $B^0 \to K_S^0 \pi^+ \pi^-$ and put it in the wider context of particle physics today. We start by introducing some fundamental concepts relevant to B physics, such as CP violation and tagging.

1.1 Matter, antimatter and CP violation

When matter is created from energy, it is always done via a "pair production" process, in which both a particle and its antiparticle are formed. If the same mechanisms were involved in the Big Bang, equal amounts of matter and antimatter would have been generated, and they would have annihilated completely back to photons. In addition, the observation that no antimatter seems to be present today in the Universe, except in events energetic enough to produce it, raises an important question: How did the early Universe evolve from its initial matter-antimatter symmetric state to the one that exists today?

This question was addressed by Andrei Sakharov in the 1960's in [1], where he concluded that three conditions were necessary: baryon number non-conservation, departure from thermal equilibrium and the so called "CP violation". The latter formalizes the concept of matter-antimatter asymmetry (see Chapter 2 for more details) and means that the combined symmetry of exchanging a particle by its antiparticle (C) and reflection through the origin (P) is not respected.

Evidence for the breaking of such combined symmetry had also been found experimentally in the decays of the neutral strange mesons [2]. Kobayashi and Maskawa, compelled to introduce a source of CP violation into the current theories of particle physics, introduced a complex matrix that couples the quarks with the carriers of the weak force [3]. By choosing the dimensionality of the matrix to be 3×3, implying three generations of quark flavours, they found that an irreducible phase could be introduced, which would lead to CP violating effects in processes where interference occurs among several amplitudes.

By the end of the 1970's and early 1980's, Bigi, Carter, Sanda and others noticed that the additional amplitudes required could be provided by the phenomenon of "mixing" [4, 5, 6], in which a neutral meson turns into its antiparticle via second-order weak interactions (see Chapter 2). These would lead to particularly large asymmetries in the case of the B^0 mesons (composed of a bottom antiquark and a light d quark) decaying to states common to both B^0 and $\overline{B}^0$. Their conjecture was supported by the discovery that, for these mesons, the period of these particle-antiparticle oscillations was comparable to their lifetimes [7, 8], which meant that the interfering amplitudes would be of the same order, and made the B system the best choice to study CP violation. However, the same phenomenon also impeded the experimental determination of the flavour (i.e. B^0 or $\overline{B}^0$) of the particle in decay, a *sine qua non* to compare the decay rates of both flavours. The proposed solution involved using pairs of B^0-$\overline{B}^0$ mesons produced in the quantum mechanical state of entanglement, one of them decaying to the channel of interest, and the other one to a final state that uniquely determines its flavour (a *flavour eigenstate*). The entanglement implied that, at the precise moment of the decay of the latter, the former meson would be its antiparticle, and that its flavour at decay time could be determined as well by evolving its known state in time (*tagging*). Thus, measuring CP asymmetries between the B mesons required simultaneous reconstruction of the decay products of both B mesons and measurement of the time difference between their decays (see Figure 1.1 and Chapter 4 for more details).

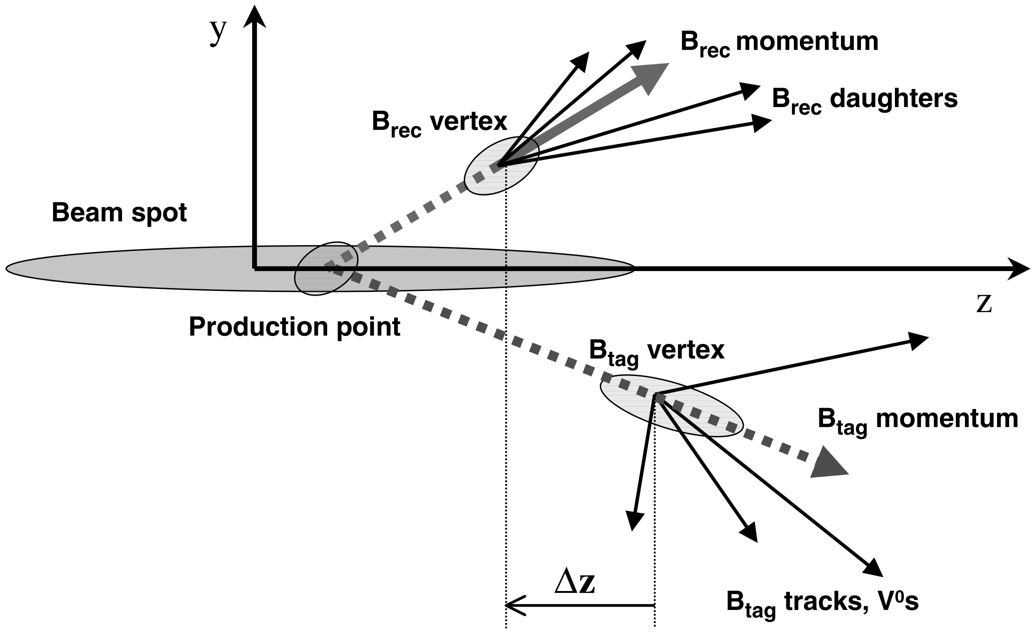


Figure 1.1: In many processes of interest, B_{rec} decays to a channel that can be reached by both a B^0 or a $\overline{B}^0$. Evaluation of the CP violating asymmetries requires knowing the flavour of the B meson at its decay time. This is achieved by identifying the decay channel of the other B meson, B_{tag}, often only partially reconstructed. If B_{tag} decays into a state that uniquely determines the flavour of its parent particle, e.g. $B^0 \to D^* + X$, then the knowledge of the difference between the decay times of both B's allows us to infer the flavour of B_{rec}. The time elapsed between the two decays can be calculated from the distance Δz measured between the two B decay vertices along the beam direction, z.

The lifetime of the B mesons ($\sim$ 1.5 ps) made impractical any such measurements using existing (symmetric) colliders, since the small distances travelled in the detector would be washed out by the resolution of the instrument. Pier Oddone had the idea of building an *asymmetric* e^+e^- collider, in which the centre-of-mass, and hence any particles produced, would move in the laboratory frame with a relativistic boost large enough to extend those distances into the measurable range [9]. The concept of a B Factory was born, and would soon grow into the *BABAR* and *Belle* experiments. (A detailed discussion of the requirements that such an unusual situation, and the physics goals, place on the detector can be found in Chapter 3).

1.2 Flavour physics, charmless 3-body B decays and physics beyond the Standard Model

The flavour sector is one of the unexplained features of the Standard Model (SM), and, within our current understanding of particle physics, CP violation is intimately linked to the existence of flavour, as well as to the ubiquitous Higgs field (see Chapter 2). A B Factory running on the $\Upsilon(4S)$ resonance ($\sqrt{s} = 10.58\,\mathrm{GeV}$) is also a $c\bar{c}$ and a τ factory, thus allowing a detailed probe of the flavour structure. Its main goal is to make precise measurements of the Kobayashi-Maskawa matrix elements, and in particular the elusive phase within it, for which the B mesons are the best experimental candidates.

The B-meson system is also unique in that second-order weak interactions (such as those giving rise to mixing) have relatively large amplitudes. In particular, loop (or *penguin*) diagrams have been shown to dominate some B decay modes, e.g. [10]. The interest in such amplitudes lies in the possibility that physics beyond the Standard Model could enter into those diagrams through virtual particles in the loop (see Figure 2.5).

Such contributions should arguably be small and, if present, would be best seen in modes where the SM contribution is itself small. Charmless B decays, which proceed through suppressed $b \to s$ transitions, are a good example and have been studied extensively in *BABAR*. Among them, 3-body decays have the added advantage of providing a larger phase space that allows for interference between different resonances, and therefore the possibility of extracting directly any phase differences involved. This feat is achieved by stepping back to the fundamental quantum mechanical postulate that what is observed in any process in nature is an intensity resulting from one or several amplitudes. Therefore, rather than modelling the intensity directly, an additional level of detail is gained if the amplitudes are taken as the most fundamental objects (*Dalitz-plot* or *amplitude analysis*). The interferences between these amplitudes, now taken into account correctly, give access to the relative phases between them. (A more thorough discussion follows in Section 2.5.)

For example, in $B^0 \to K_S^0\pi^+\pi^-$, the channel studied in this thesis, this analysis technique and the use of flavour symmetry allow the extraction of the phase γ [11] of Kobayashi and Maskawa's matrix (see Section 2.1.2), in addition to the more common β, involved in B^0-$\overline{B}^0$ mixing and accesible with many neutral decay modes.

Some experimental prerequisites for the amplitude analysis of $B^0 \to K_S^0\pi^+\pi^-$ are examined in Chapter 4, whereas the analysis itself and its results are discussed in Chapters 5 and 6, respectively.

Chapter 2

Theory

The theoretical foundations of *CP* violation in the Standard Model, briefly outlined in the previous chapter, are examined in detail here (Section 2.1). These are followed by a review of the current experimental status of the decay mode $B^0 \to K_S^0 \pi^+ \pi^-$, and the theoretical interest in it (Section 2.4). We finish by discussing the kinematic and dynamic peculiarities of three-body decays, pointing out why these are best suited for the physics goal of extracting a phase and motivating the parameterizations used in their description.

2.1 *CP* violation and the Standard Model

2.1.1 Weak interactions, mass and the flavour sector

The origin of *CP* violation in the Standard Model lies, as mentioned in Chapter 1, in an irreducible phase appearing in the Lagrangian describing the charged current decays of quarks. This phase results from the complex interplay between the structure of the flavour sector in the weak interactions, and the mass terms for the quarks, brought about by the Higgs mechanism in the Standard Model.

The charged current interactions that arise in the Glashow-Weinberg-Salam electroweak model are described by the following terms in the SM Lagrangian [12]:

$$\begin{aligned}\mathcal{L}_{\rm cc} &= \frac{ig}{\sqrt{2}}\left[\bar{\nu}_m\gamma^\mu\left(\frac{1-\gamma^5}{2}\right)e_m W_\mu^+ + \bar{e}_m\gamma^\mu\left(\frac{1-\gamma^5}{2}\right)\nu_m W_\mu^- + \right.\\ &\left.\bar{u}_m\gamma^\mu\left(\frac{1-\gamma^5}{2}\right)d_m W_\mu^+ + \bar{d}_m\gamma^\mu\left(\frac{1-\gamma^5}{2}\right)u_m W_\mu^-\right]\end{aligned} \tag{2.1}$$

where the index m runs over the three generations that make up the particle content of the model. The first line accounts for the coupling of the leptons to the W boson, and the second line for that of the quarks. Note that the expression is written in terms of the *flavour* basis, in which the three generations are not connected.

These particles acquire mass through their Yukawa couplings to the Higgs condensate ϕ:

$$\mathcal{L}_{\rm Y} = -Y_{ij}^u\bar{q}_{L,i}\phi u_{R,j} - Y_{ij}^d\bar{q}_{L,i}\phi d_{R,j} \tag{2.2}$$

where the $Y^{u,d}$ are 3×3 complex matrices, the indices i and j label the generations, and the terms for the leptons have been dropped for simplicity, being analogous to those of the quarks. These terms couple the left-handed component of each quark q with the right-handed projections of all flavours, and are thus non-diagonal in the flavours. Since the states that propagate in space and time are the mass eigenstates, the Yukawa terms must be diagonalized. This is done by changing from the flavour to the mass basis:

$$u_m = \mathcal{U}_{mn}^u u_n', \quad d_m = \mathcal{U}_{mn}^d d_n'. \tag{2.3}$$

Substituting in the charged current lagrangian gives

$$\mathcal{U}_{mi}^{u\dagger}\mathcal{U}_{mj}^d\bar{u}'_i\gamma^\mu\left(\frac{1-\gamma^5}{2}\right)d_j' W_\mu^+ + \mathcal{U}_{mj}^{d\dagger}\mathcal{U}_{mi}^u\bar{d}'_j\gamma^\mu\left(\frac{1-\gamma^5}{2}\right)u_i' W_\mu^- \tag{2.4}$$

where quarks from different flavours i and j are coupled through the complex matrix $V_{ij}^{\rm CKM} \equiv \mathcal{U}_{mi}^{u\dagger}\mathcal{U}_{mj}^d$, named after Cabbibo, Kobayashi and Maskawa [13, 3].

This matrix would be the unit matrix if the rotations needed to diagonalise the up-type and the down-type quark mass terms, $\mathcal{U}^u$ and $\mathcal{U}^d$, were identical. **The up- and down-type mass matrices turn out to be misaligned and the Standard**

Model is then endowed with generation-changing transitions via charged currents, **and *CP*-violation** if three or more generations exist, as irreducible phases appear in the CKM matrix (see next section).

A similar argument leads to an analogous matrix in the lepton sector, known as the PMNS matrix (for Pontercorvo, Maki, Nakagawa and Sakata) [12].

2.1.2 The CKM matrix and the Unitarity Triangle

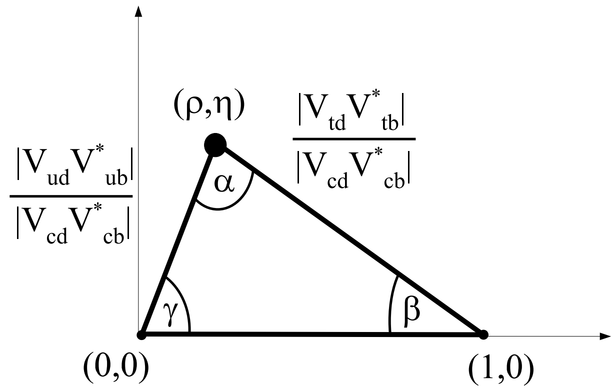
The $2N^2$ real parameters of the complex $N \times N$ matrix V^{CKM}, as Kobayashi and Maskawa realized [3], are reduced to N^2 by the unitarity conditions ($V^{\mathrm{CKM}\dagger} V^{\mathrm{CKM}} = 1$), and by another $2N - 1$ degrees of freedom when individual phase redefinitions for the $2N$ quark fields are considered. (One of those just changes the overall phase, and cannot be used to eliminate one of the free parameters.) The number of irreducible phases is found by comparing how many free parameters remain with the number expected in a $N \times N$ real, orthogonal matrix ($\frac{1}{2}N(N-1)$), and turns out to be 0 for $N = 2$ and 1 for $N = 3$. As they also noticed, this irreducible phase in a three-generation model would be the origin of all *CP* violation in the SM (see Section 2.1.3).

The CKM matrix can, therefore, be parameterized with a complex phase and three angles, that describe the rotation between each generation. Surprisingly, these angles take rather small values in Nature, favouring a nearly diagonal matrix in which little mixing between the three generations is produced. The matrix elements containing a complex phase vary among parameterizations, while the physics, of course, remains the same. A relative phase arises between transitions from the third to first generations, and those that proceed from third to second or from second to first generations. A useful parameterization, that emphasizes the hierarchical, nearly diagonal structure of the matrix, and between which flavours the relative phases appear, is due to Wolfenstein [14]:

$$V^{\mathrm{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \simeq \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & \lambda^3 A\left(\rho - i\eta\right) \\ -\lambda & 1-\frac{1}{2}\lambda^2 & \lambda^2 A \\ \lambda^3 A\left(1-\rho-i\eta\right) & -\lambda^2 A & 1 \end{pmatrix} + \mathcal{O}\left(\lambda^4\right)$$

$$\sim \begin{pmatrix} 1 & \lambda & \lambda^3 e^{-i\gamma} \\ -\lambda & 1 & \lambda^2 \\ \lambda^3 e^{-i\beta} & -\lambda^2 & 1 \end{pmatrix}, \tag{2.5}$$

where $\lambda \simeq |V_{us}| \sim 0.22$ is the expansion parameter, and A, ρ and η are all of order 1.



$$\begin{aligned} \alpha &= \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right) \\ \beta &= \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right) \\ \gamma &= \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right) \end{aligned} \tag{2.6}$$

Figure 2.1: Pictorial representation, dubbed "Unitarity Triangle", of the relation $V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$ derived from the unitarity of the CKM matrix. Each of the terms in the LHS can be viewed as a vector in the complex plane. Their sum being equal to zero then implies that they have to form a (closed) triangle in the Argand plane. The base of the triangle has been aligned with the horizontal (real) axis by dividing the equation by $V_{cd}V_{cb}^*$. The definitions of the angles of the triangle in terms of the CKM matrix elements are given on the right.

The physics goal of the B Factories is to improve the experimental constraints on the elements of V^{CKM}, particularly on the phase. (A good review of the experimental and theoretical effort towards that aim is [15]). Of course, not all the matrix elements are easily accesible to *BABAR* or *Belle*. The typical combinations of CKM factors affecting B decays also appear in one of the relations implied by the unitarity condition $V^{\text{CKM}\dagger}V^{\text{CKM}} = 1$:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \tag{2.7}$$

Regarding each term in the LHS as a vector in the complex plane, this equation can be visualized as a triangle whose apex is determined by the ρ and η parameters, or equivalently, the irreducible phase (see Figure 2.1). Once this pictorial representation, known as the *Unitarity Triangle*, has been introduced, the goals of the B Factories can be stated more accurately to be making as many independent measurements of the sides and the angles of the triangle as possible. Any direct inconsistencies between them, or indirect, such as the angles not adding up to 180° or the shapes inferred from the angles and the sides being incompatible, would indicate that the Standard Model picture of the flavour sector and of CP violation is no more than a good approximation (Figure 2.2). The analysis described in this thesis aims to measure the angle β, and contribute towards another measurement of γ.

2.1.3 CP violation in the SM and beyond

In previous sections, the origin of the non-trivial phases in the Standard Model coupling constants has been discussed with the observation that they are responsible for all CP-violating effects. Such phases, which exist in the CKM matrix and may also appear in the PMNS matrix, arise as constants multiplying otherwise CP-invariant terms of the Lagrangian (see Eq. (2.4)). They are called *CP-odd* because they become complex-conjugated under CP transformations [17], thus breaking the CP symmetry of the terms they form part of. In contrast, the phases generated by CP-invariant terms in the Lagrangian, such as those from QCD, are equal for a process and its CP-conjugate, and are therefore known as *CP-even* phases. As we shall see, it is the interplay between both behaviours that brings about CP violation.

Consider a quantum-mechanical amplitude with CP-odd and CP-even phases ϕ and δ, respectively. According to the remarks above, the amplitude transforms under CP as

$$Ae^{i(\delta+\phi)} \xrightarrow{CP} Ae^{i(\delta-\phi)} \,. \tag{2.8}$$

However, phases are not directly observable, so even though these two amplitudes are different, the rates for the processes they describe are the same. Since only relative phases have physical meaning, more than one amplitude is needed for CP-violating

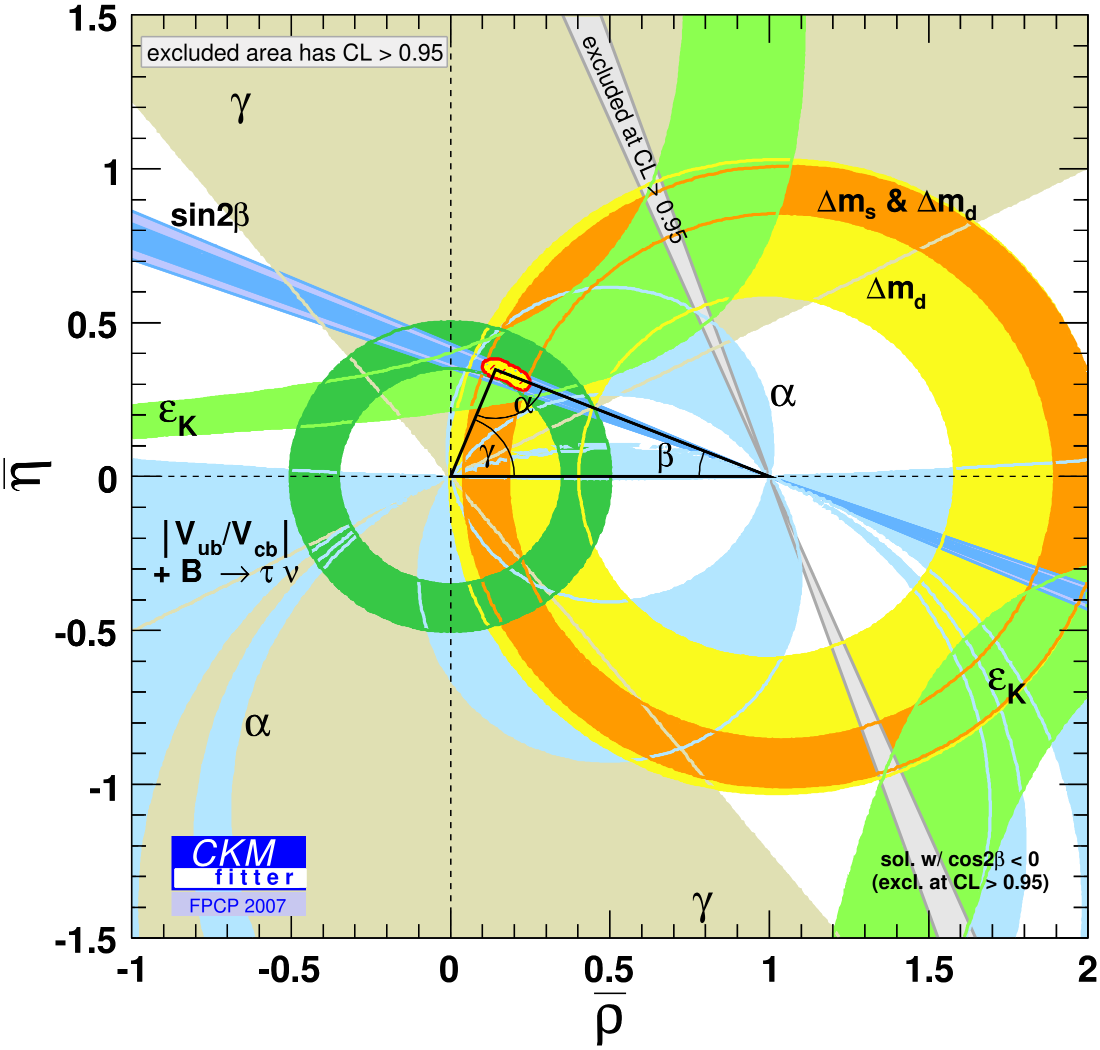


Figure 2.2: Pictorial representation of the current experimental constraints on the sides and angles of the unitarity triangle, by the CKMfitter group [16], updated with the results available in May 2007. The bands representing each experimental measurement or bound are seen to be in very good agreement with each other, all of them overlapping around the apex of the triangle drawn. The variables plotted on the axes, $\bar{\rho}$ and $\bar{\eta}$, are defined by $\bar{\rho} + i\bar{\eta} \equiv V_{ud}V_{ub}^{*}/V_{cd}V_{cb}^{*}$, and are related to the quantities ρ and η from Wolfenstein's parameterization. They are preferred over the latter because they are independent of the order at which the expansion in λ is truncated.

asymmetries. Consider now the amplitude

$$A = A_1 e^{i(\delta_1+\phi_1)} + A_2 e^{i(\delta_2+\phi_2)} \tag{2.9}$$

and its *CP*-conjugate:

$$\bar{A} = A_1 e^{i(\delta_1-\phi_1)} + A_2 e^{i(\delta_2-\phi_2)} . \tag{2.10}$$

The difference in the rates is

$$|A|^2 - |\bar{A}|^2 = -4|A_1||A_2| \sin(\delta_1 - \delta_2) \sin(\phi_1 - \phi_2) . \tag{2.11}$$

Thus, for such *direct CP* asymmetries to appear, two or more amplitudes must contribute to the process with different *CP*-odd and *CP*-even phases. Processes that give rise to non-zero direct asymmetries are said to be affected by *CP violation in the decay* or *direct CP violation*.

There are two other kinds of *CP* violation, both related to the phenomenon of mixing (see Section 2.2.1), by which a neutral meson turns into its antiparticle through a second-order weak transition. One of the possibilities, called *CP-violation in the mixing*, is that the physical states, that propagate in space and time, are not *CP* eigenstates, or equivalently, that they are not composed of equal amounts of particle and antiparticle. The first observation of *CP* violation [2], made in 1964 thanks to very precise measurements of the neutral kaon system, falls into this category. The other possibility, named *mixing-induced CP violation*, is the one foreseen by Bigi, Carter, Sanda and others to be of primary importance in the decays of neutral B mesons, in which *CP* violation arises as a consequence of the interference between decays with and without mixing, i.e. $B^0 \to f_{CP}$ and $B^0 \to \overline{B}^0 \to f_{CP}$. A requisite for these is that the final state must be accessible from both flavours, a condition that is satisfied if it is a *CP* eigenstate. This case is discussed in detail in the next section. Before the start of the B Factories, *CP* violation had been observed in the mixing and in the decay of the $K^0 - \overline{K}^0$ system [2, 15].

A further source of *CP* violation in the Standard Model is an extra term in the QCD Lagrangian, allowed by the gauge principle, of the form $\propto \vec{E}_a \cdot \vec{B}_a$, where a is the colour index, and $\vec{E}$ and $\vec{B}$ are the colour electric and magnetic fields. It has been found experimentally that the constant that determines the relative contribution of this term, $\theta_{\rm QCD}$, either vanishes or is very close to zero: $\theta_{\rm QCD} \lesssim 10^{-10}$ [15].

Beyond the Standard Model, almost any unification scheme introduces more *CP* violating sources as a consequence of the possibly complex couplings of the new particle fields incorporated. Supersymmetry is also well known for the large increase in the particle content of the model that it postulates. The minimally supersymmetric extension of the SM (MSSM) requires five physical fields (three neutral Higgs, and two charged), which are good candidates to host more *CP*-violating phases [17]. Currently, all ideas for physics beyond the Standard Model have found no support from experiment beyond the realm of speculations.

2.2 The B meson system

2.2.1 Mixing

A neutral meson can turn into its antiparticle via a process involving two charged currents. The phenomenon has been observed in kaons [18], B_d and B_s mesons [7, 19], and very recently, D mesons [20, 21], and is central to the generation of such large time-dependent *CP* asymmetries in the neutral B system. Whereas in kaons the transition is dominated by long-distance, hadronic effects, in the B mesons it is short-distance contributions that are most prominent, and hence, the mixing frequency is amenable to perturbative calculations[1].

Consideration of the relevant diagrams leads, in principle, to the inclusion of the three up-type quarks in the loop (see Figure 2.3), since all of them couple via a λ^6 CKM factor. However, integration of the internal degrees of freedom (the up-type quarks and the W) yields an expression that weighs the contribution of each quark by the ratio of its mass to that of the weak boson [17]. Thus, the top quark contribution

[1] In D mesons both kinds of effects are important, and the theoretical estimations suffer from large uncertainties.

is enhanced, and evaluation of the corresponding CKM factors (see Eq. (2.5)) readily gives a phase of $(V_{td}V_{tb}^*)^2 \sim e^{-i2\beta}$. In other words, the B^0 states that oscillate into $\overline{B}^0$ pick up an extra -2β phase, called the *mixing angle*, with respect to the B^0 states that do not oscillate. This phase is measurable whenever both flavours decay to a common state.

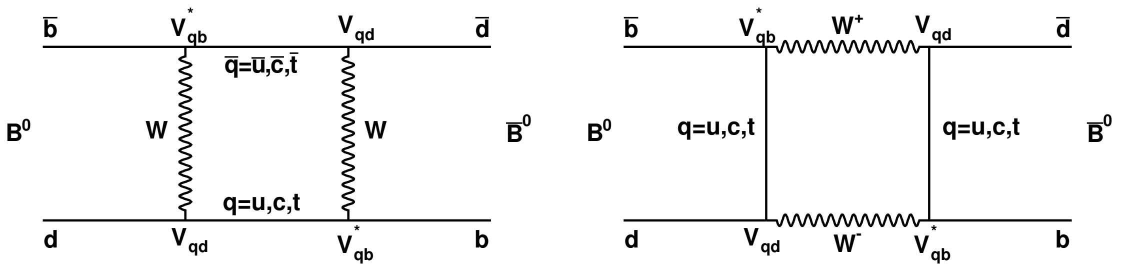


Figure 2.3: Diagrams describing the mixing phenomenon for neutral B mesons via a second-order weak transition. Although the three up-type quarks may appear in the loop, the u and c contributions turn out to be negligible, and only the t quark needs to be considered, which implies that the oscillation results in a phase shift of -2β.

In the neutral B mesons, the amplitude for the mixing process is rather large due to the relative proximity of the values for their lifetime, $\tau = 1.530(9)\,\mathrm{ps}$, and the frequency of oscillation, $\Delta m_d = 0.507(5)\,\mathrm{ps}^{-1}$ [22]. They conspire to give an integrated probability of oscillation of order 1 [23]:

$$\chi_d = \frac{(\tau \Delta m_d)^2}{2(1 + (\tau \Delta m_d)^2)} = 0.188(3)\,, \tag{2.12}$$

hence the prominent role of mixing. An immediate consequence are the large time-dependent CP asymmetries that mixing produces, and that are discussed in the next section.

2.2.2 CP violation in B mesons

Whereas for the neutral kaons CP violation appears mainly in the mixing and in the decay only as a fairly minute effect ($\sim 10^{-3}$ [22]), in the B mesons CP violation is rather large (reaching 70% [24] in time-dependent asymmetries, and $\sim 30\%$ [25, 10] in the charged modes), although mostly in the forms of mixing-induced and direct *CP* violation. Indeed, the shift in phase due to mixing, which provides the necessary *CP*-odd phase, and the high probability of oscillation join forces to produce large time-dependent *CP* asymmetries through the interference between $B^0 \to f_{CP}$ and $B^0 \to \overline{B}^0 \to f_{CP}$. The large direct *CP* asymmetries observed in charged modes are explained by the interference of tree-level and loop diagrams (see Section 2.3). *CP* violation in the mixing, however, is believed to be below the 10^{-3} level [15] for neutral B mesons.

We shall now consider the mixing-induced *CP* violation in more detail. Direct *CP* asymmetries are usually quantified experimentally by

$$\mathcal{A}_{CP} = \frac{\mathcal{B}\left(\overline{B} \to \bar{f}\right) - \mathcal{B}\left(B \to f\right)}{\mathcal{B}\left(\overline{B} \to \bar{f}\right) + \mathcal{B}\left(B \to f\right)}\,. \tag{2.13}$$

Such an observable is well defined for *charged* mesons, since the flavour of the B is readily found from the charge of the candidate, or minus the charge of the rest of the event. The flavour of neutral B mesons, however, is not straightforward to determine, as their oscillating nature means that their flavour changes over time. The design of the *BABAR* and *Belle* experiments, in which the $\Upsilon(4S)$ resonance is produced and allowed to decay into a $B\overline{B}$ system, eases the task by providing us with an *entangled* pair of mesons. Indeed, in another instance of the celebrated "EPR paradox" [26, 27], quantum mechanics predicts that the instantaneous probability of one of the mesons being a particle equals the probability for the other meson to be an antiparticle. Hence, when one of them is observed to decay in a flavour-dependent way, the flavour of the other meson can be inferred to be the opposite at that exact same instant, and evolves afterwards indepedently by means of mixing. Thus, the ingredients to the determination of the flavour of a given B meson, $B_{\rm rec}$ (see Figure 1.1), are two: observation of the other meson ($B_{\rm tag}$) decaying into a flavour state that uniquely establishes its flavour (*tag flavour*), and measurement of the time elapsed between the two decays, Δt, so that the oscillations of the meson under consideration can be accounted for. The hierarchy of the CKM matrix happens to help for the former, since the decays of a B meson most often involves a single c quark, e.g $B^0 \to D^-\pi^+ \to K^+\pi^-\pi^+$, thereby revealing the flavour of the b.

With the previous considerations in mind, it makes sense to construct the observable

$$\mathcal{A}_{CP}(\Delta t) = \frac{\mathcal{B}\left(B_{\text{tag}=B^0}(\Delta t) \to f_{CP}\right) - \mathcal{B}\left(B_{\text{tag}=\overline{B}^0}(\Delta t) \to f_{CP}\right)}{\mathcal{B}\left(B_{\text{tag}=B^0}(\Delta t) \to f_{CP}\right) + \mathcal{B}\left(B_{\text{tag}=\overline{B}^0}(\Delta t) \to f_{CP}\right)} . \tag{2.14}$$

An expression for the distributions of events as a function of the tag q_{tag} and the time difference Δt, and hence the asymmetry, can be found [28] in terms of the ratio of the amplitudes for the decay of the B^0 and $\overline{B}^0$ mesons to the final state f_{CP}, respectively $\mathcal{A} = \langle B|\mathcal{H}|f_{CP}\rangle$ and $\bar{\mathcal{A}} = \langle \bar{B}|\mathcal{H}|f_{CP}\rangle$:

$$f_{q_{\text{tag}}}(\Delta t) = \frac{e^{-|\Delta t|/\tau}}{4\tau}\left[1 + q_{\text{tag}}\frac{\Delta\mathcal{D}}{2} + q_{\text{tag}}\langle\mathcal{D}\rangle\left(\mathcal{S}\sin\left(\Delta m_d \Delta t\right) - \mathcal{C}\cos\left(\Delta m_d \Delta t\right)\right)\right] \tag{2.15}$$

from which the asymmetry is readily evaluated:

$$\mathcal{A}_{CP}(\Delta t) = \langle\mathcal{D}\rangle\left(\mathcal{S}\sin\left(\Delta m_d \Delta t\right) - \mathcal{C}\cos\left(\Delta m_d \Delta t\right)\right) \tag{2.16}$$

where

$$\mathcal{S} = \frac{2\text{Im}\lambda}{1+|\lambda|^2}\,, \quad \mathcal{C} = \frac{1-|\lambda|^2}{1+|\lambda|^2}\,, \quad \lambda = e^{-i\phi_{\text{mix}}}\frac{\bar{\mathcal{A}}}{\mathcal{A}}\,, \tag{2.17}$$

$q_{\text{tag}} = +1(-1)$ when the B_{tag} is identified as a B^0 ($\overline{B}^0$) and ϕ_{mix} is the mixing phase. The terms $\langle\mathcal{D}\rangle$ and $\frac{\Delta\mathcal{D}}{2}$ have been introduced to account for a non-zero probability of mistagging (see Section 4.1 for definitions and a detailed discussion).

The coefficient $\mathcal{S}$ is different from zero when there is mixing-induced *CP* violation, whereas $\mathcal{C} \neq 0$ indicates direct *CP* violation ($|\bar{\mathcal{A}}| \neq |\mathcal{A}|$). Also note that $\mathcal{C} = -\mathcal{A}_{CP}$. If there is only one SM contribution to the amplitudes $\mathcal{A}$ and $\bar{\mathcal{A}}$, the expectations are that $\mathcal{S} = -\eta_{CP}\sin(2\beta)$ and $\mathcal{C} = 0$, where η_{CP} is the *CP* eigenvalue of the final state f_{CP}. Deviations from that imply the existence of unaccounted amplitudes that, depending on the characteristics of the mode, could originate from theoretical uncertainties in the Standard Model contributions, or constitute a first glimpse at physics beyond the SM.

In order to facilitate the comparisons of the experimentally measured time-dependent asymmetries among themselves and with the thoeretical predictions, it is common (see Figure 2.6) to quote the sine of the "effective mixing angle" $2\beta_{\text{eff}}$, defined by:

$$\mathcal{S} = \sqrt{1-\mathcal{C}^2}\eta_{CP} \sin 2\beta_{\text{eff}} \, . \tag{2.18}$$

For three-body decays it is useful to write Eq. (2.15) in terms of the amplitudes rather than their ratios, and keep track of their dependencies. In the case of $B^0 \to K^0_S\pi^+\pi^-$, it yields [29]:

$$\begin{aligned}\frac{d\Gamma\left(m^2_{K^0_S\pi^+}, m^2_{K^0_S\pi^-}, \Delta t, q_{\text{tag}}\right)}{d\Delta t\, dm^2_{K^0_S\pi^+}\, dm^2_{K^0_S\pi^-}} &= \frac{1}{(2\pi)^3}\frac{1}{32m^3_{B^0}}\frac{e^{-|\Delta t|/\tau_{B^0}}}{4\tau_{B^0}} \times \\ &\quad \Bigg[\left(|\mathcal{A}|^2+|\overline{\mathcal{A}}|^2\right)\left(1+q_{\text{tag}}\frac{\Delta\mathcal{D}}{2}\right) \\ &\quad - q_{\text{tag}}\langle\mathcal{D}\rangle\left(|\mathcal{A}|^2-|\overline{\mathcal{A}}|^2\right)\cos\left(\Delta m_d\Delta t\right) \\ &\quad + q_{\text{tag}}\langle\mathcal{D}\rangle 2\text{Im}\left[\overline{\mathcal{A}}\mathcal{A}^* e^{-i\phi_{\text{mix}}}\right]\sin\left(\Delta m_d\Delta t\right)\Bigg],\end{aligned} \tag{2.19}$$

where the B^0 and $\overline{B}^0$ Dalitz-plot amplitudes $\mathcal{A} = \mathcal{A}\left(m^2_{K^0_S\pi^+}, m^2_{K^0_S\pi^-}\right)$ and $\overline{\mathcal{A}} = \overline{\mathcal{A}}\left(m^2_{K^0_S\pi^+}, m^2_{K^0_S\pi^-}\right)$ exhibit a dependence on the phase space variables of the final state, the invariant masses squared of two pairs of particles cons(see Section 2.5).

Finally, the distribution for the decays of neutral B mesons to flavour (i.e. non-CP) eigenstates (the so-called *flavour* or *mixing* sample) is considered, since they are experimentally the cleanest way to measure any parameters related to the time dependence, such as the lifetime of the mesons or their mixing frequencies, and constitute a common source of background for decays to CP eigenstates [23]:

$$f_{q_{\text{tag}}}(\Delta t) = \frac{e^{-|\Delta t|/\tau}}{4\tau}\left[1+q_{\text{tag}}\frac{\Delta\mathcal{D}}{2} \pm \langle\mathcal{D}\rangle\cos\left(\Delta m_d\Delta t\right)\right] \tag{2.20}$$

where the $\pm$ signs correspond to the unmixed and mixed possibilities, respectively.

2.3 Trees and Penguins

The interpretation of $\mathcal{S}$ as $\sin 2\beta$ is accurate only when, apart from the oscillation, just one amplitude dominates the decay $B \to f_{CP}$, or, in the case in which several are present, they all have the same *CP*-odd phases. That is the case in the so-called "golden mode" $B^0 \to J/\psi K^0_S$, in which two Feynman diagrams contribute. One of them, in which the W boson is emitted and absorbed by the same quark line, is called a "loop" or, more often, a "penguin" diagram, whereas the other, is said to be a "tree-level", or simply "tree", diagram (see Figure 2.4).

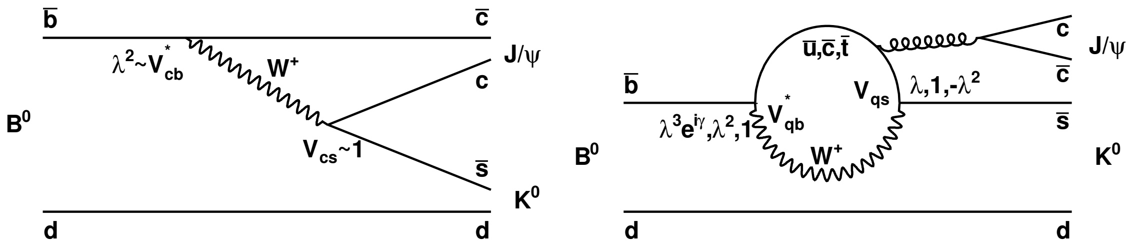


Figure 2.4: Feynman diagrams for the amplitudes contributing to the $B^0 \to J/\psi K^0_S$ decay. The one on the right, with the internal loop formed by the W boson and a quark is a *penguin* diagram; the one on the left, a *tree* diagram. Since the top quark dominates the loop, the CKM factors in both amplitudes carry the same phases, as can easily be seen thanks to the powers of λ and phases that have been written next to the CKM terms (recall Eq. (2.5)). Thus, a clean interpretation for the time-dependent asymmetry is obtained: $\mathcal{A}_{CP}(\Delta t) = \sin 2\beta \sin(\Delta m_d \Delta t)$. In fact, this decay channel is considered to be a "golden mode" due to its theoretical and experimental cleanliness, playing the role of a benchmark with which to compare measurements of $\sin 2\beta$ possibly affected by new physics.

Tree diagrams are usually cleaner in their theoretical interpretation than penguin diagrams. However, the latter are an open window to physics at the electroweak scale, as hypothetical new, heavy particles could traverse the loop and provide new *CP*-odd phases or dramatically enhance the amplitude (Figure 2.5 shows two examples). This is analogous to the constraints on the top quark mass that are derived from $B^0 - \overline{B}^0$ mixing. There is some theoretical interest [15] as to whether any of those possibilities is realized in charmless $b \to s$ decays, where the absence of the large $b \to c$ amplitudes will not mask small deviations due to unknown physics (see Section 2.4.2 for a brief review of the experimental state of affairs).

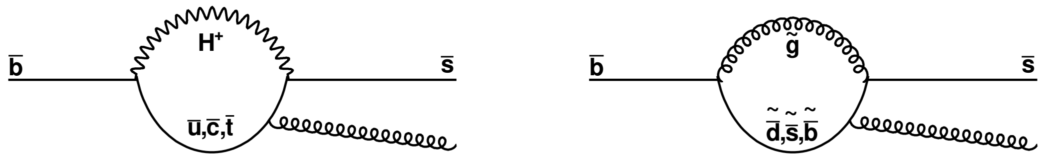


Figure 2.5: Diagrams demonstrating the potential for new physics sensitivity in $b \to s$ penguin diagrams. On the left, a charged Higgs, predicted by the MSSM, enters the loop possibly carrying a complex coupling constant. On the right, the loop is formed by a gluino and (anti)squarks, enhancing the SM amplitude. Supersymmetry has potentially large effects on flavour observables, since squarks may change flavour while propagating, e.g, the loop on the right could be produced by a gluino and a bottom squark that turns into a strange squark in the middle of the loop [17].

Within the Standard Model framework, penguin diagrams are important because they often provide amplitudes of an order of magnitude matching those of tree diagrams, hence setting the scene for interference sensitive to CKM phases.

2.4 $B^0 \to K^0_S \pi^+ \pi^-$ and charmless 3-body B decays

In this section we briefly review the current experimental knowledge on the $B^0 \to K^0_S \pi^+ \pi^-$ decay channel, and then introduce the main theoretical motivations for its study: the determination of $\sin 2\beta_{\text{eff}}$ in a $b \to s\bar{q}q$ mode, and the extraction of a constraint on the Unitarity Triangle angle γ.

2.4.1 Experimental and theoretical status

Over the past few years, the B Factories have thoroughly explored the area of charmless quasi-two body (the two bodies being a stable particle and a resonance, hence the "quasi") and three-body B decays. Early results established that the branching fractions of the main hadronic three-body modes were in the range of 5×10^{-5} to 10^{-6} (see, for instance, [30]). Later studies with larger data samples allowed for Dalitz-plot analyses of the most prominent modes or those with higher reconstruction efficiencies, such as $B^+ \to \pi^+\pi^-\pi^+$ [31], $B^0 \to \rho\pi$ [29], $B^0 \to K^+\pi^-\pi^0$ [32], $B^+ \to K^+K^+K^-$ [33] and $B^+ \to K^+\pi^-\pi^+$ [10]. The latter allowed *Belle* to find the first evidence for direct *CP* violation in a charged mode [25] in the channel $B^+ \to \rho^0 K^+$, in which the *BABAR* result also showed a hint of large direct *CP* violation.

The $B^0 \to \rho\pi$ analysis distinguishes itself by being the first time-dependent Dalitz-plot analysis. More recently, $B^0 \to K^+K^-K^0$ [34] has followed. Although a similar goal is pursued in the $B^0 \to K_S^0\pi^+\pi^-$ channel, the lack of sufficient statistics had prevented it until now, and only quasi-two-body [35] and tag-independent, time-integrated Dalitz-plot studies [36] had been performed. The quasi-two-body approach tries to isolate each resonance and measure its yield independently. Corrections for the interferences with other resonances are needed, severely limiting the accuracy with systematic effects. These interferences are taken into account by the time-integrated, tag-independent Dalitz analysis, which enables us to calculate the phases between resonances and their relative magnitudes. Still, it ignores the time-dependence and its associated asymmetries, and, although direct *CP* asymmetries for neutral modes can be calculated, the extra sensitivity provided by the tagging is lost. The aim of the analysis described in this thesis is to study simultaneously the Dalitz-plot and time dependencies, thereby gaining sensitivity on all fronts. In particular, the Dalitz-plot analysis is a technique especially well-suited for the extraction of phases, and has therefore a large impact on measurements of the CKM phase (see Section 2.4.2 and Section 2.4.3).

The works mentioned above measure a branching fraction of $\mathcal{B}\left(B^0 \to K^0\pi^+\pi^-\right) =$

Table 2.1: Previous measurements concerning B^0 decays to the $K_S^0\pi^+\pi^-$ final state. The branching fractions are in units of 10^{-6}. The first of the errors is statistical and the second systematic. Measurements from the CLEO Collaboration are excluded as they have been clearly superseded by the B Factories. The reference for each measurement is indicated. Note that the small branching fraction of $\omega \to \pi^+\pi^-$ makes its observation in $B^0 \to K_S^0\pi^+\pi^-$ unlikely. The so-called "non-resonant" (NR) branching fraction corresponds to decays that cannot be ascribed to any of the resonant structures in the Dalitz plot.

	BABAR	Belle
$\mathcal{B}(B^0 \to \rho^0(770)K^0)$	$4.9 \pm 0.8 \pm 0.9$ [37]	$6.1 \pm 1.0^{+1.1}_{-1.2}$ [36]
$\mathcal{B}(B^0 \to f_0(980)(\to \pi^+\pi^-)K^0)$	$5.5 \pm 0.7 \pm 0.6$ [35]	$7.6 \pm 1.7^{+0.9}_{-1.3}$ [36]
$\mathcal{B}(B^0 \to \omega K^0)$	$6.2 \pm 1.0 \pm 0.4$ [38]	$4.4^{+0.8}_{-0.7} \pm 0.4$ [39]
$\mathcal{B}(B^0 \to K^{*+}(892)\pi^-)$	$11.0 \pm 1.5 \pm 0.7$ [35]	$8.4 \pm 1.1^{+1.0}_{-0.9}$ [36]
$\mathcal{A}_{CP}(B^0 \to K^{*+}(892)\pi^-)$	$-0.11 \pm 0.14 \pm 0.05$ [35]	–
$\mathcal{B}(B^0 \to K_0^{*+}(1430)\pi^-)$	–	$49.7 \pm 3.8^{+6.8}_{-8.2}$ [36]
$\mathcal{B}(B^0 \to K^0\pi^+\pi^- \quad NR)$	–	$19.9 \pm 2.5^{+1.7}_{-2.0}$ [36]
$B^0 \to K^0\pi^+\pi^-$	$43.0 \pm 2.3 \pm 2.3$ [35]	$47.5 \pm 2.4 \pm 3.7$ [36]

$(44.8 \pm 2.5) \times 10^{-6}$, and identify the resonances $K^*(892)$, $K_0^*(1430)$, $f_0(980)$, $\rho^0(770)$, as well as a state in the $\pi^+\pi^-$ spectrum around $\sim 1.3\,\mathrm{GeV}/c^2$ whose nature is yet to be determined (see Table 2.1). Both the branching fraction and the Dalitz structure agree well with those found in $B^+ \to K^+\pi^-\pi^+$. For the time-dependent parameters, see Table 2.2.

Several QCD factorization [40] and $SU(3)$ flavour symmetry [41] predictions exist for the branching fractions and CP asymmetries of the submodes $B^0 \to \rho^0(770)K_S^0$ and $B^0 \to K^*(892)^+\pi^-$, that contribute to $B^0 \to K_S^0\pi^+\pi^-$. The branching fractions obtained using the latter approach agree well with experiment, whereas QCD factorization reliably calculates the branching fraction of $B^0 \to \rho^0(770)K_S^0$ while importantly underestimating that of the $K^{*+}(892)\pi^-$ submode. More recent calculations tackle

$B^0 \to f_0(980)K_S^0$ [42], but are harder to interpret due to the lack of consensus on the proportion of $f_0(980)$ that decays into two pions. All of the above computations predict rather small direct *CP* asymmetries for the channels mentioned, of the order of a few percent, reaching 7.5% for $\rho^0(770)K_S^0$ in one of the factorization scenarios. Larger asymmetries could be observed in $K_0^{*+}(1430)\pi^-$, since the diagrams that make up its amplitude are formally the same as in $B^0 \to K^+\pi^-$, where $\mathcal{A}_{CP} \simeq -11\%$ has been measured [43].

2.4.2 $\sin 2\beta$ in $b \to s\bar{q}q$ decays and New Physics

The $B \to K\pi\pi$ modes are penguin dominated and therefore, as pointed out in Section 2.3, are good candidates to display sizable deviations from the Standard Model predictions should any new physics enter the process through the loop. Indeed, the prominent role that comparisons between the Standard Model value of sin2β, obtained from the "golden channel" (see Section 2.3), and those measured in the penguin-dominated modes should play in the search for new physics has long been recognised [44]. The latter, decaying as $b \to s\bar{q}q$ with $q = u, d, s$ through a loop diagram, involve the same CKM matrix elements as the golden mode, and should therefore exhibit, up to hadronic effects, the same time-dependent asymmetries. Any significant differences in these, or between the measured direct *CP* asymmetries and their theoretical evaluation, would be a tell-tale sign of unknown physics appearing in the process.

Currently, the discrepancies between the charmonium and charmless time-dependent asymmetries are not statistically significant, but show a trend that has aroused some attention (see Figure 2.6). Most of the charmless modes seem to exhibit lower values of $\sin 2\beta_{\text{eff}}$ than the golden mode. The discrepancy cannot be accounted for by the theoretical estimations of corrections due to hadronic effects, such as final state interactions, which tend to push the theoretical charmless asymmetries in the opposite direction [48, 49].

The final state $K_S^0\pi^+\pi^-$ is a *CP*-eigenstate, thus allowing for a measurement of

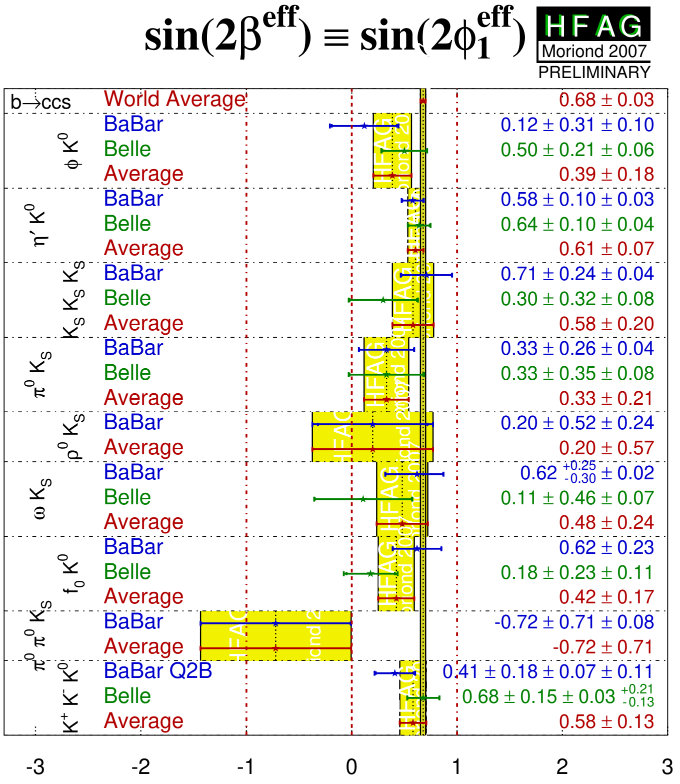


Figure 2.6: $\sin 2\beta_{\rm eff}$ (the notation ϕ_1 is also used to designate the Unitarity Triangle angle β, notably by the $\mathcal{B}elle$ Collaboration) from penguin modes compared to the golden mode, according to the Heavy Flavour Averaging Group [45] after the 2007 Winter conferences. The penguin modes tend to lie on the left of the value for the golden channel, contrary to what the theoretical estimates on hadronic corrections suggests. The statistical significance of the trend is hard to assess, since the corrections are mode-dependent. However, a naïve average is less than 3.0σ away from the charmonium value.

sin2β in the channels $B^0 \to f_0(980)K^0_S$ and $B^0 \to \rho^0(770)K^0_S$. Such measurements have been performed previously on smaller data samples by isolating each resonant mode (quasi-two-body approach). A Dalitz analysis of the larger sample is justified,

Table 2.2: Previous time-dependent measurements of B^0 decays to the $K_S^0\pi^+\pi^-$ final state. The first of the errors is statistical and the second systematic. The reference for each measurement is indicated.

	BABAR	*Belle*
$\mathcal{S}(B^0 \to \rho^0(770)K^0)$	$0.20 \pm 0.52 \pm 0.24$ [37]	–
$\mathcal{C}(B^0 \to \rho^0(770)K^0)$	$0.64 \pm 0.41 \pm 0.20$ [37]	–
$\mathcal{S}(B^0 \to f_0(980)(\to \pi^+\pi^-)K^0)$	$-0.95^{+0.32}_{-0.23} \pm 0.10$ [46]	$0.18 \pm 0.23 \pm 0.11$ [47]
$\mathcal{C}(B^0 \to f_0(980)(\to \pi^+\pi^-)K^0)$	$-0.24 \pm 0.31 \pm 0.15$ [46]	$0.15 \pm 0.15 \pm 0.07$ [47]
$\mathcal{S}(B^0 \to f_0(980)(\to K^+K^-)K^0)$	$-0.32 \pm 0.30 \pm 0.6$ [34]	–
$\mathcal{C}(B^0 \to f_0(980)(\to K^+K^-)K^0)$	$-0.45 \pm 0.28 \pm 0.10$ [34]	–

since it can improve on the quasi-two-body measurements shown in Table 2.2, by properly accounting for interferences between resonances. This makes it unnecessary to apply any harsh cuts or systematic errors that intrinsically limit the precision of the quasi-two-body approach. Moreover, whereas the quasi-two-body analyses are sensitive only to the interference of the state with its oscillated counterpart, thus reporting just $\sin 2\beta$, a Dalitz analysis can exploit the interference of other resonances with the oscillation amplitude, thus determining β itself and removing the four-fold ambiguity resulting from the sine.

2.4.3 Constraints on γ from $B \to K\pi\pi$ modes

The current methods to measure γ rely on the interference between the colour-allowed $B^- \to D^0 K^-$ and the colour-suppressed $B^- \to \overline{D}^0 K^-$ decay modes resulting in direct *CP* violation. They are theoretically very clean, as only tree amplitudes are involved, but their sensitivity to γ is governed by the rather small relative magnitude of the two amplitudes, denoted r_B: $0.05 \lesssim r_B \lesssim 0.3$, depending on the D meson decay channel. As a consequence, γ is the most poorly determined angle of the unitarity triangle, $(83 \pm 19)^\circ$ (c.f $\beta = \left(21.0^{+0.9}_{-1.1}\right)^\circ$) [50]. Therefore, any independent determinations of the angle γ should be exploited in order to reduce the statistical uncertainty.

Although a first proposal on using the charmless three-body decays $B \to K\pi\pi$ to extract the unitarity triangle angle γ via isospin relations was made in 2002 [51], the more recent ideas in [11, 52] are both far more accurate in their estimations of the theoretical uncertainties of their methods, and more convenient experimentally. We will therefore concern ourselves only with the latter two.

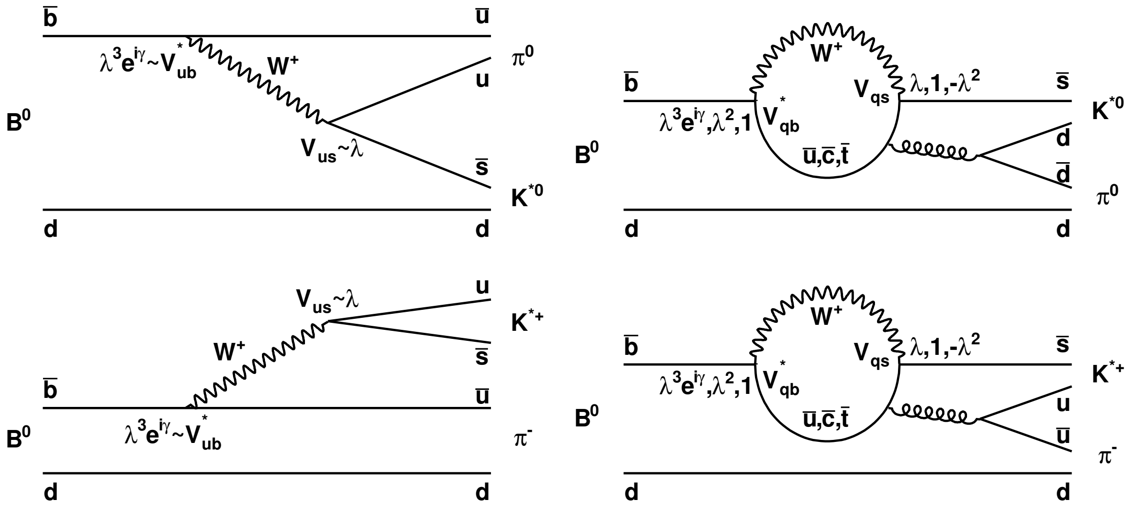


Figure 2.7: Diagrams contributing to the amplitudes for $B^0 \to K^{*0}\pi^0$ (top row) and $B^0 \to K^{*+}\pi^-$ (bottom row), with the tree diagrams on the left, and the penguin diagrams on the right. The tree diagram for $B^0 \to K^{*+}\pi^-$ is dubbed "external emission", and for $B^0 \to K^{*0}\pi^0$, "internal emission". Powers of λ and phases have been written next to the CKM matrix elements to clarify the hierarchical and phase structures.

The paper by Ciuchini, Pierini and Silvestrini [11] recovers the theoretical use of the phase-extraction capabilities of the Dalitz-plot analysis technique, which had been left unexploited since the Lipkin-Nir-Quinn-Synder method to measure α in $B^0 \to \rho\pi \to \pi^+\pi^-\pi^0$ [53]. They start by relating the ratio of the amplitudes for the decays $B^0 \to K^{*+}\pi^-$ and $B^0 \to K^{*0}\pi^0$ and their CP conjugates to γ through isospin, and then cleverly take advantage of the Dalitz plot to determine the phase difference between the two B flavours.

More explicitly, looking at the diagrams (see Figure 2.7) we can see that isospin ensures that the amplitudes with the same topology are approximately equal for the two modes, up to factors of $-1/\sqrt{2}$ and $1/\sqrt{2}$ for the penguin and tree amplitudes in $K^{*0}\pi^0$ [2]:

$$\begin{aligned} A\left(K^{*+}\pi^-\right) &= \tilde{P} + \tilde{E}_e, && (2.21) \\ A\left(K^{*0}\pi^0\right) &= \frac{-1}{\sqrt{2}}\tilde{P} + \frac{1}{\sqrt{2}}\tilde{E}_i && (2.22) \end{aligned}$$

where $\tilde{P}$ stands for the penguin amplitude, and $\tilde{E}_i$ and $\tilde{E}_e$ for the internal and external emission tree amplitudes. If we now factor out their CKM elements and group them accordingly, we get

$$\begin{aligned} A\left(K^{*+}\pi^-\right) &= V_{tb}^* V_{ts} P - V_{ub}^* V_{us}\left(E_e - P^{\mathrm{GIM}}\right), && (2.23) \\ \sqrt{2}A\left(K^{*0}\pi^0\right) &= -V_{tb}^* V_{ts} P - V_{ub}^* V_{us}\left(E_i + P^{\mathrm{GIM}}\right) && (2.24) \end{aligned}$$

where use has been made of the unitarity triangle relation $V_{tb}^* V_{ts} + V_{cb}^* V_{cs} + V_{ub}^* V_{us} = 0$ to separate the penguin amplitude into CKM-favoured (P) and CKM-suppresed (P^{GIM}) contributions[3]. Recalling that the amplitude for the CP-conjugate $\overline{B}^0$ process is obtained simply by complex-conjugating the CP-odd phases (i.e. the CKM factors), we can use the previous isospin relations to cancel out the penguin terms:

$$\begin{aligned} A^0 &= A\left(K^{*+}\pi^-\right) + \sqrt{2}A\left(K^{*0}\pi^0\right) && (2.25) \\ &= -V_{ub}^* V_{us}\left(E_e + E_i\right), && (2.26) \\ \bar{A}^0 &= A\left(K^{*-}\pi^+\right) + \sqrt{2}A\left(\overline{K}^{*0}\pi^0\right) && (2.27) \\ &= -V_{ub} V_{us}^*\left(E_e + E_i\right), && (2.28) \end{aligned}$$

whose ratio is

$$R^0 = \frac{\bar{A}^0}{A^0} = \frac{V_{ub} V_{us}^*}{V_{ub}^* V_{us}} = e^{-i2\gamma}. \qquad (2.29)$$

[2]These factors of $-1/\sqrt{2}$ and $1/\sqrt{2}$ in the $K^{*0}\pi^0$ amplitudes can be regarded as stemming from the quark wavefunction of the π^0, $\left(u\bar{u} - d\bar{d}\right)/\sqrt{2}$

[3]The u- and c- loop amplitudes entering in P^{GIM} nearly cancel each other due to the near equality of the masses of the two virtual quarks when compared to the W mass, in another instance of the celebrated "GIM mechanism" [54].

At this point it must be noted that with quasi-two-body approaches, only the magnitudes of the amplitudes involved in the A^0 quantities can be measured. In contrast, with a Dalitz analysis, their relative magnitudes and phases can be extracted from the 3-body decay $B^0 \to K^+\pi^-\pi^0$, which can proceed through $B^0 \to K^{*+}\pi^- \to K^+\pi^0\pi^-$ and $B^0 \to K^{*0}\pi^0 \to K^+\pi^-\pi^0$. The relative phase between A^0 and $\bar{A}^0$, however, cannot be measured simultaneously, since the charges of the kaon and the pion tag the flavour of the parent meson, thereby preventing any interference. This obstacle can be surmounted with the help of a time-dependent Dalitz analysis of another 3-body mode, $B^0 \to K^0_S\pi^+\pi^-$, from which the phase between the $K^{*+}\pi^-$ and $K^{*-}\pi^+$ resonances can be inferred. Indeed, even though these are flavour-eigenstates, they both decay to the same final state $K^0_S\pi^+\pi^-$, where they can interfere with common resonances. For example, the phase between $K^{*+}\pi^-$ and $K^{*-}\pi^+$ can be resolved through their mutual interference with $\rho^0 K^0_S$.

In the above discussion, the so-called electroweak penguins (obtained by exchanging the gluon in the penguin diagrams by a photon) have been ignored. These result in isospin-breaking effects due, among other things, to the different electric charges of the u and d quarks, and precision measurements must take these into account. By considering the full (weak, strong and electromagnetic) effective Hamiltonian for the transition, the authors of [11] give the following final expression:

$$R^0 = e^{-i(2\gamma+\arg(1+\kappa_{\mathrm{EW}}))} \times (1+\Delta) \tag{2.30}$$

where Δ is theoretically bound ($\lesssim 0.05$) and κ_{EW} is

$$\kappa_{\mathrm{EW}} = \frac{3}{2}\frac{C_+^{\mathrm{EW}}}{C_+}\left(1 + \frac{1-\lambda^2}{\lambda^2(\rho+i\eta)} + \mathcal{O}\left(\lambda^2\right)\right) \tag{2.31}$$

with C_+^{EW} and C_+ being, respectively, the coefficients of the electroweak and normal QCD 4-quark operators in the effective theory.

The second paper, due to Gronau, Pirjol, Soni and Zupan [52], extends the previous work and thoroughly studies the isospin structure of all the amplitudes and effective field theory operators involved in the decay. These also involve a more detailed evaluation of the electroweak penguin amplitudes, thus providing similar formulae for more general cases.

As a closing remark, it is worth emphasizing that the only as yet undetermined experimental input to the above method is the Dalitz analysis of $B^0 \to K_S^0\pi^+\pi^-$, since data from $B^0 \to K^+\pi^-\pi^0$ are already available [32]. Further isospin relations permit the use of measurements from all the $B \to K\pi\pi$ modes, increasing the statistical reach.

2.5 Three-body decays

2.5.1 Introduction

The aim of the present analysis is to study the structures arising in the three-body decay $B^0 \to K_S^0\pi^+\pi^-$. In this section we explore some of the consequences of the kinematics of the decay and discuss the parametrization employed. It is important to spell out the latter in detail, as a single sign or a different choice in the angular convention will change the numerical values obtained for the phases in a Dalitz analysis.

The general features of the decay of a particle can be discussed based on elementary concepts of Quantum Mechanics. Formulae for the transitions of an initial state into a final state can be found by application of the standard time-independent perturbation theory and are explicitly presented in most textbooks, e.g. [27]:

$$c_{fi} = \frac{\langle\psi_f|V_{\rm int}|\psi_i\rangle}{E_i - E_f} \tag{2.32}$$

Eq. (2.32) gives the amplitude c_{fi} to find a system, whose initial state is i, in a final state f when an interaction potential $V_{\rm int}$ connecting them is introduced, where ψ_i and ψ_f describe the initial and final states in the absence of the interaction and E_i and E_f are their energies. The transition can happen directly, $\langle\psi_f|V_{\rm int}|\psi_i\rangle \equiv V_{fi} \neq 0$, or it may involve intermediate, "virtual" states j, also called resonances, in which case the transition amplitude can be approximated by:

$$c_{fi} = \sum_{j\neq i,f} \frac{V_{fj}V_{ji}}{(E_f - E_j)(E_j - E_i)} - \frac{V_{fi}V_{ff}}{(E_f - E_i)^2}. \tag{2.33}$$

Both equations display a similar form, involving the *vertex factors* V_{jk} and the *propagators* $(E_j - E_k)^{-1}$. The former represents the "strength" with which the interaction connects the two states, while the latter is related to the overlap of the two states:

$$\langle E'|E\rangle = \int dt \langle E'|t\rangle\langle t|E\rangle = \int dt e^{iE't} e^{-iEt} \propto \frac{1}{E' - E}. \tag{2.34}$$

If the final state is degenerate, the probability to observe the transition has to be summed over all the states sharing the same quantum numbers:

$$c_{fi} = \int c_{fi}(E_f)\,\rho(E_f)\,dE_f \tag{2.35}$$

where $\rho(E_f)$ is the *density of final states* or *phase space factor*.

In the following subsections we discuss in detail the peculiarities of the densities of states for three-body decays (Section 2.5.2), the vertex factors (Section 2.5.3) and the propagators (Sections 2.5.4 and 2.5.6).

2.5.2 Kinematics of three-body decays

In the decay of the pseudo-scalar B meson, with mass m_B, into three more pseudo-scalar particles with masses $m_{1,2,3}$ and 4-momenta $p_{1,2,3}$, there are several kinematical constraints that reduce to two the number of degrees of freedom needed to specify the final state. Defining the invariant mass squared of a pair of particles as $m_{ij}^2 \equiv (p_i + p_j)^2$ we get

$$m_{12}^2 + m_{23}^2 + m_{31}^2 = m_B^2 + m_1^2 + m_2^2 + m_3^2 \tag{2.36}$$

so that one of the m_{ij}^2 is linearly dependent on the other two. Furthermore, in the B rest frame,

$$\begin{aligned} m_{ij}^2 &= (p_B - p_k)^2 = m_B^2 + m_k^2 - 2m_B E_k & (2.37)\\ &= (p_i + p_j)^2 = m_i^2 + m_j^2 + 2E_iE_j - 2|\vec{p}_i||\vec{p}_j|\cos\theta_{ij}. & (2.38)\end{aligned}$$

Eq. (2.37) shows that the energies depend solely on the invariant masses of the pairs of particles, whereas Eq. (2.38) indicates that the angles between the momenta of the final state particles are known once their energies are determined. Together, Eqs (2.37)-(2.38) imply that the knowledge of two quantities, customarily chosen from among the m_{ij}^2, are sufficient to specify the state of the system, up to its overall orientation.

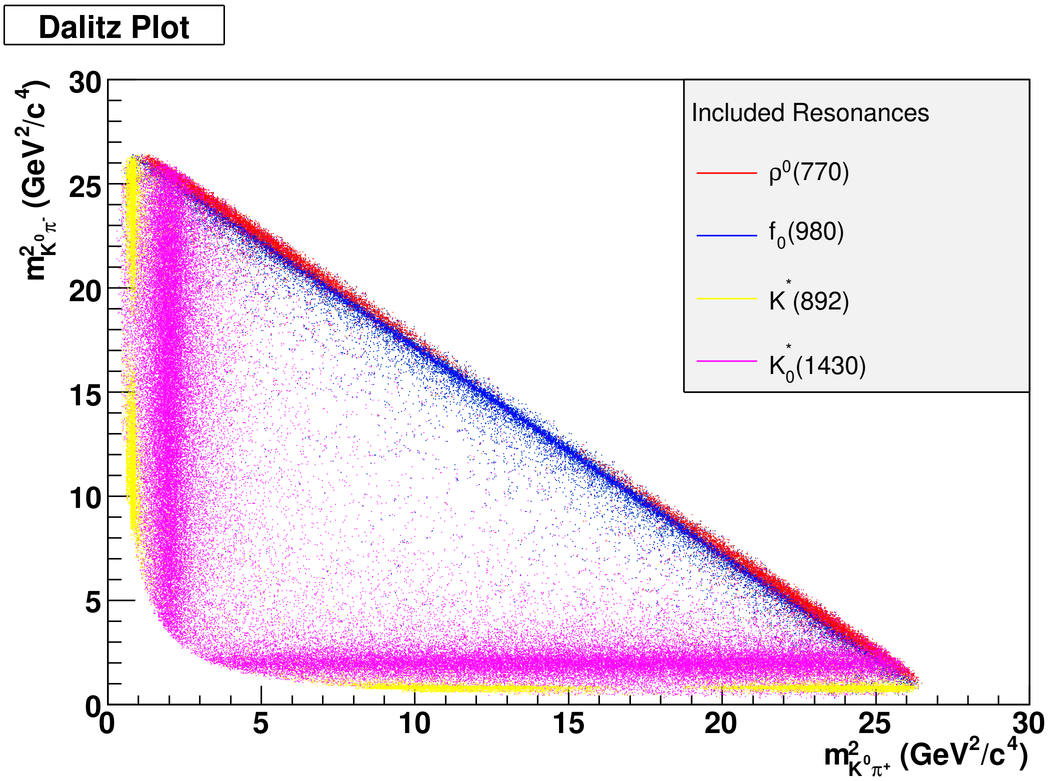


Figure 2.8: Toy Monte Carlo simulation of $B^0 \to K^0_S\pi^+\pi^-$. The resonances $f_0(980)$, $\rho^0(770)$, $K^*(892)$ and $K_0^*(1430)$ have been included, approximately in the proportions found by $\mathcal{B}elle$ [55], although without interference. Note the two wide strips parallel to the axes, and the fine band close to the diagonal edge - they correspond to the scalars $K_0^*(1430)$ and $f_0(980)$, respectively. In contrast, the two separate, elongated blobs on each axis and on the diagonal belong to the $K^*(892)$ and $\rho^0(770)$, which, being vector resonances, have their corresponding bands unpopulated in the centre. Observe as well that the sensitivity to interference between different states is greater for events on the corners, as that is where the resonance bands overlap.

A Dalitz plot [56] is produced when a two-dimensional scatter plot is made in two of the m_{ij}^2 variables, say $m_{13}^2 \equiv x$ and $m_{23}^2 \equiv y$ (see Figure 2.8). The kinematical boundaries for such a plot can be read from Eq. (2.38): for a given value of m_{jk}^2, the maximum of m_{ij}^2 is attained when the particles i and j are flying back-to-back, and the minimum when they are at rest in the ij centre-of-mass system, i.e. their directions are parallel.

Similarly, for events close to the edges of the Dalitz plot, one of the m_{ij}^2 takes a small value while the other two have rather large values (Eq. (2.36)), whereas in the centre the invariant masses of the three pairs of particles take approximately the same values. This implies that in the latter case, the directions of the three particles are distributed quite isotropically, and that they carry similar energies, whereas in the former case, one of the particles in the final state is back to back to the other two, which move in parallel, giving the event a strong directionality. It is also worth noting that, for an event lying near the corners of the Dalitz plot, one of the particles is slow, as can be seen from Eq. (2.37).

We are now ready to discuss the phase space factor. The summation should be done over all momenta in the final state, but application of the kinematical constraints noted before should enable us to write it as a function of only two of the energies or squared invariant masses:

$$\rho\left(m_{13}^2, m_{23}^2\right) dm_{13}^2 dm_{23}^2 = \frac{d^3p_1 d^3p_2 d^3p_3}{E_1 E_2 E_3} \delta\left(p_1 + p_2 + p_3 - p_B\right) \tag{2.39}$$

where the energies in the denominator of the RHS have been introduced to ensure Lorentz invariance. Integration over p_3 yields

$$\frac{d^3p_1 d^3p_2}{E_1 E_2 E_3} \delta\left(E_1 + E_2 + E_3 - m_B\right) = \frac{p_1^2 dp_1 d\Omega_1 p_2^2 dp_2 d\Omega_{1-2}}{E_1 E_2 E_3} \delta\left(E_1 + E_2 + E_3 - m_B\right) \tag{2.40}$$

where Ω_1 and Ω_{1-2} are the solid angles for the direction of $\vec{p}_1$, and for the direction of $\vec{p}_2$ with respect to $\vec{p}_1$. Since the B is a scalar, the angles should be integrated over, giving

$$\frac{p_1^2 dp_1\left(4\pi\right) p_2^2 dp_2 \left(2\pi d\cos\theta_{1-2}\right)}{E_1 E_2 E_3} \delta\left(\Sigma E_i - m_B\right) = 8\pi^2 \frac{dE_1 dE_2 p_1 p_2 d\cos\theta_{1-2}}{E_3} \delta\left(\Sigma E_i - m_B\right). \tag{2.41}$$

Noting

$$E_3^2 = p_1^2 + p_2^2 + 2p_1p_2\cos\theta_{1-2} + m_3^2 \Longrightarrow E_3dE_3 = p_1p_2d\cos\theta_{1-2}, \tag{2.42}$$

substituting it and integrating the δ-function gives

$$\int 8\pi^2 dE_1dE_2dE_3\delta\left(E_1 + E_2 + E_3 - m_B\right) = 8\pi^2 dE_1dE_2 = \frac{4\pi^2}{m_B^2}dm_{23}^2dm_{13}^2. \tag{2.43}$$

Therefore, the density of final states is constant when expressed in terms of the m_{ij}^2 variables. In other words, the decay rate (the probability of decay per unit time) has the form (c.f. Eq. (2.35))

$$d\Gamma \propto |\mathcal{M}|^2 dm_{13}^2 dm_{23}^2 \tag{2.44}$$

where $\mathcal{M}$ encodes all the dynamical information about the decay, containing the vertex factors and the propagators.

We observe that, according to Eq. (2.44), a constant $|\mathcal{M}|^2$ results in a uniform distribution over the Dalitz plot, and that any departure is due to dynamical effects, i.e. a non-trivial $|\mathcal{M}|^2$.

2.5.3 The Isobar Model

The Isobar Model [57, 58] approximates $\mathcal{M}$ as a sum of terms with individual couplings and propagators, each representing a resonance in one pair of particles:

$$\mathcal{M}\left(m_{13}^2, m_{23}^2\right) = \sum_{j=1}^{N} c_j F_j\left(m_{13}^2, m_{23}^2\right) \tag{2.45}$$

where N is the number of intermediate states considered, c_j are the complex amplitudes describing the coupling of the B meson to the particular resonant final state (i.e. the vertex factors) and $F_j\left(m_{13}^2, m_{23}^2\right)$ are the propagators. These are a product of several terms:

$$F_j\left(m_{13}^2, m_{23}^2\right) = R_j\ B_L^B\ B_L^{\text{res}}\ Z_j^L, \tag{2.46}$$

with the different terms being: first, the mass-dependent part of the propagator, second and third, factors that account for the difficulty of slow decay products to conserve the angular momentum due to the spin of the resonance and last, the term that describes the angular distribution. The conventions adopted for these are described in detail in the following sections, a good reference being [59]. World averages [22] are used for the parameters characterizing each resonant state (e.g. mass, width).

A *Dalitz* or *amplitude analysis* aims to extract the complex couplings c_j from the data, when a given model for the resonant structure has been proposed. Note that, since the decay rate depends on $|\mathcal{M}|^2$ (2.44), Eq. (2.45) implies that bilinear terms in $F_j\left(m_{13}^2, m_{23}^2\right)$ will appear in the model of the distribution over the Dalitz plot. These terms, $\propto F_j^* F_k$, represent, and are sensitive to, the interference between two resonances j and k, thus allowing for the relative phase between c_j and c_k to be determined.

In neutral decays to *CP* eigenstates, one expects the same resonances to be present in the B^0 and the $\overline{B}^0$ decays in the same amounts, up to direct *CP* violating effects. Therefore a parameterization of the complex couplings c_j and $\bar{c}_j$ appearing in the B^0 and $\overline{B}^0$ amplitudes ($\mathcal{A}$ and $\bar{\mathcal{A}}$, respectively) that reflects that fact is preferred instead of, for example, using separate magnitudes and phases for each flavour. This analysis adopts

$$c_j = (x_j + \Delta x_j) + i\,(y_j + \Delta y_j)\ , \tag{2.47}$$

$$\bar{c}_j = (x_j - \Delta x_j) + i\,(y_j - \Delta y_j)\ , \tag{2.48}$$

where the parameters free to vary in the fit are the x, y, Δx and Δy, although the alternative

$$c_j = a_j e^{i(\delta_j + \phi_j)} \left(1 + \frac{b_j}{a_j}\right)\ , \tag{2.49}$$

$$\bar{c}_j = a_j e^{i(\delta_j - \phi_j)} \left(1 - \frac{b_j}{a_j}\right)\ , \tag{2.50}$$

proposed by CLEO [60], is also found useful. Note that the latter breaks the phase of each component into their *CP*-even (δ) and *CP*-odd (ϕ) contributions.

A Dalitz analysis extracts all non-trivial information from the data – no physically meaningful aspect of the decay is left unmodelled. Therefore, all the observables mentioned in earlier sections must be expressible in terms of the c_j[4]:

$$\begin{aligned} S_j &= 2\,\mathrm{Im}\left[\bar{c}_j c_j^* e^{-i\phi_{\mathrm{mix}}}\right] / \left(|c_j|^2 + |\bar{c}_j|^2\right) , && (2.51) \\ \mathcal{A}_{CPj} &= \left(|\bar{c}_j|^2 - |c_j|^2\right) / \left(|c_j|^2 + |\bar{c}_j|^2\right) . && (2.52) \end{aligned}$$

The relative weight of a given resonance in the decay is usually quoted in terms of the *fit fraction*:

$$FF_j = \frac{\int\int_{DP} \left|c_j F_j\left(m_{13}^2, m_{23}^2\right)\right|^2 dm_{13}^2 dm_{23}^2}{\int\int_{DP} \left|\sum_j c_j F_j\left(m_{13}^2, m_{23}^2\right)\right|^2 dm_{13}^2 dm_{23}^2} , \qquad (2.53)$$

where the DP integration domain means the integral must be calculated over the whole phase space.

Finally, let us acknowledge that the approximation in Eq. (2.45) neglects rescattering of the final state particles and is known to lead to unitarity violation whenever the overlapping of two resonances is sizable. Alternatives, such as the so-called K-matrix [61], exist but their complexity outweighs their merits for use in the present analysis. Apart from the validity of the approximations, the main source of systematic uncertainties in the model is the term R_j from Eq. (2.46), whose precise functional form is not well known for some components, notably the higher K^* resonances (see the LASS subsivision in Section 2.5.4 and Section 6.4).

2.5.4 Mass term description

Breit-Wigner

The most common parametrization of the mass term is the Breit-Wigner formula, that arises from the overlap between a state of energy E and a resonant state with mass m_R and decay width Γ_R, and therefore gives the amplitude for a system in the first

[4] ϕ_{mix} for instance, being a relative phase, can be absorbed in the c_j and extracted from them. However, in this particular case we choose to keep it as an explicitly distinct, global parameter (see Section 6.3).

state to be in the second state as well:

$$\langle E|R\rangle = \int dte^{iEt}e^{-t(im_R+\Gamma_R/2)} \propto \frac{1}{(E-m_R)-i\Gamma/2}. \tag{2.54}$$

Relativistic Breit-Wigner

An obvious improvement is making Eq. (2.54) relativistic [62]:

$$R_j(m) = \frac{1}{(m_R^2-m^2)-im_R\Gamma(m)}, \tag{2.55}$$

in which the variation of the width with the energy is taken into account via

$$\Gamma(m) = \Gamma_R\left(\frac{q}{q_0}\right)^{2L+1}\left(\frac{m_R}{m}\right)B_L^2(|\vec{q}|r). \tag{2.56}$$

where L is the angular momentum quantum number of the resonance, $\vec{q}$ is the momentum of one of the daughters in the resonance rest frame and $q_0 = q\,(m = m_R)$.

Flatté

The conditions under which a Breit-Wigner lineshape is a good description are rather restrictive. A particular phenomenon not accounted for, and that is relevant to the present analysis, is the change in shape resulting from the opening of a threshold. The best example is the $f_0(980)$ state, whose main decay modes are $f_0(980) \to \pi\bar{\pi}$ and $f_0(980) \to K\bar{K}$. The mass of the $K\bar{K}$ system at rest is $\sim 990\,\mathrm{MeV}/c^2$, within a fraction of the width of the mass of the resonance. This means that events decaying through the $f_0(980)$ state with masses below that threshold can proceed essentially only to the two-pion final state, whereas events with masses above the threshold also have the possibility of producing the two-kaon state. In short, a very asymmetric shape should be expected.

The Flatté lineshape [63] correctly handles the problem:

$$R_j(m) = \frac{1}{(m_R^2-m^2)-im_R(\Gamma_{\pi\pi}(m)+\Gamma_{KK}(m))} \tag{2.57}$$

with

$$\Gamma_{\pi\pi}(m) = g_\pi \left(\frac{1}{3}\sqrt{1 - 4m_{\pi^0}^2/m^2} + \frac{2}{3}\sqrt{1 - 4m_{\pi^\pm}^2/m^2} \right), \tag{2.58}$$

$$\Gamma_{KK}(m) = g_K \left(\frac{1}{2}\sqrt{1 - 4m_{K^\pm}^2/m^2} + \frac{1}{2}\sqrt{1 - 4m_{K^0}^2/m^2} \right), \tag{2.59}$$

with the couplings to the $\pi\bar{\pi}$ and $K\bar{K}$ channels needing to be determined experimentally. As for the rest of the properties of the $f_0(980)$ resonance, there is no consensus on their values. We take those measured by BES [64]:

$$g_\pi = (0.165 \pm 0.010 \pm 0.015)\ \mathrm{GeV}/c^2,\ g_K/g_\pi = 4.21 \pm 0.25 \pm 0.21. \tag{2.60}$$

LASS

Another poorly understood area of the spectrum is that of the higher K^* resonances. In particular, the $K_S^0\pi$ mass range in which the state $K_0^*(1430)$ appears seems to be correctly described only when a non-resonant effective range component is included and allowed to interfere with the resonance. Following the LASS collaboration [65], we parameterize it as:

$$R_j(m) = \frac{m}{q\cot\delta_B - iq} + e^{2i\delta_B}\frac{m_R\Gamma_R\frac{m_R}{q_0}}{(m_R^2 - m^2) - im_R\Gamma_R\frac{q}{m}\frac{m_R}{q_0}}, \tag{2.61}$$

where m_R and Γ_R are the mass and the width of the $K_0^*(1430)$ resonance and $\cot\delta_B$ is given by

$$\cot\delta_B = \frac{1}{aq} + \frac{1}{2}rq. \tag{2.62}$$

The parameters a and r are the scattering length and the effective range, respectively, and have been measured by LASS to be

$$a = (2.07 \pm 0.10)\,(\mathrm{GeV}/c)^{-1},\ r = (3.32 \pm 0.34)\,(\mathrm{GeV}/c)^{-1}. \tag{2.63}$$

The non-resonant part of the amplitude is cut off at $1.8\,\mathrm{GeV}/c^2$.

2.5.5 Barrier factors

The Blatt-Weisskopf barrier or penetration factors [66] (the B_L terms in Eq. (2.46)) are motivated by the consideration of the Schrödinger equation in spherical polar coordinates. An effective potential, dubbed "centrifugal barrier", arises from the vanishing of the wavefunction at the origin when the orbital angular momentum is non-zero, both in the decay of the B meson to a $J \neq 0$ resonance and in the subsequent decay of the resonance to two pseudoscalar particles. Physically, it means that particles emitted very close to the centre need too large momenta to account for all the angular momentum of the resonance. A correction is thus needed to the usual Breit-Wigner lineshapes, that can be derived from the transmission coefficients for the centrifugal potential.

Empirically, an effective radius r is needed to describe the shape of the barrier correctly. We take the value of $r = 4.0\,\mathrm{GeV}^{-1} \approx 0.8$ fm. The factors are

$$B_{L=0}(z) = 1, \tag{2.64}$$

$$B_{L=1}(z) = \sqrt{\frac{1+z_0^2}{1+z^2}}, \tag{2.65}$$

$$B_{L=2}(z) = \sqrt{\frac{z_0^4+3z_0^2+9}{z^4+3z^2+9}}, \tag{2.66}$$

where $z = (|\vec{q}|r)^2$ and z_0 is the value that z takes when $\vec{q}$ is evaluated at the resonance pole mass.

2.5.6 Angular dependence and helicity angles

It can be argued that the distribution of events decaying through a scalar resonance will uniformly populate the band of mass associated to the intermediate state, since the lack of spin means there is no preferred direction for the daughters of the resonance. For a vector intermediate state, however, a privileged direction exists, and their distribution is not obvious. It can be calculated though, by evaluating the propagator for $B \to Rc \to abc$, where R is the resonance. The vectorial nature of the intermediate state is accounted for by the sum over its helicity states λ [59] (c.f. Eq. (2.33)):

$$\sum_{\lambda} \langle ab|R_{\lambda}\rangle\langle cR_{\lambda}|B\rangle. \tag{2.67}$$

The first factor represents the probability of finding the decay daughters a and b in a given state of relative motion:

$$\langle ab|R\rangle \sim (p_a - p_b)_{\nu} \ . \tag{2.68}$$

The second factor can be regarded as the probability of B turning into c by emitting a vector particle R. Since the emission of hard particles (large momentum) is suppressed, states with the momenta of c and B as parallel as possible are favoured:

$$\langle cR|B\rangle \sim (p_B + p_c)_{\mu} \ . \tag{2.69}$$

Finally, using standard techniques documented in any Quantum Field Theory textbook [12], the sum over the helicity states can be performed, giving

$$\begin{aligned} Z^1 &= (p_B + p_c)_{\mu} \left(-g^{\mu\nu} + \frac{p_R^{\mu} p_R^{\nu}}{m_{ab}^2} \right) (p_a - p_b)_{\nu} && (2.70) \\ &= \left(m_{bc}^2 - m_{ac}^2 \right) + \frac{\left(m_B^2 - m_c^2\right)\left(m_a^2 - m_b^2\right)}{m_{ab}^2} && (2.71) \\ &= -2\vec{p}\cdot\vec{q} && (2.72) \\ &= -2|\vec{p}||\vec{q}|\cos\theta_{ac} && (2.73) \end{aligned}$$

where $\vec{p}$ and $\vec{q}$ are, respectively, the momenta of c and a in the resonance rest frame. The angle θ is the *helicity* angle of the resonance. For completeness, we also quote the expression for tensor resonances:

$$Z^2 = \frac{4}{3}\left[3(\vec{p}\cdot\vec{q})^2 - (|\vec{p}||\vec{q}|)^2\right]. \tag{2.74}$$

The formulae for Z^j adopted here are collectively known as *Zemach tensors*.

It is important to note that there is a convention-dependent sign in the Zemach tensor for the vector case (Eq. (2.73)). In a resonance $R \to ab$, the helicity angle can be defined as the angle between a and c, or between b and c. For a given event, changing the choice results in a sign flip. In a Dalitz analysis, that entails shifts of π radians in the fitted phases, so it is essential to state the conventions clearly.

In the channel $B^0 \to K_S^0\pi^+\pi^-$, resonances can be found in the pairs of particles $K_S^0\pi^+$ and $\pi^+\pi^-$ for B^0 events, and in $K_S^0\pi^-$ and $\pi^-\pi^+$ for $\overline{B}^0$. In other words, if the variables labelling the axes of the Dalitz plot are chosen to be $m^2_{K_S^0\pi^+}$ and $m^2_{K_S^0\pi^-}$, the plane will be symmetric with respect to the diagonal in the absence of direct *CP* violation. In order to respect this symmetry in our helicity conventions, the particles $K_S^0\pi^-$ are chosen to define the helicity angle for all intermediate states in the B^0 decay, and $K_S^0\pi^+ = CP\,(K_S^0\pi^-)$ are used for the $\overline{B}^0$. Note that it implies a different helicity angle for neutral resonances like $\rho^0(770)$ or $f_0(980)$ depending on whether the parent meson is a particle or an antiparticle. If the same pair were used in both cases, the constructive interference between the $f_0(980)$ (scalar, no dependence on helicity) and the $\rho^0(770)$ (vector, amplitude proportional to $\cos\theta$) would happen on the same side of the Dalitz plot for both flavours of the parent, thus spoiling the symmetry (Figure 2.9). The value of the angle can be worked out at all points in the Dalitz plot with the help of

$$\cos\theta = \frac{m^2_{bc,\max} + m^2_{bc,\min} - 2m^2_{bc}}{m^2_{bc,\max} - m^2_{bc,\min}} \tag{2.75}$$

which gives the helicity for a resonance in m_{ac} in terms of m_{bc} and allows us to confirm the symmetry of our conventions under *CP* (see Figure 2.10).

2.5.7 Square Dalitz-plot

B decays proceed mostly through low mass resonances, such as $\rho^0(770)$, $f_0(980)$, $K^*(892)$ and $K_0^*(1430)$. That implies that the most populated areas of the Dalitz plot are close to the edges, where the resonances recoil against energetic bachelor particles. Furthermore, the combinatoric nature of background events (see Section 4.5) means that their density also peaks around the edges. Clearly, the binning of the histograms used to characterize the latter will be problematic, as fine binning is needed around the edges, and coarse binning around the centre. Instead of using variable binning, we introduce another set of variables to parameterize the final state phase

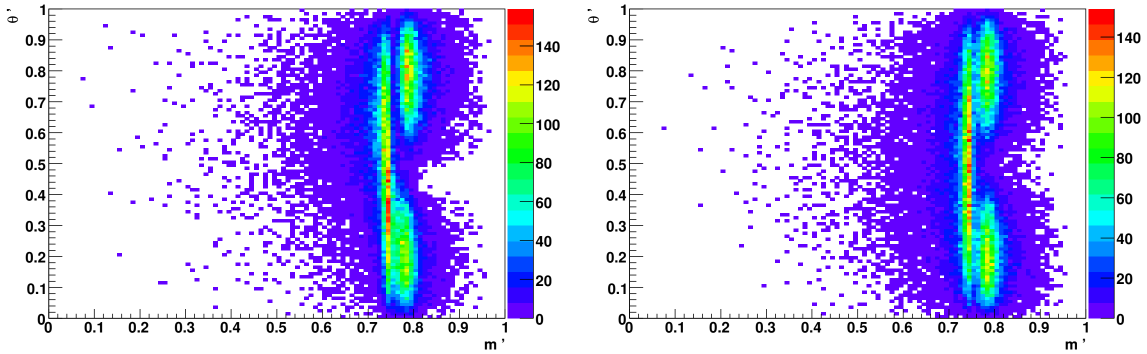


Figure 2.9: Toy Monte Carlo simulation demonstrating the asymmetry produced by the wrong convention in the calculation of the helicity angles of neutral resonances, a $f_0(980)$ and a $\rho^0(770)$ in this case. On the left, the same pair of particles has been used to calculate the helicity for resonances originating in neutral B mesons of both flavours. On the right, a CP symmetric convention has been used, resulting in a symmetric Dalitz plot distribution. Both Dalitz plots are shown in terms of the m' and θ' variables (Section 2.5.7), that expand the areas of interest, in this case the diagonal of the conventional Dalitz plot, while contracting the generally unpopulated centre of the figure.

space. All input histograms will be expressed in terms of these variables:

$$\begin{aligned} m' &\equiv \frac{1}{\pi}\arccos\left(2\frac{m_{\pi^+\pi^-} - m_{\pi^+\pi^-,\mathrm{min}}}{m_{\pi^+\pi^-,\mathrm{max}} - m_{\pi^+\pi^-,\mathrm{min}}} - 1\right), \qquad (2.76) \\ \theta' &\equiv \frac{1}{\pi}\theta_{\pi^+\pi^-}, \end{aligned}$$

where $m_{\pi^+\pi^-}$ is the invariant mass of the two pions and $\theta_{\pi^+\pi^-}$ is the angle between the π^+ and the K^0_S in the $\pi^+\pi^-$ rest frame. Their effect is a magnification of the areas of interest, as can be seen in Figures 2.9 and 2.11.

Transforming histogram bin contents from one set of variables to another involves the calculation of the jacobian:

$$dm^2_{K^0_S\pi^+}\, dm^2_{K^0_S\pi^-} \longrightarrow |\det J|\, dm'\, d\theta' , \qquad (2.77)$$

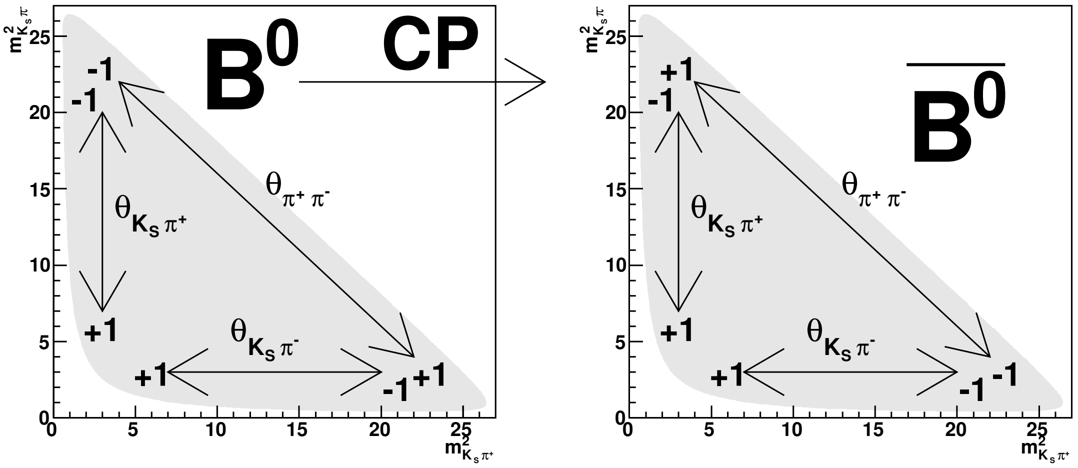


Figure 2.10: The helicity angles for resonances appearing on a given pair of particles are drawn, with the values they take at each point of the Dalitz plot. On the left, the values they take for B^0 decays, and on the right, for $\overline{B}^0$. Note that application of the CP symmetry is equivalent to folding over the lower right corner of the $\overline{B}^0$ plot to the top left corner of the B^0 plot, and that the signs of the helicities on the two corners are consistent.

where

$$|\det J| = 4\,|\mathbf{p}^*_{\pi^+}||\mathbf{p}^*_{K^0_S}|\,m_{\pi^+\pi^-}\cdot\frac{\partial m_{\pi^+\pi^-}}{\partial m'}\cdot\frac{\partial\cos\theta_{\pi^+\pi^-}}{\partial\theta'}\,,$$

$$\frac{\partial m_{\pi^+\pi^-}}{\partial m'} = -\frac{\pi}{2}\sin(\pi m')\left(m^{min}_{\pi^+\pi^-} - m^{max}_{\pi^+\pi^-}\right)\,, \tag{2.78}$$

$$\frac{\partial\cos\theta_{\pi^+\pi^-}}{\partial\theta'} = -\pi\sin(\theta'\pi)\,, \tag{2.79}$$

and all frame-dependent quantities are evaluated in the $\pi^+\pi^-$ rest frame.

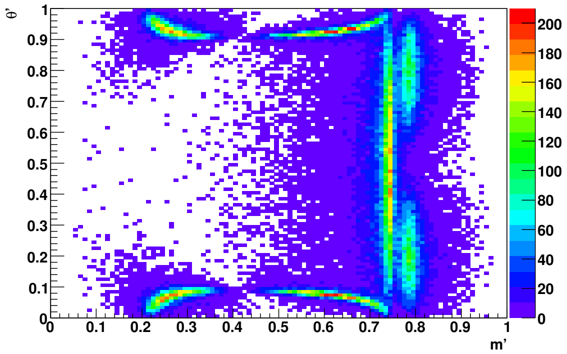


Figure 2.11: Example of a square Dalitz plot clearly showing the shape and location of typical resonances. The vertical continuous and discontinuous bands are the $f_0(980)$ and $\rho^0(770)$, respectively, whereas the discontinuous bands on top and bottom correspond to the K^{*+} and K^{*-} resonances.

Chapter 3

BaBar and PEP-II

3.1 An asymmetric e^+e^- collider as a B factory

In Chapter 1 we briefly sketched the physics-driven design requirements of a B Factory: since b quarks are produced in particle-antiparticle pairs, one of the B mesons produced can be used to determine the flavour of the other one, thus enabling the measurement of asymmetries in the decay of a B^0 and its $\overline{B}^0$.

The ideal conditions for such measurements are met by an e^+e^- collider with the centre-of-mass energy tuned to the $\Upsilon(4S)$ resonance, which decays almost uniquely to $B^0_d\bar{B}^0_d$ and $B^+_u B^-_u$ pairs with equal probabilities. The advantages over a hadronic collider are substantial, a number of factors contributing to the much cleaner environment and hence to better event reconstruction: the low multiplicity of the events (with an average of 11 charged tracks per event), the relatively large signal-to-background ratio ($\sigma_{B\overline{B}}/\sigma_{\text{total hadronic}} \simeq 0.28$), the possibility of reconstructing photons and π^0 and a physics rate low enough ($\sim$ 10 Hz) for the detector not be overwhelmed by data during its dead time [28]. Furthermore, the precise knowledge of the kinematical state of the e^+e^- system allows for a complete reconstruction of the event and naturally provides us with background discriminating variables (Section 4.5). In addition, running at the $\Upsilon(4S)$ resonance implies that the neutral B mesons are produced in a quantum mechanical state of entanglement, which greatly helps in the tagging process by forcing the two particles to remain a particle-antiparticle pair for as long as both exist. In a hadronic environment, such a statement is no longer true and the tagging efficiencies (Section 4.1) fall dramatically for two reasons: first, the higher track multiplicity is an obstacle to finding the particle whose charge uniquely tags the flavour of B_{tag}; second, determining the flavour of B_{rec} requires establishing the flavour of B_{tag}, evolving it back to the production point, and then forward to the B_{rec} decay vertex, thus depending strongly on the accuracy in the inference of the production point.

These properties, together with a large data sample consisting of $\sim 10^8$ $B\overline{B}$ pairs, are the necessary ingredients for *BABAR*'s main physics goals: precisely measuring the *CP* violating time-dependent asymmetries and constraining the CKM matrix elements. Secondary physics interests are rare B decays, such as the subject of this thesis, charm and τ physics, and QCD and two-photon physics. Other quantum electrodynamic processes, such as muon pair production, are mostly filtered due to their large cross-sections, some of them being used for calibration and luminosity measurement purposes. Specifically, the integrated luminosity is calculated to great accuracy by examining the accumulated samples of $e^+e^- \to e^+e^-(\gamma)$, $e^+e^- \to \mu^+\mu^-(\gamma)$ and $e^+e^- \to \gamma\gamma$, processes for which the cross-sections are extremely well understood thanks to QED.

The constraints posed to the accelerator configuration by the difficulties of time-dependent analyses and the high luminosities needed to achieve the desired sample size are discussed in the next section.

3.2 PEP-II and the B Factory

For the time-dependent *CP* violating asymmetries arising in neutral B mesons to be measured, an asymmetric collider is required. The boost of the B mesons in the laboratory frame allows the distance between the decay vertices of the two mesons to be measured, from which the time between the decays of the two particles can be inferred.

The asymmetry is achieved by injecting into the PEP-II storage rings 9.0 GeV e^- and 3.1 GeV e^+ beams which, upon collision, result in a boost of $\beta\gamma = 0.56$ along the e^- beam direction in the laboratory frame for the centre-of-mass of the particles produced. The centre-of-mass energy is tuned to the $\Upsilon(4S)$ mass, 10.58 GeV, for 90% of the running time, the remaining 10% being set 40 MeV below the resonance peak. The first sample, known as the *on-peak sample*, contains the $B\overline{B}$ events, whereas the second one, called the *off-peak sample*, is recorded for background characterization purposes. The light quark processes $e^+e^- \to q\overline{q}$, $q = u, d, s, c$ which constitute the most prominent background to the $B\overline{B}$ events are the only hadronic reactions allowed below the $\Upsilon(4S)$ threshold.

The injection is carried out using the two mile long Stanford Linear Accelerator, which diverts a fraction of the accelerated electrons to produce positrons in collision with a high-Z stationary target. These are then directed back to the linear accelerator to be brought to their nominal energy before entering the storage rings. Once there, the electrons and the positrons, which circulate in bunches along physically separated rings, are collided in one Interaction Region, in which the *BABAR* detector is located.

3.2.1 The Interaction Region

The Interaction Region is heavily instrumented with magnets that focus the beams before the collision, directs them so that there is no crossing angle between them, and finally separates them before a given bunch of particles collides with a second bunch from the other beam (see Figure 3.1). The quadrupole magnets labelled QD and QF, situated outside the *BABAR* detector, focus the high and low energy beams. The dipoles labelled B1 are responsible for bringing the beams together and separating them immediately afterwards. This is the reason why they need to be close to the interaction point; in fact, within the detector volume.

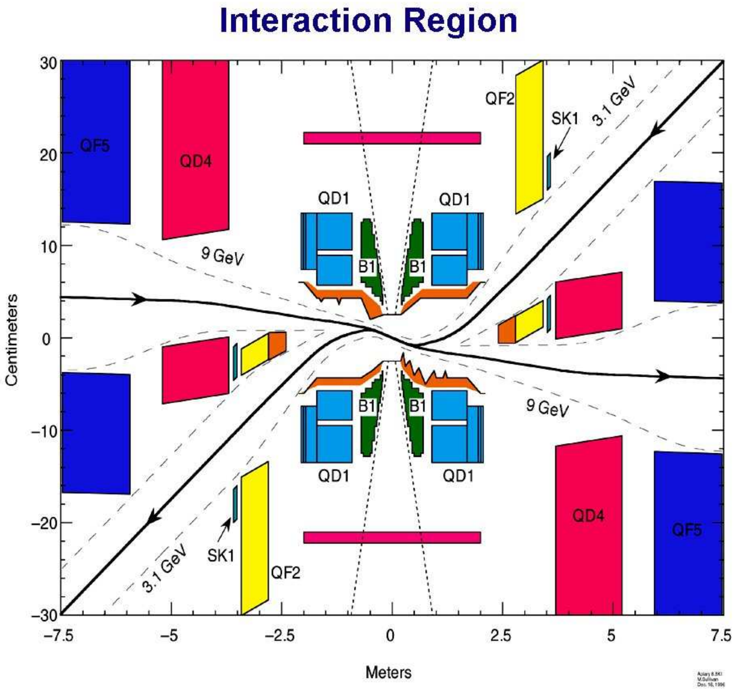


Figure 3.1: Schematic view of the Interaction Region. Observe the strong bending needed to make the two beams collide head-on, that results in the most prominent machine background: synchroton radiation.

3.2.2 Machine backgrounds

By "machine backgrounds" we refer to the radiation that reaches different parts of the detector as a consequence of being immersed in an accelerator environment. They should be avoided as they lead to degradation of the performance of the detector, due to the sustained radiation damage, and they produce large dead times in which the different systems of the detector are unable to accept any real physics event because they are processing and flushing out the spurious signals. In *BABAR*, there are three main sources of machine backgrounds. In order of decreasing importance these are: synchroton radiation, beam-gas interactions and radiative Bhabha scatterings.

The synchroton radiation is a direct consequence of the rather complex optics discussed before, and especially of the dipoles inside the detector, as it is caused by the bending of the beams so close to the interaction point. The geometry of the crossing has been designed to minimize the probability of those photons, and the electromagnetic showers that they produce, impacting the detector. Copper masks are also used to prevent them from interacting with the beam pipe and creating further debris. This kind of background scales with the currents going through each of the storage rings. Increases in luminosity stemming from higher currents bring this background as an undesired side effect.

Interactions of the beam with gas molecules present in the beam pipe often lead to particles in the beam acquiring a momentum outside the range that can be focused by the optics. These lost particles may then hit the beam pipe and produce an electromagnetic shower that spreads over the detector. Collimators are employed to prevent any of these from occurring close to the detector.

Finally, radiative Bhabha scatterings of electrons and positrons may cause one of them to hit the pipe inside the detector and produce an electromagnetic shower that spreads over our measuring instrument. Such a background scales with luminosity, and may prove to be important as the luminosity is increased towards the end of the lifetime of the experiment.

3.2.3 Trickle injection

The currents stored in the rings can be topped up once every two or three hours, during which the voltage is ramped down in the detector to prevent any background from affecting it. A new injection is only arranged when the instantaneous luminosity falls below a pre-established threshold. It can also be made continuously at a low rate, in which case the resulting machine backgrounds must be dealt with. In succesive tests during late 2003 and early 2004, it was shown that these backgrounds could be kept to a manageable level, and the default operation mode has involved such trickle injection ever since, greatly helping to improve the delivered luminosity (see Figure 3.2).

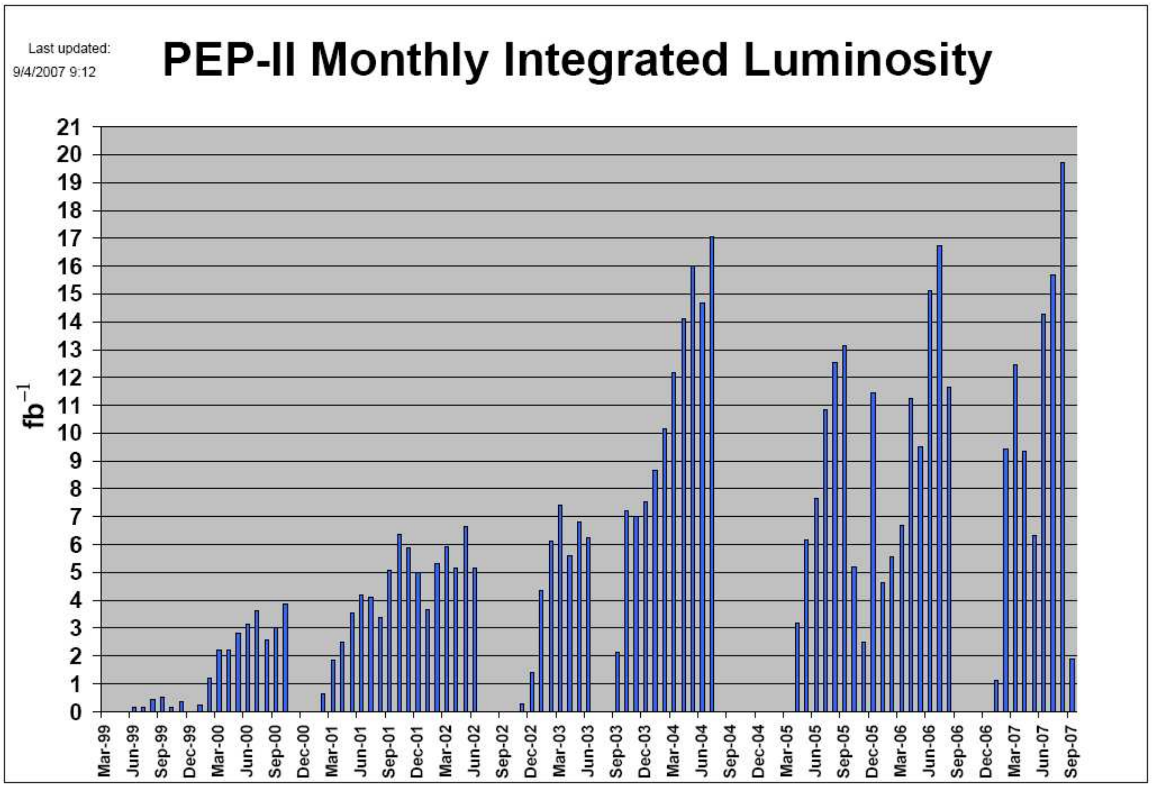


Figure 3.2: Plot showing the evolution of the monthly integrated luminosity as delivered by PEP-II. Note the increase in the first half of 2004 - partly due to the trickle injection. During the last months of running, instantaneous luminosities have been above $10^{34}\,\mathrm{cm}^{-2}s^{-1}$.

3.2.4 Performance

The design luminosity and accelerator parameter goals were met by PEP-II within the first year of running, and they have been improving ever since (Figure 3.2). Table 3.1 compares the design goals with the latest records achieved. $500\,\mathrm{fb}^{-1}$ have been delivered so far, and $750\,\mathrm{fb}^{-1}$ are projected for the end of data taking, in September 2008. The latest instantaneous luminosities are above $10^{34}\,\mathrm{cm}^{-2}s^{-1}$, and they are hoped to reach $2 \times 10^{34}\,\mathrm{cm}^{-2}s^{-1}$ in 2008.

Table 3.1: PEP-II design parameters, and best achieved.

Parameter	Design	Best achieved
HER current (A)	0.75	1.96
LER current (A)	2.14	3.03
Luminosity ($10^{33}\,\mathrm{cm}^{-2}s^{-1}$)	3.0	12.07

3.3 The *BABAR* detector

In this section we state the main requirements on the *BABAR* detector resulting from the physics under study, describing each subsystem and its performance in later sections. A detailed description of the detector can be found in [67].

The physics goals stated in Section 3.1 lead to an asymmetric collider, and therefore, since a uniform acceptance in the centre-of-mass system is preferred, to an asymmetric detector as well, with its centre displaced from the interaction point by 37 cm. It should also comply with the following requirements:

- A high reconstruction efficiency for charged and neutral particles of momenta above 60 MeV/c and 20 MeV/c, respectively.

- Good momentum resolution or charged particles in the momentum range 60 MeV/c to 4 GeV/c. Low momentum particles are important, among other reasons, because they are copiously produced by D^* mesons decaying to D mesons.

- Good photon energy and angular resolutions, in order to reconstruct π^0 and η particles.

- Excellent particle identification capabilities, as these are crucial to the tagging procedure. In particular, electrons and muons should be reliably identified, and hadrons such as p, K, π should be distinguished.

- Excellent vertex resolution, so that the typical distances associated to a time-dependent analysis, $\beta\gamma c\tau_{B^0} \lesssim 250\mu m$, are resolved. It is also desirable for D and τ physics, since these particles also exhibit displaced vertices.

- Dead times as short as possible, so that higher luminosities can be handled without problems.

- Radiation resistance, in order for the efficiencies of the subsystems not to degrade badly over the lifetime of the experiment.

The final design of the detector consists of five subsystems: the silicon vertex tracker (SVT), which provides the accuracy needed to reconstruct the displaced vertices of the B mesons and other particles with similar lifetimes; the drift chamber (DCH), the main tracking device; the detector of internally reflected Čerenkov light (DIRC), whose input for particle identification is essential; the electromagnetic calorimeter (EMC), that allows for exclusive studies of final states containing π^0 and other neutral particles; and the flux return (IFR), instrumented first with resistive plate chambers (RPC) and more recently with limited streamer tubes (LST), which serve as muon detectors and as a primitive hadronic calorimeter e.g. for the K_L^0. All but the last two subsystems are immersed in a uniform axial 1.5 T magnetic field necessary to measure the transverse momentum from the curvature of the tracks. Figure 3.3 shows the whole detector, specfiying each subsystem.

The convention adopted in *BABAR* for the coordinate system follows a standard spherical-polar coordinate system centred on the interaction point (IP), the z axis being parallel to the e^- beam direction, and θ and ϕ being the usual polar and azimuthal angles. The cartesian axes form a right-handed system with the x axis pointing outwards from the PEP-II ring and the y axis pointing upwards.

3.4 Silicon Vertex Tracker

3.4.1 Physics requirements

The Silicon Vertex Tracker is located just outside the beam pipe, at around 3 cm from its centre. Its position makes it crucial in the determination of decay vertices of B and D mesons and τ leptons. To achieve the necessary resolution in Δt, a resolution of

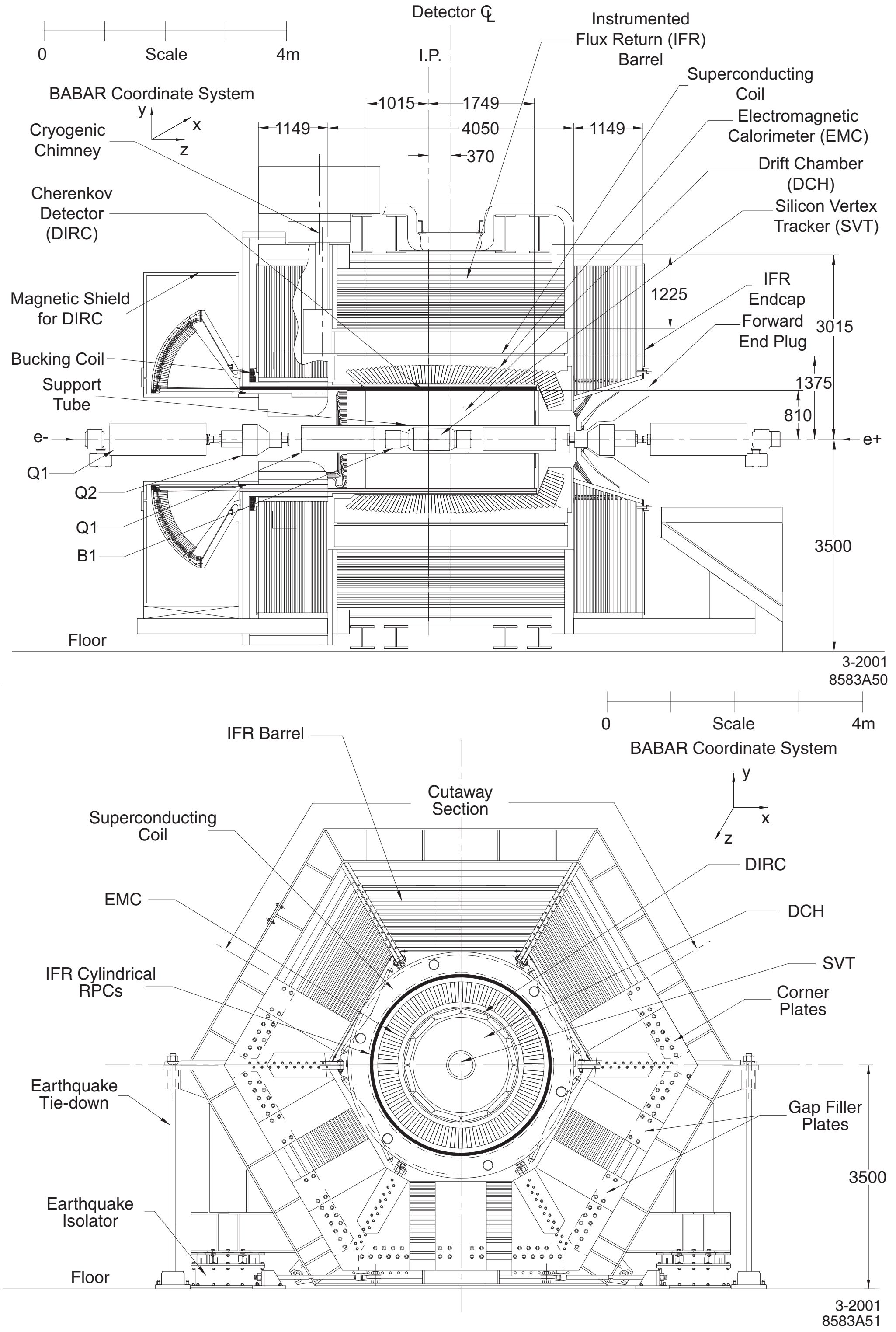


Figure 3.3: Lateral (top) and front (bottom) view of the detector, with all the subsystems clearly indicated.

80 μm must be attained in z for single-vertex measurements. In the xy plane, distances of $\sim 100\,\mu m$ must be resolved for the correct reconstruction of secondary vertices such as those from D and τ decays.

The SVT is also responsible for the tracking of low momentum particles, since for $p_t < 120\,\mathrm{MeV}/c$ they are unlikely to reach the drift chamber or produce enough hits in it. This is particularly relevant to the reconstruction of the D^* decay into a low momentum pion and a D meson, which is important in itself and for tagging purposes.

Finally, the SVT also plays a role in particle identification through its own measurements of the rate of energy loss, and by giving the best determination of the polar angle of high momentum tracks, which is a necessary input to fully exploit the DIRC.

3.4.2 Design

Although maximum coverage is desirable, the B1 dipoles situated inside the detector (see Figure 3.1) and some support structures limit the SVT acceptance to the polar angles $20.1^\circ < \theta < 150.2^\circ$, which still comprises 90% of the solid angle in the centre-of-mass system. Another constraint on the SVT design is that it must be able to withstand the irradiation associated with being so close to the beam pipe, while still keeping the amount of material as low as possible to avoid multiple Coulomb scattering.

The SVT is composed of five layers of double-sided silicon strips (see Figure 3.4), with the strips on the outside being parallel to the beam and on the inside perpendicular, thus providing simultaneous measurements of ϕ and z, respectively, for each hit.

The first three layers are composed of 6 modules each, slightly tilted to provide complete coverage. Layers 4 and 5, having 16 and 18 modules respectively, produce the overlap between neighbouring strips by alternating the radii at which thay are located. The strips in the two outermost layers are arch-shaped, in contrast to the three innermost, in order to reduce the material a track goes through while providing complete coverage. The two innermost layers are particularly important in determining the polar angle of a track, while the role of the two outermost is to help in matching

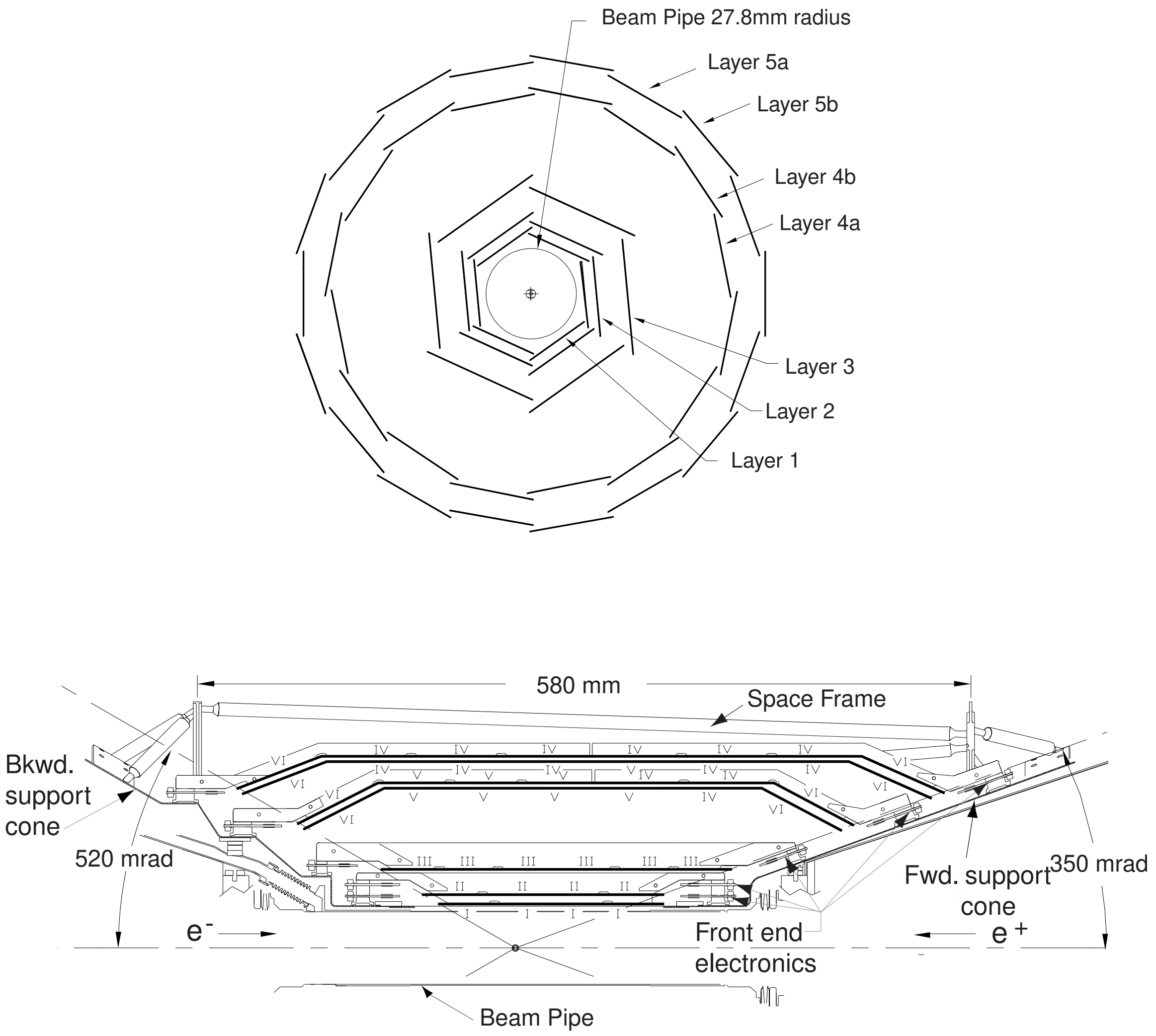


Figure 3.4: End (top) and side (bottom) views of the Silicon Vertex Tracker, in which the five layer structure and the arch shape of the outer layers can be appreciated.

tracks to those found by the DCH. The third layer provides extra information for low momentum tracks that may not reach the drift chamber.

To be able to meet the resolution goals, the local and global alignment of the SVT is crucial. The local alignment, of the different modules relative to each other, is only necessary after accesses to the detector. It is carried out by fitting tracks from cosmic rays and $e^+e^- \rightarrow \mu^+\mu^-$ events. Global alignment, of the SVT with respect to the rest of the detector, is done at the beginning of each run, by minimizing the differences between the SVT and DCH tracks in a small sample of events.

3.4.3 Performance

The efficiency in track reconstruction of the SVT as measured in data on dimuon events is 97%, after excluding the defective strips. A good example of the efficiency of low momentum tracking in the SVT is *BABAR*'s recent evidence for $D^0 - \overline{D}^0$ mixing [20], in which extraordinarily large samples of D^* mesons decaying to a charged pion and a D meson are needed in order to tag the flavour of the neutral meson and detect the minute oscillation.

The spatial resolution of the SVT hits can be evaluated by fitting high momentum tracks without the hit in the layer under inspection and comparing the hit with the intersection of the fitted track. The residuals are divided by the uncertainty on the track determination to get the resolution. This is found to be better than $40\mu m$, implying a vertex resolution better than $70\mu m$.

3.5 Drift Chamber

3.5.1 Physics requirements

The drift chamber is the main tracking system in the *BABAR* detector and it is therefore expected to measure the momenta and polar angles of the tracks efficiently and precisely over a wide range of momenta, $0.12 < p_t < 5.0\,\mathrm{GeV}/c$. It provides one of the main inputs to the Level-1 trigger and plays a key role in the extrapolation of tracks into the DIRC, EMC and IFR. Thus, the solid angle coverage must be as complete as possible whilst minimizing the amount of material that the particles have to traverse.

The DCH must achieve a resolution of $\sigma_{p_t}/p_t < 0.3\%$ in order to reconstruct B and D candidates. Furthermore, a spatial resolution of $140\mu m$ is needed. The vertexing of long-lived particles such as the K_S^0, present in the golden mode and in the channel analyzed in this thesis, as well as in many other final states studied with time-dependent analyses, requires the drift chamber to measure longitudinal positions to better than 1 mm.

Finally, the DCH also bears the responsibility for particle identification for momenta $p_t < 700\,\mathrm{MeV}/c$, for which the DIRC is is not effective, and for tracks that fall outside the acceptance of the latter in the forward region. K/π separation is attained with a precision of 7% on $\mathrm{d}E/\mathrm{d}x$ measurements.

3.5.2 Design

The DCH is a 276 cm long cylinder located immediately outside the SVT, with inner and outer radii of 23.6 and 80.9 cm respectively and displaced towards the forward direction to increase the centre-of-mass acceptance. A mixture of helium and isobutane in a ratio 4:1 fills the chamber, with additional small amounts of water vapour (0.3%) to extend the lifetime of the device.

The DCH is formed from 40 layers of hexagonal drift cells, with each group of 4 layers organized into a superlayer. The cells consist of a sense wire in the centre, with a diameter of $20\mu m$ and kept at 1930 V, surrounded by 6 field wires, of 80 and $120\mu m$ diameters, that are grounded. Each layer of cells is staggered with respect to the previous one, which allows the left-right ambiguity in the measurements to be resolved. Longitudinal positions are measured by orienting of the layers at a small angle to the z axis. Their arrangement is as follows: superlayers alternate between axial (A) and stereo (U,V) superlayers forming a pattern AUVAUVAUVA, with the angle of the stereo layers with respect to the z axis increasing outwards from 45 to 76

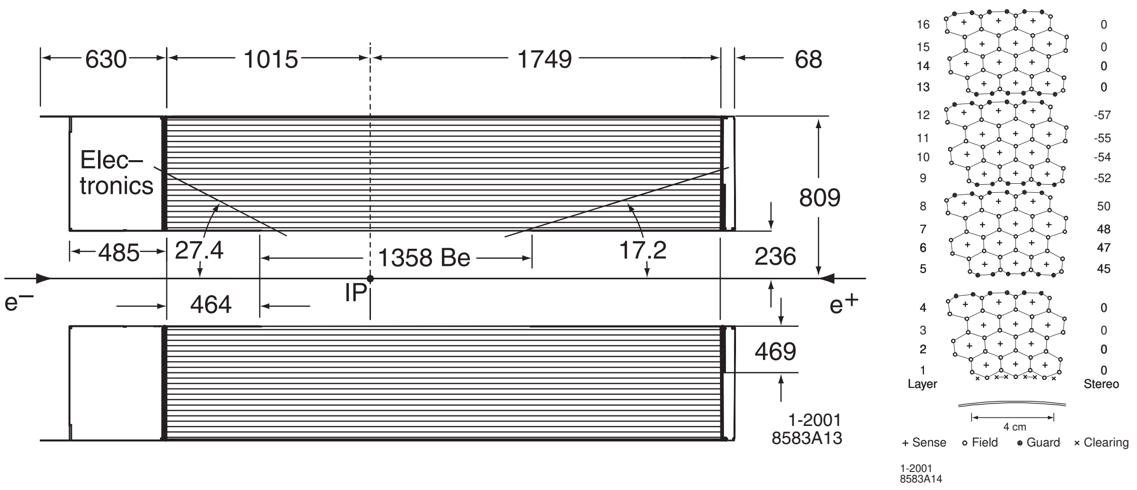


Figure 3.5: Side view of the drift chamber (left) and cell layout in first four superlayers (right). The angle in mrad of the stereo layers with the z axis is written on the right of each layer.

mrad. Altogether, the whole volume of gas and the wires represent only 0.28% of a radiation length for tracks with normal incidence.

3.5.3 Performance

When a charged particle goes through the chamber, it ionises some of the molecules in the gas. The charges then start drifting due to the electric field, which is locally quasi-cylindrical around each sense wire. Collisions with further gas molecules result in a gain of $\sim 5 \times 10^4$. The time taken for the charge to arrive at the wire translates into a distance from the wire. That drift time, however, requires cell-by-cell calibration, which is performed by fitting high momentum $\mu^+\mu^-$ and e^+e^- tracks while omitting the cell being calibrated. The total charge deposited, which is used to calculate $\mathrm{d}E/\mathrm{d}x$, also needs calibration.

The track reconstruction efficiency can be estimated from the sample of tracks that traverse both tracking devices, the DCH and the SVT. After correcting for fake SVT tracks, the ratio of the number of tracks reconstructed in the DCH over those observed in the SVT is found to be $(96 \pm 1)\%$.

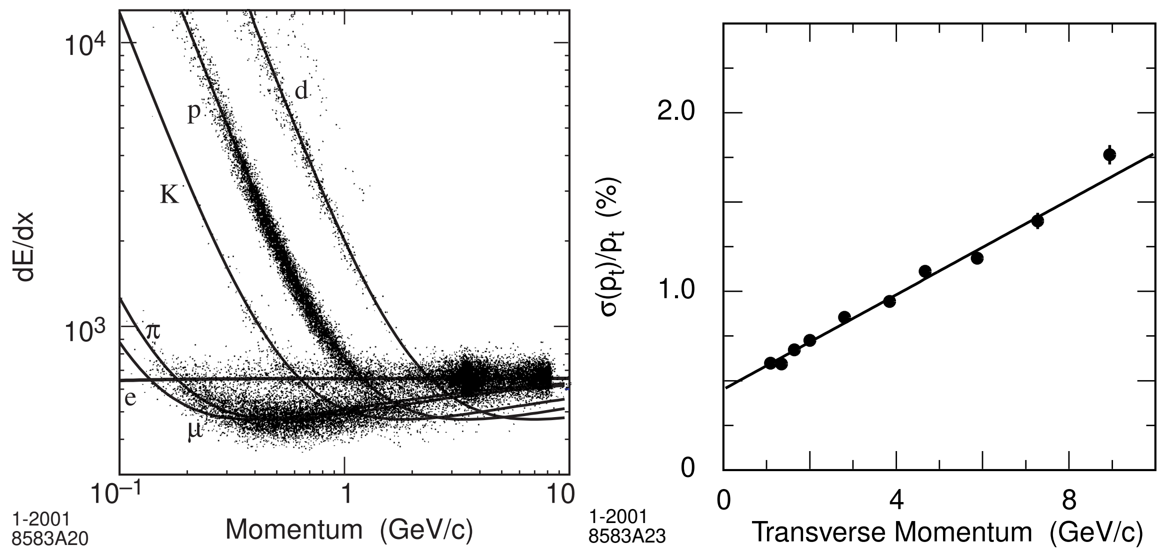


Figure 3.6: On the left, DCH dE/dx measurements and the Bethe-Bloch predictions [67]. On the right, the resolution on p_t as calculated from cosmic ray events that fall within the acceptance of the SVT and the DCH [67].

The dE/dx value for a given track is calculated as the truncated mean of the 80% lowest measurements dE/dx measurements for the track, since these follow a Landau distribution whose mean diverges. Figure 3.6 shows the dE/dx measurements in the DCH as a function of the momentum, and the corresponding Bethe-Bloch expectations [22]. A good separation between pions and kaons is achieved below 0.7 GeV/c momentum, above which the DIRC has the main responsibility for particle identification. This is also demonstrated in practice in *BABAR*'s last measurement of $B^0 \to h^+h^-$ (where $h = K, \pi$) [43], where the DCH is used to provide particle identification of forward tracks falling outside the acceptance of the DIRC[1]. The dE/dx resolution on electrons is 7.5%, almost at the design value (7%). The resolution on p_t is, as shown in Figure 3.6, very close to the design value too.

[1] Unlike in this example, most analyses in *BABAR* which require particle identification do not use the information from each subdetector separately, and instead draw on a combination of information from the DIRC and the DCH and SVT dE/dx measurements, as described in Section 4.3.3.

3.6 Detector of Internally Reflected Čerenkov light

3.6.1 Physics requirements

Particle identification is essential to *BABAR*, since its physics programme consists of measuring *CP* violating asymmetries in a variety of channels. In neutral modes, the measurement of these asymmetries necessitates the determination of the flavour of the other B in the event, which is done through the correlation of the charges of certain particles with the flavour of the parent meson. These correlations depend on the particle species, making their identification mandatory (see Section 4.1). It is also crucial to avoid contamination in the isolation of final states, since similar channels, like $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$, have different asymmetries.

More specifically, above $700\,\mathrm{MeV}/c$, the drift chamber is no longer able to distinguish kaons from pions, which the DIRC aims to separate at 4σ significance up to a momentum of $4.2\,\mathrm{GeV}/c$. For the muons, the DIRC must complement the IFR, whose effectiveness falls for momenta below $750\,\mathrm{MeV}/c$.

Finally, given its location between the drift chamber and inside the calorimeter, it must be small to minimize the size of the most expensive part of the detector, the calorimeter, and it amounts to only a fraction of a radiation length (see below).

3.6.2 Design

When a particle travels faster than the speed of light in the medium that surrounds it, $v/c = \beta \geq 1/n$, it emits Čerenkov photons at an angle $\cos\theta_C = 1/n\beta$ with the direction of the particle. Hence, provided that its trajectory is known accurately enough, a measurement of the direction of these photons establishes the speed of the particle. Given the space constraints sketched above, the instrumentation to detect them must lie outside the main body of the detector. Internal reflection on a plane surface is used to preserve the angle of these photons while directing them towards the photomultiplier tubes (see Figure 3.7). Forward moving photons are reflected in a mirror, allowing the DIRC instrumentation to occupy only the less populated

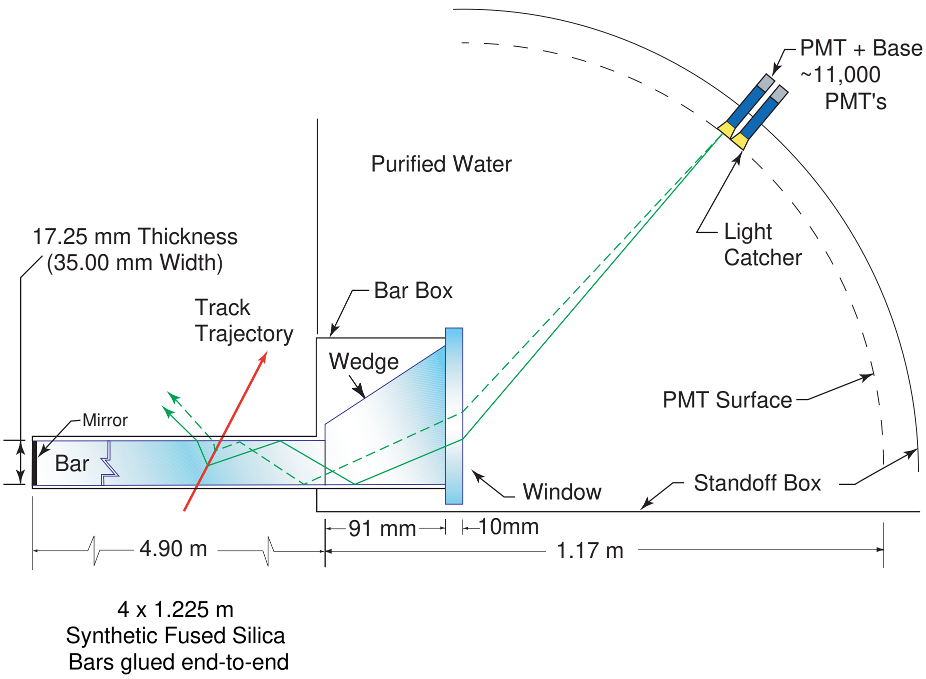


Figure 3.7: Diagram illustrating the operating principles of the DIRC.

backward end of the detector.

The material used to confine the photons is quartz ($n = 1.474$) shaped in bars that are only 17 mm thick and 35 mm wide, but as long as 4.9 m. They amount to 17% of a radiation length for a normally incident particle. In the backward end of the detector, the photons go through a wedge-shaped quartz piece and then into a water filled expansion region, known as the standoff box, after which they meet the photomultiplier tubes. The role of the wedge is to reflect photons arriving at large angles, thereby reducing the area of the standoff box that needs to be instrumented at the cost of introducing ambiguities in the angle.

The photomultiplier tubes, of which there are 10752, are surrounded by "light catchers", increasing the detection area. Also, the standoff box is magnetically shielded to avoid disturbances in the tubes.

3.6.3 Performance

The angle and time resolution can be calibrated from dimuon events. The Čerenkov angle resolution for a track turns out to be 2.5 mrad, giving over 4σ separation at 3 GeV/c (see Figure 3.8). Figure 3.8 also shows the mass peak of the decay $\overline{D}^0 \to K^+\pi^-$ with and without the kaon/pion separation provided by the DIRC.

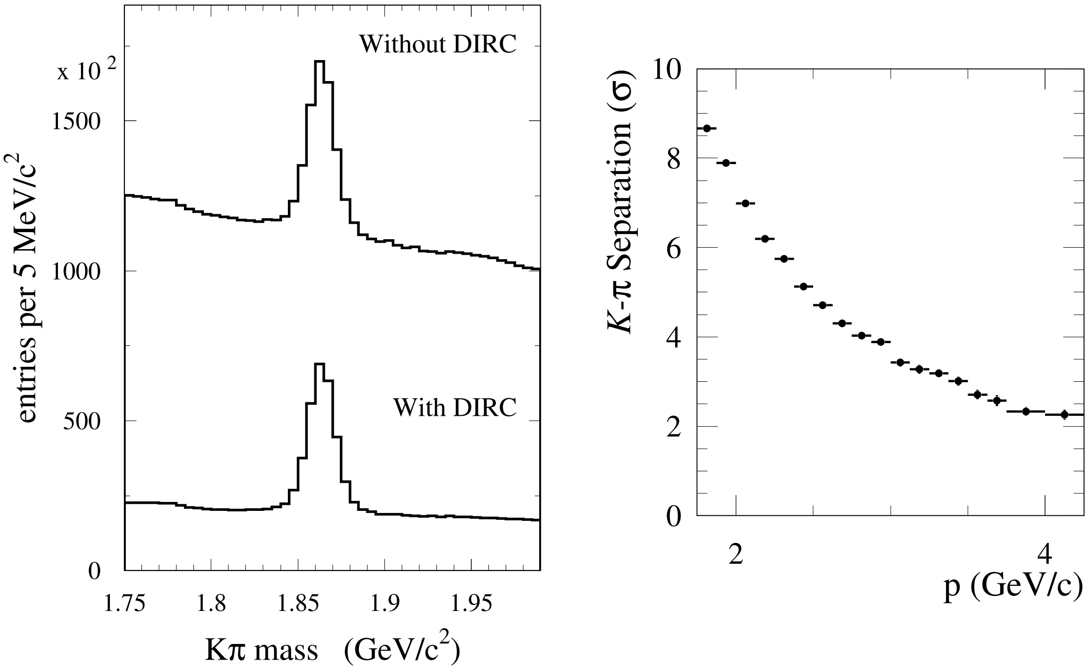


Figure 3.8: Plots illustrating the DIRC capabilities for kaon-pion separation [67]. On the left, the reconstructed $\overline{D}^0$ mass with and without the PID information provided by the DIRC. On the right, the separation of kaons and pions achieved by the DIRC in standard deviations.

3.7 Electromagnetic Calorimeter

3.7.1 Physics requirements

A number of *CP* eigenstates within *BABAR*'s physics goals contain π^0's in the final state. Many others involve η particles or photons directly, such as $b \to s\gamma$, in which the spectrum is quite hard. Some QED processes, such as $e^+e^- \to e^+e^-\gamma$ or $e^+e^- \to \gamma\gamma$ are also important for calibration or luminosity measurement purposes. Therefore, *BABAR* must be able to reconstruct photons over a wide range of energies, from 20 MeV up to 4 GeV.

The EMC must also be efficient in identifying electrons, as they are important for flavour tagging and semi-leptonic B decays, and no other system can provide accurate particle identification information for them.

3.7.2 Design

The EMC is formed from 6580 Thallium doped Caesium Iodide crystals arranged in a barrel and a forward endcap . The material was chosen due to its high light yield and small Molière radius, which imply good energy and angular resolutions, respectively. The crystal size varies from 16 radiation lengths in the backward direction to 17.5 radiation lengths in the forward endcap, since these receive impacts from the more energetic Lorentz-boosted particles. The crystals are tilted in such a way that they face the interaction point. Their exposed area is $\sim 5\,\mathrm{cm}^2$ (c.f. their Molière radius, 3.8 cm), so a typical electromagnetic shower will spread over several crystals. Figure 3.9 shows a schematic view of the subdetector.

3.7.3 Performance

The calibration of the electromagnetic calorimeter involves determining the relation between the light yield and the energy deposited in the crystal, and between the cluster energy and the total energy of the incident particle. The light yield dependence on the energy varies from crystal to crystal, and may change over time due to radiation damage. It is calibrated for low energies using 6.13 MeV photons from a radioactive source, and in the high energy range by using Bhabha scattering events, for which the polar angle precisely determines the energy of the particle. Crystal leakage and absorption of energy by the material at the front of the crystals or between them leads

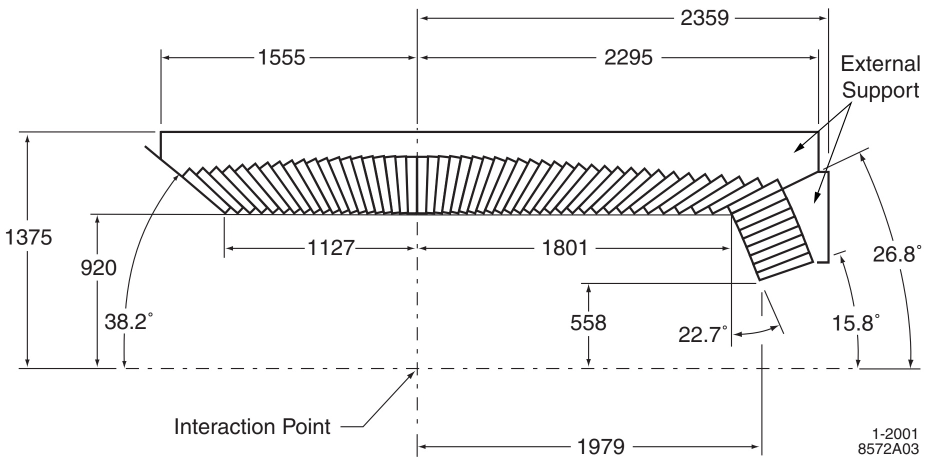


Figure 3.9: Side view on the Electromagnetic Calorimeter.

to the need for the calibration of the cluster energy. This is applied during the offline reconstruction and is derived from samples of π^0 and η mesons.

The photon energy and angular resolutions of the EMC are also extracted from the calibrations, and are found to be parameterized, respectively (see Figure 3.10), by

$$\frac{\sigma_E}{E} = \frac{a}{(E(\text{ GeV}))^{1/4}} \oplus b \tag{3.1}$$

$$\sigma_\theta = \sigma_\phi = \frac{c}{E(\text{ GeV})} + d\,, \tag{3.2}$$

where the first sum is in quadrature, and $a = (2.3 \pm 0.3)\%$, $b = (1.85 \pm 0.12)\%$, $c = 3.87 \pm 0.07$ and $d = 0.00 \pm 0.04$.

3.8 Instrumented Flux Return

3.8.1 Physics requirements

The golden mode, $J/\psi\, K_S^0$, involves muons, as the J/ψ is reconstructed in the channels e^+e^- and $\mu^+\mu^-$. Their detection is also essential for semi-leptonic physics and for the tagging algorithms. Particle identification information on muons is desirable for momenta from about $1\,\text{GeV}/c$.

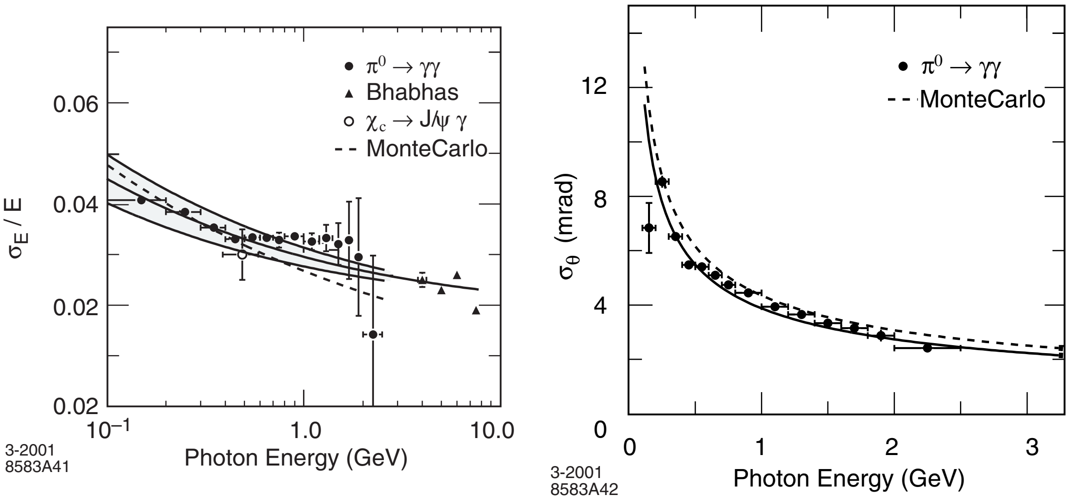


Figure 3.10: Photon energy (left) and angular (right) resolutions achieved by the electromagnetic calorimeter [67].

Muons are heavier than electrons, making bremsstrahlung a far less effective energy loss mechanism for them. Since they have relatively long lifetimes and do not participate in nuclear interactions either, they are very penetrating particles. Therefore, the best choice is to place a dedicated subdetector outside the rest of the instruments.

In *BABAR*, the outer part of the detector plays the role of the flux return for the solenoid, at the same time as it provides a support structure. Interleaved between the steel plates of the flux return, instruments can be placed to turn it into a muon detector and a primitive hadron calorimeter, in charge of detecting neutral hadrons, mainly K_L^0. These feature in a number of modes of interest, due to them having an opposite *CP* eigenvalue to the best experimentally suited modes containing a K_S^0.

3.8.2 Design

The steel of the flux return, which is distributed in layers of increasing thickness from the inner to the outer sides, serves the purpose of filtering the muons and absorbing the neutral hadrons. Between the steel sheets, in the barrel and the endcaps, there are 19 and 18 gaps, respectively, which host the instrumentation. These are shown in Figure 3.11, where the almost complete coverage of the detector is also apparent. The

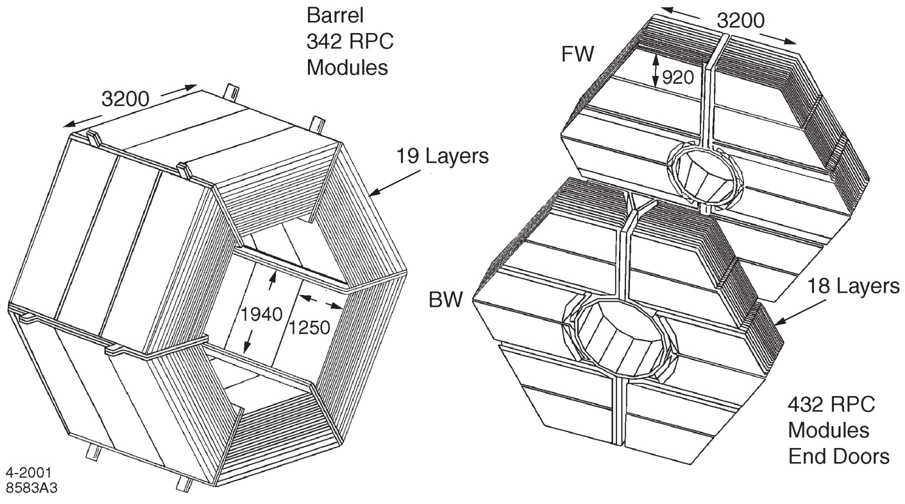


Figure 3.11: Schematic view of the IFR, with the barrel on the left and the forward (FW) and backward (BW) endcaps on the right.

arrangement of these gaps, and the thickness of each of the layers was carefully chosen after dedicated MC studies to optimize the physics capabilities.

In these gaps, resistive plate chambers were installed. Two cylindrical RPCs were also placed between the EMC and the magnet to detect particles leaving the EMC and link any EMC clusters to IFR energy deposits.

The resistive plate chambers (see Figure 3.12) consist of two graphite electrodes separated by two 2 mm thick sheets of bakelite, and in between these, another 2 mm gap filled with a mixture of gases: argon, freon and isobutane in the proportions 57:39:4. Readout strips are located next to the graphite, separated from it only by a film of insulator. They are placed orthogonally (hence the labels “X strip ” and “Y strip” in the figure), providing three-dimensional positional information when combined with the distance of the RPC to the interaction point. The apparatus works as a capacitor, with one of the graphite electrodes grounded and the other one set to an 8 kV voltage. The passage of a charged particle or a hadronic shower do not cause a discharge, but induce temporary changes in the charge accumulated at each electrode, that are capacitatively read by the readout strips.

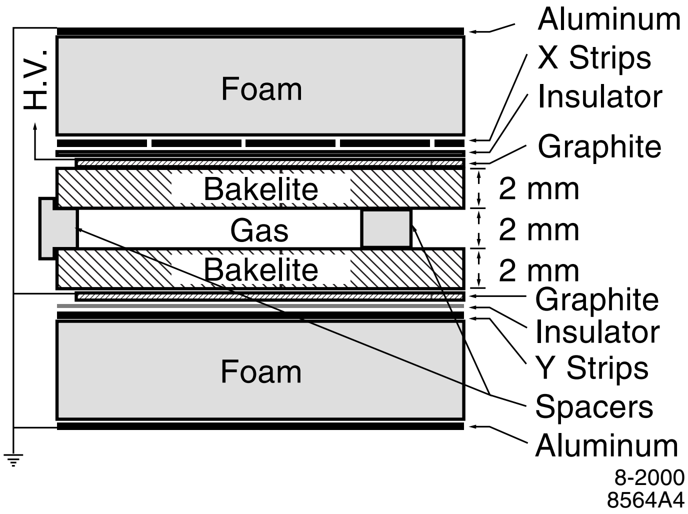


Figure 3.12: Resistive plate chamber design.

3.8.3 Performance

During *BABAR*'s first year of running, an 8% pion misidentification probability was found for a 90% muon efficiency.

The calibration of the angular resolution and efficiency of the detection of neutral hadrons was studied through the process $e^+e^- \to \phi\gamma \to K^0_S K^0_L \gamma$, and yielded efficiencies between 20 and 40%, and angular resolutions around 60 mrad for K^0_L that did not interact in the EMC. When the latter also provided information, the resolution was twice as good.

However, shortly after installation, the performance of the RPC was observed to degrade quickly, with the muon efficiency dropping at an average rate of 1.2% per month and growing numbers of plates being declared "dead" (efficiencies less than 10%). The RPCs in the endcaps were replaced by new RPCs built with more stringent quality constraints, except for the first five, in whose place brass was placed to improve pion rejection. A different solution was adopted for the barrel RPCs. These were substituted by limited streamer tubes.

3.8.4 Limited streamer tubes

The principle of operation of limited streamer tubes is similar to that of the RPCs. In the case of *BABAR*, a conducting wire with a $100\mu m$ diameter is placed in a long resistive cell (the "tube"), with a section of $15 \times 17\,\mathrm{mm}^2$, the wire playing the role of the anode, and the tube, of the cathode. The volume between them is filled with a gas that is ionized upon the passage of a charged particle or the spread of a hadronic shower, altering the charge distribution in the cylindrical capacitor. The signal can then be read either by external strips attached to both sides, or from the wires directly. In *BABAR*, the latter method is used to measure the ϕ coordinate, and the former, to read the z coordinate.

Their efficiency is monitored by using dimuon events and cosmic rays, finding an average of 90%, without any noticeable degradation trend over time.

3.9 Trigger

The aim of the trigger is to reduce the potential number of events per second reconstructed by the detector, which is essentially determined by the frequency of bunch crossing, to a manageable level of events that can be recorded. Of course, the goal of the trigger is to reject badly reconstructed events and background while retaining as much signal as possible.

In *BABAR*, that is achieved through a two-stage trigger, composed of the level-1 trigger (L1T), implemented in hardware, and the level-3 trigger, which is carried out by software, and after which all surviving events are recorded.

3.9.1 Level-1 trigger

The level-1 trigger consists of a global trigger (GLT) that combines the input from several individual triggers linked to the different subsystems of the detector, and accepts events at rates around 1 kHz, its limit being 2 kHz. The indivual triggers feeding the global one are the DCH trigger (DCT), the EMC trigger (EMT) and the IFR trigger (IFT). These are continuously producing abstract data (*primitives*) describing the objects found by the subdetectors they are associated to, and are passed to the GLT. The global trigger then tries to match them to any of 24 *trigger lines* that represent events of interest, and if the timing of the trigger signal coincides with one bunch crossing, the fast control and timing system issues an *accept* signal. It is at this point that some classes of physics events, such as typical QED processes that are used only in calibration, are scaled down, making their acceptance less likely.

The DCT produces its primitives by looking for sequential DCH hits in neighbouring cells. These are then joined if possible to construct either short tracks, that traverse only a few superlayers, and long ones, that reach the end of the chamber. Axial superlayers are also examined looking for segments consistent with tracks with transverse momenta greater than $800\,\mathrm{MeV}/c$.

The EMT sums the energy deposited on 40 strips along the ϕ polar angle, and finds a peak whose energy is compared to thresholds for different physics processes: minimum ionizing particle cluster ($E > 120\,\mathrm{MeV}$), intermediate energy cluster ($E > 307\,\mathrm{MeV}$), high energy electron or photon ($E > 768\,\mathrm{MeV}$), minimum ionizing particle in forward endcap ($E > 100\,\mathrm{MeV}$) and backward high energy cluster ($E > 922\,\mathrm{MeV}$).

The IFT primitives are just single clusters or back-to-back coincidences. These select cosmic ray events for calibration purposes, and $\mu^+\mu^-$ events.

The different trigger subsystems are optimized to select high multiplicity, multi-hadronic events, resulting in efficiencies over 99% for $B\overline{B}$ events for both the DCT and the EMT individually, and over 99.9% when combined.

3.9.2 Level-3 trigger

The level-3 trigger must reduce by $\gtrsim 10$ the number of events accepted by the L1T. It is implemented in software and run in computing farms, which allow the use of information from all the subdetectors. Examples are the rejection events with tracks not originating from the interaction point, as these are likely to be machine background, or events whose timing does not match a bunch crossing.

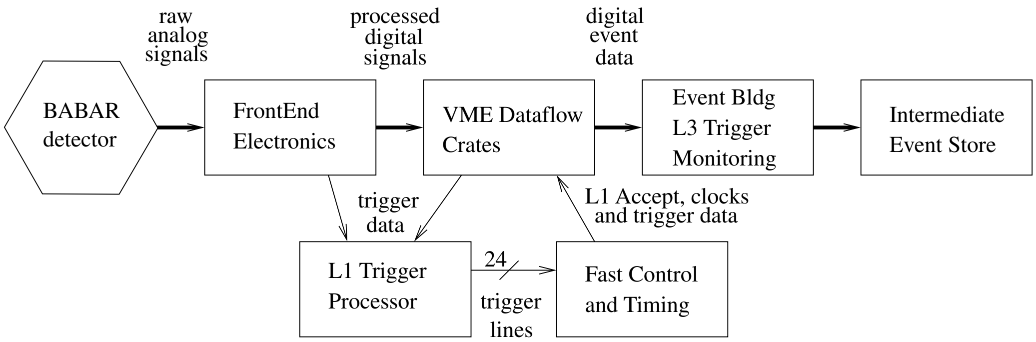


Figure 3.13: Schematic explanation of the interplay between the detector, the triggers and the first stages of the reconstruction process, known globally as the data acquisition system.

Level-3 trigger lines may also be prescaled to reduce the rate of less interesting physics events, such as Bhabha scatterings. Calculation of efficiencies requires accepting events that do not satisfy any of the level-3 criteria. These are known as *L1 passthrough* events.

3.10 Data Aqcuisition

By data acquisition (DAQ) system we refer to the overall architecture by which the detector, the triggers and the computing structure are governed. The diagram in Figure 3.13 schematically depicts it. The front end electronics process and digitize the signals coming from the detector and passes them to the the level-1 trigger and the data flow buffers. If an accept signal is issued by the fast control and timing, the event is passed to the level-3 trigger, which also performs some basic data quality monitoring. Finally, if the event is accepted by the L3T, it is written to disk, where it will be passed to the online prompt reconstruction software in a matter of days. The DAQ is also responsible for recording the detector conditions, that will be used in the production of simulated data to better reproduce the running conditions (see Section 4.2).

3.11 Online Prompt Reconstruction

After a data sample, typically consisting of around an hour of experiment running, has been logged on to disk, it goes through a prompt calibration processing, during which some of the calibration methods mentioned earlier in the chapter are run, and part of the data quality monitoring is performed. Following that, the data continue to event reconstruction, where tracks and clusters are found, and particle identification information (PID) is calculated. An event, by then essentially a collection of tracks, EMC clusters and IFR clusters, is stored in a database that will be accessed by the analysts reconstruction code to form *candidates* for events of a given decay channel (see Section 4.3). Once all these quantities have been calculated, a more detailed data quality check is made.

Chapter 4

Analysis Techniques

In this Chapter, we describe in detail some tools and methods commonly used in High Energy Physics analyses, such as simulated data (see Section 4.2), the approach to reconstructing data events (Section 4.3), signal and background discriminating variables (Section 4.5) and maximum likelihood fits (Section 4.6). We also discuss some other techniques that are particular to B physics, such as tagging (Section 4.1), the measurement of Δt and the modelling of its resolution (Section 4.4). A good, published reference for the latter ones is [23].

4.1 Flavour Tagging

As discussed in Chapter 2, an essential ingredient to the measurement of time-dependent CP asymmetries is the determination of the flavour of one of the two B mesons in the event, referred to as B_{tag}. This is achieved by examining the decay products of the meson when it decays into a flavour-specific state. A good example is the process $b \to c l^- \bar{\nu}$, in which the charge of the lepton l unambiguously identifies that of the b quark and hence its flavour, b or $\bar{b}$. The importance of such a procedure, known as *tagging*, stems from the fact that the statistical error on measurements extracted from time-dependent analyses like $\sin 2\beta$ or Δm_d depends strongly on the *mistag rate*, w, the probability of wrongly assigning a flavour to B_{tag}. Specifically, the error σ on such quantities is inversely proportional to the square-root of the so-called “effective tagging efficiency”, Q [23]:

$$Q = \epsilon_{tag}\,(1-2w)^2\,, \qquad \sigma \propto \frac{1}{\sqrt{Q}}, \tag{4.1}$$

where ϵ_{tag} is the fraction of events that are assigned a tag, and $\mathcal{D} = 1 - 2w$ is the *dilution* factor. The efficiency ϵ_{tag} is calculated with respect to the sample of events that satisfy the requirements for the tag-side vertexing (see Section 4.4) and that have at least one fully reconstructed candidate. In such conditions, the effective tagging efficiency of *BABAR* is $Q = 30.5 \pm 0.3\%$ [24] with the latest version of the tagging protocol, `Tag04`, that we describe later in this section.

Denoting by w ($\bar{w}$) the probability of incorrectly reporting a $B_{tag} = B^0$ as $\overline{B}^0$ ($B_{tag} = \overline{B}^0$ as B^0), the following quantities, which are more convenient from the experimental point of view, are defined:

$$\langle w\rangle = \tfrac{1}{2}(w+\bar{w}) \quad , \qquad \Delta w = (w-\bar{w}) \tag{4.2}$$

$$\mathcal{D} = 1-2w \quad , \qquad \bar{\mathcal{D}} = 1-2\bar{w} \tag{4.3}$$

$$\langle\mathcal{D}\rangle = \tfrac{1}{2}(\mathcal{D}+\bar{\mathcal{D}}) = 1-(w+\bar{w}) \quad , \qquad \Delta\mathcal{D} = (\mathcal{D}-\bar{\mathcal{D}}) = -2(w-\bar{w}), \tag{4.4}$$

where $\Delta\mathcal{D}$ (or Δw) parameterizes a possible difference in performance of the tagging procedure for the two tags, B^0 and $\overline{B}^0$.

4.1.1 *BABAR*'s flavour tagging algorithm

The *BABAR* tagging algorithm [68] first removes from the event the tracks and neutral objects belonging to the fully reconstructed B meson. The remnants are analyzed by a Neural Net (NN, see Section 4.5.3), referred to as `Tag04`, which assigns to the event an overall (signed) probability, the magnitude representing the confidence in the estimation, and the sign indicating the flavour of the meson (NN=+1 $\rightarrow B_{\rm tag} = B^0$, $q_{\rm tag} = +1$; NN=$-1 \rightarrow B_{\rm tag} = \overline{B}^0$, $q_{\rm tag} = -1$). The inputs to `Tag04` are themselves the results of other NNs, which are optimized to find any of nine distinct processes that would uniquely identify the flavour of their parent B meson, and are hence known as "sub-taggers".

The algorithm also classifies the event in the most stringent of seven disjoint, hierarchical categories matching the event's `Tag04` output value. In order of decreasing dilution, these categories are: `Lepton`, `KaonI`, `KaonII`, `Kaon-Pion`, `Pion`, `Other` and `Untagged`, where the last one is deemed to provide no reliable tagging information, $\mathcal{D}_{untagged} = 0$. Although the categories are defined solely according to the range of NN values they span (and in the exceptional cases of `Lepton` and `KaonI`, PID information), their names are related to the sub-tagger that contributes the most.

The algorithm is trained and checked against simulations (MC), and is later validated in data on a large sample of fully reconstructed, flavour-specific decays, $B^0 \to D^{(*)\pm}\pi^{\mp}$, $B^0 \to D^{(*)\pm}\rho^{\mp}$ and $B^0 \to D^{(*)\pm}a_1^{\mp}$ (together known as B_{flav}), where a fit to the Δt distribution (see Eq. (2.20)) allows the extraction of the mistag rates (see Table 4.1).

Table 4.1: Performance of *BABAR*'s `Tag04` tagging algorithm, described by the tagging efficiencies and mistag rates and broken down by tagging category, as measured on the B_{flav} sample. $\Delta\epsilon_{tag}$ and ΔQ are defined analogously to Δw. The algorithm is described in detail in [69] and references therein, and the efficiencies presented here are compatible with those in the cited reference.

Category	$\epsilon_{tag}(\%)$	$\Delta\epsilon_{tag}(\%)$	$w(\%)$	$\Delta w(\%)$	$Q(\%)$	$\Delta Q(\%)$
`Lepton`	8.69 ± 0.07	-0.0 ± 0.2	3.1 ± 0.3	-0.1 ± 0.6	7.66 ± 0.12	0.04 ± 0.41
`KaonI`	10.96 ± 0.08	0.2 ± 0.2	5.2 ± 0.4	-0.1 ± 0.7	8.78 ± 0.16	0.21 ± 0.50
`KaonII`	17.23 ± 0.10	0.1 ± 0.3	15.4 ± 0.4	-0.5 ± 0.6	8.26 ± 0.18	0.29 ± 0.54
`Kaon-Pion`	13.78 ± 0.09	-0.3 ± 0.3	23.5 ± 0.5	-1.8 ± 0.7	3.88 ± 0.14	0.43 ± 0.38
`Pion`	14.37 ± 0.09	-0.7 ± 0.3	32.9 ± 0.5	5.1 ± 0.7	1.67 ± 0.10	-1.08 ± 0.26
`Other`	9.57 ± 0.08	0.3 ± 0.2	41.8 ± 0.6	4.6 ± 0.9	0.26 ± 0.04	-0.28 ± 0.10
Total	74.61 ± 0.12	-0.4 ± 0.6			30.5 ± 0.3	-0.4 ± 1.0

We now proceed to describe each of the nine sub-taggers in more detail. It should

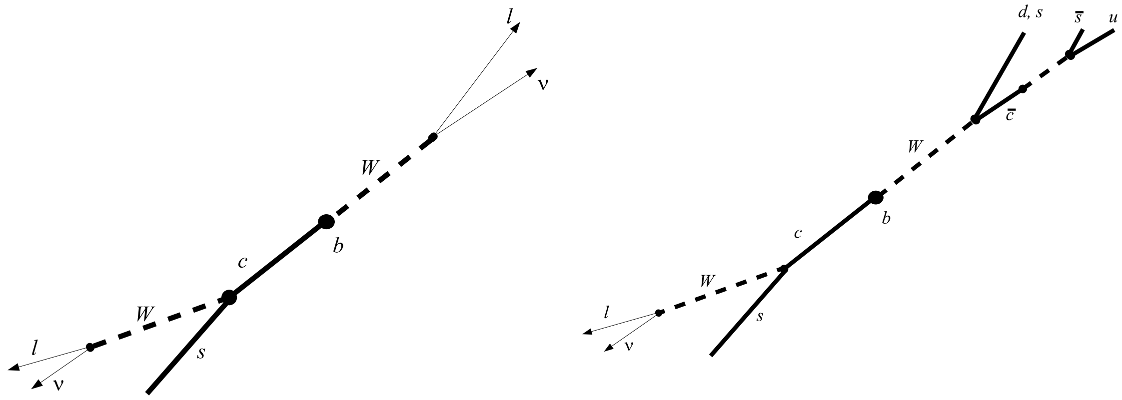


Figure 4.1: Diagrams representing b decays likely to produce a `Lepton` tag (left) and a `Kaon` tag (right). The effectiveness of the variable E_W^{90} in identifying primary leptons is illustrated, since, unlike secondary ones, they are likely to be isolated from the rest of the charged tracks involved. The right-hand diagram demonstrates how right- and wrong-sign kaons may appear in a b decay.

be born in mind that the classification of events into tagging categories is performed according to their overall numerical value of the NN, to which all nine sub-taggers contribute, in spite of the names of some of the sub-taggers and tagging categories suggesting a direct link between the two.

Lepton sub-taggers

The lepton sub-taggers aim to exploit the semileptonic decays of the B meson (with a branching fraction of 10.4(4)% [22]), in which the $b(\bar{b})$ emits a virtual $W^-(W^+)$ boson, and the latter decays to an electron or a muon and an (anti)neutrino. The charge of the lepton, called the *primary* lepton, is unambiguously linked to the flavour of the parent meson. The issue, then, is to isolate these primary leptons from *secondary* leptons that could arise further along the decay chain, e.g. in a semileptonic decay of a daughter D^0 meson.

Several discriminating variables are used, an obvious candidate being the CM momentum of the track, p^*. The momentum spectrum of a primary lepton will be harder than that of a secondary lepton. The cosine of the angle between the missing momentum (which approximates that of the neutrino) and the lepton's momentum, $\cos\theta_{miss}$, is also used, since their directions are expected to be anticorrelated in the CM frame. Finally, the energy contained in the hemisphere defined by the direction of the virtual $W^{\pm}$, E_W^{90}, is found to be useful (Figure 4.1). For primary leptons, the W recoils against a c quark in the CM frame, leading to a virtually empty hemisphere, whereas in the case of secondary leptons, the c quark that emits the W has recoiled from the decay of the b with an appreciable boost, and all its decay products will be boosted in the same direction.

There are 3 sub-taggers that are trained independently using the discriminating variables described above:

- the electron sub-tagger, whose inputs are only tracks that satisfy the `VeryTight` (see Section 4.3.3) electron PID criteria.
- the muon sub-tagger, whose input tracks must be at least `Tight` muon candidates.
- the kinematical lepton sub-tagger, whose goal is to recover primary leptons based solely on the kinematics, for the cases in which the stringent PID criteria of the previous sub-taggers are not satisfied. It is therefore run over all the remaining charged tracks.

Kaon sub-tagger

The purpose of this sub-tagger is to select charged kaons, formed mainly in cascade decays, $B^0 \to \bar{D}\,(\to K^+Y)\,X$. Again, one must distinguish between kaons whose charges correlate differently with the flavour of $B_{\rm tag}$. The most likely production mechanism, $b \to c \to s$, gives rise to *right-sign* kaons, whereas the *wrong sign* kaons are, for instance, the decay products of the W^-, e.g. $b \to W^- c$ with $c \to s$ (right sign) and $W^- \to \bar{c}q$ with $\bar{c} \to \bar{s}$ (wrong sign). This is iluustrated in Figure 4.1.

The kaon sub-tagger is the highest contributor to *BABAR*'s Q value, due to the high branching fraction for inclusive $B^0 \to K^{\pm}X$ decays (78(8)% [22]) and the good signal-to-background ratio. This is yet another example of the close link between the physics goals of a B Factory and its need for excellent particle identification capabilities, and especially kaon-pion separation.

The input variables to the NN are in this case the charge and PID likelihood of the best three kaon candidates of the event, the number of reconstructed K_S^0 candidates and the sum of transverse momentum squared, Σp_t^2. The latter helps to discriminate kaons originating from a W rather than from a charmed object, whereas a non-zero number of K_S^0 decreases the certainty of the tag, since the strange quark from the cascade $b \to c \to s$ could have formed a neutral rather than a charged kaon, providing no information on the $B_{\rm tag}$ flavour.

Slow Pion sub-tagger

Slow pions provide another source of flavour identification, their charge being the key that unlocks the copious $B^0 \to D^{*+}X$ decays, where $D^{*\pm} \to D^0\pi^{\pm}$ about $\frac{2}{3}$ of the time. The small mass difference between the D^* and D^0 mesons means that the pion will carry an unusually low momentum (hence the "slow" label), with an average of around 100 MeV/c. It also means that the pion and the D^0 are emitted almost at rest in the D^* CM frame, and that the latter's boost will strongly correlate the directions of the tracks originating from the D^0 to that of the pion.

This leads to the slow pion's momentum, p^*, and $\cos\theta_{thrust}$, the cosine of the angle between the pion's track and the thrust axis of $B_{\rm tag}$ (see Section 4.5), as natural choices to be input variables. PID information about the slow pion candidate, $\mathcal{L}_K$, is also used to reject slow kaons from the D^0 decay.

Kaon-Slow Pion sub-tagger

In a high proportion of $D^{*\pm}$ decays we can do better than the previous sub-tagger, recalling that the favoured decay for a D^0 is $D^0 \to K^-X$ (53(4)% [22]), and that the

D^0 decay products fly along the same direction as the slow pion. Thus, using the slow pion candidates given by the previous sub-tagger, PID information for the kaon candidates and the angle between their tracks, $\cos\theta_{K\pi}$, a cleaner tag can be obtained.

Highest p^* sub-tagger

Another source of information about the flavour of B_{tag} is the charge of the decay products of the virtual W, which can be quite energetic. These decay products may be leptons not properly identified, or fast particles from the hadronization of the W, such as the pion in $B^0 \to D^*\pi$. Apart from the momentum of the track, p^*, and the cosine of the angle between the fast particle and the thrust axis of B_{tag}, another useful variable is the track impact parameter in the xy plane, since the W decays so rapidly that its daughter particles originate from the B_{tag} vertex.

Fast-Slow correlation sub-tagger

An aspect of the decay $b \to cW^-$ that has not been exploited by the sub-taggers described so far is the correlation between the directions of the decay products of the W and of the charmed meson formed by the c. In particular, a slow pion originating from a $D^{*\pm}$ decay should fly roughly back-to-back with the fast decay products of the $W^{\mp}$, and should also be oppositely charged. The natural discriminating variables are, thus, the momenta of the slow and fast tracks, the cosine of their angles with the thrust axis of B_{tag}, the cosine of the angle between the tracks, and PID information about the slow track (to ensure that it is not a kaon).

Lambda sub-tagger

In a small fraction of events, the process $b \to c \to s$ ends with the strange quark forming a Λ baryon. In spite of its scarcity, the cleanliness of the signature makes it worth considering. A number of standard variables for neutral, long-lived objects are used (mass of the candidate, m_Λ, cosine of the angle between its momentum and its flight direction, $\cos\alpha$, flight length, $c\tau_\Lambda$, momentum of the candidate p_Λ and probability for the fit of the $\Lambda \to p\pi$ decay vertex, χ^2) as well as PID information for the proton.

4.2 Monte Carlo

Simulated data, usually referred to as Monte Carlo data, or simply MC, are essential to understand detector effects (e.g. efficiency, misreconstruction of signal), backgrounds and any systematic effects that could afflict our analysis procedure. The simulation of the physics mechanisms that operate in e^+e^- collisions within *BABAR*'s energy regime and the way that their products interact with the detector and are handled by the reconstruction software does not need to be perfect in order to make a measurement. However, the more detailed and faithful it is, the more effective we are at discarding any systematic problems in the data analysis, e.g. efficiency evaluations.

Often, it is most useful to be able to trace the behaviour of single particles within an event through the whole process, and for this reason information about each of them is carried along all the phases of simulation, from the production, to the later stages of track-fitting, cluster-matching and vertexing. The comparison between the reconstructed information about the event and its generator-level counterpart (*truth-matching*) can be realized in a variety of ways. In *BABAR*, the approach adopted consists of assessing the fidelity of the reconstructed data only after the full reconstruction has been completed, without any such assessment in the intermediate stages, e.g. after the track-finding.

The `EvtGen` package [70] is responsible for the physics simulation, providing an accurate representation of subtle phenomena such as mixing and interference (necessary for the correct modelling of *CP*-violation) or the angular distributions of the decay products in non-trivial situations like pseudoscalar-to-vector-vector decays, for instance. The vast majority of B decays are generated by `EvtGen`, with the remaining generic B decays to hadronic final states (for which there is no specific model) and the continuum events ($e^+e^- \to q\bar{q}, \quad q = u, d, s, c$) being produced via an interface to `JETSET` [71]. Charmless 3-body decays are modelled in detail by `EvtGen` using the Isobar approximation (Section 2.5.3) with the lineshapes and angular dependencies described in Sections 2.5.4 and 2.5.6. The output of this stage is a list of particles, the 4-vectors specifying their kinematics, and the (potentially displaced) vertices for the products of the decays.

The simulation of the interaction of these decay products with the detector as they propagate through it is carried out by software based on `GEANT4` [72], and requires a detailed model of the instrument, both in geometric and material terms. Processes like rescattering or photon conversions, for instance, as well as a detailed account of the energy lost and deposited by the particles in the different parts of the detector (e.g. the gas that fills the Drift Chamber, or a crystal in the calorimeter) are the concern at this stage. Each of these interactions with the detector is recorded as a "gHit".

In the following stage, these "gHits" are used to simulate the data read out from the electronics of the detector, the trigger and the data acquisition system. Typical electronic noise and machine backgrounds characterizing a certain period of running of the experiment are then added. These are obtained by recording the detector's state at regular intervals (~ 1 Hz) during normal operations, and, due to their essentially random nature, they are unlikely to represent any physics event.

Finally, the simulated detector's electronic output is run through the same version of *BABAR*'s reconstruction software that is used on real data.

A related, widely used term, is *toy MC*, by which we refer to events simulated with the highly simplified model employed in the analysis to extract the relevant physics quantities from signal and background. Only a few of the relevant variables are usually taken into account and many of their possible correlations are neglected. Detector response effects are also often ignored, or modelled in a highly abstract manner, avoiding all the details of the passage of the particles through the detector and the reconstruction software.

4.3 Reconstruction

The reconstruction of events is performed in two stages. The first one, known as Offline Prompt Reconstruction, consists of finding and reconstructing tracks and calorimeter clusters from hits in the Drift Chamber and the Silicon Vertex Tracker, and crystals with energy deposits in the Electromagnetic Calorimeter, respectively. Čerenkov photons and $\mathrm{d}E/\mathrm{d}x$ information are also processed at this stage and abstracted into "particle identification selectors". The second part of the process deals with the reconstruction of *composites*, objects that are not directly observed in the detector but can be inferred from the properties and correlations of their decay products, the best example being a B meson. "Candidates" for composites are formed from combinations of tracks and neutral objects, allowing the important vertexing of the B meson and the Δt measurement to be made at this point.

4.3.1 Tracking algorithms

Due to the axial magnetic field in which the inner parts of the detector are immersed, charged tracks follow helices and are described by five parameters, which we take to be defined at the point of closest approach (POCA) to the z-axis:

- d_0, the distance in the xy plane to the z-axis
- z_0, the coordinate along the z-axis
- ϕ_0, the azimuthal angle of the POCA
- λ, the dip angle of the track with respect to the transverse (xy) plane. It is related to the cylindrical polar angle θ via $\theta = \pi/2 - \lambda$
- ω, the (signed) curvature of the track, whose sign and magnitude are related, respectively, to the charge of the associated particle and its transverse momentum, $\omega \propto 1/p_t$.

The tracks are fitted using a Kalman filter technique [73], that essentially fits each vertex independently, and iteratively and recursively propagates the changes in the parameters to the neighbouring vertices. Although the result is a global fit, the local character of each step of the algorithm allows corrections to be made that model the fine detail of the material distribution of the detector, the slight inhomogeneities of the magnetic field or the energy loss of low momentum tracks.

The algorithm starts from the DCH hits found by the Level 3 Trigger to form a track, and further hits are added if they are observed to be consistent with that track. Once the process is finished, the remaining hits in the DCH are searched for tracks that may not have originated at the beamspot (like K_s^0 or Λ, that live long enough to have their decay vertices outside the SVT), or may not be energetic enough to traverse the whole chamber. Afterwards, SVT hits are examined and added to the existing DCH tracks if possible, and are otherwise searched to locate any low momentum, SVT-only tracks.

The reconstructed tracks are then classified and stored in lists according to different selection criteria. In the present analysis, the pion candidates are required to meet the conditions of the `GoodTracksLoose` list:

- $p_t > 0.1\,\mathrm{GeV}/c$
- $p < 10.0\,\mathrm{GeV}/c$
- at least 12 hits in the Drift Chamber
- $d_0 < 1.5$ cm
- $|z_0| < 10$ cm

The K_s^0 candidates are formed from any two oppositely charged tracks (often not meeting the `GoodTracksLoose` requirements above), assumed to be pions, and whose mass, once vertexed, is within $25\,\mathrm{MeV}/c^2$ of the PDG value [22].

4.3.2 Calorimeter algorithms

The interactions of particles in the Electromagnetic Calorimeter typically result in *showers*, with the deposited energy spreading over neighbouring crystals. Each group of crystals, known as a *cluster*, might be due to the impact or passage of more than one particle and hence present energy distributions with several maxima. The aim of the calorimeter reconstruction routines is to locate and extract the right shape of the clusters, and to identify and correctly assign the energy to all the maxima within them.

The algorithm first looks for crystals with energies greater than 10 MeV, that will be used as 'seeds' for cluster formation. Surrounding crystals containing above 1 MeV themselves, or being neighbours of other crystals with more than 3 MeV are added to the cluster.

Local maxima are found by standard methods, and are assigned a fraction of the energy of each crystal in the cluster that depends on the ratio of the distance from the crystal to the maximum, and the Molière radius.

Finally, tracks are projected onto the calorimeter, and if their position and entrance angle are consistent with one of the maxima, they are linked and considered as a single particle in the following reconstruction routines. The remaining maxima are assumed to be neutral objects and placed in lists analogous to the `GoodTracksLoose` list described above.

4.3.3 Particle Identification

There are five common types of charged, long-lived particles that can be tracked in the *BABAR* detector: electrons, muons, pions, kaons and protons. Their correct identification is paramount for the physics goals of a B Factory, and this can be achieved thanks to the different ways in which those particles interact with each part of the detector. Information from all the sub-detectors (SVT, DCH, EMC, DIRC and IFR) is gathered to form PDFs (see Section 4.6) that represent the likelihood of a track belonging to a certain species. Since electrons and muons can often be separated easily from the other types of particles by their behaviour in the Electromagnetic Calorimeter and the Instrumented Flux Return, respectively, and protons are quite scarce, we will focus on the kaon-pion separation.

The likelihood for kaon and pion hypotheses is constructed as the product of the PDFs from the SVT, the DCH and the DIRC for the given particle hypothesis, where the first two contribute with measurements of the rate of energy loss ($\mathrm{d}E/\mathrm{d}x$) and the last one, with an estimation of the angle with respect to the track at which photons are emitted in the quartz bars of the DIRC. For both the DCH and the SVT, the measured energy loss of each track is compared with the Bethe-Bloch [22] expectation[1] by forming the pull (see Section 4.6), and is parameterized with a gaussian and a gaussian with asymmetric widths, respectively. The DIRC suffers from long non-gaussian tails that prevent the use of a similar method, so a binned likelihood is calculated instead with the help of MC. This likelihood depends on the angle of the Čerenkov photons with respect to the track, and also on the number of photons, since the latter is a function of the momentum and type of the particle, and it helps to improve the identification of low momentum tracks.

Once the likelihoods for the different particle hypotheses have been calculated, cuts on their values are applied, and the track is entered into different lists according to the criteria satisfied: `VeryLoose, Loose, Tight` and `VeryTight` for pions, and `NotPion, VeryLoose, Loose, Tight` and `VeryTight` in the case of kaons. For the analysis described in this thesis, `piLoose` was chosen (see Section 5.4). The efficiency of such a requirement on pions is above 95% in most of the kinematical range, and around $5-15\%$ for kaons is $\sim$ 10% depending on the momentum. The higher intrinsic fraction of pions over kaons in hadronic events, and the additional cuts on background discriminating variables, mean that we expect a negligible kaon contamination in our sample.

[1]In practice, rather than using the theoretical expectation, an empirical parameterization of the expected value is used.

4.3.4 Vertexing of candidates

Candidates for composite (not observable) particles are first formed from all the possible combinations of tracks and neutral objects matching the decay daughters of the particle. These candidates are then required to meet some kinematical criteria, and are subsequently vertexed. For instance, in the reconstruction of the charmless decays $B^0 \to K_S^0\pi^+\pi^-$, any intermediate states that may appear, like ρ^0 or $K^{*+}(892)$, are governed by the strong force and have such short lifetimes that their decay daughters may be assumed to originate from the B meson. Hence, there can only be candidates for the K_S^0 and the B mesons, and those of the latter are chosen as all the combinations of any two oppositely charged tracks in the event and a K_S^0 candidate, whose total centre-of-mass energy is in the range $(4.99, 5.59)$ GeV.

Once a kinematical candidate has been found, its decay vertex is calculated by means of a geometric fit, in which the tracks of the daughters are required to emerge from a common vertex. Such is the task of the `TreeFitter` package, which performs a global fit to the whole decay chain by applying the Kalman filter technique. An estimation of the interaction point, obtained from a fit to all tracks in the event, is used as the first guess in the iterative procedure for the B vertex reconstruction.

Constraints may be applied in the fit to reduce detector resolution effects. Indeed, the impact of these is higher in Dalitz analyses, since it may cause events to lie outside the kinematical boundaries of the Dalitz plot (i.e. the 4-momenta of the three daughters do not add up to that of the parent B meson). Final-state radiation can result in similar problems. To address these issues, and improve the resolution of structures in the Dalitz plot, two vertexing fits are done to the B candidates. The first one is unconstrained, and the kinematical background-rejecting variables m_{ES} and ΔE are extracted from it (see Section 4.5). In the second fit, in which the composite is constrained to have the nominal B meson mass [22], the Dalitz-plot coordinates and the related event shape variables are calculated.

4.4 Δt measurement and resolution

4.4.1 Measurement of Δz

As mentioned earlier, the measurement of the time difference between the decays of the two B mesons is made by first determining the separation along the z-axis between the decay vertices of the two particles. This sets a constraint on the resolution that should be achieved, as the distance $\beta\gamma c\tau_{B^0} = 257\mu m$ should be able to be measured comfortably. The vertex of the fully reconstructed meson, $B_{\rm rec}$, is fairly well determined by vertexing its daughters (see Section 4.3.4). The remaining tracks are used to reconstruct the other B, $B_{\rm tag}$, but in contrast with $B_{\rm rec}$, this meson has to be reconstructed inclusively, in a general way, in order to keep the efficiencies at a reasonable level. This inability to reconstruct $B_{\rm tag}$ fully is the cause of the poor resolution for the $B_{\rm tag}$ vertex, which dominates the overall resolution in Δz. Indeed, although the RMS of the resolution in z for $B_{\rm rec}$ is $\sim 65\mu m$ for more than 99% of the candidates (and as good as $\sim 45\mu m$ for more than 80%), the RMS of the residual in Δz (fitted minus true value) is $\sim 190\mu m$ ($\sim 150\mu m$ for 99% of the events). Note that these numbers compare well with the necessary resolution quoted before. We have quoted the resolutions in this way because a few of the decay vertices are not correctly determined, leading to long, very widely spread tails in the residuals distribution, comprising a very small fraction of events ($< 1\%$).

The algorithm used for the reconstruction of the $B_{\rm tag}$ vertex replaces the daughters of long-lived particles, such as K^0_S and Λ, by the trajectories of the composites in an effort to avoid the bias and tails that using tracks coming from a displaced vertex would cause. Tracks consistent with photon conversions are also removed. The daughters of D mesons and other charmed particles, whose lifetimes are comparable to those of the B mesons, can still introduce those undesired effects. In an effort to avoid them, an iterative procedure is used to find the $B_{\rm tag}$ vertex, in which the track that brings in the highest increment in the χ^2 of the fit is excluded, and the fit is redone if the increase in χ^2 is greater than 6 units. This procedure is repeated until all tracks satisfy the criterion. The adverse effects of the tracks originating from charmed particles and other displaced vertices are thus reduced but not completely removed, and they have to be properly accounted for in the modelling of the Δt resolution (see Section 4.4.3).

In cases like $B^0 \to K_S^0\pi^+\pi^-$, where there are no neutrinos and so B_{rec} is fully reconstructed, further use can be made of the quantities extracted from its fit to improve the measurement of Δz. In particular, its momentum and the beam spot position can be used to form a pseudo-trajectory for B_{tag}, thus providing an additional constraint to fit the tagging candidate daughters to a common vertex. This is demonstrated in the pictorial explanation of the tagging process given in Figure 1.1.

Finally, it is worth stating that, even though Δz is constructed as the difference between the positions of two vertices, it is mostly insensitive to the displacements of the vertices with respect to the beam spot.

4.4.2 Determination of Δt

In the approximation of negligible B momenta in the $\Upsilon(4S)$ rest frame, Δt is given by

$$\Delta t = \frac{\Delta z}{\beta\gamma c}\,, \tag{4.5}$$

where $\beta\gamma$ is the boost factor of the $\Upsilon(4S)$ resonance in the laboratory frame. This factor is calculated from the beam energies, which are monitored continuously, and has a value of $\beta\gamma = 0.56$ for the current configuration of *BABAR*.

However, the B mesons do have a small momentum in the $\Upsilon(4S)$ rest frame, $p_B^* = 340\,\mathrm{MeV}/c$. Furthermore, this momentum can be measured in the case of a fully reconstructed B_{rec} meson, as discussed in the previous section, and used to correct Eq. (4.5):

$$\Delta z = \beta\gamma\gamma^*_{rec}c\,(t_{rec} - t_{tag}) + \gamma\beta^*_{rec}\gamma^*_{rec}\cos\theta^*_{rec}c\,(t_{rec} + t_{tag}) \tag{4.6}$$

where $\gamma^*_{rec} = 1.002$, $\beta^*_{rec} = 0.064$ and θ^*_{rec} are, respectively, the boost factor of the reconstructed B meson in the centre-of-mass frame, its speed, and its angle with respect to the z axis. The problem with the last expression is that the quantity $t_{rec} + t_{tag}$ is not directly measured. It can be related to the transverse displacement of one B meson with respect to the other, but its value is rather small ($\sim$ 35-40μm) compared to the resolution achieved for the B_{tag} vertex. Instead, we take its average value[2]:

$$\langle t_{rec} + t_{tag} \rangle = \tau_B + |\Delta t| \tag{4.7}$$

giving

$$\Delta z = \beta\gamma\gamma^*_{rec} c \left(t_{rec} - t_{tag}\right) + \gamma\beta^*_{rec}\gamma^*_{rec} \cos\theta^*_{rec} c \left(\tau_B + |\Delta t|\right) \tag{4.8}$$

which can be solved for Δt. The value of Δt is corrected by only $\sim 0.02\,\text{ps}$ relative to Eq. (4.5). The use of Eq. (4.8) improves the resolution for Δt by about 5% and removes a correlation existing in signal data between the true value of Δt and its resolution, although at the cost of introducing a similar correlation for the less understood continuum background events.

4.4.3 Δt resolution model

The resolution in Δt is modelled as the sum of three gaussians, known as the core, tail and outlier components:

$$\begin{aligned} \mathcal{R}_{sig}(\delta t,\ \sigma_{\Delta t}) &= (1 - f_{\text{tail}} - f_{\text{outlier}}) G\left(\delta t,\ b_{\text{core}}\sigma_{\Delta t},\ s_{\text{core}}\sigma_{\Delta t}\right) \\ &\quad + f_{\text{tail}} G\left(\delta t,\ b_{\text{tail}}\sigma_{\Delta t},\ s_{\text{tail}}\sigma_{\Delta t}\right) + f_{\text{outlier}} G\left(\delta t,\ b_{\text{outlier}},\ s_{\text{outlier}}\right) \end{aligned} \tag{4.9}$$

where $\delta t = \Delta t - \Delta t_{\text{true}}$ is the residual, $\sigma_{\Delta t}$ is the event-by-event error on Δt extracted from the fit and the G functions represent normalized gaussians:

$$G\left(\delta t,\ \mu,\ \sigma\right) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\left(\delta t - \mu\right)^2}{2\sigma^2}\right). \tag{4.10}$$

The dependence of the means and the widths on the error $\sigma_{\Delta t}$, introduced via the scaling factors $b_{\text{core(tail)}}$ and $s_{\text{core(tail)}}$, is due to an observed correlation between $\sigma_{\Delta t}$ and the means and RMS of the Δt residual distributions caused by charmed decay

[2]Its conditional average for a fixed Δt, actually (see [74] for details). Note that, as expected from the LHS, the RHS of Eq. (4.7) is also approximately equal to $2\tau_B$, since $|\Delta t|$ is the time that one of the mesons takes to decay after the other one has decayed, and therefore satisfies $|\Delta t| \sim \tau_B$.

daughters. The effect is bigger when the charmed meson flies along the z direction than in the xy plane, since the error and the bias brought about by vertexing particles coming from a secondary displaced vertex affects the determination of the z component of the $B_{\rm tag}$ vertex directly only in the first case. Thus, a correlation between larger values of $\sigma_{\Delta z}$ and the bias in Δz is produced. This trend is observed to be less strong in events with a lepton tag, so the b and s parameters are taken to be tagging category-dependent quantities.

The outlier gaussian, which is indepedent of $\sigma_{\Delta t}$, describes the small fraction of events ($< 1\%$) for which at least one of the two vertices is badly reconstructed.

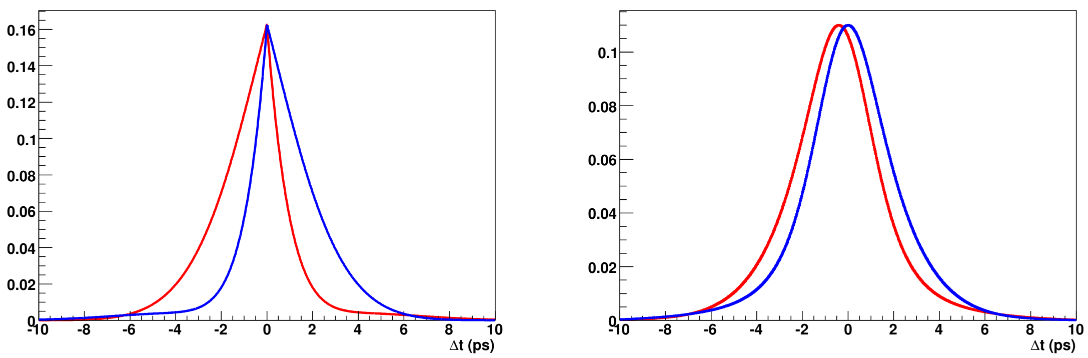


Figure 4.2: Δt distributions for B^0 (red) and $\overline{B}^0$ (blue) tagged events with perfect tagging and resolution (left), and with typical *BABAR* mistag rates and resolution effects incorporated (right). A simple one gaussian model has been used in the demonstration of the resolution effects.

Although the parametrization described above has been developed in comprehensive studies for the charmonium sin2β analyses [23], we adopt the same model for our work in the decay mode $B^0 \rightarrow K_S^0\pi^+\pi^-$. This is justified because the dominant factor in the Δt error is the poor determination of the $B_{\rm tag}$ vertex, and the algorithm to find this is independent of the channel into which $B_{\rm rec}$ decays, so it is a good approximation to consider the resolution in Δt to be the same for all channels. The effects that a typical resolution model has on the Δt distribution have been illustrated in Figure 4.2.

See also Table 4.2 for the values of the resolution parameters used in $B^0 \to K^0_S\pi^+\pi^-$. As for the mistag rates (Section 4.1.1), these are extracted from a fit to the B_{flav} sample.

Table 4.2: Parameters that describe the resolution in Δt for signal events, extracted from the B_{flav} sample for the charmonium sin2β analyses [24], and used for $B^0 \to K^0_S\pi^+\pi^-$. The parameters from all tagging categories except `Lepton` were found to be consistent. Extensive MC studies also suggested which parameters to keep fixed in the fit. The scaling constants $s_{\text{core(tail)}}$ (> 1) correct for an overall underestimation of $\sigma_{\Delta t}$.

	`Lepton`	Other categories
b_{core}	-0.0789 ± 0.0304	-0.1850 ± 0.0138
s_{core}	1.0610 ± 0.0459	1.1059 ± 0.0223
f_{core}	0.8916 ± 0.0088	
b_{tail}	-1.1186 ± 0.1354	
s_{tail}	3.0 fixed	
f_{outlier}	0.0034 ± 0.0006	
b_{outlier}	0.0 fixed	
s_{outlier}	8.0 ps^{-1} fixed	

4.4.4 Implementation of resolution effects in Δt

Once an accurate model of the Δt resolution has been produced, its implementation requires convolving the expected PDFs with the resolution to obtain the *observed* distributions:

$$\begin{aligned}
\mathcal{P}^{\text{observed}}(x,\ y,\ \Delta t) &= \left(\mathcal{P}^{\text{theory}} \otimes \mathcal{R}_{sig}\right)(x,\ y,\ \Delta t) && (4.11)\\
&= \int_{-\infty}^{+\infty} \mathcal{P}^{\text{theory}}(x,\ y,\ \Delta t_{\text{true}})\, \mathcal{R}_{sig}(\Delta t - \Delta t_{\text{true}}) d\Delta t_{\text{true}}\ . && (4.12)
\end{aligned}$$

Since the resolution is modelled as a sum of three gaussians, inspection of the expected signal distribution Eq. (2.19), which has the form

$$f_{q_{\rm tag}}(\Delta t) = \frac{e^{-|\Delta t|/\tau}}{4\tau}\Big(1 + q_{\rm tag} S \sin(\Delta m_d \Delta t) - q_{\rm tag} C \cos(\Delta m_d \Delta t)\Big), \tag{4.13}$$

leads us to consider the following convolution integrals:

$$e^{-|x|/\tau} \otimes G(x,\mu,\sigma), \tag{4.14}$$

$$e^{-|x|/\tau} e^{ikx} \otimes G(x,\mu,\sigma). \tag{4.15}$$

These can be calculated analytically once the following definitions have been introduced:

$$\mathrm{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^{+\infty} e^{-t^2} dt \tag{4.16}$$

$$\mathrm{w}(z) = e^{-z^2}\mathrm{erfc}(-iz)\ . \tag{4.17}$$

These are, respectively, the "complementary error function" and the "complex error function" [75]. Writing Eq. (4.14) and (4.15) in term of these definitions gives

$$\begin{aligned} e^{-|x|/\tau} \otimes G(x,\mu,\sigma) &= \frac{1}{2}\exp\left(\frac{\sigma^2}{2\tau^2} + \frac{x-\mu}{\tau}\right)\mathrm{erfc}\left(\frac{\sigma}{\tau\sqrt{2}} + \frac{x-\mu}{\sigma\sqrt{2}}\right) \\ &+ \frac{1}{2}\exp\left(\frac{\sigma^2}{2\tau^2} - \frac{x-\mu}{\tau}\right)\mathrm{erfc}\left(\frac{\sigma}{\tau\sqrt{2}} - \frac{x-\mu}{\sigma\sqrt{2}}\right) \end{aligned} \tag{4.18}$$

and

$$\begin{aligned} e^{-|x|/\tau} e^{ikx} \otimes G(x,\mu,\sigma) &= \frac{1}{2}e^{-(x-\mu)^2/2\sigma^2}\left[\mathrm{w}\left(-\frac{k\sigma}{\sqrt{2}} + i\left(\frac{\sigma}{\tau\sqrt{2}} + \frac{x-\mu}{\sigma\sqrt{2}}\right)\right)\right. \\ &\qquad \left. + \,\mathrm{w}\left(+\frac{k\sigma}{\sqrt{2}} + i\left(\frac{\sigma}{\tau\sqrt{2}} - \frac{x-\mu}{\sigma\sqrt{2}}\right)\right)\right]. \end{aligned} \tag{4.19}$$

Computational calculations of these convolutions rely on the implementation of the complementary error function from `ROOT` and on our own implementation of the algorithm [76] for the complex error function. However, numerical problems due to large cancellations arise for parts of the range of interest. Following `RooFit` [77], they are taken care of by using

$$\mathrm{w}(z) \simeq 2e^{-z^2}\left(1 + \frac{e^{z^2}}{iz\sqrt{\pi}}\right) \tag{4.20}$$

when Im $(z) \leq 4.0$, thus cancelling explicitly with the exponential on the RHS of Eq. (4.19) the divergent behaviour of the factor $\exp\left((x-\mu)^2/2\sigma^2\right) = \exp - \left(i^2 \Im^2(z)\right)$ that appears when writing the complex error functions from Eq. (4.19) in terms of the complementary error function (Eq. (4.16)).

Finally, the normalization can also be derived analytically. Although in the limit of an infinite range in Δt, the normalization reduces to the constant 2τ and can be dropped, for finite intervals the tails due to resolution effects get cut off. This issue would only give a constant correction to 2τ, were it not for the dependence of the resolution on the per-event quantity $\sigma_{\Delta t}$, which means that the normalization must also be calculated on a per-event basis.

To ease the task, we observe that the form of the Δt distribution, Eq. (4.13), suggests the following reasoning: since q_{tag} can take the values ± 1, and the normalization of the total PDF involves a sum over q_{tag} as well as the integration over Δt, we calculate the total normalization, in which the terms multiplied by q_{tag} cancel and only the convolved exponential term remains[3]. Hence, if $(-t_0, t_0)$ is the allowed range for Δt,

$$\mathrm{norm} = \int_{-t_0}^{+t_0} e^{-|x|/\tau} \otimes G(x, \mu, \sigma)\, dx \tag{4.21}$$

which gives

$$\begin{aligned}
\mathrm{norm} = {} & \tau \left\{ \mathrm{erf}\left(\frac{t_0+\mu}{\sigma\sqrt{2}}\right) + \mathrm{erf}\left(\frac{t_0-\mu}{\sigma\sqrt{2}}\right) \right\} \\
& + \frac{\tau}{2} e^{\frac{\sigma^2}{2\tau^2}+\frac{\mu}{\tau}} \left\{ e^{\frac{t_0}{\tau}} \mathrm{erfc}\left(\frac{\sigma}{\tau\sqrt{2}} + \frac{\mu+t_0}{\sigma\sqrt{2}}\right) - e^{-\frac{t_0}{\tau}} \mathrm{erfc}\left(\frac{\sigma}{\tau\sqrt{2}} + \frac{\mu-t_0}{\sigma\sqrt{2}}\right) \right\} \\
& + \frac{\tau}{2} e^{\frac{\sigma^2}{2\tau^2}-\frac{\mu}{\tau}} \left\{ e^{\frac{t_0}{\tau}} \mathrm{erfc}\left(\frac{\sigma}{\tau\sqrt{2}} + \frac{-\mu+t_0}{\sigma\sqrt{2}}\right) - e^{-\frac{t_0}{\tau}} \mathrm{erfc}\left(\frac{\sigma}{\tau\sqrt{2}} + \frac{-\mu-t_0}{\sigma\sqrt{2}}\right) \right\}.
\end{aligned} \tag{4.22}$$

These formulae and their implementations were tested extensively against numerical calculations over a range of values for the lifetime τ and resolution width σ parameters, and in particular for the typical values in *BABAR*.

[3]In other words, the probability of observing an event regardless of its tag is given by an exponential, resolution effects apart, as we expect due to the exponential law for the decay of particles. We can therefore calculate the overall normalization by integrating the exponential rather than adding the normalizations of the PDFs for each tag separately.

4.5 Discriminating variables

The small branching fractions of charmless B decay modes like the one studied in this thesis (for $B^0 \to K_S^0 \pi^+ \pi^-$, $(2.24 \pm 0.13) \times 10^{-5}$ [22]) and the high cross sections for the undesired processes $e^+e^- \to q\bar{q}, \quad q = u, d, s, c$ ($\sigma_{udsc} \sim 3.2 \times \sigma_{b\bar{b}}$), make the use of background-rejecting variables unavoidable. By making use of the differences between the distributions of signal and background events in these variables, statistical separation of the two species can be attained. These differences can be taken advantage of in two ways. If the densities of events for the two types peak at different points of the range in the variable considered, a "cut" on the variable may be imposed, rejecting all events that lie on one side of the cut value, and enriching the sample with signal events. The other approach consists of accepting all events, and assigning each of them a weight or probability of belonging to each species based on their value for the discriminating variable. In the present analysis, a mixed strategy has been followed: loose cuts are applied on the three discriminating variables, m_{ES}, ΔE and $\mathcal{F}$, but their distributions are also used in the fit to help determine the number of events of each species.

We now proceed to describe in detail the variables mentioned above.

4.5.1 Kinematic variables

Two kinematic variables, largely uncorrelated [78], are defined to help discriminate signal and background: the energy-substituted mass, m_{ES}, and the difference between the reconstructed centre-of-mass energy of the B candidate and half the total centre-of-mass energy, ΔE:

$$m_{\mathrm{ES}} = \sqrt{(s/2 + \vec{p}_i \cdot \vec{p}_B)^2 / E_i^2 - \vec{p}_B^{\,2}}\,, \qquad \Delta E = E_B^* - \sqrt{s}/2\,, \tag{4.23}$$

where $\vec{p}_B$ is the momentum of the B candidate, $\sqrt{s}$ denotes the centre-of-mass energy and $(E_i, \vec{p}_i) \equiv p_i$ is the four-momentum of the initial state (the electron-positron system). The mass of the B candidate calculated from the kinematic constraints, m_{ES}, is used rather than simply $\sqrt{E_B^2 - \vec{p}_B^{\,2}}$. The reason is that the candidate is formed from a number of tracks and neutral objects whose energies are not as accurately measured as are the beam conditions. Therefore, a great improvement in the mass resolution is achieved by using our knowledge of the initial kinematics. Since the electron and the positron annihilate creating a $B\overline{B}$ pair, and the masses of the particle and of the antiparticle are equal,

$$p_B^2 = p_{\overline{B}}^2 = (p_i - p_B)^2 = p_i^2 + p_B^2 - 2p_i p_B \Rightarrow 0 = p_i^2 - 2p_i p_B \Rightarrow s/2 = -E_i E_B + \vec{p}_i \cdot \vec{p}_B$$
$$\Rightarrow m_B = \sqrt{E_B^2 - \vec{p}_B^{\,2}} = \sqrt{\left(\frac{s/2 - \vec{p}_i \cdot \vec{p}_B}{E_i}\right)^2 - \vec{p}_B^{\,2}} = m_{\mathrm{ES}}.$$

Thus, for signal events, m_{ES} yields the mass of the B meson and shows a clean peak. For continuum events, composed of light quarks, the only way of reaching the B rest mass is by artificially associating random tracks. As a consequence, their distribution displays the slowly varying shape that one could expect from their combinatoric nature.

The idea behind ΔE is different and complementary to that of m_{ES}. Whereas the latter is by construction independent of the mass hypotheses for each of the tracks, ΔE depends strongly on them. If, for example, a kaon is misidentified as a pion, its energy $\sqrt{\vec{p}^{\,2}_{measured} + m^2_{hypothesis}}$, and consequently that of the B candidate, will be smaller than its due share, and the event will be shifted towards negative values of ΔE. In contrast, the distribution for signal events peaks at zero as expected, making ΔE especially helpful in isolating backgrounds from misreconstructed B decays.

Plots of both variables for signal and background can be seen in Figures 5.7 and 5.22.

4.5.2 Event-shape variables

Event-shape variables, also known as topological variables, aim to exploit the angular correlations among the decay products in $B\overline{B}$ and $q\overline{q}$ events to further help the separation of the two species.

In $q\overline{q}$ events ($e^+e^- \to q\bar{q}, \quad q = u, d, s, c$), known as *continuum*, the small amount of energy invested in the rest masses of the quarks means that most of the available centre-of-mass energy will be carried as kinetic energy. This, in turn, implies that the event will have a two-jet-like structure, roughly following a $(1 + \cos^2 \theta)$ dependence, where θ is the centre-of-mass angle of a jet with respect to the beam axis. This is as predicted by lowest order Feynman diagram for the quantum electrodynamical (QED) process of annihilation of an electron-positron system to produce a fermion-antifermion pair. Indeed, since the typical energy scale for strong interactions is far smaller than the available kinetic energy, $\Lambda_{\mathrm{QCD}} \ll m_B - m_{q\bar{q}}$, hadronization and other QCD effects are not expected to alter greatly the expected QED angular dependence.

The $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ process, in contrast, is charaterized by the decay of the vector resonance $\Upsilon(4S)$ into two pseudo-scalars, resulting in a $\sin^2 \theta$ distribution, where θ is the angle between the momentum of one of the B mesons and the beam axis. Furthermore, the reaction is barely allowed kinematically, with very little of the centre-of-mass energy converting into kinetic energy of the B mesons. Their average momenta, $\sim 340\,\mathrm{MeV}/c$, are in fact smaller than the typical momenta of their daughters, $\sim$ 1-2 GeV/c, which means that the decay products of a bottom meson will not be boosted enough to follow the flight direction of their parent. Hence the angular distribution of the decay products will be roughly preserved. Since the B mesons are pseudo-scalars, they decay isotropically and the distributions of their daughters in the $\Upsilon(4S)$ centre-of-mass frame will be approximately spherical.

We will now describe a few variables that put to good use the differences explained above. In the definitions that follow, it is useful to distinguish between the reconstructed B side of the event, and the Rest Of the Event (ROE), that comprises all tracks, composites and neutral objects that do not make up the B candidate.

Angle between the B momentum and the beam axis

As noted before, the distribution of the B momentum direction with respect to the beam axis for $B\overline{B}$ events has a parabolic shape, $\sin^2 \theta_{B_{mom}} = 1 - \cos^2 \theta_{B_{mom}}$. For continuum events, in contrast, kinematically appropriate B candidates can only be formed from random combinations of tracks (often referred to as *combinatoric* background) and as a consequence, $\cos \theta_{B_{mom}}$ will also take random values. Hence, the distribution

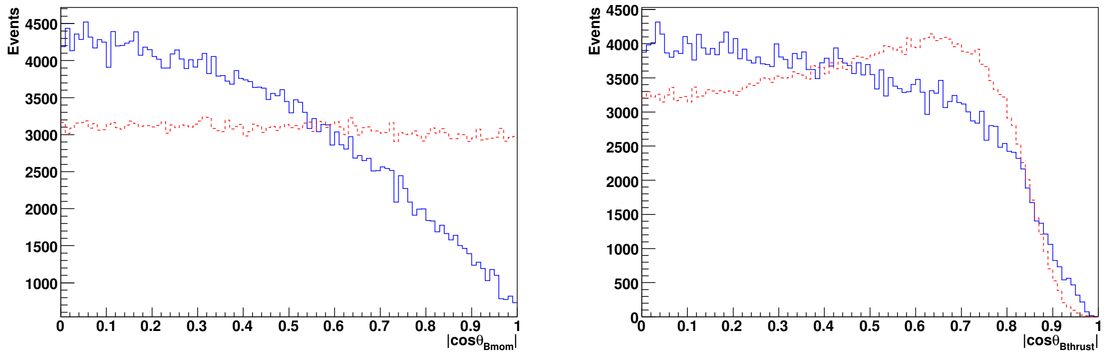


Figure 4.3: Distributions of $|\cos\theta_{B_{mom}}|$ and $|\cos\theta_{B_{thrust}}|$ for non-resonant $B^0 \to K^0_S\pi^+\pi^-$ signal MC (blue, solid) and off-peak data (red, dashed). The shapes seen in the figure on the right differ somewhat from those discussed in the text due to detector acceptance effects.

is expected to be uniform (see Figure 4.3).

Thrust Axis variables

The thrust axis of a collection of particles is defined as the direction in which the sum of the projections of the momenta of the particles is maximized:

$$\text{thrust axis } \hat{n}\colon \ \max \sum_i |\hat{n}\cdot\vec{p_i}|,\ |\hat{n}| = 1\,, \tag{4.24}$$

where the i index runs over all the particles in the collection. Given the spherical nature of B decays, the thrust axis of a true B candidate is essentially random. For continuum events, however, which are strongly collimated, the above definition ensures that the thrust axis approximates the direction along which the pair of quarks was emitted, even when the tracks are selected artificially to form a kinematical B candidate.

Several variables can be defined employing the thrust axis, such as the cosine of the angle between the thrust axis of the B candidate and the z axis, $\cos\theta_{B_{thrust}}$, or the cosine of the angle between the thrust axes of the B candidate and the rest of the event (Figure 4.3). Comprehensive studies [79] showed the latter to be strongly correlated with the variables described in the following subsection, so only $\cos\theta_{B_{thrust}}$ is used in this analysis.

The ROE and Legendre polynomials

We can take further advantage of the marked differences in the angular distributions of the momentum flow in signal events ($\propto \sin^2\theta$) and continuum events ($\propto 1+\cos^2\theta$) to refine the selection criteria. Since our analysis explores the whole allowed phase space of a three-body decay, we cannot make use of the signal-side angular information on that aspect without biasing our sample, but we can exploit the fact that the other B in the event behaves statistically, but independently, in the same way. Furthermore, a good way of characterizing the angular correlations of the rest of the event is to calculate the components of the momentum distribution in the basis formed by the Legendre polynomials.

Indeed, a calculation of the expectation values of the Legendre polynomials for the signal and background momentum distributions enables us to identify the order of the polynomials with the largest separation power. These turn out to be the zeroth and the second order, whose expectation values are non-zero and different. The remaining orders have vanishing expectation values (see Figure 4.4). They are defined as follows:

$$L_0 = \sum_i^{ROE} p_i \tag{4.25}$$

$$L_1 = \sum_i^{ROE} p_i \cos\theta_i \tag{4.26}$$

$$L_2 = \sum_i^{ROE} p_i \frac{1}{2}\left(3\cos^2\theta_i - 1\right) \tag{4.27}$$

where p_i and θ_i are the momentum and the angle with respect to the beam axis of the i-th track or neutral object in the rest of the event. The ratio of the two polynomials, L_2/L_0, is used, instead of their individual values separately, as suggested by the studies performed in [79].

It should be noted that the intrinsic symmetry inherent to continuum events, due to their jet-like structure, leads to correlations between the values of their ROE and signal-side quantities. Since the signal side state is completely characterized by giving its Dalitz-plot coordinates, correlations between these and some discriminating variables, such as the Legendre polynomials evaluated from the ROE, should be expected for these events.

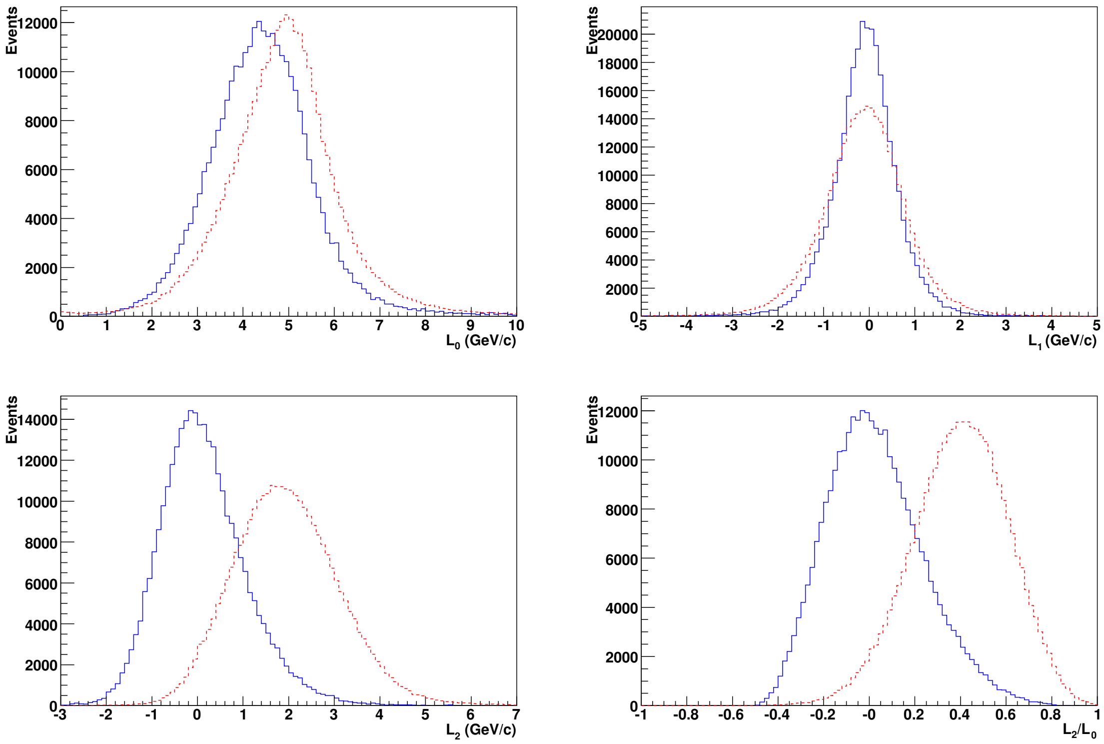


Figure 4.4: Distributions of the zeroth, first and second order components in the Legendre polynomials of the momentum flow in the rest of the event. The plot in the bottom, right-hand corner shows the distribution of the ratio of the second-order to the zeroth-order polynomials. Non-resonant $B^0 \to K^0_S\pi^+\pi^-$ signal MC is displayed in blue (solid) and off-peak data in red (dashed). It can be seen that the expectation value of L_1 is zero for both species.

4.5.3 Fisher discriminant

If several discriminating variables are used, they are likely to be correlated to some degree, especially if they exploit different aspects of the same physical principle, such as those mentioned above. Cutting on one of them will then imply a cut on the others, making it hard to find the optimal combined set of cuts. Instead, a unique variable may be defined which takes full advantage of all the discriminating variables and their correlations to maximize the separation power. The Fisher discriminant ($\mathcal{F}$) [80] is designed to do precisely that, although the maximization is only guaranteed in the case in which the two populations can be separated by means of a linear boundary. It is defined as a linear combination of the variables x_i under consideration:

$$\mathcal{F} = \sum_i a_i x_i = \vec{a}^T \vec{x}\,, \tag{4.28}$$

where the coefficients $\vec{a}$ are found by maximizing the separation between the signal and background distributions. This is defined to be

$$D(\vec{a}) = \frac{(\bar{\mathcal{F}}_S - \bar{\mathcal{F}}_B)^2}{\sigma_S^2 + \sigma_B^2}\,, \tag{4.29}$$

$\bar{\mathcal{F}}_j$ being the means of the distributions, and σ_j^2 the variances. Writing these in terms of the means $\vec{\mu}_j$ and covariance matrices E_j of the variables,

$$\bar{\mathcal{F}}_j = \vec{a}^T \vec{\mu}_j, \quad \sigma_j^2 = \vec{a}^T E_j \vec{a}\,, \tag{4.30}$$

we get

$$D(\vec{a}) = \frac{\vec{a}^T (\vec{\mu}_S - \vec{\mu}_B)(\vec{\mu}_S - \vec{\mu}_B)^T \vec{a}}{\vec{a}^T (E_S + E_B) \vec{a}}\,, \tag{4.31}$$

which we can maximize by differentiating and equating to zero, giving

$$\vec{a} = (E_S + E_B)^{-1} (\vec{\mu}_S - \vec{\mu}_B). \tag{4.32}$$

Thus, by using signal and background control samples, such as signal MC and off-resonance data, for example, the coefficients $\vec{a}$ can be calculated.

Studies were carried out to find the optimal choice of background discriminating variables. The purpose was two-fold: firstly, to identify the optimal set of variables; and secondly, to establish whether a linear or a non-linear method would perform better. During the course of those studies, a number of additional variables were tried (more event-shape variables, the cosine of the angle between the thrust axes of the B candidate and the rest of the event, CLEO cones [81]) and rejected in favour of $\cos\theta_{B_{mom}}$, $\cos\theta_{B_{thrust}}$ and L_2/L_0 (see Figures 4.3 and 4.4 for the chosen input variables, and Figure 4.5 for $\mathcal{F}$) due to strong correlations among the variables themselves or with the Dalitz-plot position. Comparisons between the Fisher discriminant, a Neural Net [82] and a configuration of Boosted Decision Trees [83] were also made. The performance of the Fisher was seen to be similar to that of the Neural Net, and both of them were superior to the Boosted Decision Trees, although a smaller number of configurations of the tunable parameters were explored in the last two approaches.

4.6 Maximum Likelihood fits

Maximum likelihood fitting is a powerful method to estimate the parameters that characterize a given statistical distribution from a data sample representing it. A more complete discussion can be found in [84] and [82], for example.

The distribution taken by the outcomes of a sampling experiment (i.e. the values of a random variable x) can usually be described by a functional form $\mathcal{P}(x, \vec{a})$ whose shape is determined by some parameters $\vec{a}$. If the function $\mathcal{P}(x, \vec{a})$ is normalized, it is said to be a Probability Density Function (PDF) for x. Given a set of N measurements of the random variable, the problem consists of having the best possible estimations for the values of the parameters $\vec{a}$ that charaterize the PDF. The estimations provided by the maximum likelihood method are attained by seeking the values of $\vec{a}$ that maximize the so-called *likelihood function*:

$$\mathcal{L}\left(\vec{a}\right) = \prod_{i=1}^{N} \mathcal{P}\left(x_i, \vec{a}\right) \ . \tag{4.33}$$

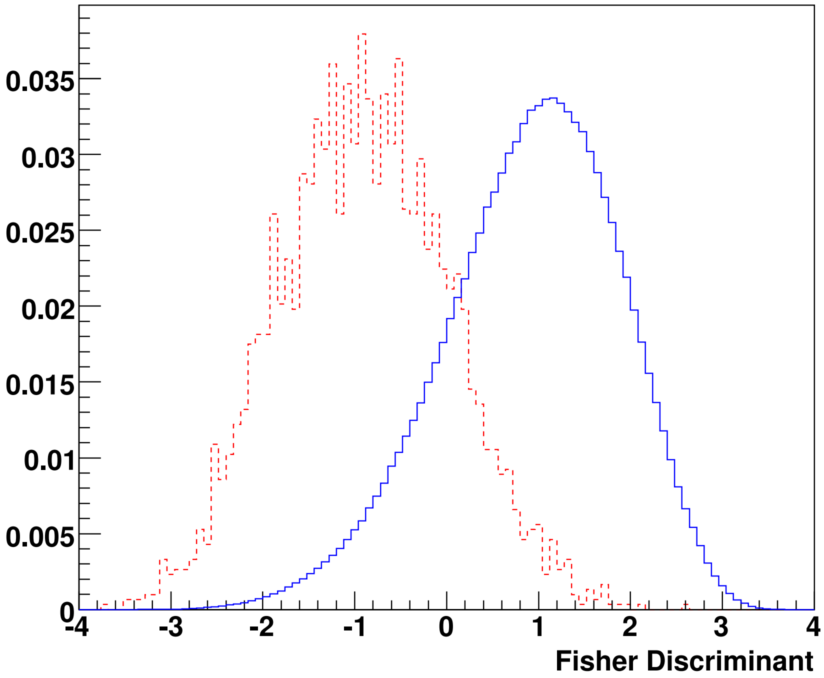


Figure 4.5: Normalized distributions of the Fisher discriminant for signal (blue, solid) and continuum background (red, dashed). The variables from which it is constructed are $|\cos\theta_{B_{mom}}|$, $|\cos\theta_{B_{thrust}}|$ and L_2/L_0. As stated in Section 5.4, a cut is applied on the discriminant ($\mathcal{F} > -0.365$) that retains 90.0% of the signal.

Intuitively, the likelihood function represents the probability of drawing the N measurements of the random variable given a certain set of values for the parameters $\vec{a}$, so optimizing this probability should yield the parameter values that best describe the sample. In the simple case in which the data follow gaussian distributions, it can be proved that the methods of maximum likelihood and the time-honoured least squares are equivalent. The former, though, is not limited to binned distributions.

The PDFs can be quite complicated, reflecting several hypotheses for the source of the measurement (e.g. whether it is signal or background), or the fact that the outcome of the experiment requires several random variables to be described:

$$\mathcal{P}\left(i, \vec{a}\right) = \sum_{j=1}^{M} \mathcal{P}_j\left(x_i, y_i; \vec{a}\right) = \sum_{j=1}^{M} \mathcal{Q}_j\left(x_i; \vec{a}_{\mathcal{Q}}\right) \mathcal{R}_j\left(y_i; \vec{a}_{\mathcal{R}}\right) \tag{4.34}$$

where M is the number of hypotheses, x_i and y_i are the outcomes of the i-th experiment and in the last equality it has been assumed that the two random variables are uncorrelated and, therefore, that their joint PDF $\mathcal{P}$ can be written as a product of their individual PDFs $\mathcal{Q}$ and $\mathcal{R}$.

A crucial point to the maximum likelihood method is the assumption that the PDFs are normalized. If this were not the case, the results would be distorted or meaningless, since changes to the parameters could increase the normalization without increasing the probability, leading the optimization process to converge on incorrect values, or even pushing the global maximum to infinity. Therefore, recalculation of the norms of the PDFs is often necessary on each iteration of the fit.

Eq. (4.33) is not usually applied as such, but with a slight modification to ease its computation. Taking logarithms, it can be rewritten as

$$\ell = -\log \mathcal{L} = -\sum_{i=1}^{N} \log \mathcal{P}\left(x_i, \vec{a}\right), \tag{4.35}$$

where the sum of logarithms is far more manageable in terms of machine precision than the previous product. The minus sign has been introduced so that the optimization of the likelihood function is performed by minimizing ℓ. In our likelihood fitting package, `Laura++` [85], this is carried out numerically via an interface to `Minuit` [86, 87] through `ROOT` [88].

4.6.1 Extended Maximum Likelihood fits

In particle physics, the number of events observed in an experiment is often unknown *a priori* and can be considered to be one of the outcomes of the measurement. In that case, the likelihood function must be appropriately modified to include the probability, given by the Poisson distribution, of having N occurrences when ν is the expected value:

$$\mathcal{L}(\nu,\vec{a}) = \frac{e^{-\nu}\nu^N}{N!}\prod_{i=1}^{N}\mathcal{P}(x_i,\vec{a}) = \frac{e^{-\nu}}{N!}\prod_{i=1}^{N}\nu\mathcal{P}(x_i,\vec{a}) \ , \tag{4.36}$$

which, dropping constant factors, is generalized for M species or hypotheses by

$$\mathcal{L}(\vec{n},\vec{a}) = e^{-\sum_{k=1}^{M} n_k}\prod_{i=1}^{N}\left(\sum_{j=1}^{M} n_j\mathcal{P}_j(x_i,\vec{a})\right) \ , \tag{4.37}$$

where n_j is the number of events for the hypothesis j.

4.6.2 Error estimation

There are several ways of calculating the errors on the estimations of parameters returned by a maximum likelihood fit, each of them involving different assumptions. Usually, the standard deviation σ, calculated as the square-root of the variance, is taken as the error on a parameter. In the large sample limit, the covariance matrix, and hence the errors, can be computed by inverting the matrix of the second derivatives of the likelihood function with respect to the parameters evaluated at the maximum:

$$\left(V^{-1}\right)_{ij} = -\left.\frac{\partial^2 \log\mathcal{L}}{\partial a_i \partial a_j}\right|_{\vec{a}=\vec{a}^0} \tag{4.38}$$

where $\vec{a}^0$ are the values returned by the fit.

Another method consists of defining the errors σ_i by the points $a_i^0 \pm \sigma_i$ in which the logarithm of the likelihood drops by 1/2:

$$\ell\left(a_i^0 \pm \sigma_i\right) = \ell\left(a_i^0\right) + \tfrac{1}{2} = \ell_{\max} + \tfrac{1}{2}. \tag{4.39}$$

This prescription is inspired by the fact that, when the second derivatives of the likelihood can be considered constant in the range given by $a_i^0 \pm \sigma_i$, the shape of the function at the minimum is well approximated by a gaussian, as can easily be seen by making a Taylor expansion of ℓ. The definition of the error then reduces to that of the width of the gaussian.

Finally, an assumption-free procedure consists of generating a large sample of MC experiments using the values returned from the fit, fitting them again, and calculating the standard deviation of the results for the estimated parameter.

4.6.3 Toy Monte Carlo

A good way to identify and assess potential problems in the maximum likelihood fit is to generate a large number of MC experiments with given PDFs, and fit them using the same PDFs, in line with the procedure described at the end of the previous section. That simple check enables us to evaluate any bias due to low statistics, a defective likelihood or a mistake in the calculations, and to correct it in the latter cases.

This toy MC (see Section 4.2) is generated using Von Neumann's acceptance-rejection method [82] with our fitting package `Laura++` [85] and should result in gaussian distributions around the true value. Furthermore, the so-called *pull* distribution can be constructed by evaluating

$$\text{pull} = \frac{a_i^{true} - a_i^{fit}}{\sigma_i^{fit}}\,, \tag{4.40}$$

where the numerator is referred to as the "residual". These pull distributions should be gaussian in shape, centred around zero, and with a unit width.

Chapter 5

Analysis

The next two chapters discuss the time-dependent amplitude analysis of $B^0 \to K^0_S\pi^+\pi^-$ in detail. This chapter revolves around the construction of the likelihood, and develops the methodology. A careful examination of all the species of events and ways of characterizing them is made, as well as of all the variables used to differentiate them. The following chapter presents the tests performed to ensure that the complex fit is handled correctly, describes how the model for the Dalitz structure is explored, and finally comments on the results obtained, comparing them with previous measurements.

5.1 Overview

A Dalitz or amplitude analysis aims to extract the relative magnitudes and phases of all the structures contributing to a three-body decay. In this analysis, it is achieved by performing an unbinned (extended) maximum likelihood fit to the data sample with the `Laura++` software package [85]. The likelihood function for a signal event is rather similar to the expected decay rate for signal, Eq. (2.19), although with the normalization and the dependence of the Δt resolution and the dilutions on the tagging category accounted for (see Section 2.2.2 for a reminder on the notation):

$$
\begin{aligned}
\mathcal{L}\left(m^2_{K^0_S\pi^+}, m^2_{K^0_S\pi^-}, \Delta t, q_{\rm tag}\right) &= \frac{1}{\mathcal{N}}\sum_c f_c \frac{e^{-|\Delta t|/\tau_{B^0}}}{4\tau_{B^0}} \times \\
&\quad \left[\left(|\mathcal{A}|^2 + |\overline{\mathcal{A}}|^2\right)\left(1 + q_{\rm tag}\frac{\Delta\mathcal{D}^c}{2}\right) \right. \\
&\quad - q_{\rm tag}\langle\mathcal{D}\rangle^c \left(|\mathcal{A}|^2 - |\overline{\mathcal{A}}|^2\right)\cos\left(\Delta m_d \Delta t\right) \\
&\quad + q_{\rm tag}\langle\mathcal{D}\rangle^c 2{\rm Im}\left[\overline{\mathcal{A}}\mathcal{A}^* e^{-i\phi_{\rm mix}}\right]\sin\left(\Delta m_d \Delta t\right) \\
&\quad \left. \vphantom{\frac{1}{2}} \right] \otimes \mathcal{R}^c_{sig}(\delta t,\ \sigma_{\Delta t}),
\end{aligned} \tag{5.1}
$$

$$
\mathcal{N} = \int_{DP} \left(|\mathcal{A}|^2 + |\overline{\mathcal{A}}|^2\right) dm^2_{K^0_S\pi^+}\, dm^2_{K^0_S\pi^-}, \tag{5.2}
$$

where c labels the tagging categories, f_c are the fractions of each of them and

$$
\mathcal{A}\left(m^2_{K^0_S\pi^+}, m^2_{K^0_S\pi^-}\right) = \sum_{j=1}^{N} c_j F_j\left(m^2_{K^0_S\pi^+}, m^2_{K^0_S\pi^-}\right) \tag{5.3}
$$

$$
\bar{\mathcal{A}}\left(m^2_{K^0_S\pi^+}, m^2_{K^0_S\pi^-}\right) = \sum_{j=1}^{N} \bar{c}_j F_j\left(m^2_{K^0_S\pi^+}, m^2_{K^0_S\pi^-}\right) \tag{5.4}
$$

are the B^0 and $\overline{B}^0$ amplitudes, the sum is over all the resonances included in the model, whose Dalitz plot PDF shapes are described by the F_j (essentially Breit-Wigners, see Section 2.5.3), and c_j and $\phi_{\rm mix}$ are the parameters that characterize the signal model and whose values we intend to extract from the fit. The number of these parameters (22 in this analysis) is a good indication of the complexity of the fit. From now on, we adopt the notation $m^2_{K^0_S\pi^+} \equiv x$ and $m^2_{K^0_S\pi^-} \equiv y$, unless clarity advises otherwise.

The analysis was performed on an integrated luminosity of $349.0\,{\rm fb}^{-1}$, which translates into a data sample of $(383.6 \pm 4.2)\times 10^6$ $B\overline{B}$ pairs. A further $36.8\,{\rm fb}^{-1}$ of off-peak data were used for background characterization purposes.

Events are selected if the candidates found in them satisfy a number of deliberately moderately loose kinematic and event-shape requirements. Essentially, they are demanded to have a mass close to that of the B^0 meson, decay in a rather isotropic way, and pass some loose PID requirements (Section 5.4).

The efficiency of these selection criteria and of the reconstruction process is modelled carefully, as it varies over the Dalitz plot and thus distorts the observed shapes of any structures.

Another reconstruction effect, the migration of the recorded position of badly reconstructed signal events over the Dalitz plot, is examined in Section 5.5. Its importance grows in the corners, where the resonances overlap and hence where most of the sensitivity to the relative phases between resonances lie. In an attempt to partially recover the lost information, simulated signal events are used to characterize the misreconstruction and statistically track down their point of origin (Section 5.6).

Of course, there are background events in addition to signal in our sample, and they need to be accounted for. To that end, a detailed study of the number of background events and their distributions in the Dalitz plot, Δt and all other variables has been carried out (Section 5.7, Section 5.8). Two types of background are considered:

- Continuum events. They are by far the most numerous, in fact outweighing the signal due to the loose selection criteria. These are, however, very different kinematically and in shape from signal, and an effective separation between the two is attained with the help of the Fisher discriminant described in Section 4.5.3.

- $B\overline{B}$ background. By this generic name we refer to true $B\overline{B}$ events decaying to different channels from $K^0_S\pi^+\pi^-$ that get misreconstructed as signal events. Their numbers are expected to be only a fraction of those for signal, but they share many of the characteristics of signal, their distributions peaking close to or at the same point as that of true $B^0 \to K^0_S\pi^+\pi^-$ events. In particular, unlike continuum background, they exhibit a behaviour very akin to that of signal in the Δt variable and can represent a real obstacle in correclty retrieving the value of the mixing angle $\phi_{\rm mix}$. A careful modelling is needed, and $B\overline{B}$ background events are further split according to their Δt properties:

a) charged $B\overline{B}$ events. As continuum events, they decay with a lifetime of the same order of magnitude as that of signal, but do not mix.

b) neutral decays to flavour eigenstates. They do oscillate, but their final state determines the flavour of the B.

c) neutral decays to CP eigenstates. Their Δt dependence mimics exactly that of signal.

As mentioned above, to maximize our ability to separate them from signal, and hence reduce the statistical uncertainty of our measurements, background discriminating variables are used. Rather than a harsh cut that optimizes the signal significance in the sample, loose selection requirements were applied with the idea of including the variables in the fit. As was sketched in Section 4.5.2, correlations between these and the Dalitz-plot coordinates may appear in background events. Since the construction of PDFs depends strongly on whether a given set of variables is assumed to be correlated or uncorrelated, thorough checks are in order (Section 5.2).

5.2 Dependence of the discriminating variables on the Dalitz-plot coordinates, tag and tagging category

As mentioned in Section 4.6, the joint PDF for two uncorrelated variables can easily be constructed as the product of their individual PDFs. If the variables are correlated, the joint PDF does not factorize, and a more complicated dependence has to be considered.

The distributions for signal and continuum background of the variables m_{ES}, ΔE and $\mathcal{F}$, that do not appear in the signal likelihood Eq. (5.1) but are included in the fit for background discriminating purposes, were examined for any dependence on the Dalitz-plot coordinates.

For signal, any dependence of m_{ES} was found to be negligible, as can be seen in Figure 5.1. ΔE, however, exhibited some dependence on the Dalitz-plot position in the

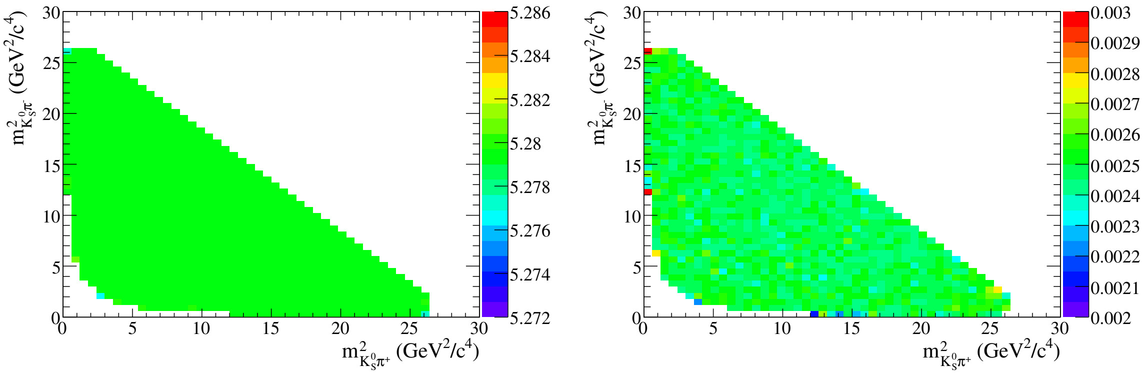


Figure 5.1: Value of the mean (left) and RMS (right) of the variable $m_{\rm ES}$ at each Dalitz-plot position. Both distributions are extremely uniform, demonstrating the independence of $m_{\rm ES}$ on the Dalitz-plot coordinates, and hence that they are uncorrelated.

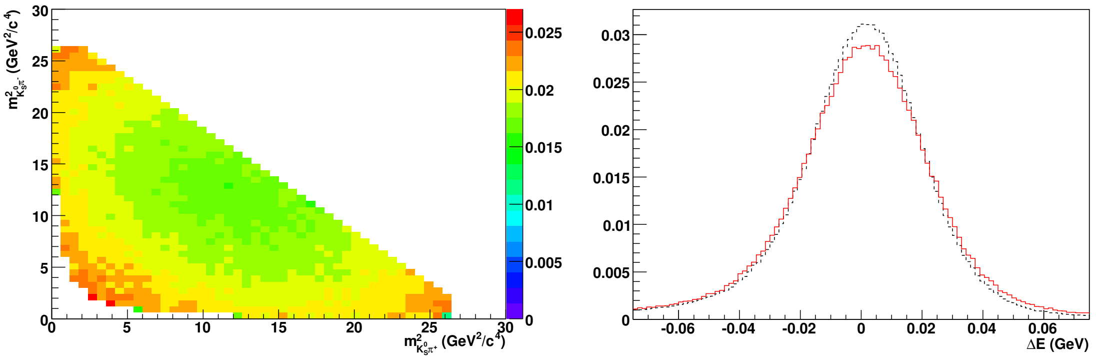


Figure 5.2: On the left, value of the RMS of ΔE at each Dalitz-plot position. On the right, the normalized ΔE distributions for signal *phase space* MC (black, dashed, events generated uniformly across the Dalitz plot), and *resonant* MC (solid, red, model in which the Dalitz plot is endowed with a resonant structure similar to that found in previous measurements [36]). In the latter, the events cluster around the edges of the Dalitz plot, emphasizing the overall difference in the ΔE distributions.

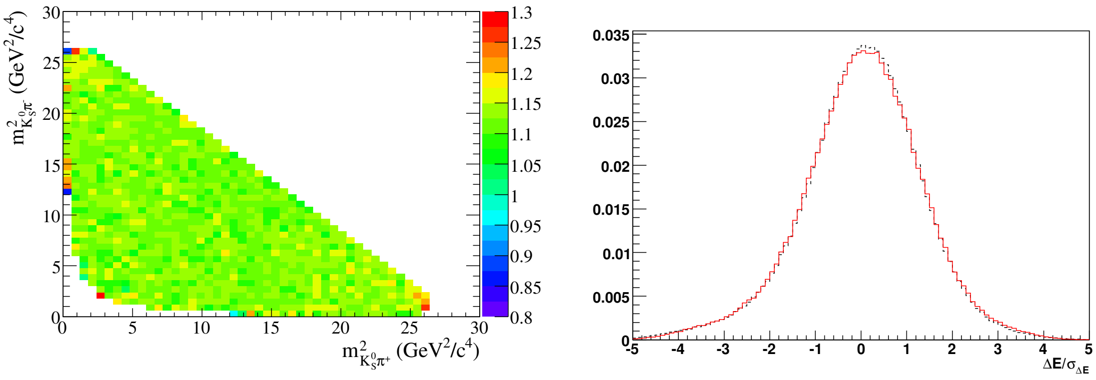


Figure 5.3: On the left, value of the RMS of $\Delta E/\sigma_{\Delta E}$ at each Dalitz-plot position. On the right, the normalized $\Delta E/\sigma_{\Delta E}$ distributions for phase space MC (black, dashed), and resonant MC (solid, red). No dependence on the Dalitz coordinates is observed.

width of its distribution, which was found to be twice as large in the corners as it was in the centre (see Figure 5.2). This is explained by the difficulty in correctly determining the energy of the low momentum particles appearing in the corners. The effect is amplified by the known Dalitz structure of the decay mode (see Section 2.4), that causes most of the events to lie close to the kinematical boundaries. These arguments suggest that the deviation from the mean might be quantified by the error on ΔE. Indeed, the ratio of the variable and its error, *the* ΔE *significance*, shows no variation across the Dalitz plot while retaining all its discriminating power. We therefore adopt it instead of ΔE (Figure 5.3).

Examination of the $m_{\rm ES}$ distribution for continuum events in the off-peak sample (see Section 5.7) did not show any dependencies either. However, the Fisher distribution was observed to vary significantly for events in the centre of the Dalitz plot compared to those by the edges. This is expected, as explained in Section 4.5.2: the event-shape variables used in the Fisher quantify the jet-like structure of the tag side of the event. In continuum events, both sides of the event are qualitatively expected to be "jetty". Three-body kinematics imply that the "jettier" an event is, the closer it is to the boundary of the Dalitz plot (Section 2.5.2). Hence, significant changes in the Fisher discriminant across the Dalitz plot are seen. These changes translate into an observed dependence of the mean and the width of the gaussian-like shape of the Fisher distribution on the distance of the event to the centre of the Dalitz plot. However, our selection criteria include a cut on the Fisher that rejects the main body of the gaussian, keeping only the right tail. The resulting distribution is far less sensitive to the dependence on the Dalitz coordinates (see Figure 5.4 and Figure 5.5), although it is still accounted for in our model to avoid the systematic uncertainties associated with neglecting it [1].

As reasoned in Section 4.5.2, a dependence of $\mathcal{F}$ on the Dalitz plot position is not found in signal MC, although different shapes are required for each tagging category (see Figure 5.6). This is yet another consequence of the event-shape variables being evaluated on the tag-side of the event, since assigning a tag and a tagging category to a given event is a statement about the type of the decay of the tagging B meson. A different parameterization is therefore employed for each tagging category in the signal hypothesis.

Finally, it is worth noting that no tag-dependence was found in the shapes of the distributions of the discriminating variables, neither for signal nor for the different kinds of backgrounds. Therefore, identical PDFs are used to model m_{ES}, $\Delta E/\sigma_{\Delta E}$ and $\mathcal{F}$ for both tags in each species.

5.3 Total likelihood and the background discriminating variables

Previous studies have shown that the m_{ES} and ΔE variables are mostly uncorrelated [78]. Under the additional assumption that they are also uncorrelated with the

[1]Figure 5.5 also suggests another posible, and simpler, treatment that would be statistically indistinguishable: since it is only the bin corresponding to the Dalitz plot centre that has a value markedly different from the rest, a two-bin histogram, composed of the central bin and another bin representing the rest of the Dalitz plot, could be used to model the Dalitz dependence of the Fisher.

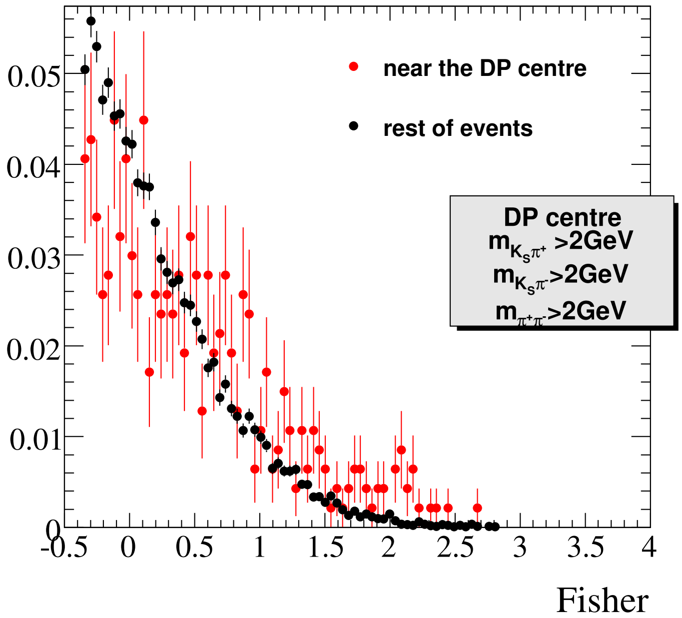


Figure 5.4: Continuum background Fisher shape dependence on the distance from the centre of the Dalitz plot. In red, events contained in the centre of the Dalitz plot, defined as $m_{ij}^2 > 2\,\mathrm{GeV}/c^2$ for all three pairs of particles. In black, events outside that region. The red points tend to lie below the black ones towards the left of the picture. The difference in the trends is rather small and played down by the low number of continuum events in the centre of the Dalitz plot.

Dalitz-plot coordinates, including them in the likelihood is trivial, as the joint PDF simply factorizes. Thus, for $B\overline{B}$ events (both signal and background):

$$\mathcal{P}_{B\overline{B}}\left(x, y, \Delta t, m_{\mathrm{ES}}, \Delta E/\sigma_{\Delta E}, \mathcal{F}\right) = \mathcal{P}\left(x, y, \Delta t\right) \mathcal{P}\left(m_{\mathrm{ES}}\right) \mathcal{P}\left(\Delta E/\sigma_{\Delta E}\right) \mathcal{P}\left(\mathcal{F}\right) \quad (5.5)$$

where $\mathcal{P}\left(x, y, \Delta t\right)$ is the joint PDF for the Dalitz coordinates and the time difference Δt. For the continuum background, a joint DP-$\mathcal{F}$ PDF is employed to account for the dependence of the Fisher on the Dalitz coordinates:

$$\mathcal{P}_{q\overline{q}}\left(x, y, \Delta t, m_{\mathrm{ES}}, \Delta E/\sigma_{\Delta E}, \mathcal{F}\right) = \mathcal{P}\left(x, y, \mathcal{F}\right) \mathcal{P}\left(\Delta t\right) \mathcal{P}\left(m_{\mathrm{ES}}\right) \mathcal{P}\left(\Delta E/\sigma_{\Delta E}\right) \quad (5.6)$$

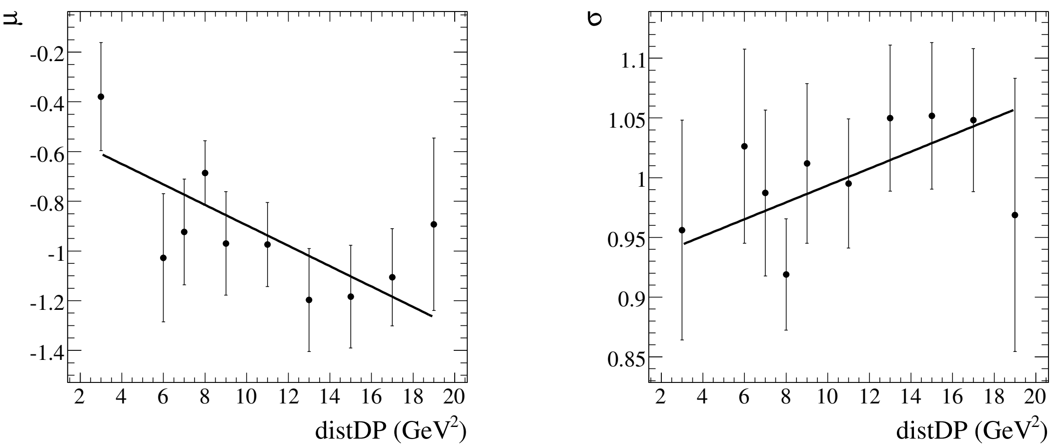


Figure 5.5: Dependence of the mean (left) and width (right) of the continuum Fisher distribution on the distance to the centre of the Dalitz plot, assumed to be linear. The correlation is rather weak, as can be deduced from the fitted slopes: -0.041 ± 0.015 for the mean, and 0.007 ± 0.005 for the width.

since in this case the DP and Δt PDFs do factorize (see Section 5.7). In these conditions, then, including the $m_{\rm ES}$ and $\Delta E/\sigma_{\Delta E}$ variables in the fit simply requires obtaining the PDFs for those variables and including a (trivial) multiplication in the calculation of the likelihood. These are shown in Figure 5.7 for signal. With the previous expressions for the likelihoods of the signal and background species in mind, the total likelihood is trivially constructed (recall Eq. (5.7)):

$$\mathcal{L}\left(\vec{n}, \vec{a}\right) = e^{-\left(n_{\rm sig}+n_{q\bar{q}}+n_{B^+B^-}+n_{B_{flav}}+n_{B_{CP}}\right)} \prod_{e=1}^{N} \mathcal{L}_e^c \, , \tag{5.7}$$

$\mathcal{L}_e^c$ being

$$\begin{aligned} \mathcal{L}_e^c \; = \; & n_{\rm sig} f_{\rm sig}^c \mathcal{P}_{{\rm sig},e} \\ & + n_{q\bar{q}} f_{q\bar{q}}^c \mathcal{P}_{q\bar{q},e} + \sum_{j=1}^{n_{\rm class}^{B^+B^-}} n_j f_j^c \mathcal{P}_{B^+B^-,j,e} + \sum_{k=1}^{n_{\rm class}^{B_{flav}}} n_k f_k^c \mathcal{P}_{B_{flav},k,e} + \sum_{l=1}^{n_{\rm class}^{B_{CP}}} n_l f_l^c \mathcal{P}_{B_{CP},l,e} \, , \end{aligned} \tag{5.8}$$

where c is the tagging category, the n_i represent the numbers of events of each species and the sums for each of the three types of $B\overline{B}$ background run over the different

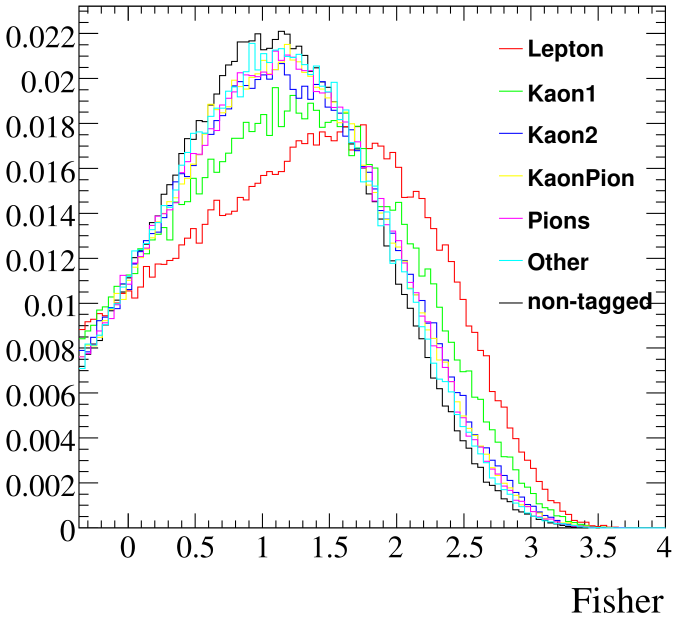


Figure 5.6: Fisher distributions for signal MC split by tagging categories, black being untagged events, and red those with leptonic tags. A separate parameterization for each tagging category is used to evaluate the Fisher PDF in our software package.

classes of background channels within that type. After a careful exposition of the selection criteria, we revisit the signal likelihood (Section 5.6), and proceed to discuss the form of the likelihoods for the background species (Sections 5.7, 5.8).

5.4 Event selection

The selection criteria are applied to the data in several stages. In the first stage, B candidates are formed by requiring two oppositely charged particles belonging to the `GoodTracksLoose` list and a K_s^0 candidate as described in Section 4.3.1. They are asked to pass three very basic cuts:

- $m_{\rm ES} > 5.2\,{\rm GeV}/c^2$,

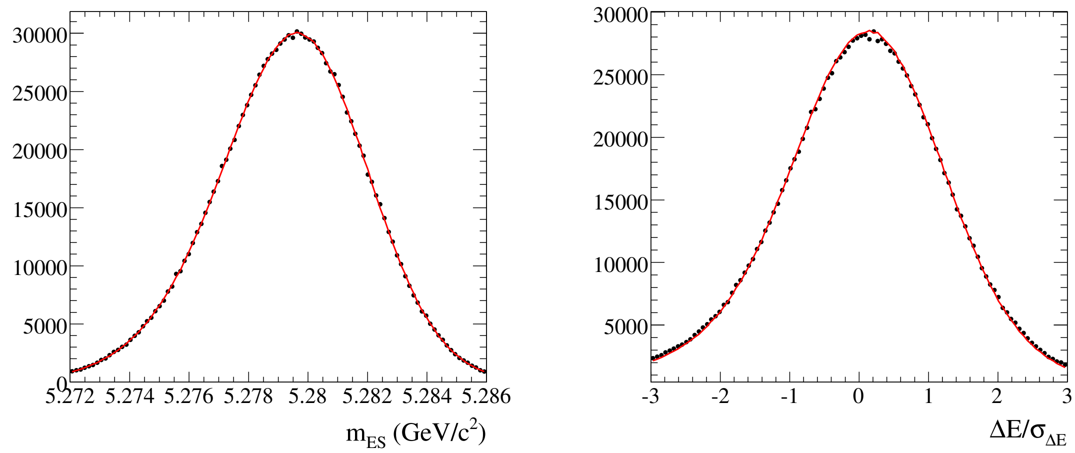


Figure 5.7: Signal $m_{\rm ES}$ (left) and $\Delta E/\sigma_{\Delta E}$ (right) PDFs, extracted from MC, in the signal region (see Section 5.4). Since these variables are uncorrelated among themselves or with the Dalitz-plot coordinates, including their discriminating power in the fit is as straightforward as multiplying the Dalitz-plot PDF by their PDFs. The absence of any bias is then easily checked with toy MC studies.

- $4.99 < E^* < 5.59$ GeV (see Section 4.3.4),
- total energy of the event $E_{\rm TOTAL} < 20.0$ GeV.

In the following stage, the PID likelihoods for the tracks, some K^0_S-quality related variables and the background discriminating variables are calculated and the following cuts are applied:

- The K^0_S lifetime significance, $\tau_{K^0_S}/\sigma_{\tau_{K^0_S}}$ is required to be greater than 5.0, thus rejecting combinatorial background.
- The cosine of the angle between the momentum of the K^0_S and the line that joins its decay vertex with that of the B candidate must be greater than 0.999.
- The mass requirement on the K^0_S candidate is tightened to $|m_{\pi^+\pi^-} - m_{K^0_S}| < 15\,{\rm MeV}/c^2$.

- The charged tracks are demanded to satisfy the `PiLoose` requirements (Section 4.3.3), since a study [79] showed that such a choice maximizes the signal sensitivity. Also, candidates likely to be electrons are rejected.

- Very loose cuts on Δt and its error are applied, following the standard time-dependent analyses: $|\Delta t| < 20\,\mathrm{ps}$ and $\sigma_{\Delta t} < 2.5\,\mathrm{ps}$.

- The Fisher variable is calculated and required to be $\mathcal{F} > -0.365$, thereby reducing the continuum background to 30% of its original size and decreasing the number of signal events by only around 10% (see Figure 4.5 and Section 5.7).

- Cuts on the kinematical variables m_{ES} and $\Delta E/\sigma_{\Delta E}$ are applied to select the three regions of interest in the m_{ES}-$\Delta E/\sigma_{\Delta E}$ plane (see Figure 5.8):

  a) the *signal region*, where true $B^0 \to K^0_S\pi^+\pi^-$ decays are expected, is defined as a three-standard deviation window around the m_{ES} and $\Delta E/\sigma_{\Delta E}$ peaks, $5.272 < m_{\mathrm{ES}} < 5.286\ \mathrm{GeV}/c^2$ and $-3.0 < \Delta E/\sigma_{\Delta E} < 3.0$,

  b) the *grand sideband*, $5.20 < m_{\mathrm{ES}} < 5.26\ \mathrm{GeV}/c^2$ and $-3.0 < \Delta E/\sigma_{\Delta E} < 3.0$,

  c) the *upper sideband*, $5.2 < m_{\mathrm{ES}} < 5.286\ \mathrm{GeV}/c^2$ and $4.0 < \Delta E/\sigma_{\Delta E} < 15.0$,

  of which the last two are used to characterize the continuum distributions (see Section 5.7).

- Events are rejected (*vetoed*) if they lie inside one of four narrow strips that are heavily populated by $B\overline{B}$ backgrounds originating from charmed and charmonium resonances (see Section 5.8 for more details).

The efficiency of each of these cuts, as well as the overall efficiency, have been evaluated from phase space and resonant MC, and are shown in Table 5.1.

Among the events from the on-peak data sample that presented at least one candidate satisfying the above selection requirements, about 6% had more than one. In this case, the candidate with the highest B-vertex probability was chosen. This procedure was shown not to bias any physical quantity.

Table 5.1: Summary of cut efficiencies evaluated from MC with no structure across the Dalitz Plot (left) and with a resonant structure taken from [36] (right). The efficiency for each selection criterion is calculated relative to the sample of events passing the previous requirement. The last three lines show the absolute efficiency for the three regions of interest in the m_{ES}-$\Delta E/\sigma_{\Delta E}$ plane. Note that the overall efficiency depends on the resonant content of the Dalitz plot, which is only known after the fit to data is performed, so the uncertainty quoted is merely indicative of the statistical uncertainties due to the size of the samples involved in its evaluation.

	Efficiency (%) for $B^0 \to K^0_S\pi^+\pi^-$	
Selection requirement	Phase space MC	Resonant MC
First stage selection, vertexing and reconstruction	39.9	40.9
Pion PID requirements	93.8	93.9
Electron veto	97.1	97.2
$E_{\text{tot}} < 20\,\text{GeV}$	99.2	99.1
Fisher > -0.365	90.1	89.8
$\|m_{\pi^+\pi^-} - m_{K^0_S}\| < 15\,\text{MeV}/c^2$	95.8	96.2
$\tau_{K^0_S}/\sigma_{\tau_{K^0_S}} > 5$	91.2	92.0
$\cos\alpha_{K^0_S} > 0.999$	98.3	98.3
$\|\Delta t\| < 20\,\text{ps}$	97.7	97.7
$\sigma_{\Delta t} < 2.5\,\text{ps}$	97.7	97.7
$5.20 < m_{\text{ES}} < 5.286\,\text{GeV}/c^2$	99.1	99.1
$-3.0 < \Delta E/\sigma_{\Delta E} < 15.0$	93.2	93.0
Veto D, J/ψ and $\psi(2S)$	84.0	86.4
Signal Region:	19.200	18.651
$\pm$ statistical uncertainty	0.012	0.026
Upper Sideband:	0.654	1.719
$\pm$ statistical uncertainty	0.003	0.009
Grand Sideband:	0.223	0.437
$\pm$ statistical uncertainty	0.001	0.005

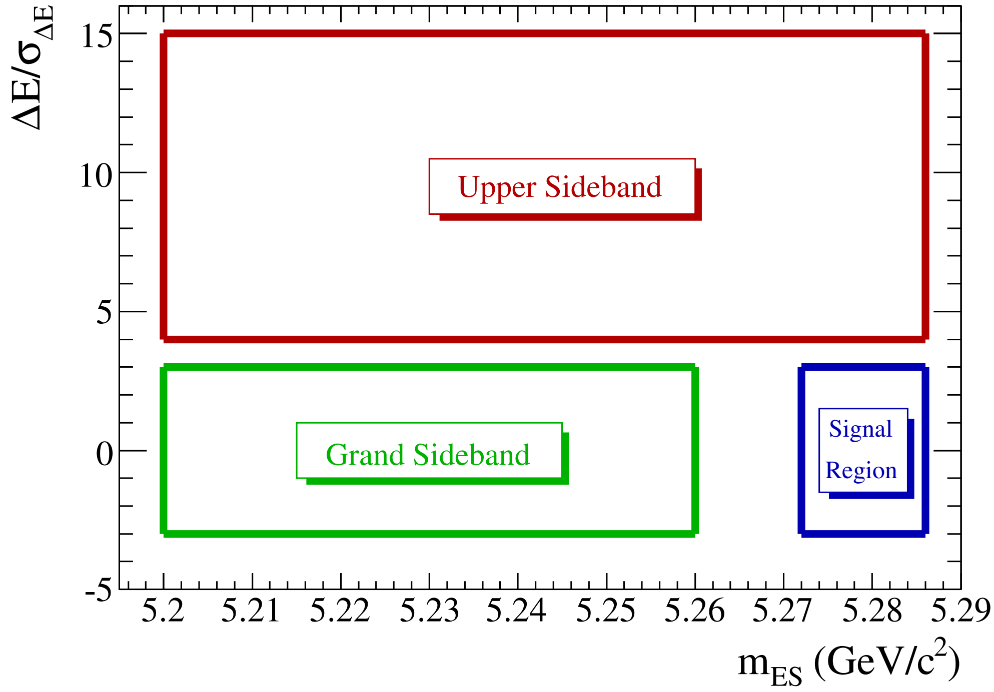


Figure 5.8: Signal region (blue), grand sideband (green) and upper sideband (red). The first is from where the events are drawn for the full time-dependent amplitude fit, whereas the last two are used to extract the continuum distributions in the variables included in the fit. A lower sideband (corresponding to $\Delta E/\sigma_{\Delta E} < -4.0$) is avoided, as it is heavily contaminated by $B\overline{B}$ backgrounds.

5.5 Efficiency and migration

The efficiency, defined as the fraction of signal events that are reconstructed and selected, varies across the Dalitz plot due to the very different kinematical properties of the three final state particles for different areas of the phase space. In particular, as noted in Section 2.5.2, at each of the corners of the Dalitz plot one of the three particles is slow. Low momentum tracks are less likely to be reconstructed, since they either do not reach the Drift Chamber and must be formed from their SVT hits only, or they only partially traverse the DCH. Hence, low efficiencies are expected at the

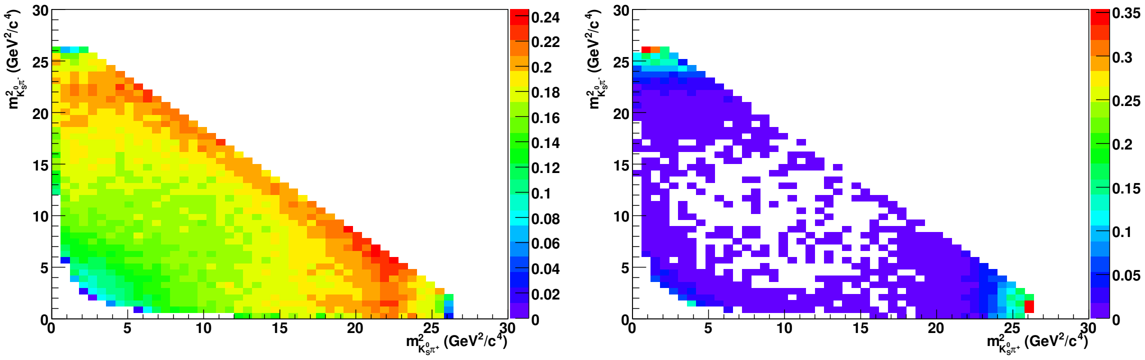


Figure 5.9: Efficiency (left) and self cross feed fraction (right) as a function of the Dalitz coordinates. Note the drop of efficiency in the three corners. The fraction of self cross feed is rather high in the slow pion corners ($m^2_{K^0_S\pi^\pm} > 20\,\mathrm{GeV}/c^2$), although not as much in the slow K^0_S corner. This, and the depletion in efficiency in the same area, is due to the selection requirement on the K^0_S lifetime significance, which reduces the combinatorial background (and hence the self cross feed), and the probability of reconstructing a slow K^0_S, since the distance travelled will be small compared to the uncertainty in the measured vertex position.

corners of the Dalitz plot. Such a trend can be clearly observed in Figure 5.9.

A related effect, which is a consequence of the necessity to have low momentum tracks at the Dalitz plot corners, is that events are likely to be misreconstructed. Slow particles may be reconstructed, but the poor quality with which they are determined increases the probability of them being assigned to the wrong B candidate. The reason is that the large uncertainties on the track parameters prevent the χ^2 of the fit for the B candidate vertex from revealing the mistake. That can either make a background event match the $K^0_S\pi^+\pi^-$ final state, and hence be selected as signal (see Section 5.8), or alter the properties of a true $B^0 \to K^0_S\pi^+\pi^-$ event. In the latter case, the assignment of an incorrect track to the candidate will change the balance of energy and momentum among the three particles in the final state, leading to differences between the reconstructed Dalitz plot position and the true one. Such an effect, called *migration*, is characteristic of misreconstructed signal, which we will refer to as *self cross feed*. The fraction of this kind of event, calculated across the Dalitz plot and defined as the number of events that are reconstructed as self cross feed divided by the total number of events that are reconstructed, is plotted in Figure 5.9.

Figure 5.9 shows rapidly falling signal efficiencies and rising self cross feed fractions on the areas most sensitive to the interference between resonances, the corners of the Dalitz plot. This promptly raises the question of whether the phases between the resonant components are rendered impossible to measure by the two reconstruction effects. Fits to MC simulated events, reconstructed in exactly the same way as real data, show without any doubt that the extraction of the phases is possible, and allow us to discard any important systematic effect in the treatment of the difficulties brought about by the reconstruction process (Section 6.2).

The total fraction of self cross feed events in resonant MC after the selection criteria is 2.8%. These can either be treated as background to the properly reconstructed signal events, or used to extract the physics information still available by statistically tracing their true positions. The method developed to achieve the latter is described in the next section.

5.6 Treatment of Self Cross Feed

5.6.1 Operational definition of Self Cross Feed

For a given process, such as $B^0 \to K^0_S \pi^+\pi^-$, the true distributions of events in certain variables as dictated by the laws of Nature, $\mathcal{P}^{\rm true}$, are modified by detector and reconstruction effects such as efficiency, migration and experimental resolution, as seen in the previous sections. These can be described by convolving the true distribution with a "detector response function", $\mathcal{R}$:

$$\mathcal{P}^{\rm observed} = \mathcal{P}^{\rm true} \otimes \mathcal{R}^{\rm total} . \tag{5.9}$$

The different contributions to $\mathcal{R}$ are conventionally separated:

$$\mathcal{R}^{\text{total}} = \epsilon\,(x,y)\,\mathcal{R}^{\text{migration}}\mathcal{R}^{\Delta t}\,, \tag{5.10}$$

where ϵ represents the (Dalitz plot-dependent) efficiency, as defined in Section 5.5, and $\mathcal{R}^{\Delta t}$ has been described in detail for signal in Section 4.4.

In Dalitz analyses where, apart from the time dependence, an event is completely characterized by its position in the Dalitz plane, it is convenient to split signal events further according to whether their migration plays a significant role (self cross feed) or not (truth-matched). Rather than identifying the different mechanisms at work in the (mis)reconstruction of events, and classifying them as truth-matched or self cross feed according to that, our approach adopts an operational definition: a given event is *defined* to be self cross feed if the difference between the true and reconstructed momentum of one of the daughters divided by the error of its reconstruction is greater than a certain value (Figure 5.10):

$$\left|\frac{p^{\text{reco}}_{K^0_S} - p^{\text{true}}_{K^0_S}}{\sigma^{\text{reco}}_{K^0_S}}\right| \;>\; 30\,, \tag{5.11}$$

$$\left|\frac{p^{\text{reco}}_{\pi^\pm} - p^{\text{true}}_{\pi^\pm}}{\sigma^{\text{reco}}_{\pi^\pm}}\right| \;>\; 20\,. \tag{5.12}$$

. These inequalities ensure that en event classified as self cross feed migrates significantly over the Dalitz plot. Those events in which migration can be neglected (*truth-matched*), are assumed to have been reconstructed with perfect experimental resolution in the Dalitz plot coordinates. This assumption is valid as long as the distances between their true and reconstructed positions are small compared to the typical widths of the Dalitz plot structures. In our case, the narrowest resonance is the $K^*(892)$, with a width of $50.8(9)\,\text{MeV}/c^2$ [22], and the previous assumption on the resolution is guaranteed to hold by the above inequalities (Figure 5.11). An effective separation between the two signal species is also attained, as can be seen Figure 5.12.

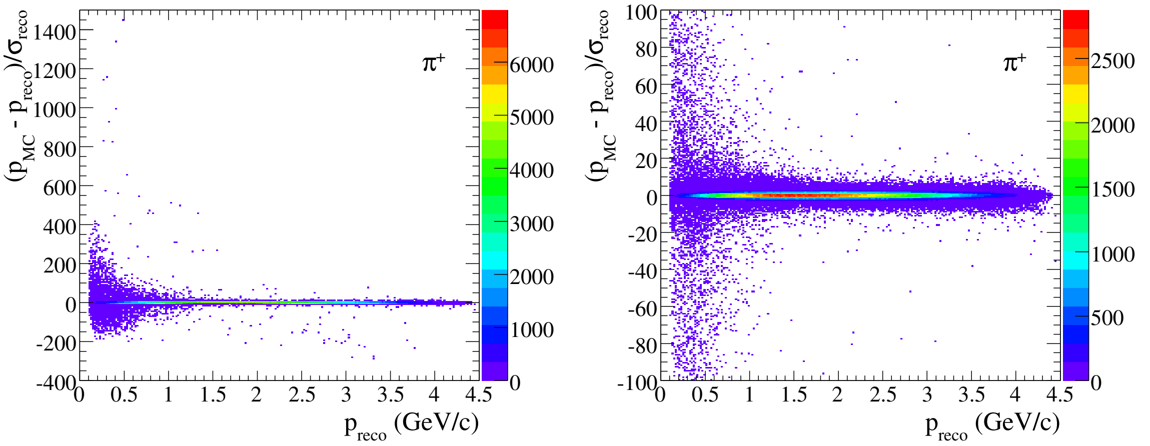


Figure 5.10: Plots showing the ratio of reconstructed minus true momentum over the reconstruction error for the pion candidates: on the left, the whole range and on the right, a zoom in the central region, where it can be seen that for most pions, the absolute value of the ratio is below 20.

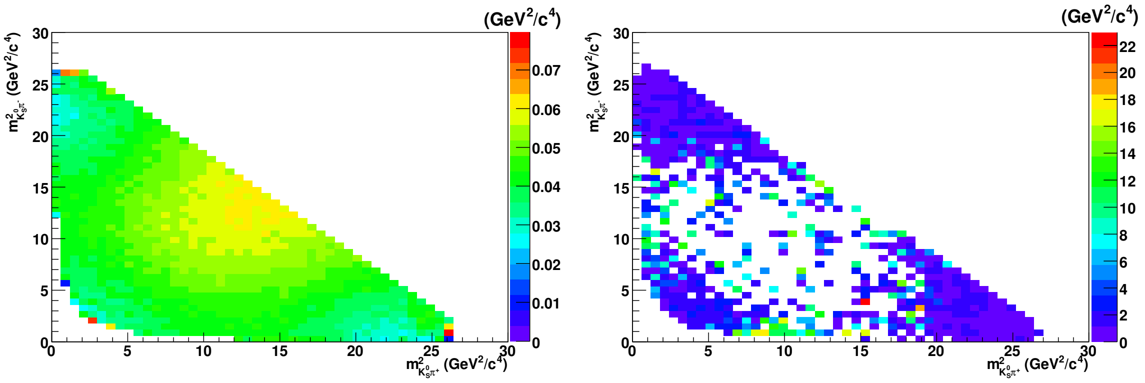
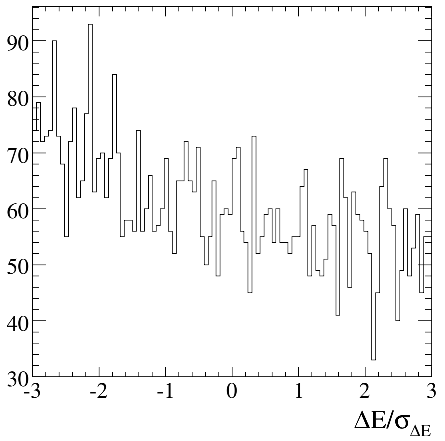


Figure 5.11: Average distance between the reconstructed and the true position of truth-matched (left) and self cross feed (right) events, plotted at the reconstructed position. For the former it is rather small, demonstrating the validity of neglecting the experimental resolution. For the latter, however, it takes on average values around $1{-}2\,\mathrm{GeV}^2/c^4$ for events originating in the corners, occassionally reaching $\sim 10\,\mathrm{GeV}^2/c^4$.



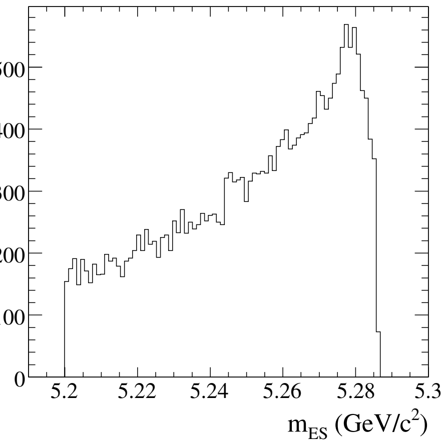


Figure 5.12: $\Delta E/\sigma_{\Delta E}$ (left) and $m_{\rm ES}$ (right) distributions for self cross feed events. The latter peaks at $5.28\,{\rm GeV}/c^2$, as signal should, but presents a long tail due to the misreconstruction. $\Delta E/\sigma_{\Delta E}$ exhibits an even more combinatorial profile, and shows no enhancement at the origin, thus demonstrating the correct separation of the two signal species.

5.6.2 Formalism

Once the self cross feed and the truth-matched events have been separated, further simplification of $\mathcal{R}^{\rm migration}$ can be attained:

$$\begin{aligned}\mathcal{R}^{\rm migration}\left(x_{\rm reco}, y_{\rm reco}; x_{\rm true}, y_{\rm true}\right) &= \left(1 - f_{\rm SCF}\left(x_{\rm true}, y_{\rm true}\right)\right)\delta\left(x_{\rm reco} - x_{\rm true}\right)\delta\left(y_{\rm reco} - y_{\rm true}\right) \\ &\quad + f_{\rm SCF}\left(x_{\rm true}, y_{\rm true}\right)\mathcal{R}^{\rm SCF}\left(x_{\rm reco}, y_{\rm reco}; x_{\rm true}, y_{\rm true}\right) \quad (5.13)\end{aligned}$$

where $f_{\rm SCF}$ is the fraction of self cross feed events defined in Section 5.5, and the quantity $\mathcal{R}^{\rm SCF}\left(x_{\rm reco}, y_{\rm reco}; x_{\rm true}, y_{\rm true}\right)$ represents the probability for an event originally at $\left(x_{\rm true}, y_{\rm true}\right)$ to migrate to $\left(x_{\rm reco}, y_{\rm reco}\right)$ (Figure 5.13). Assuming the same Δt resolution model for truth-matched and self cross feed, Eq. (5.9) implies

$$\mathcal{P}^{\rm observed}\left(x_{\rm reco}, y_{\rm reco}, \Delta t\right) = \frac{\epsilon\left(1 - f_{\rm SCF}\right)\mathcal{P}^{\rm true} + \iint_{\rm DP}\mathcal{P}^{\rm true}\epsilon f_{\rm SCF}\mathcal{R}^{\rm SCF} dx_{\rm true} dy_{\rm true}}{\iint_{{\rm DP}, \Delta {\rm t}, q_{\rm tag}}\mathcal{P} dx dy} \tag{5.14}$$

where $\mathcal{P}^{\rm true}$ is the likelihood written in Eq. (5.1) without the normalization $\mathcal{N}$, which now reduces to

$$\iint_{{\rm DP},\Delta {\rm t},q_{\rm tag}} \mathcal{P}\, dx_{\rm reco}\, dy_{\rm reco} = \iint_{{\rm DP},q_{\rm tag}} \epsilon \mathcal{P}^{\rm true}\, dx_{\rm true}\, dy_{\rm true} = \iint_{\rm DP} \epsilon \left(|\mathcal{A}|^2 + |\overline{\mathcal{A}}|^2\right) dx\, dy$$

since the Δt part is already normalized and the smearing of $\mathcal{P}^{\rm true}$ in Eq. (5.14) only redistributes the total probability, i.e.

$$\iint_{\rm DP} \mathcal{R}^{\rm SCF}\left(x_{\rm reco}, y_{\rm reco}; x_{\rm true}, y_{\rm true}\right) dx_{\rm reco} dy_{\rm reco} = 1\,. \tag{5.15}$$

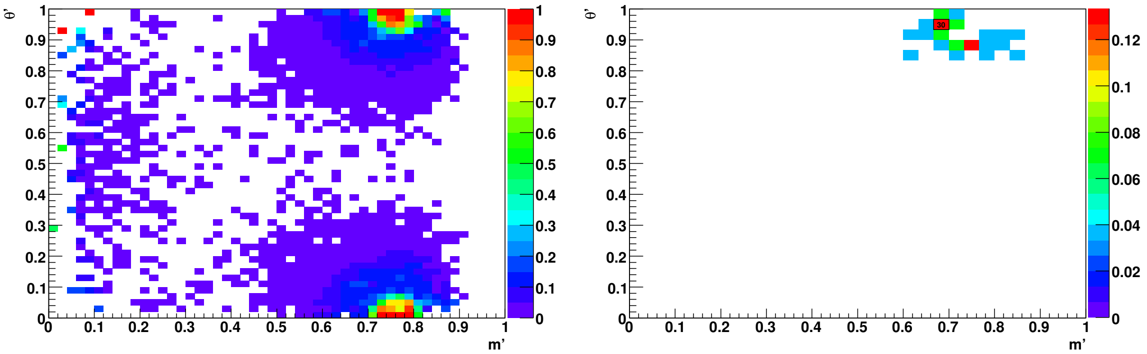


Figure 5.13: Fraction of self cross feed depicted in the square Dalitz plot (left), and probability of migration $\mathcal{R}^{\rm SCF}$ (right). Observe that the fraction of self cross feed events rises to nearly one at the tips of the slow pion corners. $\mathcal{R}^{\rm SCF}$ can be thought of as a 4×4 matrix, or as a 2×2 matrix of 2D histograms. On the right, one of these histograms is plotted, depicting the probability of migration for a single bin. The number in a bin on the top right corner is the number of self cross feed events found in the MC originating at that bin. (i.e. $x_{\rm true}, y_{\rm true}$). The coloured bins around it represent the probability for those events to migrate to each of the bins.

Eq. (5.14) requires a further modification if, as has been done in this analysis, the efficiency, self cross feed fraction and migration are modelled by 2D histograms in the square Dalitz plot variables $m' \equiv x'$ and $\theta' \equiv y'$, and the true distributions of events depend on the usual Dalitz plot coordinates, $m^2_{K^0_S\pi^+} \equiv x$ and $m^2_{K^0_S\pi^-} \equiv y$. As noted in Section 2.5.7, jacobians must be introduced due to the change of variable in the integration of the numerator:

$$\iint_{\rm DP} \left(\epsilon f_{\rm SCF} \mathcal{P}^{\rm true}\right) (x_{\rm true}, y_{\rm true})\, \mathcal{R}^{\rm SCF} \left(x_{\rm reco}, y_{\rm reco}; x_{\rm true}, y_{\rm true}\right) dx_{\rm true} dy_{\rm true}$$

$$\xrightarrow{x,y \mapsto x',y'} \tag{5.16}$$

$$\frac{1}{|\det J_{\rm reco}|} \iint_{\rm DP} \left(\epsilon f_{\rm SCF} \mathcal{P}^{\rm true}\right) (x_{\rm true}, y_{\rm true})\, \mathcal{R}^{\rm SCF} \left(x'_{\rm reco}, y'_{\rm reco}; x'_{\rm true}, y'_{\rm true}\right) |\det J_{\rm true}| dx'_{\rm true} dy'_{\rm true}$$

where $x_{\rm true} = x_{\rm true}\left(x'_{\rm true}, y'_{\rm true}\right)$ and $y_{\rm true} = y_{\rm true}\left(x'_{\rm true}, y'_{\rm true}\right)$. The jacobian at the front is needed to convert $\mathcal{R}^{\rm SCF}\left(x'_{\rm reco}, y'_{\rm reco}; x'_{\rm true}, y'_{\rm true}\right) dx'_{\rm true} dy'_{\rm true}$, which is a density in the square DP variables x', y', into a density in the usual Dalitz coordinates x and y.

5.6.3 Tests on MC and conclusion

The formalism described above was implemented in our fitting package `Laura++`, and tested with toy MC studies first (see Section 6.1 and Figure 5.14), and later by fitting signal MC (as in Section 6.2, but without the backgrounds) with and without accounting for self cross feed. Improvements were observed in the non-gaussian tails of the distibutions, which were clearly reduced. The means and especially the widths of the pull plots of about half of the variables ($K^*(892)\ \Delta X$, $\rho^0(770)\ Y$, $\rho^0(770)\ \Delta X$ and $f_0(980)\ X$) were reasonably improved when the self cross feed treatment described before was enabled. In particular, non-gaussian tails were consistently reduced. The rest of the variables showed pull plots as good in both scenarios. As can be seen in Table 5.2, the overall gain was modest.

The small sensitivity to self cross feed effects that was found is consistent with the fact that it affects only $\sim 3\%$ of the signal events, and that the statistical uncertainties on the fitted parameters c_j are far greater than that figure (see, e.g. Table 6.2). Therefore, in view of the large increase in computing time compared to the rather small improvement to the fit, it was decided not to treat self cross feed and truth-matched events separately for the present iteration of the $B^0 \to K^0_S \pi^+ \pi^-$ analysis.

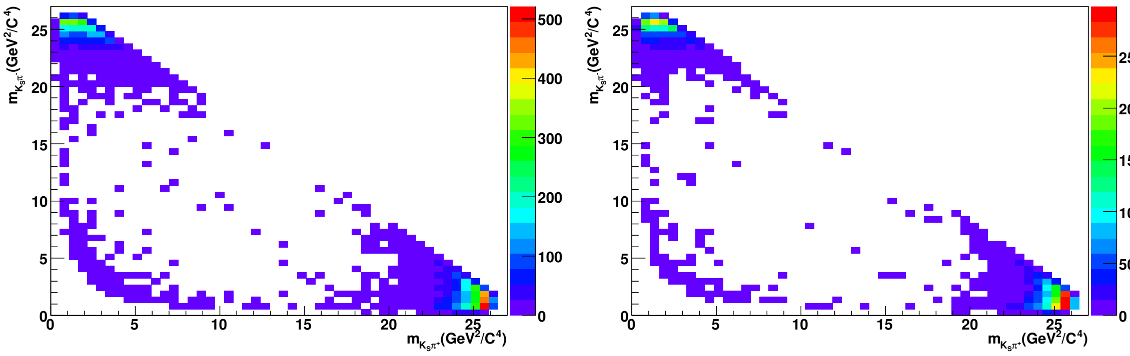


Figure 5.14: Self cross feed events in full MC (left) and in toy MC generated by our implementation of the procedure described in the text (right). There is a good agreement between them.

5.7 Continuum Background

Continuum events are by far the most prominent source of backgrounds. Excluding the cut on the Fisher discriminant, about 28100 of them are estimated to satisfy the selection criteria described in Section 5.4. In contrast, only around 1710 signal events are expected. To mildly enrich the sample in signal events, and for reasons of computational expediency and stability of the fit, a loose cut on the Fisher, designed to retain 90.0% of the signal, is applied. This additional requirement reduces the expected number of events to 1530 for signal and 8070 for continuum. Clearly, an accurate determination and modelling of the distributions of the latter is essential, given the signal-to-background ratio.

The off-peak sample is ideal to characterize this kind of background, since events collected below the $\Upsilon(4S)$ resonance can only be composed of the light quarks u, d, s, c that make up the continuum events. However, the small size of the sample gives a poor estimation of the shapes of its distributions. The grand and upper sidebands are therefore used to ameliorate the problem (see Figure 5.15). However, previous studies have shown that those regions of the m_{ES}-$\Delta E/\sigma_{\Delta E}$ plane are not exempt from $B\overline{B}$

Table 5.2: Comparison between fits to full MC with and without separating self cross feed (SCF) and truth-matched events. The MC was generated according to a three resonance model: $K^{*\pm}(892)\pi^{\mp}$, $\rho^0(770)K^0_S$ and $f_0(980)K^0_S$. The means and widths of the pull plots for each of those resonance parameters and the mixing phase are shown.

Parameter	μ (neglect SCF)	μ (with SCF)	σ (neglect SCF)	σ (with SCF)
$K^*(892)\ \Delta X$	-1.14	-0.91	1.86	1.58
$K^*(892)\ \Delta Y$	-0.44	-0.51	0.86	0.83
$\rho^0(770)\ X$	-0.95	-1.18	1.20	1.11
$\rho^0(770)\ Y$	1.32	1.23	1.83	1.64
$\rho^0(770)\ \Delta X$	0.65	0.52	1.17	1.07
$\rho^0(770)\ \Delta Y$	-0.65	-0.53	1.04	1.05
$f_0(980)\ X$	-1.74	-1.51	2.18	1.95
$f_0(980)\ Y$	-1.44	-1.56	1.35	1.28
$\phi_{\rm mix}$	-0.15	-0.13	1.01	0.98

background contamination. This contamination is studied simultaneously in the three areas of interest (the two sidebands and the signal box), as described in Section 5.8, and histograms are formed detailing its distributions in the Dalitz coordinates, $m_{\rm ES}$, $\Delta E/\sigma_{\Delta E}$ and $\mathcal{F}$. These are then employed to "subtract" the $B\overline{B}$ presence in the sidebands, either literally, thus forming a histogram with the continuum distribution in a given variable, or via a fit, in which the $B\overline{B}$ histograms are used as the PDF for the undesired contamination. The first method applied to the grand sideband yields the Dalitz distribution of continuum events. The second is employed to extract the continuum shapes in the $\Delta E/\sigma_{\Delta E}$ and $\mathcal{F}$ variables from the grand sideband, and in the $m_{\rm ES}$ variable from the upper sideband (Figure 5.16).

The two dimensional histograms describing the Dalitz distribution of continuum events, in principle one for each tag $q_{\rm tag}$ and tagging category c, are then normalized

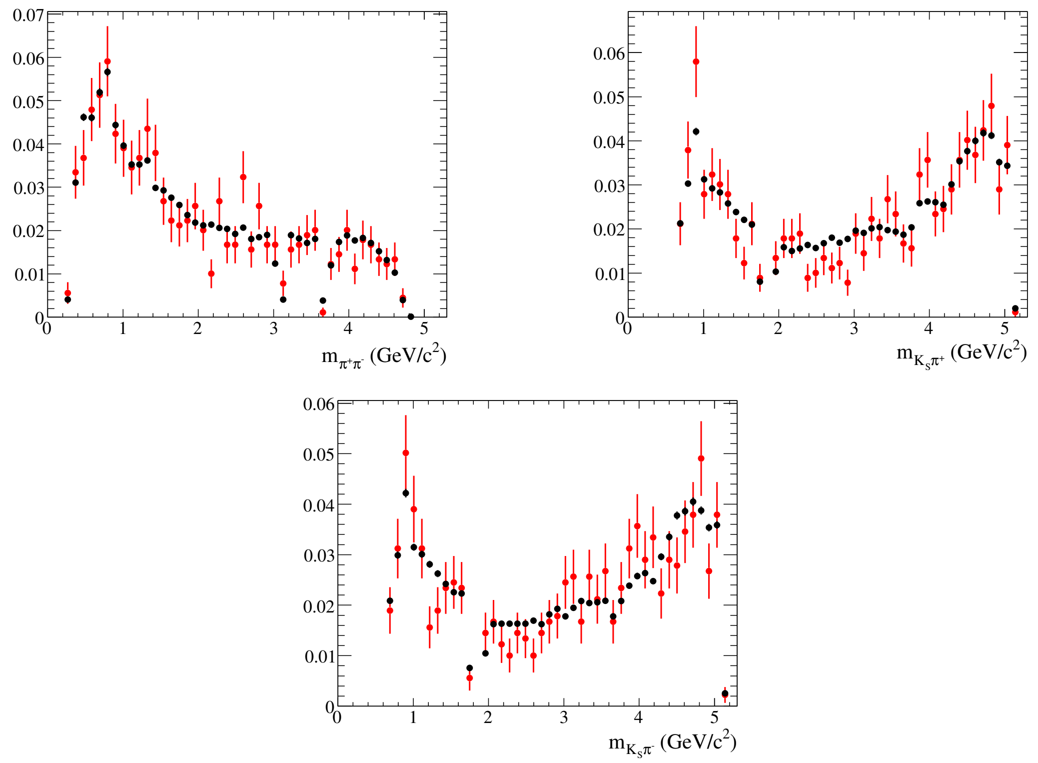


Figure 5.15: Projections on the three invariant masses of the Dalitz plot distributions of off-peak data (red) and on-peak sidebands (black). Since a good agreement is observed, both samples are used in the extraction of the continuum PDFs.

and used to build the Δt-Dalitz plot PDF:

$$\mathcal{P}_{q\bar{q}} = \frac{P_{q\bar{q}}^{c,q_{\rm tag}}(x,y)}{\int\int_{DP} P_{q\bar{q}}^{c,q_{\rm tag}}(x,y)\, dxdy} \times \left(f_{\rm prompt}\delta(\Delta t_{\rm true}) + (1-f_{\rm prompt})\frac{e^{-|\Delta t_{\rm true}|/\tau_{q\bar{q}}}}{2\tau_{q\bar{q}}} \right) \otimes \mathcal{R},$$

where the Δt behaviour is modelled as consisting of two components, with zero (*prompt*) and non-zero lifetimes, convolved with a customized resolution function $\mathcal{R}$ extracted from off-peak data and consisting of three gaussians (see Eq. (4.9) and Figure 5.17).

For a given tag flavour, the Dalitz distributions were observed to be similar for all tagging categories except for the untagged events, so a combined PDF was used for those tagging categories (see Figure 5.18). Events that are deemed to lack reliable

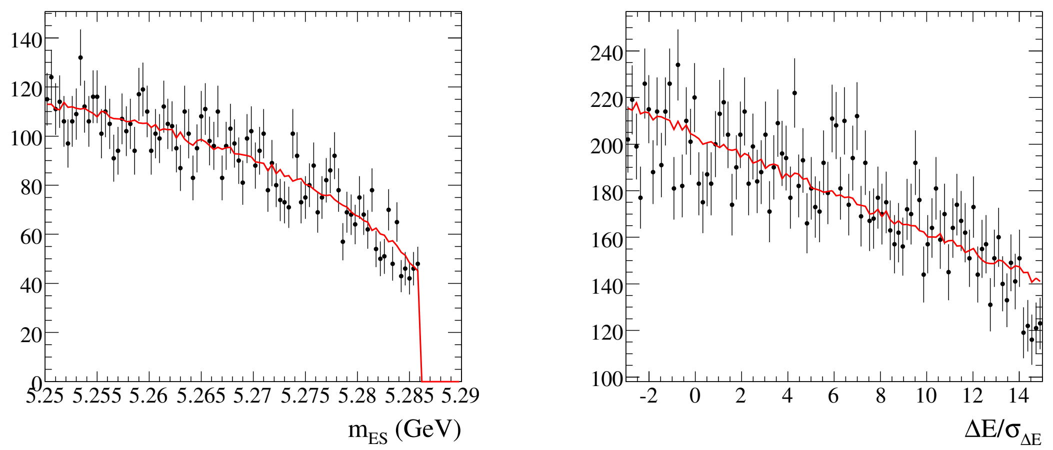


Figure 5.16: Continuum m_{ES} (left) and $\Delta E/\sigma_{\Delta E}$ (right) distributions extracted from the upper sideband and the grand sideband, respectively. The black points represent the data, and the red line, a high statistics toy MC sample generated from the fitted values. The combinatorial m_{ES} profile of the continuum background is parameterized with the empirically motivated "Argus" function [89] ($x\sqrt{1-x^2}\exp -\xi(1-x^2)$ where $x = m_{\text{ES}}/5.29\,\text{GeV}/c^2$ and ξ determines the curvature of the profile), whereas $\Delta E/\sigma_{\Delta E}$ is assumed to be distributed according to a linear PDF.

tagging information ("untagged") also displayed a 29% asymmetry in the number of events assigned to each of the tags, whereas the other tagging categories showed no statistically significant difference. A tagging category-dependent asymmetry was introduced to account for that:

$$\mathcal{P}_{q\bar{q}} = \frac{1}{2}\left(1 + q_{\text{tag}} A^{c}_{q\bar{q}}\right) \mathcal{P}_{q\bar{q}}\left(x, y, \Delta t\right) \tag{5.17}$$

5.8 $B\overline{B}$ Background

Although not as abundant as the continuum background, $B\overline{B}$ background still represents a significant fraction of events after all the selection cuts: as many as 1/3 of the estimated number of signal events. It is at this point that one of those selection

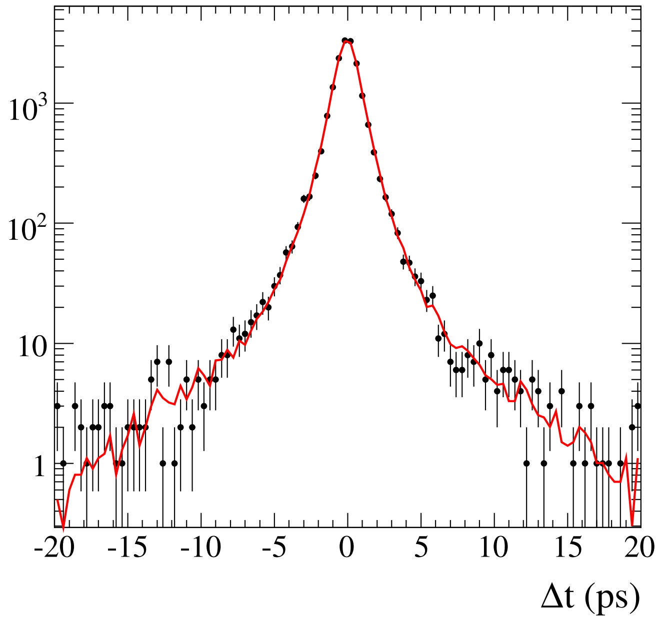


Figure 5.17: Off-peak data are used to extract a resolution function in Δt for continuum events, as well as the fraction of events that have a non-negligible lifetime and the value of that lifetime. A measurable fraction is expected, since charmed resonances are produced in the continuum with lifetimes similar to those of the B meson (e.g. $\tau_{D^\pm} = 1.04(7)\,\mathrm{ps}$). They are found to be $1 - f_{\mathrm{prompt}} = 0.11 \pm 0.01$ and $\tau_{q\bar{q}} = (1.37 \pm 0.12)\,\mathrm{ps}$.

criteria is justified.

The *raison d'être* of the vetoes, mentioned in passing in Section 5.4, is to avoid very high numbers of $B\overline{B}$ background events concentrating in small areas of the Dalitz plot, since their sheer numbers could cause any small discrepancies between our modelling of them and their behaviour in data to pull the signal quantities dramatically in the fit. The charmed mode $B^0 \to D^- (\to K^0_S \pi^-)\, \pi^+$, for instance, has a branching fraction greater than that of signal (~ 1.1 times larger [22]), and essentially the same efficiency, so it contributes to the on-peak sample with around 2000 events. In addition, other decays of the D meson are likely to be misreconstructed as signal through a variety

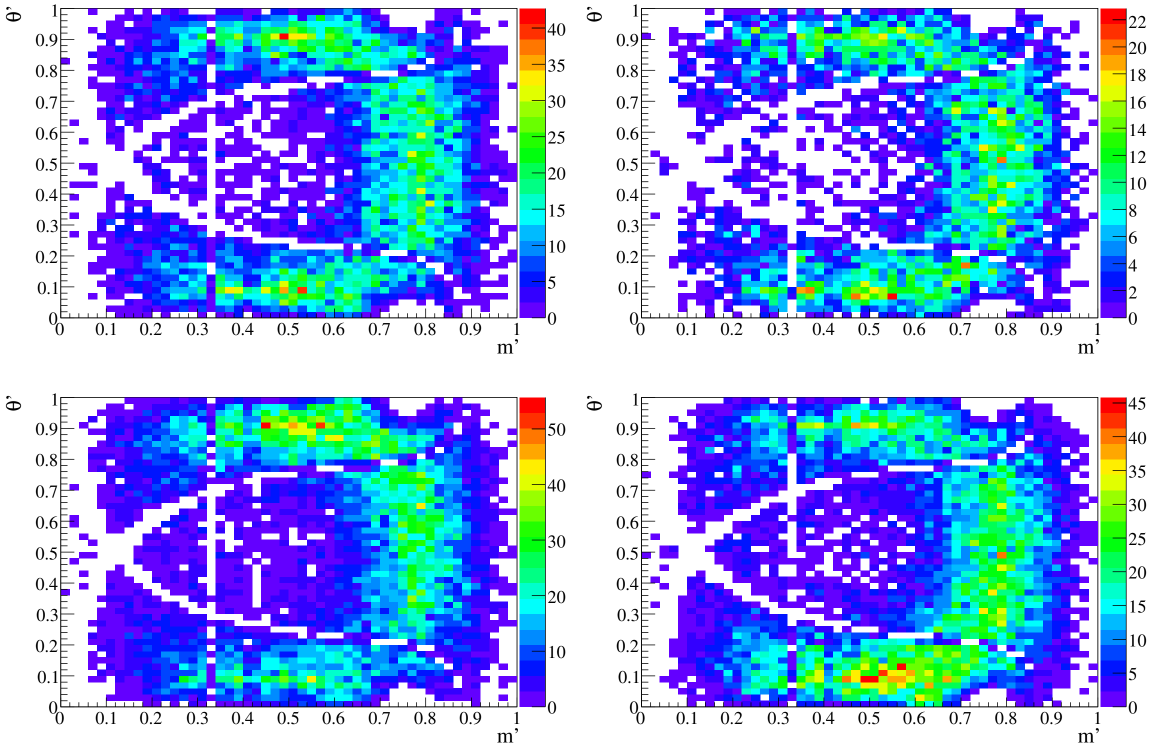


Figure 5.18: Continuum Dalitz-plot distributions split by tag ($\overline{B}^0$ left column, B^0 right column) and tagging category (untagged events, top row; tagged events, bottom). For B^0 ($\overline{B}^0$) tagged events, the K^{*+} (K^{*-}) band is more populated. In untagged events such an asymmetry is not observed; instead, a significant asymmetry in the total number of events assigned to each tag is found, demonstrating that for such events any tag assignment is an artefact rather than a statement on the underlying physical processes.

of mechanisms, such as particle misidentification or losing one of its daughters, most often a neutral particle or a low momentum track. It is probable that these events have an invariant mass close to the true D meson mass, and will therefore appear near the events of the specific decay channel discussed before. Thus, this channel and the charmonium modes $B^0 \to J/\psi\,(\to \ell^+\ell^-)\,K^0_S$ and $B^0 \to \psi(2S)\,(\to \ell^+\ell^-)\,K^0_S$, reconstructed by misidentification of the leptons as pions, are rejected by excluding all events from narrow strips on the Dalitz plot comprising a $\pm 3\sigma$ window around the peak of these distributions. The numerical values of such cuts are shown in Table 5.3, and plots of the peaks in the excluded areas, in Figure 5.19.

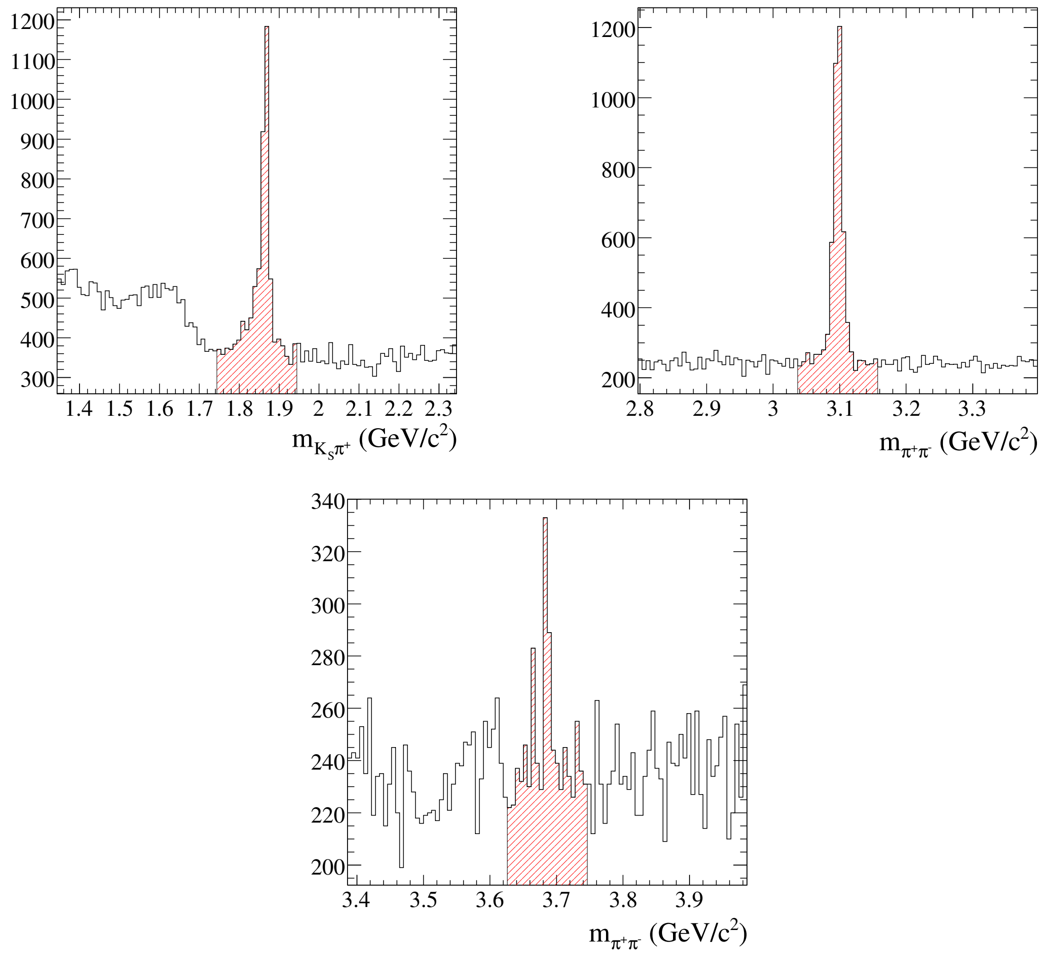


Figure 5.19: Plots showing the $m_{K^0_S\pi^+}$ and $m_{\pi^+\pi^-}$ ranges in which vetoes are applied. Hatched areas represent the excluded ranges. The first and most prominent one shows the D veto, that contains 2790 ± 280 $B\overline{B}$ events; the second one is the J/ψ veto, and the last one, the $\psi(2S)$ veto.

The method employed to identify the remaining $B\overline{B}$ modes that cause backgrounds to this analysis is based on the study of large generic B^0-$\overline{B}^0$ and B^+-B^- MC samples

Table 5.3: Regions of the Dalitz plot excluded (*vetoed*) to remove the most prominent $B\overline{B}$ backgrounds.

Decay Mode	Veto Definition (GeV/c^2)
$D^{\pm} \to K_S^0 \pi^{\pm}$	$1.744 < m_{K_S^0 \pi^{\pm}} < 1.944$
$J/\psi \to \ell^+ \ell^-$	$3.037 < m_{\pi^+\pi^-} < 3.157$
$\psi(2S) \to \ell^+ \ell^-$	$3.626 < m_{\pi^+\pi^-} < 3.746$

(their sizes being roughly equivalent to 2.5 times the analyzed on-peak data sample). All modes contributing more than one event were studied in more detail with the help of exclusive MC samples, from which all the m_{ES}, $\Delta E/\sigma_{\Delta E}$, $\mathcal{F}$, Dalitz-plot and Δt PDFs were extracted. A specific analytical model is needed for the latter depending on the nature of the background mode; the rest of the variables, including the Dalitz plot, were modelled using histograms. The remaining background events are grouped together forming a rather combinatoric-like contribution, without much structure in any of the variables (see Figure 5.23 and text below).

The background modes were divided into three categories that reflect their DP-Δt behaviour:

- B^+B^- backgrounds. Their Δt dependence follows the usual expression for particles with a lifetime that do not mix. These events are composed of true B mesons, and the tagging information is mostly accurate. Thus, the Dalitz-plot distributions for the two tags are expected to be roughly mirror images of each other (recall how a CP transformation acts on the Dalitz plot Section 2.5.6). The complete expression for their likelihood is:

$$\begin{aligned} \mathcal{P}_{B^+B^-} \quad = \quad & \Bigg[\left(\frac{1 - q_{\mathrm{tag}} A_j}{2} \right) \omega^c \frac{P_{B^+B^-}(-q_{\mathrm{tag}}; x, y)}{\int\int_{DP} \; P_{B^+B^-}(-q_{\mathrm{tag}}; x, y) \; dxdy} \\ & + \left(\frac{1 + q_{\mathrm{tag}} A_j}{2} \right) (1 - \omega^c) \frac{P_{B^+B^-}(q_{\mathrm{tag}}; x, y)}{\int\int_{DP} \; P_{B^+B^-}(q_{\mathrm{tag}}; x, y) \; dxdy} \\ & \Bigg] \times \frac{e^{-|\Delta t_{\mathrm{true}}|/\tau_j}}{2\tau_j} \otimes \mathcal{R}^c_{B^+B^-} \end{aligned} \tag{5.18}$$

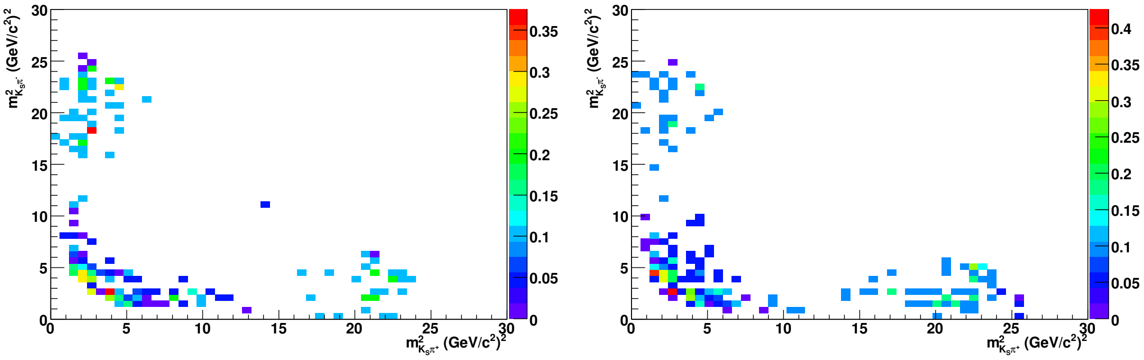


Figure 5.20: Dalitz plot distributions of the charmed $B^{\pm}$ decays for the B^0 (left) and $\overline{B}^0$ tag (right). The events lie mostly on the slow K_S^0 corner, but also sketch the $D^{\pm}$ bands, running close to both edges, and the slow pion corners. Note the slight asymmetry in the population of these, and how the pattern is reversed for the opposite tag. The distributions in the $m_{\rm ES}$ and $\Delta E/\sigma_{\Delta E}$ variables do not exhibit peaking structures.

where the mistag fractions ω^c are extracted from MC truth, and a specific Δt resolution model, taken from the sin2β analyses [24], is used for B^+B^- modes. Misreconstruction effects cause the (effective) lifetimes τ_j to be mode-dependent and mildly different from the nominal value for charged B mesons.

The main contributors to this kind of background are charmed decays of the B^+ meson, with only a handful of events coming from charmless modes (see Table 5.4 for a complete list with their estimated numbers of events). In some of them, the $K_S^0\pi^+\pi^-$ final state is attained through the loss of a photon or a π^0 during the reconstruction, but the major misreconstruction mechanism is the incorporation into the B candidate of low momentum (and therefore poorly determined) tracks arising in the corners of the Dalitz plot from the other side of the event. For instance, the reconstruction of $B^+ \to \pi^0\pi^+K_S^0$ as $K_S^0\pi^+\pi^-$ involves losing the π^0 and adding a low momentum pion from the other side. Note that the large number of events expected in this channel is only due to the uncertainty on its

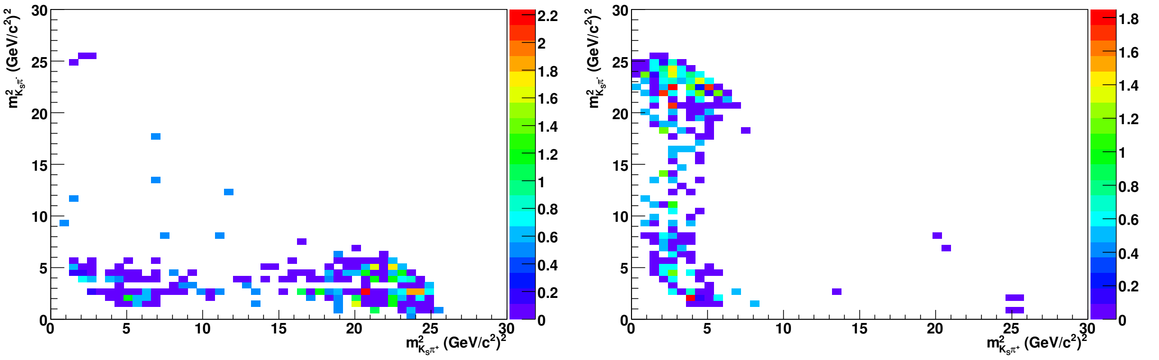


Figure 5.21: Dalitz plot distributions for the $B^0 \rightarrow (flavour\ eigenstate)$ background mode $B^0 \rightarrow D^-\pi^+$ with $D^- \rightarrow X$, separated by the flavour of the signal meson (B^0 left, $\overline{B}^0$ right), as required by the likelihood expression. Note that the events spread widely around the $D^\pm$ vetoes. For this channel, the $m_{\rm ES}$ and $\Delta E/\sigma_{\Delta E}$ distributions are rather featureless, due to the many different decay modes of the $D^\mp$ meson comprised, and hence the very different mechanisms at work in their (mis)reconstruction processes.

branching fraction, for which only an upper limit exists.

- $B^0 \rightarrow (flavour\ eigenstate)$. In this case, the effect that determines the Δt distribution is the mixing of the neutral B mesons before decaying to a final state that uniquely identifies the flavour of the parent particle. In a Dalitz analysis, the final state of the background mode being a flavour eigenstate translates into opposite regions of the Dalitz plot for each flavour, typically the bands corresponding to the $D^\pm$ mass. Hence, rather than using a histogram to model the DP distribution of each tag flavour, the appropriate treatment consists of utilising a histogram for each signal B flavour, which is inferred from MC truth.

Thus, the likelihood expression for such backgrounds is (c.f. Eq. (2.20)):

$$\mathcal{P}_{B_{flav}} = \frac{e^{-|\Delta t|/\tau_j}}{4\tau_j} \Bigg[\left(1 + q_{\rm tag}\frac{\Delta\mathcal{D}_j}{2} + \langle\mathcal{D}_j\rangle \cos\left(\Delta m_{d,j}\Delta t\right)\right) \tag{5.19}$$

$$\times \frac{P_{B_{flav}}(q_{\rm sig} = -q_{\rm tag}; x, y)}{\int\int_{DP} \; P_{B_{flav}}(q_{\rm sig} = -q_{\rm tag}; x, y) \; dxdy} \tag{5.20}$$

$$+ \left(1 + q_{\rm tag}\frac{\Delta\mathcal{D}_j}{2} - \langle\mathcal{D}_j\rangle \cos\left(\Delta m_{d,j}\Delta t\right)\right) \tag{5.21}$$

$$\times \frac{P_{B_{flav}}(q_{\rm sig} = +q_{\rm tag}; x, y)}{\int\int_{DP} \; P_{B_{flav}}(q_{\rm sig} = +q_{\rm tag}; x, y) \; dxdy} \tag{5.22}$$

$$\Bigg] \otimes \mathcal{R}^c_{B_{flav}} \tag{5.23}$$

where $q_{\rm sig} = 1(-1)$ when $B_{\rm rec} = B^0(\overline{B}^0)$, $\mathcal{D}_j$, $\Delta\mathcal{D}_j$, τ_j and $\Delta m_{d,j}$ represent mode-dependent, effective dilutions, lifetimes and oscillation frequencies that may vary from those of correctly reconstructed signal, and the possibility for a Δt resolution model different from signal has also been considered.

Although the dilutions are heavily distorted by the misreconstruction, the resolution model, lifetimes and oscillation frequency are in general rather similar to those of signal. In particular, the lifetimes lie in the range 1.3–1.8 ps, and the frequencies are set to that of signal, $0.502\,{\rm ps}^{-1}$ for all modes except for $B^0 \to D^0 K^0_S$ and $B^0 \to \rho^- K^{*+}$, that have values of $\sim 0.30\,{\rm ps}^{-1}$ (Figure 5.22).

A complete list of the modes contributing to this kind of background, and the numbers of events expected for each of them is shown in Table 5.5. As mentioned above, a moderate fraction of events clusters around the D veto strip, although gaining prominence in the corners, for identical reasons as for the B^+B^- modes (Figure 5.21). All the contributing channels are found to be decays to D and D^* states, except for $B^0 \to \rho^- K^{*+}$, whose DP dependence on the signal flavour is exactly the opposite of the one exhibited by decays to charmed mesons.

- $B^0 \to$ (*CP eigenstate*). Their expected Δt behaviour follows Eq. (2.15), although reconstruction effects may lead to effective values for the $\mathcal{S}$ and $\mathcal{C}$ coefficients. The effective direct *CP* violation was found to be consistent with zero for all the background modes in this category (Table 5.6), so the following expression was used for the per-event likelihood:

$$\mathcal{P}_{B_{CP}} = \sum_c f_c \frac{P_{B_{CP}}(x,y)}{\int\int_{DP} P_{B_{CP}}(x,y)\; dxdy} \frac{e^{-|\Delta t|/\tau_j}}{4\tau_j} \times \left[\left(1 + q_{\text{tag}} \frac{\Delta\mathcal{D}^c}{2}\right) + q_{\text{tag}} \langle \mathcal{D}\mathcal{S}_{\text{eff}} \rangle^c \sin\left(\Delta m_{d,j} \Delta t\right) \right] \otimes\; \mathcal{R}^c_{B_{CP}} \tag{5.24}$$

where $\langle \mathcal{D}\mathcal{S}_{\text{eff}} \rangle^c$ denotes a tagging category-dependent parameter that accounts for the time-dependent asymmetry $\mathcal{S}$ and the effective dilution at the same time.

A list of the background modes that belong to this category is in Table 5.6. The most prominent channel, $B^0 \to \eta' K^0_S$ with $\eta' \to \rho^0\gamma$, is reconstructed as $K^0_S\pi^+\pi^-$ by losing the photon, which gives it a characteristic $\Delta E/\sigma_{\Delta E}$ distribution with the peak shifted to negative values, but leaves the peaking m_{ES} and $\mathcal{F}$ unaffected. The neutral combinatoric has been included in the $B^0 \to$ (*CP eigenstate*) category because it does not exhibit any mixing behaviour. Its time-dependent asymmetry is also consistent with zero, and is therefore assigned $\langle \mathcal{D}\mathcal{S}_{\text{eff}} \rangle^c = 0$ for all c. Finally, the mode $B^0 \to K^0_S K^0_S$ has its Δt distribution distorted by the displaced vertices of the two B daughters, and is found to require a tailored Δt resolution model (Figure 5.23).

5.9 Control sample

Since we rely heavily on MC to extract the shapes of most PDFs, our assumptions about the MC describing the data well should be validated. A good way of doing it consists of using a channel as similar as possible to the one under study in this thesis. For those purposes, the vetoed $B^0 \to D^-\pi^+$ with $D^- \to K^0_S\pi^-$ is ideal, since its final state is identical to that of the charmless $B^0 \to K^0_S\pi^+\pi^-$. The shapes from MC and data for the m_{ES} and $\Delta E/\sigma_{\Delta E}$ variables can be extracted and compared.

The peaks in m_{ES} and $\Delta E/\sigma_{\Delta E}$ are parameterized by two gaussians. Shifts with respect to the values obtained in MC are observed in data for the means in both variables. They are rather small in m_{ES} (~ 1 and $5\,\mathrm{MeV}/c^2$), whereas in $\Delta E/\sigma_{\Delta E}$ one of the means shows no shift ($\sim 0.02 \pm 0.15$) and the other one does ($\sim 0.7 \pm 0.3$). The latter is expected from the unaccounted $B\overline{B}$ background. The ratios of the widths for data and MC are compatible with unity, and any differences are therefore neglected. In both cases, the slight disagreements in the central values are taken into account in the systematic uncertainty computations (Section 6.4). The agreement between data and MC is shown in Figure 5.24.

Table 5.4: B^+B^- background modes. The two most prominent channels, $B^+ \to \overline{D}^0\pi^+$ with $\overline{D}^0 \to K^+\pi^-$ and $\overline{D}^0 \to K^0_S\pi^+\pi^-$, yield a little more than 15 events each. The number of events originating from $B^+ \to \pi^0\pi^+K^0_S$ has been calculated by using half of the upper limit on the branching fraction with an equal error. The values of the branching fractions are taken from either [22] or [45]. The errors on the expected numbers of events arise from the uncertainties on the branching fractions and the statistically limited samples used to evaluate the efficiencies.

Mode (CP conjugate included)	Efficiency	Branching Fraction	Expected events in Signal Region
$B^+ \to \overline{D}^0\pi^+, \overline{D}^0 \to K^+\pi^-$	$(2.58 \pm 0.19) \times 10^{-4}$	$(1.87 \pm 0.08) \times 10^{-4}$	18.5 ± 1.6
$B^+ \to \overline{D}^0\pi^+, \overline{D}^0 \to K^0_S\pi^+\pi^-$	$(2.79 \pm 0.22) \times 10^{-4}$	$(1.43 \pm 0.11) \times 10^{-4}$	15.2 ± 1.7
$B^+ \to \overline{D}^0\pi^+, \overline{D}^0 \to K^+\pi^-\pi^0$	$(3.69 \pm 0.51) \times 10^{-5}$	$(6.94 \pm 0.37) \times 10^{-4}$	9.8 ± 1.5
$B^+ \to D^{*0}\pi^+, D^{*0} \to D^0\pi^0, D^0 \to K^-\pi^+$	$(1.44 \pm 0.13) \times 10^{-4}$	$(1.08 \pm 0.11) \times 10^{-4}$	6.0 ± 0.9
$B^+ \to D^0\rho^+, D^0 \to K^-\pi^+$	$(2.85 \pm 0.45) \times 10^{-5}$	$(5.09 \pm 0.69) \times 10^{-4}$	5.6 ± 1.2
$B^+ \to D^{*0}\pi^+, D^{*0} \to D^0\gamma, D^0 \to K^-\pi^+$	$(1.35 \pm 0.15) \times 10^{-4}$	$(6.66 \pm 0.78) \times 10^{-5}$	3.4 ± 0.6
$B^+ \to D^{*0}\pi^+, D^{*0} \to D^0\pi^0, D^0 \to K^0_S\pi^+\pi^-$	$(6.16 \pm 0.59) \times 10^{-5}$	$(8.26 \pm 0.98) \times 10^{-5}$	1.9 ± 0.3
$B^+ \to D^0\rho^+, D^0 \to K^-\pi^+\pi^0$	$(2.08 \pm 0.85) \times 10^{-6}$	$(1.89 \pm 0.26) \times 10^{-3}$	1.5 ± 0.6
$B^+ \to D^{*0}\pi^+, D^{*0} \to D^0\gamma, D^0 \to K^0_S\pi^+\pi^-$	$(7.10 \pm 0.70) \times 10^{-5}$	$(5.08 \pm 0.68) \times 10^{-5}$	1.4 ± 0.2
$B^+ \to D^{*0}\pi^+, D^{*0} \to D^0\pi^0, D^0 \to K^-\pi^+\pi^0$	$(7.65 \pm 2.31) \times 10^{-6}$	$(4.01 \pm 0.42) \times 10^{-4}$	1.2 ± 0.4
$B^+ \to D^{*0}\pi^+, D^{*0} \to D^0\gamma, D^0 \to K^-\pi^+\pi^0$	$(1.18 \pm 0.29) \times 10^{-5}$	$(2.47 \pm 0.30) \times 10^{-4}$	1.1 ± 0.3
$B^+ \to \pi^0\pi^+K^0_S$	$(5.60 \pm 0.04) \times 10^{-3}$	$(3.30 \pm 3.30) \times 10^{-5}$	70.8 ± 70.8
$B^+ \to K^0_S\pi^+$	$(7.50 \pm 0.51) \times 10^{-4}$	$(7.99 \pm 0.35) \times 10^{-6}$	2.3 ± 0.2
$B^+ \to \rho^+\rho^0$ (Longitudinal)	$(2.08 \pm 0.27) \times 10^{-4}$	$(1.82 \pm 0.30) \times 10^{-5}$	1.5 ± 0.3
$B^+ \to \rho^0K^{*+}, K^{*+} \to K^0_S\pi^+$ (Longitudinal)	$(3.34 \pm 0.10) \times 10^{-3}$	$(1.05 \pm 1.07) \times 10^{-6}$	1.3 ± 1.4
$B^+ \to \pi^+\pi^+\pi^-$	$(1.56 \pm 0.15) \times 10^{-4}$	$(1.62 \pm 0.15) \times 10^{-5}$	1.0 ± 0.1
Combinatorics			11.9 ± 2.7
Total Charged B backgrounds			154.4 ± 70.9

Table 5.5: Summary of the $B^0 \to$ (*flavour eigenstate*) background modes. All modes involve a D or a D^* meson except for the charmless $B^0 \to \rho^- K^{*+}$, whose branching fraction has not been measured. Whenever a generic decay, e.g. $D^- \to X$, and some of its subdecays are listed, the latter have been subtracted from the former in the calculation of the efficiencies and branching fractions.

Mode (CP conjugate included)	Efficiency	Branching Fraction	Expected events in Signal Region
$B^0 \to D^-\pi^+, D^- \to X$	$(9.86 \pm 0.75) \times 10^{-5}$	$(2.78 \pm 0.90) \times 10^{-3}$	105.0 ± 34.9
$B^0 \to D^-\pi^+, D^- \to K^0_S\pi^-\pi^0$	$(2.41 \pm 0.13) \times 10^{-4}$	$(2.38 \pm 0.65) \times 10^{-4}$	21.9 ± 6.1
$B^0 \to D^{*+}\pi^-, D^{*+} \to D^0\pi^+, D^0 \to X$	$(2.43 \pm 0.23) \times 10^{-5}$	$(1.78 \pm 0.21) \times 10^{-3}$	16.6 ± 2.5
$B^0 \to D^0 K^0_S$	$(1.64 \pm 0.03) \times 10^{-3}$	$(2.50 \pm 0.70) \times 10^{-5}$	15.7 ± 4.4
$B^0 \to D^+\rho^-, D^+ \to X$	$(5.31 \pm 1.00) \times 10^{-6}$	$(7.39 \pm 1.20) \times 10^{-3}$	15.0 ± 3.8
$B^0 \to D^-\pi^+, D^- \to K^0_S\pi^-$	$(6.73 \pm 0.22) \times 10^{-4}$	$(5.00 \pm 1.34) \times 10^{-5}$	12.9 ± 3.5
$B^0 \to D^{*+}\pi^-, D^{*+} \to D^0\pi^+, D^0 \to K^-\pi^+$	$(3.35 \pm 0.24) \times 10^{-4}$	$(7.10 \pm 0.55) \times 10^{-5}$	9.1 ± 1.0
$B^0 \to D^{*+}\pi^-, D^{*+} \to D^+\pi^0, D^+ \to X$	$(2.76 \pm 0.28) \times 10^{-5}$	$(8.35 \pm 2.10) \times 10^{-4}$	8.8 ± 2.4
$B^0 \to D^+\rho^-, D^+ \to K^0_S\pi^+$	$(1.70 \pm 0.13) \times 10^{-4}$	$(1.10 \pm 0.18) \times 10^{-4}$	7.2 ± 1.3
$B^0 \to D^-\pi^+, D^- \to K^0_S K^-$	$(1.21 \pm 0.064) \times 10^{-3}$	$(1.01 \pm 0.027) \times 10^{-5}$	4.7 ± 1.3
$B^0 \to D^-\pi^+, D^- \to K^+\pi^-\pi^-$	$(3.12 \pm 1.01) \times 10^{-5}$	$(3.23 \pm 0.86) \times 10^{-4}$	3.9 ± 1.6
$B^0 \to D^{*+}\pi^-, D^{*+} \to D^0\pi^+, D^0 \to K^0_S\pi^0$	$(4.79 \pm 0.18) \times 10^{-4}$	$(2.13 \pm 0.27) \times 10^{-5}$	3.9 ± 0.5
$B^0 \to D^{*+}\pi^-, D^{*+} \to D^+\pi^0, D^+ \to K^0_S\pi^+$	$(6.96 \pm 0.22) \times 10^{-4}$	$(1.25 \pm 0.10) \times 10^{-5}$	3.3 ± 0.3
$B^0 \to \rho^- K^{*+}, K^{*+} \to K^0_S\pi^+$ (Longitudinal)	$(1.81 \pm 0.07) \times 10^{-3}$	$(2.0 \pm 2.0) \times 10^{-6}$	1.4 ± 1.4
Total Flavour Neutral B backgrounds			229.4 ± 36.4

Table 5.6: Summary of $B^0 \to$ (CP *eigenstate*) background modes. Although $a_1^+\pi^-$ is not a CP eigenstate, it is not a flavour eigenstate either, since $a_1^-\pi^+$ can be reached from a B^0 with equal probability to $a_1^+\pi^-$. However, the sum of the two can be treated as a CP eigenstate.

Mode	Efficiency	Branching Fraction	Expected events in Signal Region
$B^0 \to \eta' K^0_S, \eta' \to \rho^0\gamma$	$(6.55 \pm 0.09) \times 10^{-3}$	$(9.56 \pm 0.59) \times 10^{-6}$	24.0 ± 1.5
$B^0 \to K^0\mu^+\mu^-$	$(1.10 \pm 0.00) \times 10^{-1}$	$(2.00 \pm 1.3) \times 10^{-7}$	8.4 ± 5.5
$B^0 \to K^0_S K^0_S$	$(5.39 \pm 0.05) \times 10^{-2}$	$(2.30 \pm 0.50) \times 10^{-7}$	4.7 ± 1.0
$B^0 \to a_1^+\pi^-$ & $B^0 \to a_1^-\pi^+$	$(2.45 \pm 0.08) \times 10^{-4}$	$(3.97 \pm 0.37) \times 10^{-5}$	3.7 ± 0.4
$B^0 \to J/\psi K^0_S$	$(1.71 \pm 0.70) \times 10^{-5}$	$(2.71 \pm 0.11) \times 10^{-4}$	1.8 ± 0.7
$B^0 \to \pi^+\pi^-\pi^0$	$(1.37 \pm 0.09) \times 10^{-4}$	$(25.8 \pm 2.6) \times 10^{-6}$	1.4 ± 0.2
$B^0 \to a_1^0 K^0_S$	$(1.03 \pm 0.08) \times 10^{-3}$	$(3.50 \pm 3.5) \times 10^{-6}$	1.4 ± 1.4
Combinatorics			29.2 ± 4.0
Total CP Neutral B backgrounds			74.6 ± 7.2

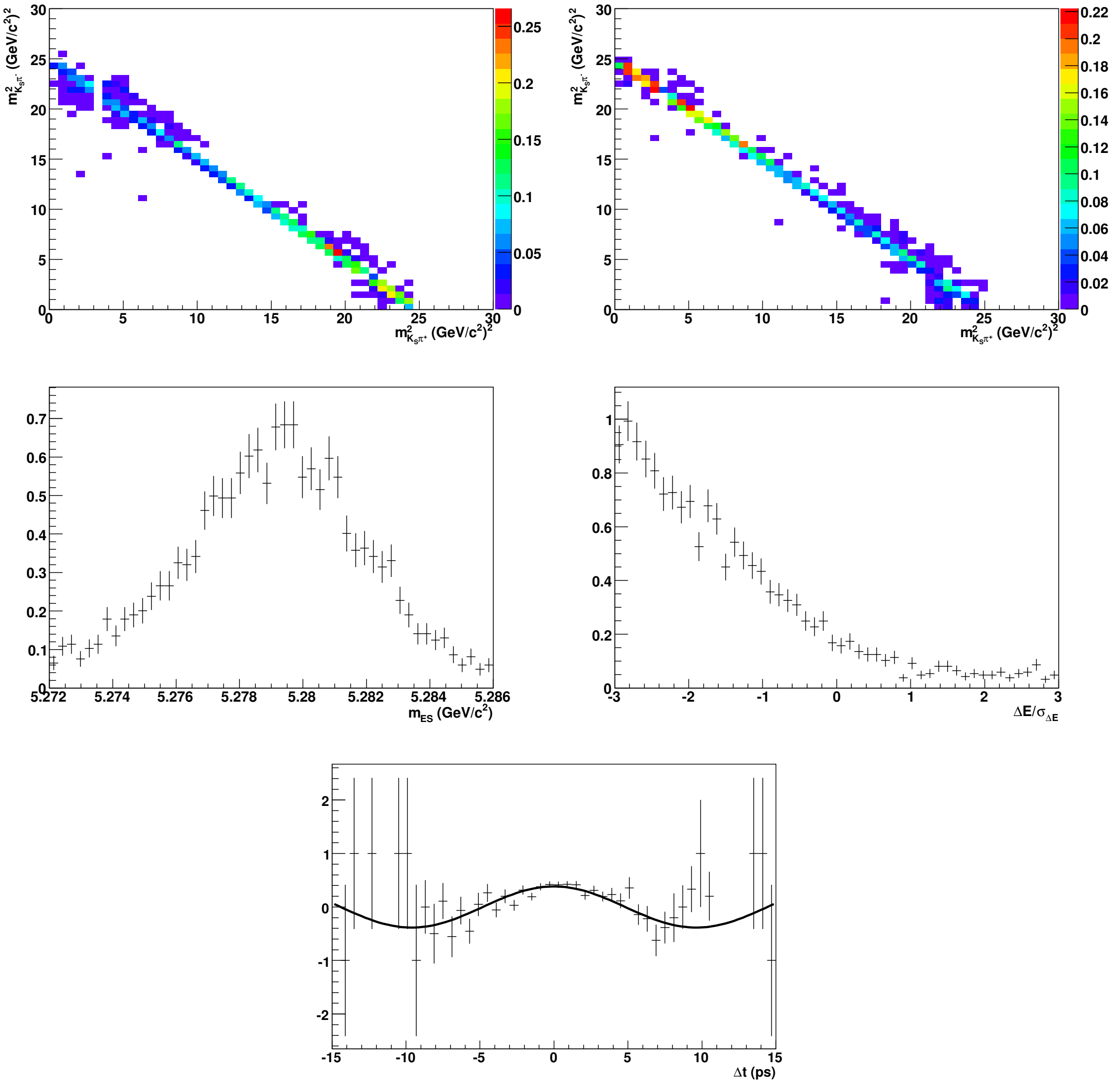


Figure 5.22: Distributions for the $B^0 \rightarrow$ (*flavour eigenstate*) background mode $B^0 \rightarrow D^0 K_S^0$. The channel has very clear peaking structures both in m_{ES} and $\Delta E/\sigma_{\Delta E}$, the latter showing that some of the D^0 daughters have been lost in the reconstruction. The bottom plot, depicting the ratio of the number of unmixed minus mixed events over their sum at each point in Δt, exhibits a profile typical of mixing. The dilution can be extracted from it simply as the amplitude of the cosine wave, and its average over all tagging categories is found to be $\langle \mathcal{D}_{\mathrm{eff}} \rangle = 0.385 \pm 0.025$ (c.f signal, $\langle \mathcal{D} \rangle \simeq 0.44$). The mixing frequency is seen to be $\Delta m_{d,\mathrm{eff}} = 0.327 \pm 0.029$, presumably distorted by the unreconstructed displaced vertex of the D^0.

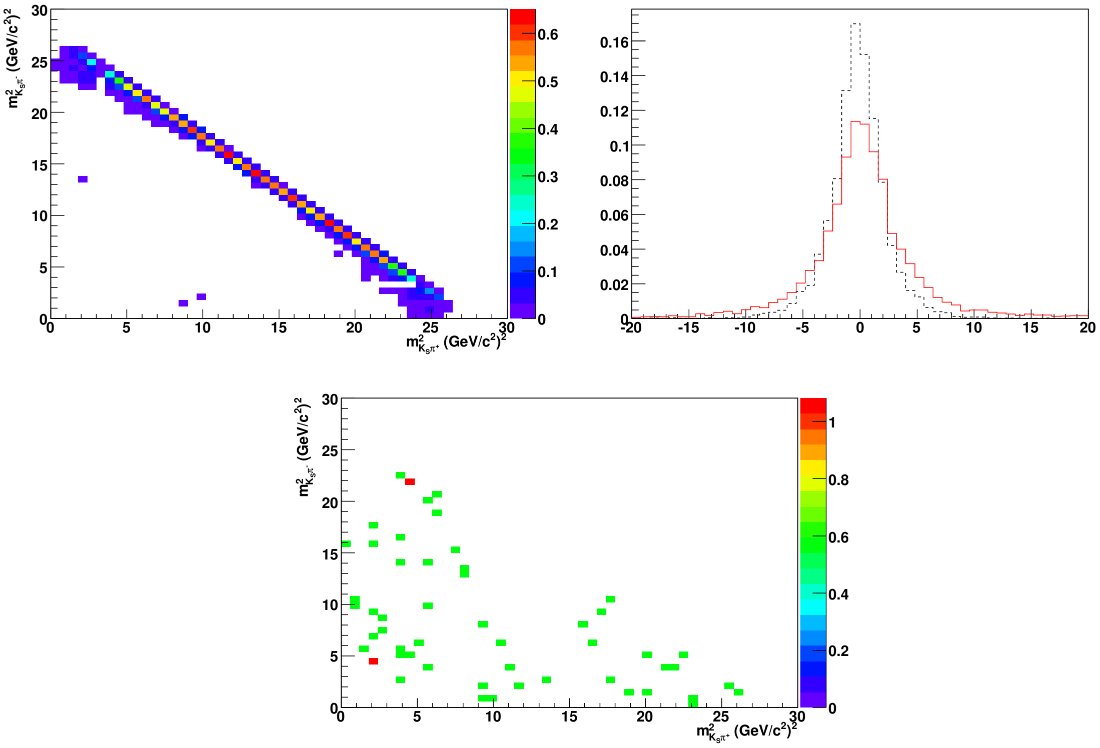


Figure 5.23: Dalitz plot distributions for the most prominent $B^0 \to$ (*CP eigenstate*) background mode, $B^0 \to \eta' K^0_S$ with $\eta' \to \rho^0\gamma$ (top left), and for the neutral combinatorics (bottom). On the top right is a plot showing the Δt distributions of $B^0 \to \eta' K^0_S$ (black, dashed), and $B^0 \to K^0_S K^0_S$ (red, solid). Whereas the signal Δt resolution model is perfectly appropriate for the former, the latter requires another one that differs significantly from that of signal: $s_{\text{core}} = 2.12$, $s_{\text{tail}} = 5.0$ and $s_{\text{outlier}} = 9.0$ (c.f. Table 4.2). With that model, however, the lifetime and the mixing frequency take values close to those of signal.

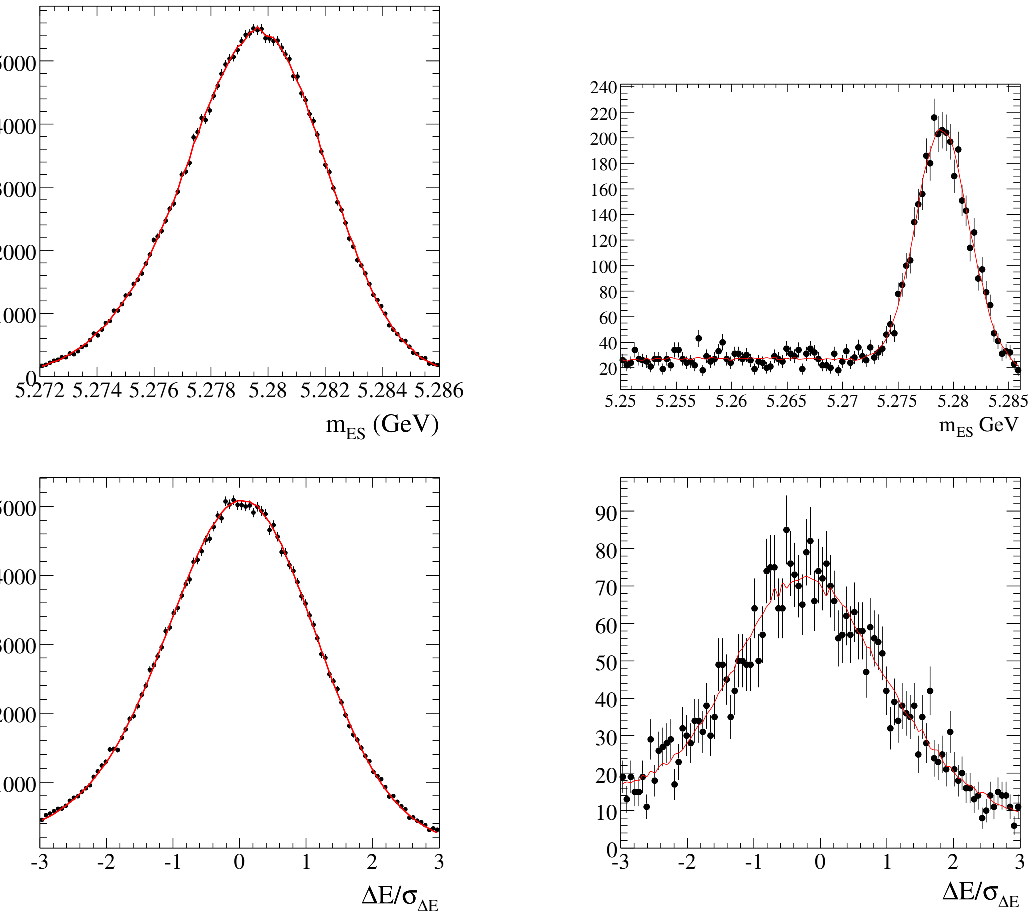


Figure 5.24: Comparisons of the m_{ES} (top) and $\Delta E/\sigma_{\Delta E}$ (bottom) PDF shapes in MC (left) and data (right) for the control sample. The $B\overline{B}$ background has been neglected in both variables, and the continuum background has been modelled as an Argus function in m_{ES}, and a first order polynomial in $\Delta E/\sigma_{\Delta E}$. Good agreement is observed in both cases, making unnecessary any correction of the PDF parameters extracted from MC.

Chapter 6

Results and Conclusions

6.1 Toy MC tests

As sketched in Section 4.6, a standard procedure to check the correct behaviour of the fit is to generate toy MC according to some given PDFs, employ the same PDFs to fit those toy simulated data, and compare the recovered values with those used to generate the data by producing pull plots for all the parameters that are estimated by the fit.

A number of different scenarios were explored in these tests. The resonance content was varied, and so were the relative magnitudes and phases between the resonances. Biases were looked for among the fitted parameters, estimated as the difference between the mean of the distribution of fitted values (which was assumed to be gaussian) and the true value. It was found that simpler models would either show no biases, or small ones (~ 0.2). As the number of resonances increased, so did the biases, for two reasons: first, by keeping the number of events per experiment constant, the quantity of events assigned to each resonance diminished; second, strong correlations often appeared between the c_j coefficients of different resonances. However, as we shall see, these correlations conspire to eliminate most of the biases from the physical observables.

Here we only present the results of a toy MC bias study in which the values of all parameters used in the simulation have been set to those extracted from data, in order to examine in detail the reliability of the fit for the region of interest of the solutions space. A careful description of the model is given in Section 6.3. Once the decay model has been established, 250 toy MC samples are generated, each one containing the amount of continuum background, $B\overline{B}$ background and signal events expected in the on-peak sample. These 250 "experiments" must then be fitted, and the results for each parameter compared to the true values.

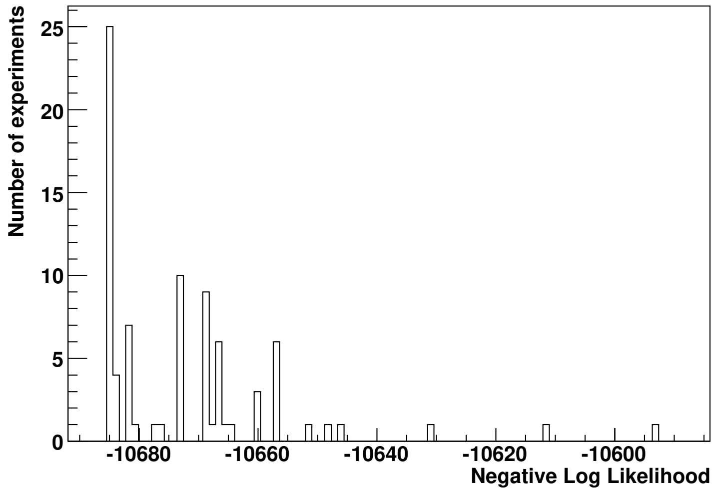


Figure 6.1: Negative Log Likelihood minima attained by 90 fits to a single experiment with randomized initial values.

Not all the degrees of freedom in the c_j are physical, since only relative magnitudes and phases can be measured. Hence, one magnitude and one phase are fixed to reference values, usually chosen as 1.0 and 0.0, respectively. In this analysis, it is the magnitude and phase of the B^0 and $\overline{B}^0$ average amplitude of a resonance that are fixed to those values. A further phase has to be fixed to zero, since the quantity $\phi_{\rm mix}$ appearing in Eq. (5.1) is just an overall phase difference between B^0 and $\overline{B}^0$, but, like the latter, it must be a *CP*-odd phase.

All in all, the number of free parameters in such fits is rather high: 24 in the case of the fit to data (two of them being the signal and continuum yields, and the other 22 being the relative magnitudes and phases of the 6 components of the decay models for B^0 and $\overline{B}^0$). As a consequence, the landscape of extrema of the likelihood hypersurface in this high-dimentionality space is rather complex, and the location of the global maximum turns into a non-trivial task. The fit often does not converge to the global maximum, so a multitude of fits with randomized initial values is needed from which the fit with the highest likelihood is selected. Figure 6.1 illustrates this by showing the different maxima to which the fits converge for a typical experiment. In this analysis, 90 fits were made to each sample. Plots showing the distribution of fitted values for the observables from the best of the fits can be seen in Figures 6.2 to 6.6.

6.2 Fully simulated MC tests

Fully simulated MC, as opposed to toy MC, is used to check whether any neglected effects, such as the experimental resolution, self cross feed or ignored correlations between variables, are more important than initially estimated and should be accounted for in the fit. It cannot be used to study the biases of the fit or the expected sizes of the statistical uncertainties because the actual resonance content of the Dalitz plot model, and the relative magnitudes and phases are only found from the later fit to data.

This full MC, in which the $B^0 \to K_S^0\pi^+\pi^-$ decay is modelled to proceed through the $K^{*\pm}(892)\pi^{\mp}$, $\rho^0(770)K_S^0$ and $f_0(980)K_S^0$ resonances with arbitrary phases but in roughly the relative fractions expected in data, is used to generate 250 samples, which are mixed in the appropriate proportions with the continuum and $B\overline{B}$ background simulated with the PDFs of the nominal fit to data. Like before, the B^0 and $\overline{B}^0$ average of the magnitude and phase of $K^{*\pm}(892)\pi^{\mp}$ are kept fixed in the fit. The additional degrees of freedom fixed in the fit are in this case taken to be the aysmmetry terms in $f_0(980)K_S^0$.

Figures 6.7 to 6.9 show the fitted distributions and the values used for the gen-

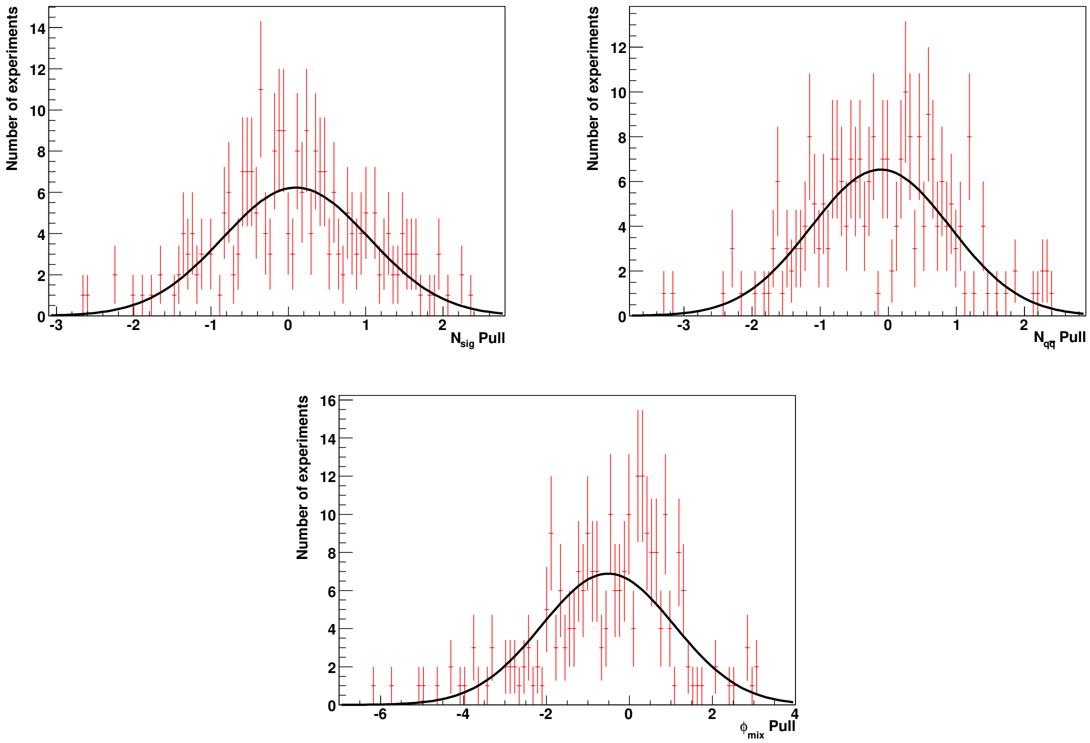


Figure 6.2: Pull plots for ϕ_{mix} and the signal and continuum yields. The black curves are gaussian fits to the pull distributions. The fitted yields are excellent, with both means and widths of the pull distributions consistent with 0.0 and 1.0, respectively. For ϕ_{mix}, the mean and width are found to be $\mu = -0.52 \pm 0.10$ and $\sigma = 1.60 \pm 0.08$, indicating a poor performance of the fit. This is due (see Section 6.5) to this parameter being dominated by a very statistically limited resonance.

eration. Noticeable, but not extremely large, differences between those values are observed in several of those plots. These are expected to be smaller in the actual fit to data, since a higher number of resonances implies fewer events per resonance, and hence less sensitivity to small effects like those mentioned above. Note, in contrast, the good performance of the fit in recovering ϕ_{mix}. The signal and continuum yields show biases of $\sim 0.75 \times \sigma$ and $\sim 0.60 \times \sigma$, which amount to ~ 38 and ~ 57 events, respectively. None of the biases are corrected for, and are instead dealt with by assigning

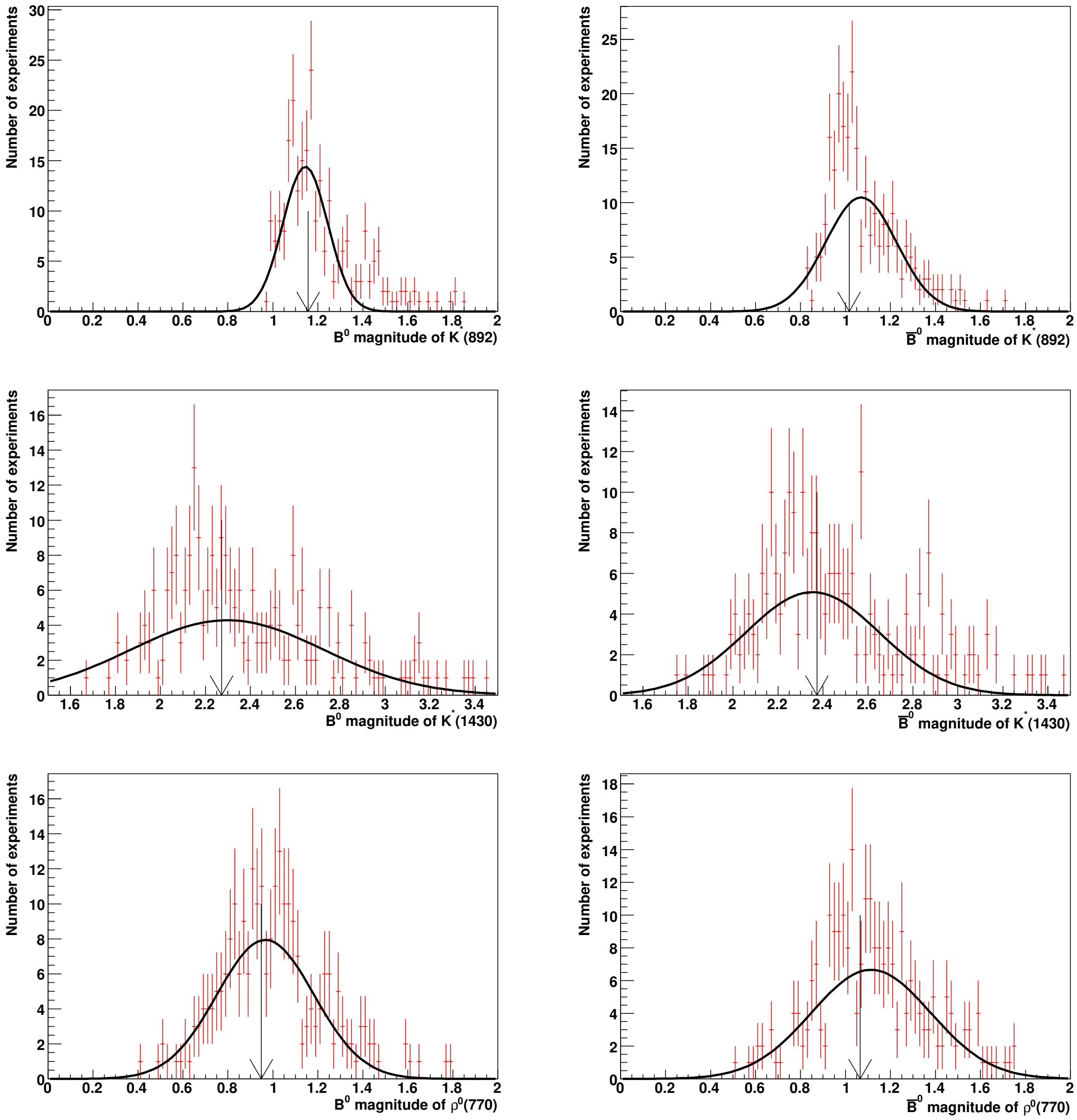


Figure 6.3: Distributions of fitted values (red points) for the magnitudes of the $K^{*\pm}(892)\pi^{\mp}$, $K_0^{*\pm}(1430)\pi^{\mp}$ and $\rho^0(770)K_S^0$ components in the B^0 and $\overline{B}^0$ decay models. The black curves are gaussian fits to the distributions, and the black arrows point to the true value. No significant bias is observed, although non-gaussian distributions are found in the first two plots. This is due to the fit configuration, in which one magnitude must be fixed. In out case, it is the average of the B^0 and $\overline{B}^0$ magnitudes of the first resonance that is fixed to 1.0.

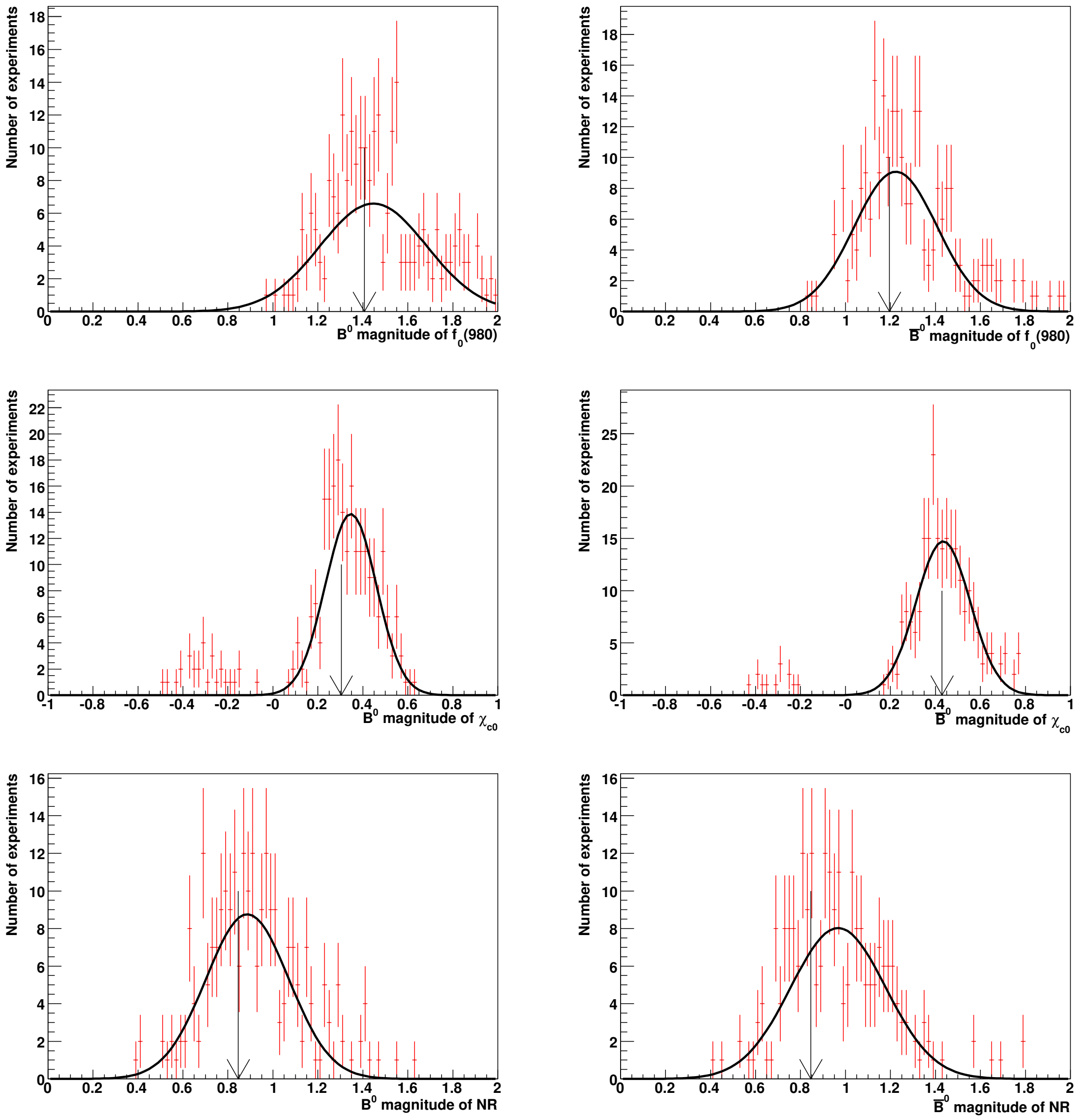


Figure 6.4: Distributions of fitted values (red points) for the magnitudes of the $f_0(980)K_S^0$, $\chi_{c0}K_S^0$ and non-resonant components in the B^0 and $\overline{B}^0$ decay models. The black curves are gaussian fits to the distributions, and the black arrows point to the true value. A secondary solution is hinted at in the plots of the middle row, that correspond to the magnitudes of the resonance with the smallest contribution to the model. The bottom row shows the distributions of the magnitude for the non-resonant component, whose $\overline{B}^0$ component exhibits a noticeable, but manageable, bias.

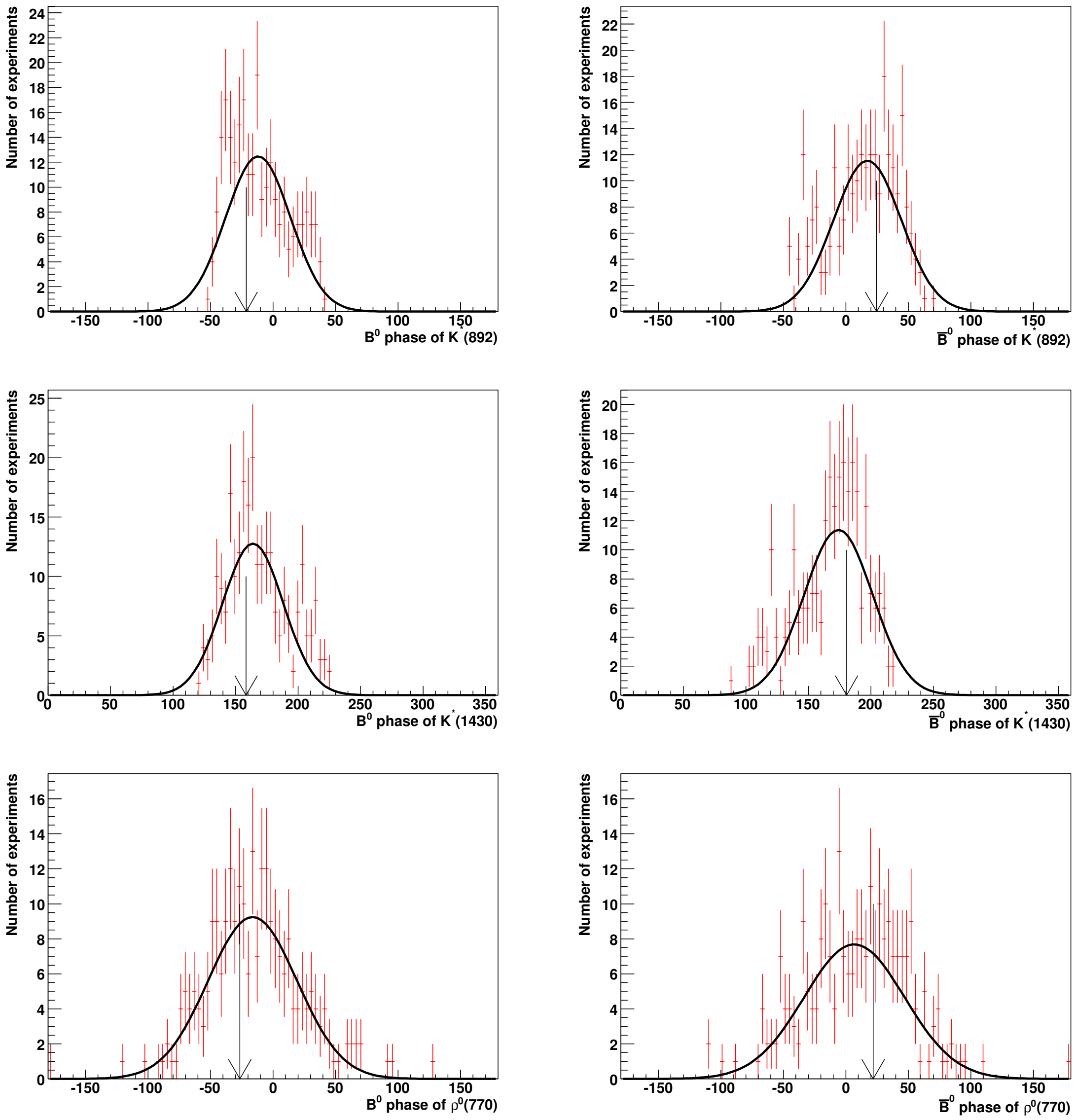


Figure 6.5: Distributions of fitted values (red points) in degrees for the phases of the $K^{*\pm}(892)\pi^{\mp}$, $K_0^{*\pm}(1430)\pi^{\mp}$ and $\rho^0(770)K_S^0$ components in the B^0 and $\overline{B}^0$ decay models. The black curves are gaussian fits to the distributions, and the black arrows point to the true value. Phases are generally not as well determined as magnitudes. In spite of that, no worrying biases are observed. Notice that, the top row exhibits once more non-gaussian profiles, this time due to the phase of the average of those components being fixed to 0.

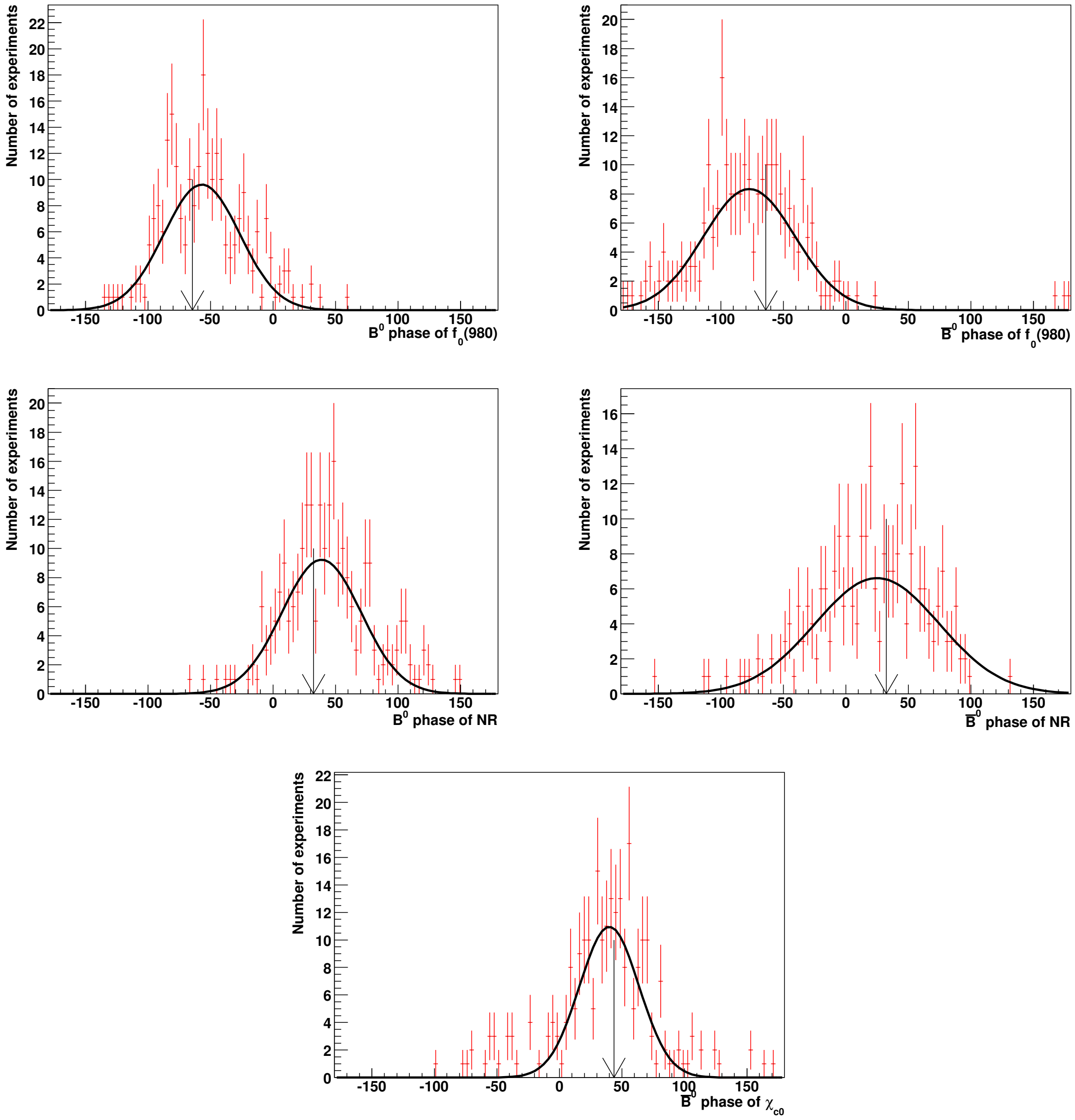


Figure 6.6: Distributions of fitted values (red points) in degrees for the phases of the $f_0(980)K^0_S$, $\chi_{c0}K^0_S$ and non-resonant components in the B^0 and $\overline{B}^0$ decay models. The black curves are gaussian fits to the distributions, and the black arrows point to the true value. No large biases are observed. Only one parameter is shown for the last resonance, as the other one is absorbed into $\phi_{\rm mix}$.

appropriate systematic uncertainties (see Section 6.4).

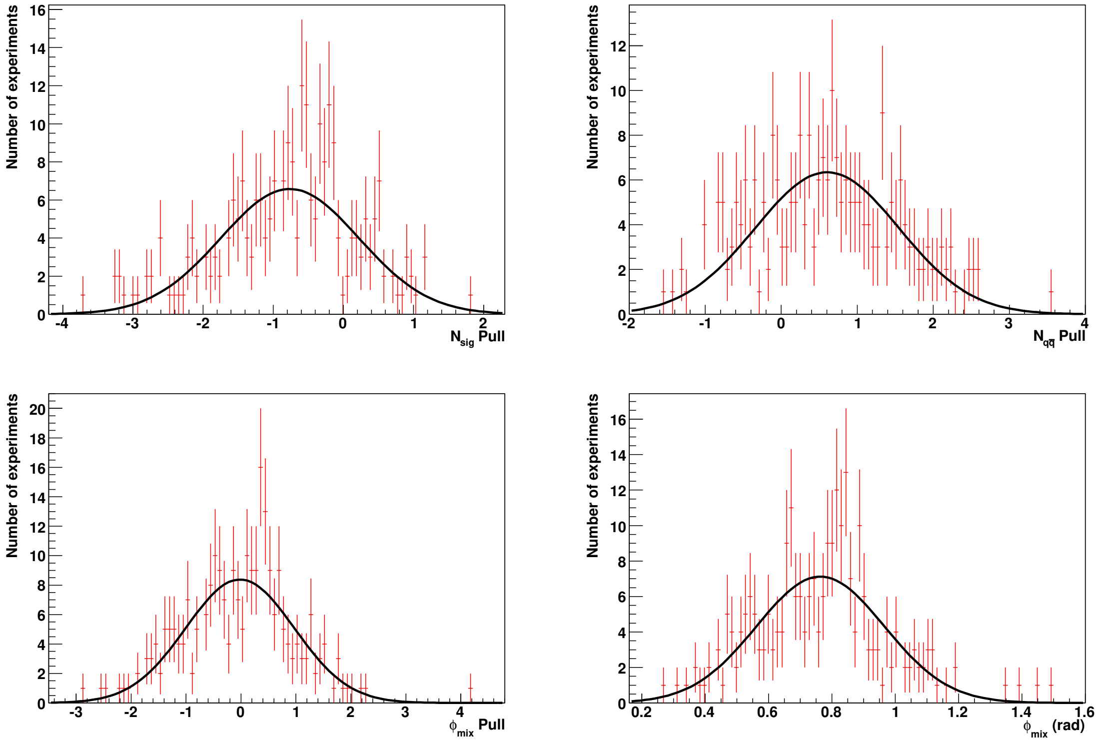


Figure 6.7: Pull plots for ϕ_{mix} and the signal and continuum yields in the full MC tests. The last two show slight biases, but the recovery of ϕ_{mix} by the fit is excellent, with the mean and width of the pull plot consistent with 0.0 and 1.0, respectively. Note that, unlike other time-dependent analyses, we determine ϕ_{mix} rather than $\sin\phi_{\text{mix}}$, and that the mirror solution $(\pi - 2\beta)$ is not observed in the pull plot. This is better appreciated in the lower right plot, where none of the fits to the 250 experiments is seen to converge to $\phi_{\text{mix}} = \pi - 2\beta \simeq 2.4$ rad. This plot demonstrates the capabilities of Dalitz-plot analyses, boasting an error on ϕ_{mix} of ~ 0.2 rad $\simeq 11°$.

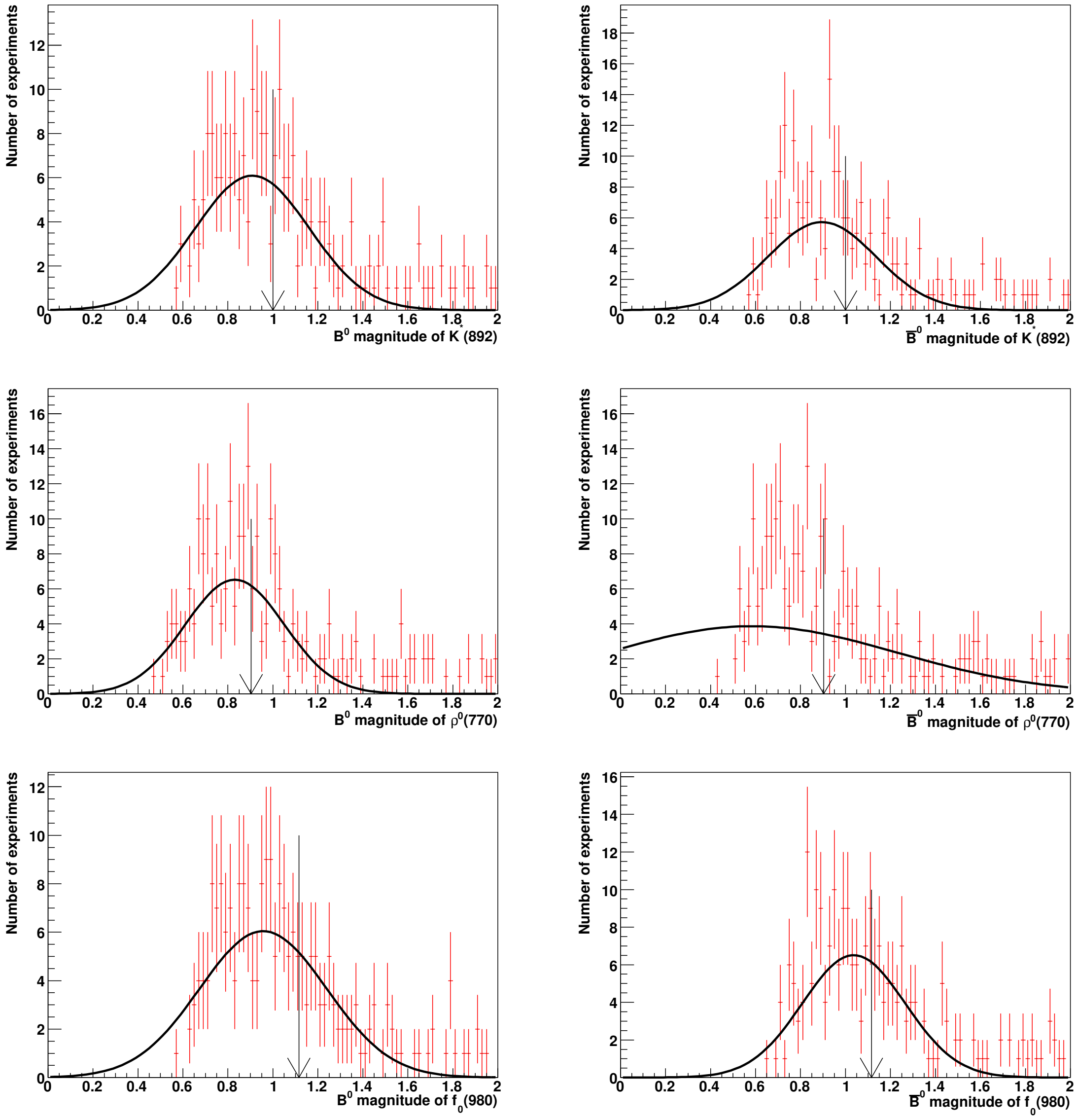


Figure 6.8: Distributions of fitted values (red points) for the magnitudes of all the components in the full MC tests. The black curves are gaussian fits to the distributions, and the black arrows point to the true value. Non-gaussian tails are observed in all three components, presumably due to the neglected resolution effects and self cross feed. Modelling of the latter was shown to improve these tails in Section 5.6.

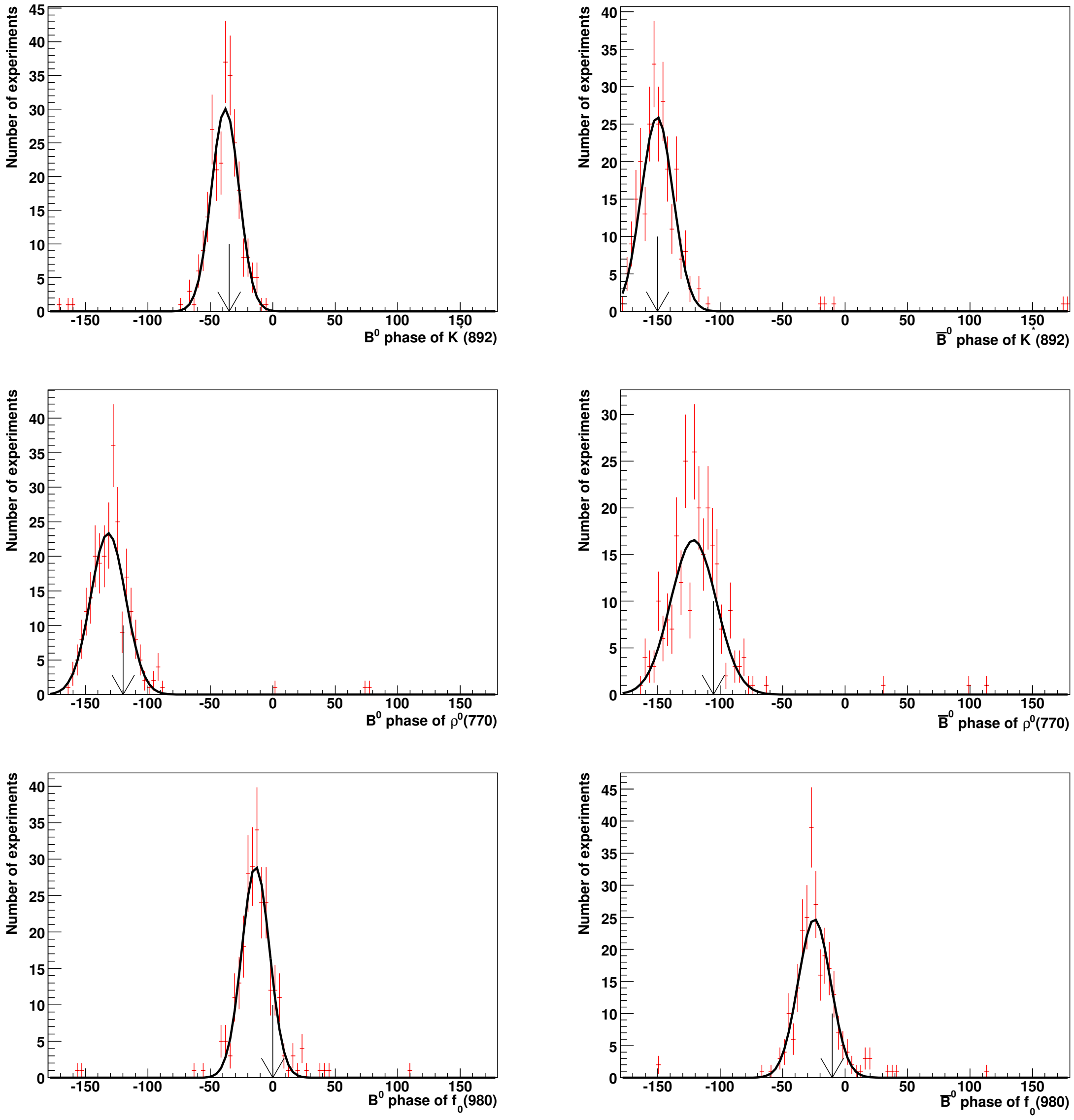


Figure 6.9: Distributions of fitted values (red points) in degrees for the phases of all the components in the full MC tests. The black curves are gaussian fits to the distributions, and the black arrows point to the true value. Phases have smaller errors than in Section 6.1 because a lower number of resonances leads to more events associated to each of them. Significant biases are observed, although the central values are still within 1σ of the true value.

6.3 Exploring the Dalitz plot structure

Since the resonant structure of the decays is *a priori* unknown, a procedure must be in place to systematically identify and assess in data the significance of each component added in the model. Following previous analyses [31, 10], we start from a minimal model, established by earlier measurements, and include one more resonance at a time. If small fit fractions are found for any of the resonances in the minimal model, fits excluding the resonance in question are performed. Either way, estimating the associated statistical errors requires lengthy toy MC studies (see Sections 6.1, 6.5), so the approach adopted consists of simply evaluating the difference in the likelihood calculated in the fit with and without the component under study, and checking that the results for the well established resonances are consistent in both cases. A spurious resonance could be assigned large fit fractions by the artefact of extremely large interferences with other components, thus changing greatly at least one of the well determined fit fractions.

In $B^0 \to K^0_S\pi^+\pi^-$, the minimal model is formed by $K^{*\pm}(892)\pi^{\mp}$, $K^{*\pm}_0(1430)\pi^{\mp}$, $\rho^0(770)K^0_S$, $f_0(980)K^0_S$, $\chi_{c0}K^0_S$ and a non-resonant component with a constant magnitude and phase over the Dalitz plot (*flat*). The model is based on *BABAR*'s findings from the charged $B^+ \to K^+\pi^-\pi^+$ [10], which is a cleaner experimental analysis, as well as *Belle*'s $B^0 \to K^0_S\pi^+\pi^-$ and $B^+ \to K^+\pi^-\pi^+$ results [25]. As for the parameterization of each component, according to the discussion in Section 2.5.4, the states $K^{*\pm}(892)\pi^{\mp}$, $\rho^0(770)K^0_S$ and $\chi_{c0}K^0_S$ are known to follow closely a relativistic Breit-Wigner lineshape, whereas $f_0(980)K^0_S$ requires a Flatté propagator. The $K^*_0(1430)$ states, however are found not be well described simply by a (relativistic) Breit-Wigner, and a LASS-type parameterization is preferred by the $K^+\pi^+\pi^-$ data. The nature of the non-resonant component is even less understood than that of the $K^*_0(1430)$. In the analyses mentioned above, fits to larger data samples than the one expected for the present study of $B^0 \to K^0_S\pi^+\pi^-$ seem to show slight improvements when non-uniform non-resonant terms are used. There are, however, no clear guiding principles either theoretical or empirical that convincingly favour one model over the others. In such a situation, the simplest option (flat) is chosen for this analysis, since it is observed to result in smaller correlations with the other intermediate states in the decay model.

A rather subtle point that needs to be discussed before proceeding, and that was only mentioned in passing in the previous sections is the conventions followed concerning which parameters are kept fixed in the fit. Clearly, one magnitude and one phase have to be fixed, since the overall normalization is related to the signal event yield, n_{sig}, and only relative phases are physically meaningful. Almost any choice will do, as long as, to avoid statistical limitations, a prominent resonance is selected that interferes strongly with the other components. The question of which other phase to fix to make up for the free parameter accounting for the mixing phase, ϕ_{mix}, is not as straightforward. In principle, the mixing phase is independent of the resonant state by which the (un)oscillated meson decays. There can be, however, weak phases that depend on the resonant submode, and these would be indistinguishable from ϕ_{mix}. $B^0 \to K^0_S\pi^+\pi^-$ is in fact a very good example, as all the charmless decays to this final state include a $b \to u$ tree contribution carrying the phase γ, as depicted in Figure 2.7. Furthermore, the level of "contamination", and hence the effective weak phase, brought by these diagrams also varies from one resonance to another, as it is immediately apparent when confronting, for instance, the $\rho^0(770)$ and $f_0(980)$ states. Whereas the former couples to $u\bar{u}$ and $d\bar{d}$, the latter also has an $s\bar{s}$ component. Ultimately, ϕ_{mix} can be defined operationally as the weak phase measured in charmonium decays. Following that spirit, we decide to use the parameterization Eq. (2.49) for the $\chi_{c0}K^0_S$ subchannel and fix its *CP*-odd phase to zero, thereby forcing ϕ_{mix} to play its role. This channel is assured to be free from Standard Model contaminations. The price to pay is the very statistically limited determination of ϕ_{mix}, due to the rather small expected fit fraction for that channel[1].

The fit fractions of all six components have been established to have at least a 3σ significance in other analyses. Our work, then, consists only of adding new resonances

[1]Alternatively, the idea of having a global ϕ_{mix} parameter can be given up, and the weak phases of each resonant channel measured independently.

and assessing their impact on the description of the decay. Introducing more free parameters can only improve the fit, which means that the changes in the total likelihood induced by the addition of a new component will have to be moderately large, on the order of dozens of units of likelihood or more, to be considered significant. Large asymmetries in the fitted fractions for the resonance under scrutiny can be considered as circumstantial evidence against its presence, as statistical fluctuations could easily produce such situations.

The resonant modes tested are: $\omega(782)K_S^0$, $K_2^{*\pm}(1430)\pi^{\mp}$, $f_0(1370)K_S^0$, $f_2(1270)K_S^0$ and $f_X(1300)K_S^0$, the scalar seen by $\mathcal{B}elle$ in [55]. Only the last one, for which new estimations of the mass and width [90] are used instead of $\mathcal{B}elle$'s, is found to induce a change in the likelihood large enough to suggest that it really appears in the decay. The difference is still rather small, so it is decided not to include this state in the nominal model. Table 6.1 shows the changes in the likelihood associated to each resonance, and their fitted branching fractions.

6.4 Systematic uncertainties

The Dalitz-plot related systematic uncertainties come, in part, from the statistical uncertainties on the bin contents of the histograms that model the efficiency variation across the Dalitz plot, the migration of self cross feed events and the distributions of $q\overline{q}$ and $B\overline{B}$ backgrounds. These uncertainties on the bin contents can be understood as an uncertainty on the shape of the corresponding PDFs. The uncertainty on the normalization of those PDFs, i.e. the number of expected background events, further contributes to the systematic error. Other parameters that are kept fixed during the fit contribute as well, such as those involved in the Δt resolution model, the fractions of signal events belonging to each tagging category, or the m_{ES} and $\Delta E/\sigma_{\Delta E}$ distributions. Fit biases are also taken care of, by assigning systematic errors. Finally, there are the so-called model errors, which arise from uncertainties in the structure of the Dalitz Plot and its parameterization. These are the harder to assess, and the

Table 6.1: Table showing the changes in the likelihood and the fit fractions induced by the addition of further resonances to the decay model. The first column shows the values of the fit fractions obtained from data for the $\overline{B}^0$ minimal model. The rest of the columns show their values when additional resonances are introduced in the model. Only the changes in the likelihood associated to the addition of the $f_X(1300)$ seem to indicate something else besides the expected improvement due to the larger number of free parameters.

	Minimal model	With $\omega(782)$	With $f_2(1270)$	With $f_0(1370)$	With $f_X(1300)$	With $K_2^{*0}(1430)$
$\Delta \log \mathcal{L}$	–	1.14	7.95	11.56	17.43	10.09
$K^{*\pm}(892)\pi^{\mp}$ fraction	9.6%	9.6%	9.5%	9.3%	9.4%	9.7%
$K_0^{*\pm}(1430)\pi^{\mp}$ fraction	53%	52%	52%	52%	52%	51%
$\rho^0(770)K_S^0$ fraction	11%	11%	9.7%	10%	11%	10%
$f_0(980)K_S^0$ fraction	13%	14%	14%	8.8%	11%	13%
$\chi_{c0}K_S^0$ fraction	1.7%	1.7%	1.8%	1.7%	1.7%	1.5%
$K_S^0\pi^+\pi^-$ NR fraction	8.7%	8.4%	8.4%	11%	11%	9.6%
Additional submode ($\overline{B}^0$)	–	0.034%	4.9%	1.4%	1.9%	2.4%
Additional submode (B^0)	–	0.37%	2.2%	3.3%	2.1%	5.0%

most important in size for some of the most interesting observables, such as the time-dependent asymmetries.

The calculation of all these uncertainties implies redoing the fit to the data with a multitude of different values for the fixed parameters or shapes, and is therefore extremely computationally intensive. For this reason, the actual evaluation of the errors has not been done. Nevertheless, a careful examination of all the possible sources of systematic uncertainties and the ways to estimate them follows, giving rough orders of magnitude as to what those uncertainties would be if extrapolated from similar analyses.

Clearly, these uncertainties must be calculated before results can be publicly presented.

6.4.1 PDF parameter fluctuations

All the signal, $q\bar{q}$ and $B\overline{B}$ m_{ES}, $\Delta E/\sigma_{\Delta E}$ and Fisher PDF parameters are fixed to the values fitted from MC or sideband data. For each of these parameters, the fit is repeated 100 times taking, instead of the nominal value of the parameter, a value generated from a gaussian whose mean and width are the central value and the error of the parameter. The RMS of the resulting distribution for each measured parameter is taken as the associated systematic uncertainty. For signal, the widths of the gaussians for the m_{ES} and $\Delta E/\sigma_{\Delta E}$ parameters are enlarged to account for the discrepancies observed between data and MC in the control sample (Section 5.9).

These partial systematic uncertainties on the total branching fraction are expected to be $\sim 1 - 2 \times 10^{-6}$.

6.4.2 Δt parameter fluctuations

The description of the resolution in Δt follows that one found in the $\sin 2\beta$ charmonium analyses, as described in Section 4.4.3. These parameters, and the signal tagging category fractions, are also fluctuated in the manner described above. In addition, the parameters used to describe the Δt distributions of the background need to be varied.

Again, the RMS of the resulting distributions are taken as the systematic errors, and are expected to be roughly $\sim 1-2\%$ for $\mathcal{S}$ and $\mathcal{C}$, and a few degrees on ϕ_{mix}.

Any systematic uncertainties arising from vertexing and misalignment are accounted for by the rather small errors ($\sim 0.1\%$) found in the charmonium analyses [24].

6.4.3 Tag side interference

A further systematic uncertainty affecting the quantities derived from any time-dependent analysis must be considered. In the tagging algorithm, the interference between the decays $b \to c + \bar{u}d$ and the doubly Cabibbo supressed $\bar{b} \to \bar{u} + c\bar{d}$ is neglected [91]. Such an effect has been shown to be $\lesssim 1\%$ in the charmonium analyses, as expected from the ratio of the amplitudes for the two processes, ~ 0.02.

6.4.4 Histogram fluctuations

A similar procedure is used to estimate the systematic errors associated with the uncertainty in the shapes of the efficiency, $q\bar{q}$ and $B\overline{B}$ histograms. 200 histograms are generated by fluctuating independently the contents of each bin within their errors according to a binomial distribution, which are then used to repeat the nominal fit.

These uncertainties are expected to amount to a few percent of the measured quantities.

6.4.5 $B\overline{B}$ background yield fluctuations

Although the $q\bar{q}$ background yield is extracted from the fit and therefore does not contribute to the systematic uncertainties, the numbers of $B\overline{B}$ background events across the different categories are fixed from the MC studies. To estimate the associated systematic uncertainty, these numbers are again fluctuated before being fed into another 100 instances of the otherwise nominal fit. That should result in a $\sim 1\%$ error on the magnitudes and phases.

6.4.6 Fit biases

Toy and full MC studies have been performed (see Section 6.1), and used to characterize the fit biases. These are not corrected for, and are instead dealt with by assigning appropriate systematic errors, which are taken to be the differences between the true and the fitted (biased) values, and are generally rather small.

6.4.7 Reconstruction efficiency systematics and $N_{B\overline{B}}$

The branching fraction calculations are affected by uncertainties on the efficiency of the reconstruction and the selection criteria, and on the number of $B\overline{B}$ pairs produced in the experiment. The latter has been estimated to have a 1.1% uncertainty [92]. Efficiency corrections due to the reconstruction of two charged tracks and the K_S^0 lead, respectively, to a 1.6% systematic error, as suggested by the *BABAR* Tracking Efficiency Task Force [93], and a $\sim$ 1% uncertainty that depends on the selection cuts [94]. PID systematics have been estimated before as 2.8% [30].

The MC simulation from which the efficiencies are calculated does not reproduce completely the behaviour of data. Therefore, a correction factor, that is obtained from the comparison of the effects of these selection requirements on data and MC for signal and the control sample, must be applied. The correction has an associated systematic uncertainty of $\sim 2-3\%$, coming from the propagation of the poissonian errors on the sample sizes.

6.4.8 Model errors

Uncertainties in the masses and widths used to characterize the various resonances, as well as coupling (Flatté) constants, centrifugal barrier factors and LASS parameters, contribute to a systematic uncertainty in the DP structure. Also, addition and omission of other resonances to/from the model further increase the uncertainties. A good example is one of the most important observables, $\mathcal{S}\left(f_0(980)\right)$, which varies significantly depending on the resonant content in the $\pi^+\pi^-$ spectrum: a value of

$\mathcal{S} = -0.84$ is obtained when the state $f_X(1300)$ is included, in contrast to the value $\mathcal{S} = -0.77$ yielded by the nominal Dalitz-plot model. Finally, the effect on the results of different parameterizations for each component in the model must also be taken into account, such as those resulting from the use of the Gounaris-Sakurai lineshape [95] instead of the relativistic Breit-Wigner for the $\rho^0(770)$, or a non-constant non-resonant component.

These uncertainties are expected to be dominant in some parameters, and rather large in others: as much as 0.1 on $\mathcal{S}$, and $\sim 5°$ on the phases, including $\phi_{\rm mix}$.

6.5 Results

Once the B^0 and $\overline{B}^0$ decay models have been established (Section 6.3), a fit is performed to data with the B^0 lifetime and mixing frequency free to vary in the fit. The values obtained for these are compared to their world averages [22] in order to validate the Δt treatment, and an excellent agreement is found: $\tau_{B^0} = 1.59 \pm 0.06$ and $\Delta m_d = 0.50 \pm 0.04$, where the errors are those returned by `Minuit`. A final fit is then performed, with τ_{B^0} and Δm_d fixed to their world average values, from which the yields, fit fractions and phase differences are extracted. The ${}_s\mathcal{P}lots$ [2] are examined first, and after checking that there is a good agreement between the fitted distribution in each variable for signal and continuum events with that inferred from the ${}_s\mathcal{P}lots$ technique (see Figures 6.10 to 6.14), the values obtained from the fit for the parameters of interest are finally looked at. These are listed in Table 6.2, the uncertainties quoted being statistical only. The correlations among the parameters varied in the fit are showed in Figure 6.16. A few plots concentrating on the areas of interest are also

[2] The ${}_s\mathcal{P}lots$ technique [96] consists of calculating weights from the covariance matrix of a fit to data in which the yields of the species (e.g. signal, background) are left to vary, and the distributions in all the variables considered but one are used. These weights give the probability for each individual event to belong to a given species, and can be used to deduce how the distributions in the excluded variable should be. A good consistency check is then obtained by comparing these with their assumed PDFs, as described in the main text.

Table 6.2: Results from the fit to data. The errors are statistical only. The B^0 and $\overline{B}^0$ average magnitude and phase of $K^{*\pm}(892)\pi^{\mp}$ were fixed to 1.0 and 0.0, respectively, and the *CP*-odd phase of $\chi_{c0}K^0_S$ was also set to 0.0, constraining the B^0 and $\overline{B}^0$ phases quoted below for this resonance to be equal. In calculating relative phases between B^0 and $\overline{B}^0$ components, the mixing phase must be substracted from the $\overline{B}^0$ phase.

Resonance	B^0 Fit Fraction	B^0 Phase	$\overline{B}^0$ Fit Fraction	$\overline{B}^0$ Phase
$K^{*\pm}(892)\pi^{\mp}$	$(12.5 \pm 1.9)\%$	$(-21 \pm 26)^\circ$	$(9.4 \pm 1.6)\%$	$(24 \pm 27)^\circ$
$K_0^{*\pm}(1430)\pi^{\mp}$	$(48.1 \pm 2.9)\%$	$(158 \pm 24)^\circ$	$(52.6 \pm 4.8)\%$	$(181 \pm 28)^\circ$
$\rho^0(770)K^0_S$	$(8.4 \pm 2.7)\%$	$(-26 \pm 36)^\circ$	$(10.6 \pm 2.9)\%$	$(22 \pm 40)^\circ$
$f_0(980)K^0_S$	$(18.5 \pm 3.1)\%$	$(-64 \pm 30)^\circ$	$(13.4 \pm 2.5)\%$	$(-64 \pm 37)^\circ$
$\chi_{c0}K^0_S$	$(0.9 \pm 0.9)\%$	$(43 \pm 24)^\circ$	$(1.7 \pm 0.9)\%$	$(43 \pm 24)^\circ$
$K^0_S\pi^+\pi^-$ NR	$(8.2 \pm 2.4)\%$	$(5 \pm 32)^\circ$	$(8.7 \pm 2.8)\%$	$(60 \pm 50)^\circ$
ϕ_{mix}	$(52 \pm 57)^\circ$			
Signal yield	1719 ± 52 events			
$q\overline{q}$ yield	8331 ± 100 events			

shown in Figure 6.15.

The estimation of the errors on these is made via toy MC studies, as detailed in Section 6.1, since the observables of interest (e.g. the fit fractions or the asymmetries) are highly non-linear functions of the fitted parameters c_j, which are often strongly correlated. Hence, simple error propagation is rendered unfeasible. The values found from the final fit are fed as the generating values in the toy MC study, and the statistical uncertainty on each of the parameters of interest is calculated as the width of the resulting distributions. These distributions are gaussian for most of the obervables, except for the asymmetries of the $\chi_{c0}K^0_S$ for which upper limits are quoted instead of central values with errors.

A rather large error is obtained on ϕ_{mix}. As explained in Section 6.3, the *CP*-

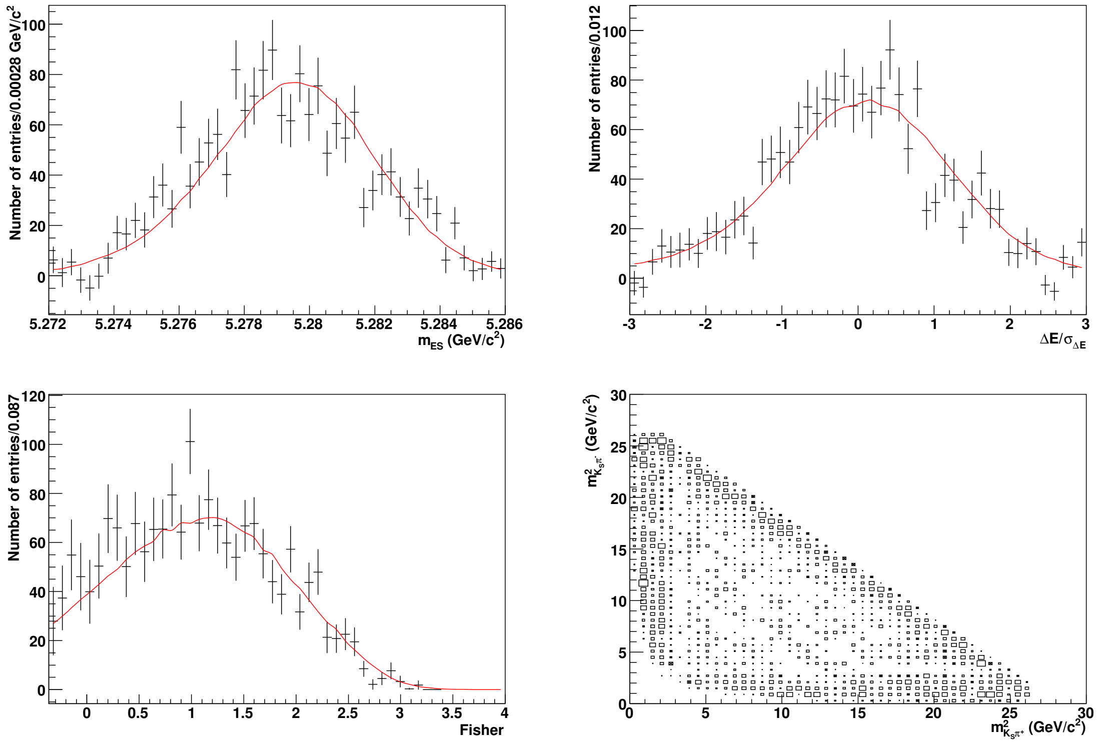


Figure 6.10: ${}_s\mathcal{P}lots$ distributions for the signal species given by the three background discriminating variables included in the fit, $m_{\rm ES}$, $\Delta E/\sigma_{\Delta E}$ and $\mathcal{F}$, and the Dalitz plot variables. The points with errors in a given variable represent the signal distribution as determined from data by a fit in which all the variables but the one plotted are used to separate signal and backgrounds, and the solid line is the distribution predicted by the results of the total fit. There is a good agreement between them. Note that the lower right plot, that shows the Dalitz plot distribution obtained when the three discriminating variables are used to separate signal from background, shows heavily populated bands in the $m_{\pi^+\pi^-}$ and $m_{K^0_S\pi}$ spectra, with hints of vector structures in each of those.

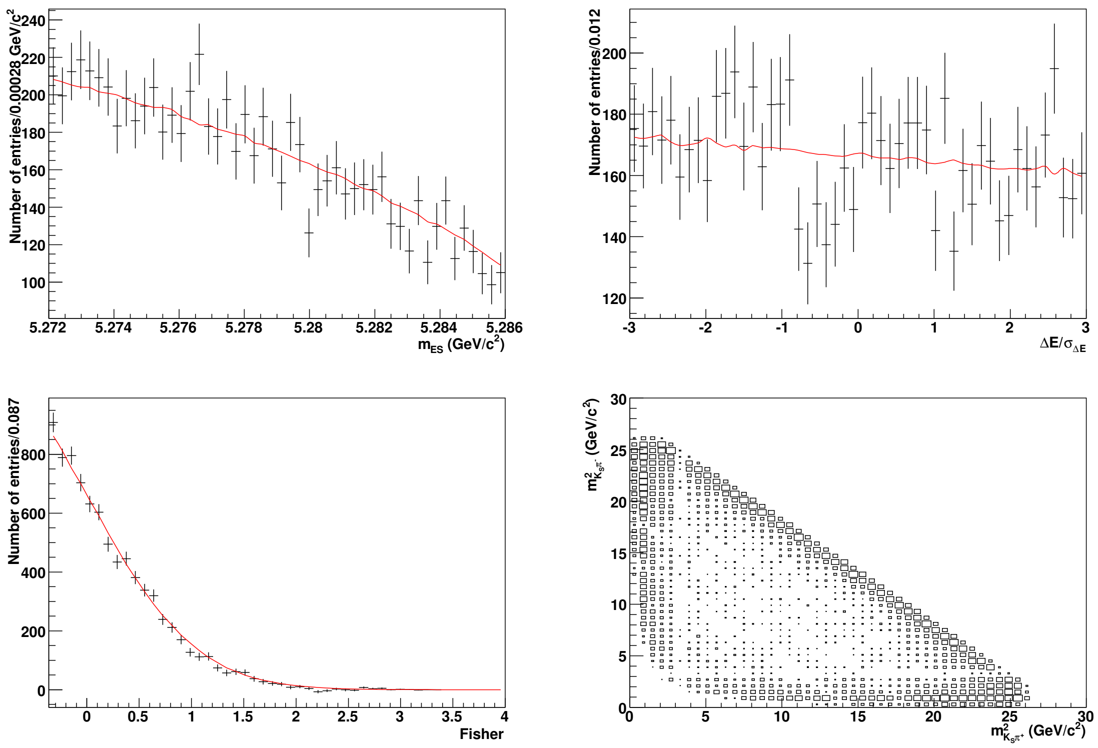
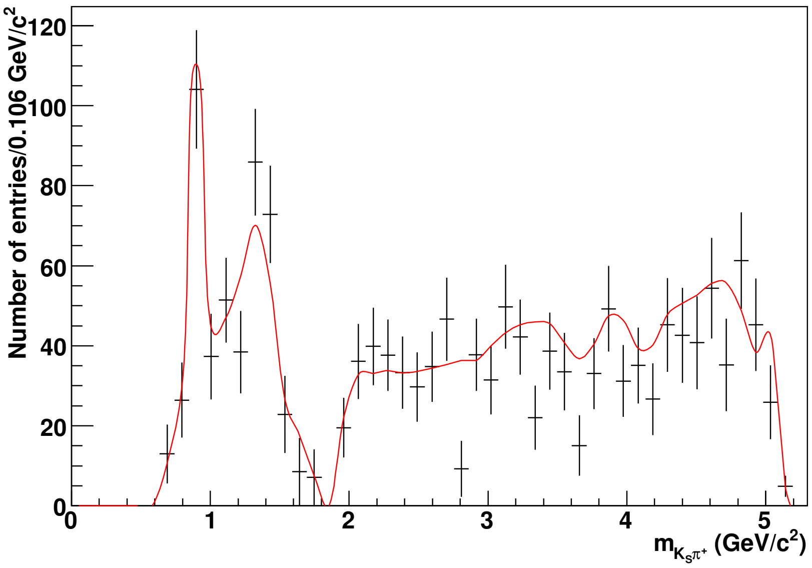


Figure 6.11: ${}_s\mathcal{P}lots$ distributions for the continuum background species given by the three background discriminating variables included in the fit, m_{ES}, $\Delta E/\sigma_{\Delta E}$ and $\mathcal{F}$, and the Dalitz plot variables. The data points, weighted by the ${}_s\mathcal{P}lot$ technique, show a good agreement with the distributions given by the values extracted from the fit.



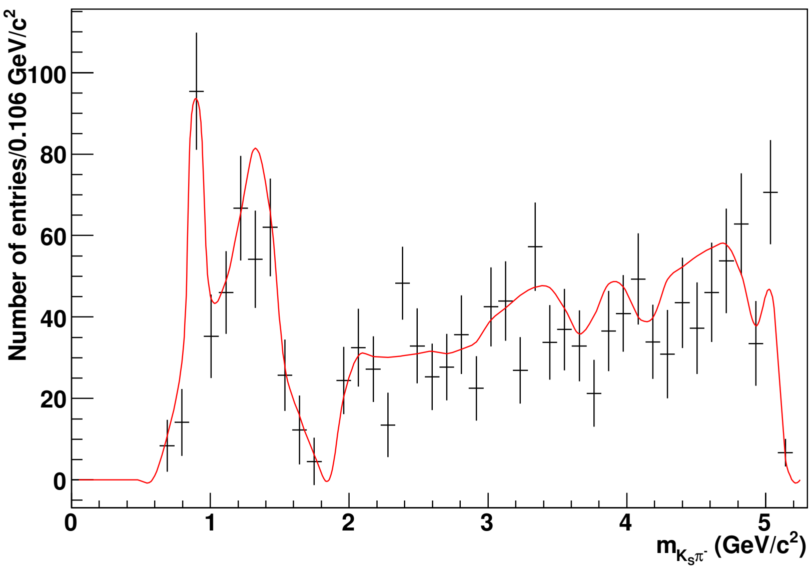


Figure 6.12: Projections on the $m_{K_S^0\pi}$ invariant masses of the ${}_s\mathcal{P}lots$ Dalitz distribution for the signal species. Good agreement between the global fit and the prediction on the Dalitz plot PDF from the fit only to the three discriminating variables $m_{\rm ES}$, $\Delta E/\sigma_{\Delta E}$ and $\mathcal{F}$ is observed.

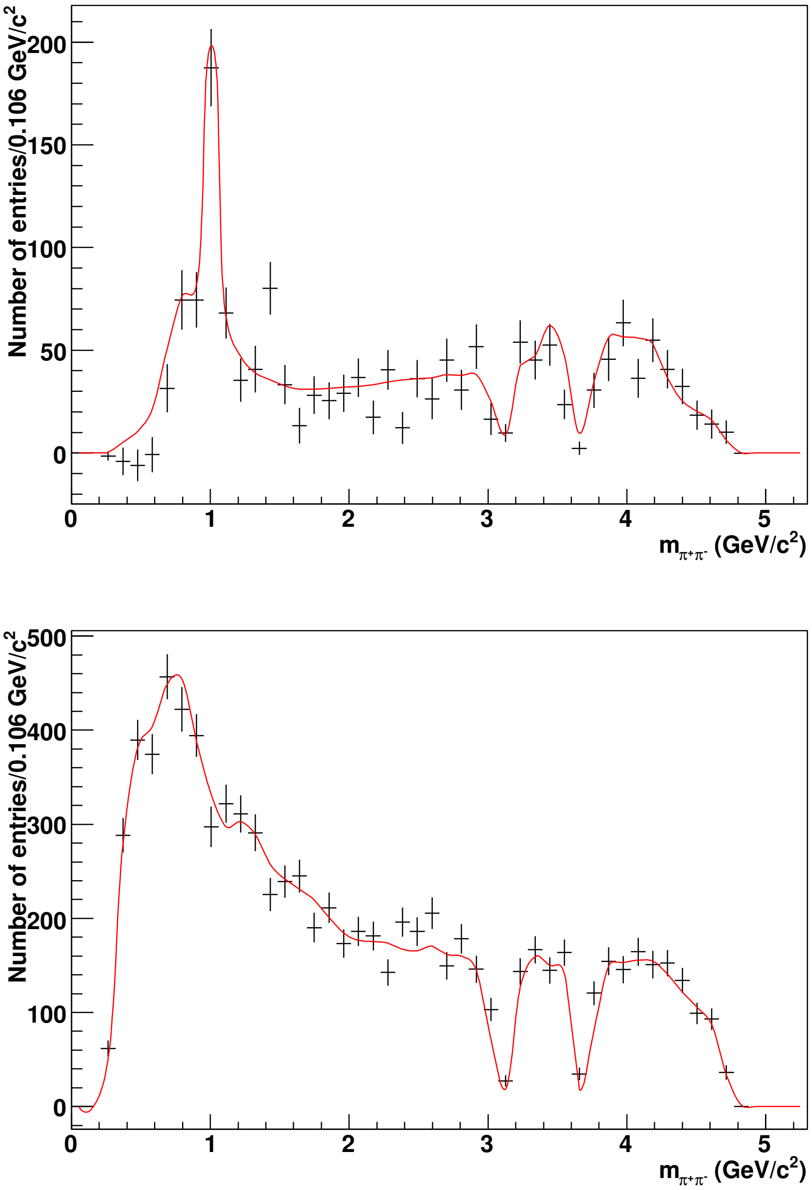


Figure 6.13: Projections on the $m_{\pi^+\pi^-}$ invariant masses of the ${}_s\mathcal{P}lots$ Dalitz distribution for the signal (top) and continuum background (bottom) species. Good agreement between the global fit and the prediction on the Dalitz plot PDF from the fit only to the three discriminating variables $m_{\rm ES}$, $\Delta E/\sigma_{\Delta E}$ and $\mathcal{F}$ is observed. Note the point at $m_{\pi^+\pi^-} \simeq 1.4\,\mathrm{GeV}/c^2$ for signal, showing a hint of an unaccounted component, most likely the $f_X(1300)$.

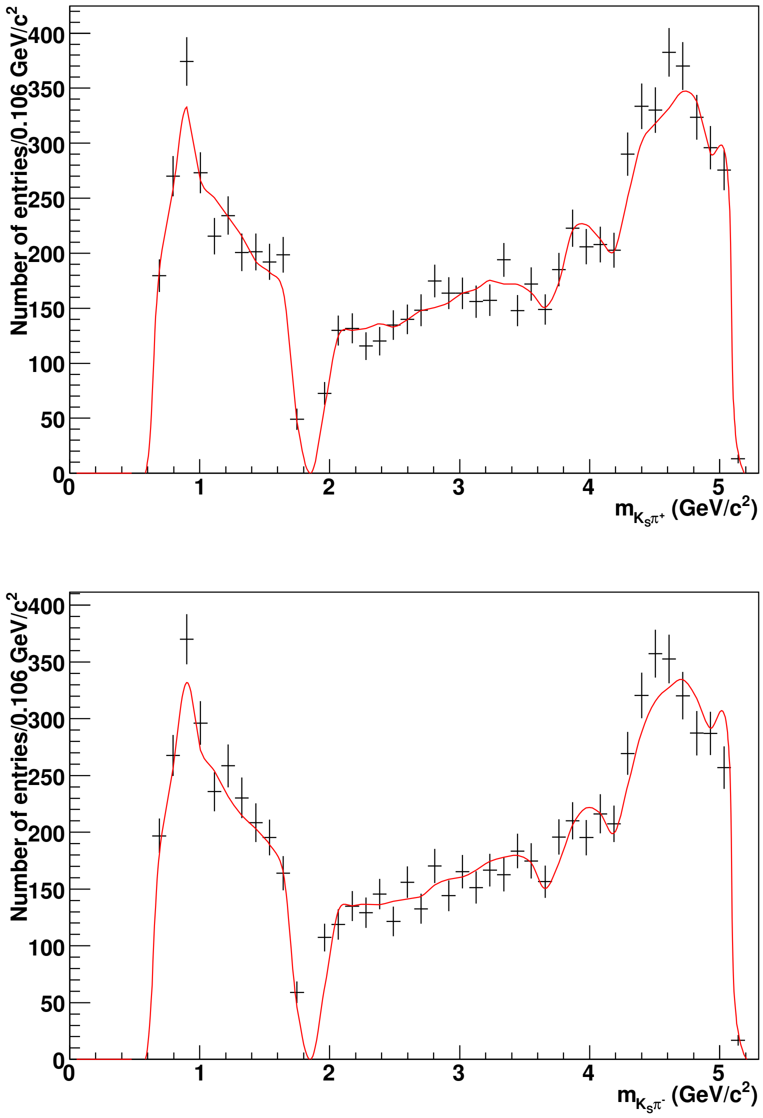


Figure 6.14: Projections on the $m_{K^0_S\pi}$ invariant masses of the ${}_s\mathcal{P}lots$ Dalitz distribution for the continuum background. There is a good agreement between the global fit and the prediction on the Dalitz plot PDF from the fit only to the three discriminating variables $m_{\rm ES}$, $\Delta E/\sigma_{\Delta E}$ and $\mathcal{F}$, although a slight underestimation of the peaks at low and high $m_{K^0_S\pi}$ is visible. This is a known feature arising from the use of on-peak sideband data to model the continuum background, in which resonances from the continuum are smoothed by the misreconstruction. Indeed, there is a hint at such behaviour in Figure 5.15, where the off-peak and the on-peak sidebands are compared.

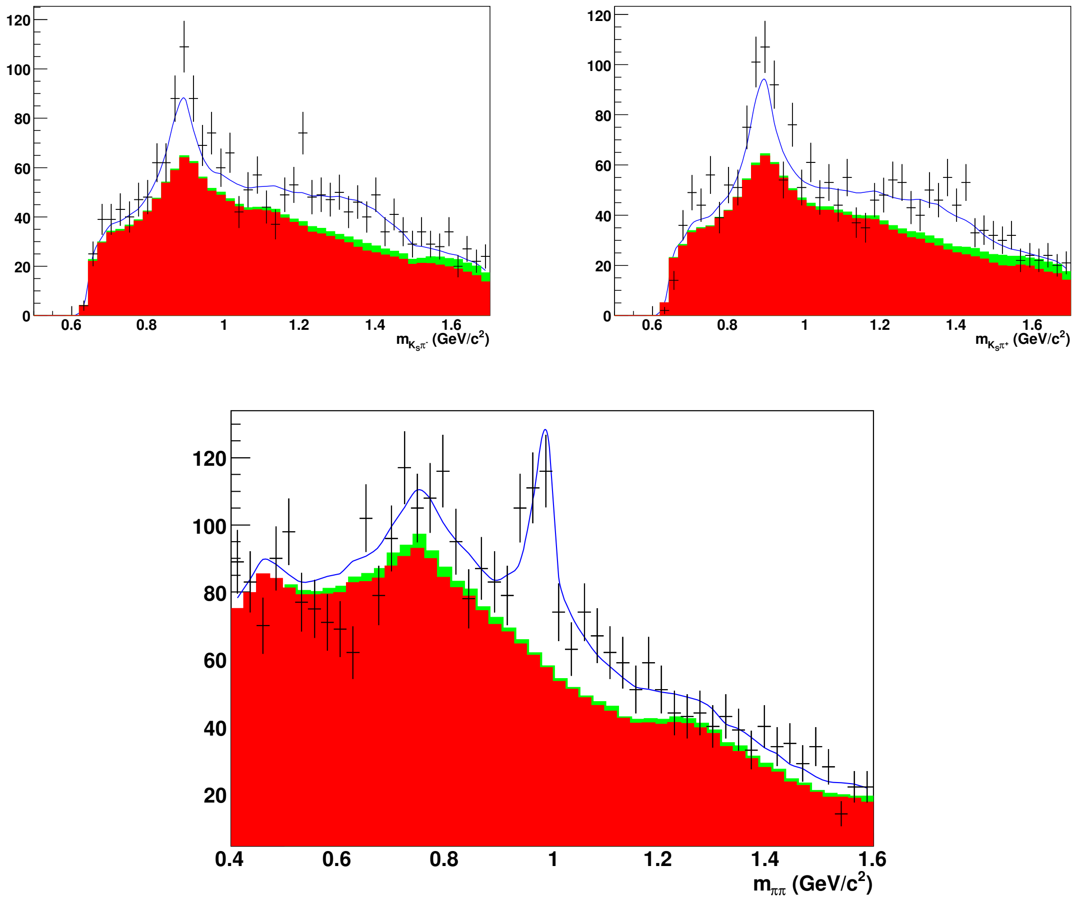


Figure 6.15: Figures showing the detail of the $\rho^0(770)$ (top), K^{*+} (middle) and K^{*-} regions. The red and green-filled areas represent continuum and $B\overline{B}$ background, respectively, and the continuous line is the prediction given by the values extracted from the fit. Clear peaks are observed for the identified resonances, $\rho^0(770)$, $f_0(980)$ and $K^{*\pm}(892)$, and an enhancement reveals the presence of $K_0^{*\pm}(1430)$.

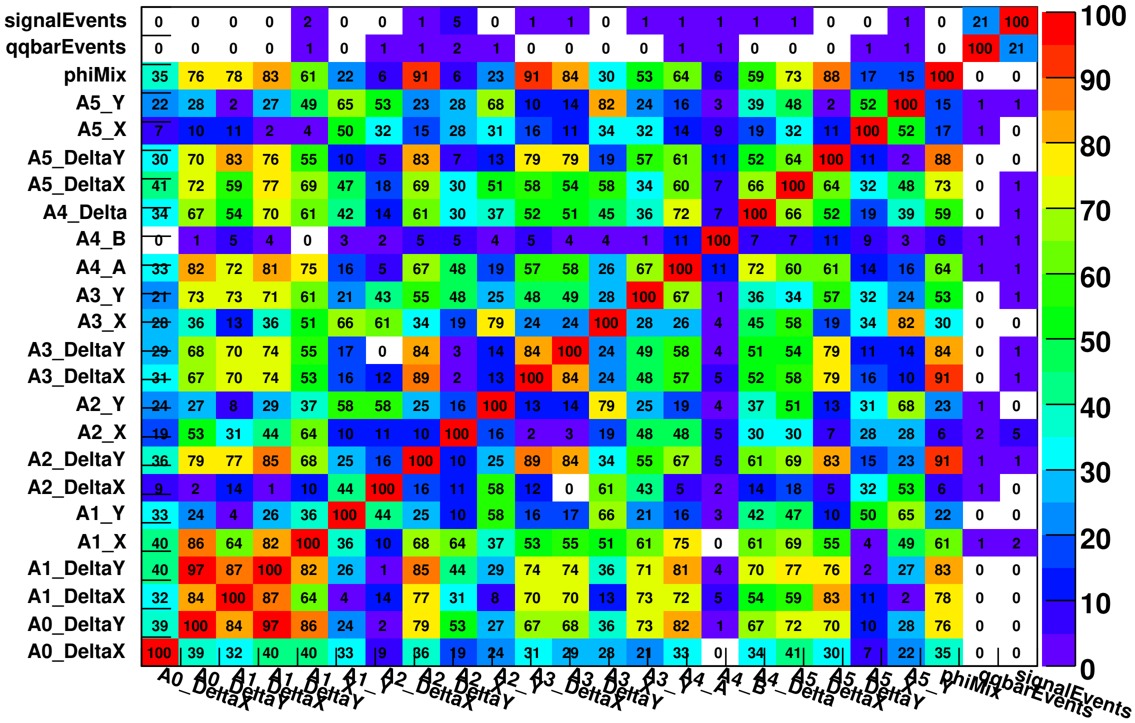


Figure 6.16: Absolute vlaues of the correlations between the parameters varied in the fit. Those corresponding to a component of the signal model are labelled with a number, starting from 0 and following the sequence $K^*(892)$, $K_0^*(1430)$, $\rho^0(770)$, $f_0(980)$, χ_{c0} and non-resonant. It can be seen that the most important correlations appear between the two K^* components themselves and with the rest of the resonances of the model. Rather high numbers of strongly correlated parameters arise often in Dalitz analyses.

Table 6.3: B^0-$\overline{B}^0$ averaged branching fractions, time-dependent asymmetries and direct CP asymmetries derived from the fit. The errors are statistical only.

Resonant mode	$\mathcal{B}(10^{-6})$	$\mathcal{S}$	$\mathcal{C}$
Total $K^0\pi^+\pi^-$	45.8 ± 1.6	–	–
$K^{*\pm}(892)\pi^\mp; K^{*\pm}(892) \to K^0\pi^\pm$	5.1 ± 0.5	–	0.13 ± 0.09
$K_0^{*\pm}(1430)\pi^\mp; K_0^{*\pm}(1430) \to K^0\pi^\pm$	23.4 ± 1.4	–	-0.04 ± 0.07
$\rho^0(770)K^0; \rho^0(770) \to \pi^+\pi^-$	4.4 ± 0.6	-0.06 ± 0.37	-0.12 ± 0.31
$f_0(980)K^0; f_0(980) \to \pi^+\pi^-$	7.4 ± 0.7	-0.77 ± 0.21	0.16 ± 0.16
$\chi_{c0}K^0; \chi_{c0} \to \pi^+\pi^-$	< 1.0 at 98%CL	< 0.0125 at 90%CL	-0.33 ± 0.58
$K^0\pi^+\pi^-$ NR	3.1 ± 0.8	–	-0.40 ± 0.37

odd phase of $\chi_{c0}K_S^0$ dominates the result, and is therefore affected by the statistical uncertainty on its small contribution. The mirror solution, $\phi_{\text{mix}} = \pi - 2\beta \simeq 2.4$ rad, however, is clearly disfavoured (see Figure 6.17).

The branching fractions of each submode, averaged over the two CP conjugate states, their time-dependent asymmetries $\mathcal{S}$ and their direct CP asymmetries $\mathcal{A}_{CP} - \mathcal{C}$ have also been calculated, and are shown in Table 6.3.

The correlations observed among the c_j parameters floated in the fit are quite varied. Although typical values are ~20% or 30%, correlations as high as ~90% are observed between the ΔY components of the $K^*(892)$ and $K_0^*(1430)$ resonances. More generally, strong correlations are often found between the parameters of the $K^*(892)$ and $K_0^*(1430)$ resonances, and the non-resonant and the $K_0^*(1430)$, and sometimes between those of the $f_0(980)$ and $\rho^0(770)$states.

More of the parameters are found to correlate strongly ($\rangle$70%) with a few other parameters, typically another parameter of the same resonance (e.g. X and ΔY or ΔX and ΔY components of the $K_0^*(1430)$), and

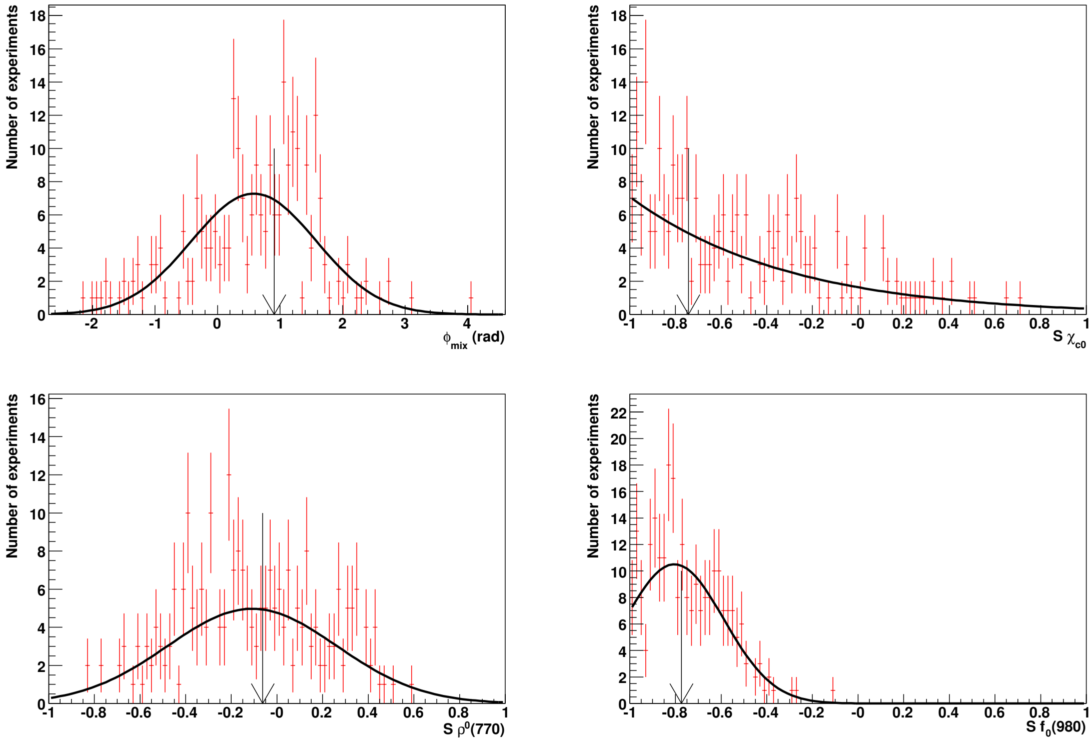


Figure 6.17: Distribution of the mixing phase and the time-dependent asymmety coefficients $\mathcal{S}$ of $\rho^0(770)K^0$, $f_0(980)K^0$ and $\chi_{c0}K^0$ obtained in the toy MC studies when the fitted values are used as input to the generation. Due to the low fit fraction of $\chi_{c0}K^0$, the errors on $\phi_{\rm mix}$ and $\mathcal{S}(\chi_{c0}K^0)$ are rather large. In spite of that, the mirror solution $\phi_{\rm mix} = \pi - 2\beta \simeq 2.4$ rad is clearly excluded. The accuracy of $\mathcal{S}(f_0(980)K^0)$ improves on previous measurements.

6.6 Conclusions and outlook

The branching fractions of $B^0 \to K^0\pi^+\pi^-$, $B^0 \to K^{*\pm}(892)\pi^\mp$, $B^0 \to K_0^{*\pm}(1430)\pi^\mp$, $B^0 \to \rho^0(770)K^0$, $B^0 \to f_0(980)K^0$, and $B^0 \to K^0\pi^+\pi^-$ non-resonant have been measured, finding results compatible with previous measurements and improving on their statistical accuracy. In particular, the fit fractions are in very good agreement with those from $B^+ \to K^+\pi^-\pi^+$ [10].

The results obtained by *Belle* for the branching ratios of $B^0 \to K_0^{*\pm}(1430)\pi^\mp$ and $B^0 \to K^0\pi^+\pi^-$ NR seem to be superficially at odds with our results. These differences originate from different parameterizations: *Belle* parameterize the $K_0^*(1430)$ with a relativistic Breit-Wigner, whereas we follow LASS (see Section 2.5.4). This also explains the difference in the non-resonant branching fraction, since the LASS parameterization includes an effective range term that behaves as a non-resonant component. *Belle*'s choice of a relativistic Breit-Wigner seems simplistic, as a variety of different sources (e.g. [65, 10]) have shown that it does not describe well the $K\pi$ spectrum. In spite of the better performance of the LASS lineshape, it is uncertain whether such parameterization will be the last word about the form of the $K_0^*(1430)$ resonance, making the associated systematic uncertainties difficult to estimate.

Direct *CP* asymmetries have been measured for all the resonant submodes mentioned above, and have been found to be consistent with zero.

Measurements of the mixing phase and the time-dependent asymmetries coefficients $\mathcal{S}$ have also been made. Although the precision on the first one is rather low, the mirror solution is disfavoured. The measurement of $\mathcal{S}(f_0(980))$ is compatible with *BABAR*'s previous measurement and slightly improves on the statistical errors. It is also in excellent agreement with the charmonium measurements of sin2β. The result obtained for $\mathcal{S}(\rho^0(770))$ is compatible with the only previous measurement, and, interestingly, provides circumstantial evidence for a value rather different from the charmonium value. This would not be completely unexpected, as the mode $B^+ \to \rho^0(770)K^+$ exhibits a $\sim 30\%$ asymmetry, indicating a strong interference. Under isospin symmetry, two of the four diagrams that contribute to the latter are also the ones that make up the amplitude of $B^0 \to \rho^0(770)K^0$: a $b \to s$ penguin loop and a $b \to u$ tree carrying the unitarity triangle phase γ (as in the diagrams for $B^0 \to K^{*0}\pi^0$ in Figure 2.7, but with a K_S^0 instead of a K^{*0} and a $\rho^0(770)$ instead of the π^0). If the tree level diagram were enhanced, and forced to counteract the effect of 2β with its weak phase, $\mathcal{S} \simeq 0$ could be obtained. Perhaps in the future this channel alone will determine γ directly, since the presence of $\chi_{c0}K^0$ allows for the measurement of 2β and $\rho^0(770)K^0$ is sensitive to a combination of 2β and γ.

Finally, the relative phase between $B^0 \to K^{*\pm}\pi^{\mp}$ and $\overline{B}^0 \to K^{*\mp}\pi^{\pm}$ has also been found, allowing for a new independent determination of the unitarity angle γ to be made. Since the method [11] to calculate the angle relies heavily on Dalitz plot measurements, the error on the new determination is expected to be of the order of the typical error on a Dalitz-plot measured phase, i.e. $\sim 20^\circ - 40^\circ$. Such a measurement is competitive and genuinely improves the overall knowledge of the angle.

Clearly, given the rich physics harvest that this decay mode yields, the analysis will be carried out again on the total dataset that *BABAR* is due to accumulate by September 2008. This will be done in the light of the knowledge on the $K\pi\pi$ Dalitz-plot resonant structure gained from the newest iterations of the better experimentally suited $B^+ \to K^+\pi^-\pi^+$, which will also be redone. Of course, the other channel needed for the determination of γ, $B^0 \to K^+\pi^-\pi^0$, will also be studied with the final datasample. Given the size of the total datasample, twice as large as the one used for the analysis presented here, it may be then necessary to use, and possibly tune, the author's implementation of the self cross feed treatment, since it is only negligible at the moment because the statistical uncertainty is comparable to the fraction of self cross feed events.

References

[1] A. D. Sakharov. Violation of *CP* invariance, *C* asymmetry, and Baryon Asymmetry of the Universe. *Pisma Zh. Eksp. Teor. Fiz.*, 5:32–35, 1967.

[2] 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.*, 13:138–140, 1964.

[3] M. Kobayashi and T. Maskawa. *CP* Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.

[4] A. B. Carter and A. I. Sanda. *CP* Nonconservation in Cascade Decays of B Mesons. *Phys. Rev. Lett.*, 45(12):952–954, Sep 1980.

[5] A. B. Carter and A. I. Sanda. *CP* violation in B-meson decays. *Phys. Rev. D*, 23(7):1567–1579, Apr 1981.

[6] I. Y. Bigi and A. I. Sanda. Notes on the Observability of *CP* Violations in B Decays. *Nucl. Phys.*, B193:85, 1981.

[7] H. Albrecht et al. Observation of B^0-$\overline{B}^0$ Mixing. *Phys. Lett.*, B192:245, 1987.

[8] N. Lockyer et al. Measurement of the Lifetime of Bottom Hadrons. *Phys. Rev. Lett.*, 51:1316, 1983.

[9] P. Oddone. An asymmetric B factory based on PEP. *Annals N. Y. Acad. Sci.*, 578:237–247, 1989.

[10] B. Aubert et al. Dalitz-plot analysis of the decays $B^{\pm} \to K^{\pm}\pi^{\mp}\pi^{\pm}$. *Phys. Rev.*, D72:072003, 2005.

[11] M. Ciuchini, M. Pierini, and L. Silvestrini. New bounds on the CKM matrix from $B \to K\pi\pi$ Dalitz plot analyses. *Phys. Rev.*, D74:051301, 2006.

[12] C. P. Burgess and G. D. Moore. *The Standard Model: A Primer*. Cambridge, UK: Cambridge University Press, 2007.

[13] N. Cabibbo. Unitary Symmetry and Leptonic Decays. *Phys. Rev. Lett.*, 10:531–532, 1963.

[14] L. Wolfenstein. Parametrization of the Kobayashi-Maskawa Matrix. *Phys. Rev. Lett.*, 51:1945, 1983.

[15] A. Höcker and Z. Ligeti. *CP* violation and the CKM matrix. *Ann. Rev. Nucl. Part. Sci.*, 56:501–567, 2006.

[16] J. Charles et al. CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories. *Eur. Phys. J.*, C41:1–131, 2005. Updated results and plots available at: http://ckmfitter.in2p3.fr.

[17] I. Y. Bigi and A. I. Sanda. *CP Violation*. Cambridge, UK: Cambridge University Press, 2000.

[18] K. Lande, E. T. Booth, J. Impeduglia, L. M. Lederman, and W. Chinowsky. Observation of Long-Lived Neutral V^0 Particles. *Phys. Rev.*, 103:1901–1904, 1956.

[19] A. Abulencia et al. Measurement of the B_s^0 - $\bar{B}_s^0$ Oscillation Frequency. *Phys. Rev. Lett.*, 97:062003, 2006.

[20] B. Aubert et al. Evidence for D^0-$\overline{D}^0$ mixing. *Phys. Rev. Lett.*, 98:211802, 2007.

[21] M. Staric et al. Evidence for D^0-$\overline{D}^0$ Mixing. *Phys. Rev. Lett.*, 98:211803, 2007.

[22] W.-M. Yao et al. Review of Particle Physics. *J. Phys.*, G33:1–1232, 2006.

[23] B. Aubert et al. A study of time dependent *CP*-violating asymmetries and flavor oscillations in neutral B decays at the $\Upsilon(4S)$. *Phys. Rev.*, D66:032003, 2002.

[24] B. Aubert et al. Improved Measurement of *CP* Violation in Neutral B decays to $c\bar{c}s$. 2007. Unpublished. hep-ex/0703021.

[25] A. Garmash et al. Dalitz analysis of $B \rightarrow Khh$ decays at Belle. *AIP Conf. Proc.*, 814:680–684, 2006.

[26] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47(10):777–780, May 1935.

[27] D. J. Griffiths. *Introduction to Quantum Mechanics*. Benjamin Cummings, second edition.

[28] P. F. Harrison and H. R. (editors) Quinn. The *BABAR* physics book: Physics at an asymmetric B factory. Papers from Workshop on Physics at an Asymmetric B Factory (BaBar Collaboration Meeting), Rome, Italy, 11-14 Nov 1996, Princeton, NJ, 17-20 Mar 1997, Orsay, France, 16-19 Jun 1997 and Pasadena, CA, 22-24 Sep 1997.

[29] B. Aubert et al. Measurement of *CP*-violating asymmetries in $B^0 \rightarrow (\rho\pi)^0$ using a time-dependent Dalitz plot analysis. *Phys. Rev.*, D76:012004, 2007.

[30] B. Aubert et al. Measurements of the Branching Fractions and Bounds on the Charge Asymmetries of Charmless Three-Body Charged B Decays. *Phys. Rev. Lett.*, 91:051801, 2003.

[31] B. Aubert et al. An amplitude analysis of the decay $B^{\pm} \rightarrow \pi^{\pm}\pi^{\pm}\pi^{\mp}$. *Phys. Rev.*, D72:052002, 2005.

[32] B. Aubert et al. $B^0 \rightarrow K^+\pi^-\pi^0$ Dalitz plot analysis. 2004. Unpublished. hep-ex/0408073.

[33] B. Aubert et al. Dalitz plot analysis of the decay $B^+ \to K^\pm K^\pm K^\mp$. *Phys. Rev.*, D74:032003, 2006.

[34] B. Aubert et al. Measurements of *CP*-Violating Asymmetries in the Decay $B^0 \to K^+K^-K^0$. *Submitted to Phys.Rev.Lett.*, 2007. arXiv:0706.3885 [hep-ex].

[35] B. Aubert et al. Measurements of neutral B decay branching fractions to $K^0_S\pi^+\pi^-$ final states and the charge asymmetry of $B^0 \to K^{*+}\pi^-$. *Phys. Rev.*, D73:031101, 2006.

[36] A. Garmash et al. Dalitz analysis of three-body charmless $B^0 \to K^0_S\pi^+\pi^-$ decay. *Phys. Rev.*, D75:012006, 2007.

[37] B. Aubert et al. Measurement of the *CP* asymmetry and branching fraction of $B^0 \to \rho^0 K^0$. *Phys. Rev. Lett.*, 98:051803, 2007.

[38] B. Aubert et al. Measurements of *CP*-violating asymmetries and branching fractions in B decays to ωK and $\omega\pi$. *Phys. Rev.*, D74:011106, 2006.

[39] C.-M. Jen et al. Improved measurements of branching fractions and *CP* partial rate asymmetries for $B \to \omega K$ and $B \to \omega\pi$. *Phys. Rev.*, D74:111101, 2006.

[40] M. Beneke and M. Neubert. QCD factorization for $B \to PP$ and $B \to PV$ decays. *Nucl. Phys.*, B675:333–415, 2003.

[41] C.-W. Chiang, M. Gronau, J. L. Luo, Z.and Rosner, and D. A. Suprun. Charmless $B \to VP$ decays using flavor $SU(3)$ symmetry. *Phys. Rev.*, D69:034001, 2004.

[42] H.-Y. Cheng, C.-K. Chua, and K.-C. Yang. Charmless hadronic B decays involving scalar mesons: Implications to the nature of light scalar mesons. *Phys. Rev.*, D73:014017, 2006.

[43] B. Aubert et al. Observation of *CP* violation in $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$. *Phys. Rev. Lett.*, 99:021603, 2007.

[44] D. London and A. Soni. Measuring the *CP* angle β in hadronic $b \to s$ penguin decays. *Phys. Lett.*, B407:61–65, 1997.

[45] E. Barberio et al. Averages of B-hadron properties at the end of 2006. 2007. Unpublished. arXiv:0704.3575 [hep-ex].

[46] B. Aubert et al. A measurement of *CP* violating asymmetries in $B^0 \to f_0(980)K_S^0$ decays. 2004. Unpublished. hep-ex/0408095.

[47] K. Abe et al. Measurements of time-dependent *CP* violation in $B^0 \to \omega K_S^0$, $f_0(980)K_S^0$, $K_S^0\pi^0$, and K^+ K^- K_S^0 decays. 2006. Unpublished. hep-ex/0609006.

[48] M. Beneke. Corrections to $\sin(2\beta)$ from *CP* asymmetries in $B^0 \to (\pi^0, \rho^0, \eta, \eta', \omega, \phi)\, K_S^0$ decays. *Phys. Lett.*, B620:143–150, 2005.

[49] H.-Y. Cheng, C.-K. Chua, and A. Soni. Effects of final-state interactions on mixing-induced *CP* violation in penguin-dominated B decays. *Phys. Rev.*, D72:014006, 2005.

[50] M. Bona et al. The 2004 UTfit Collaboration Report on the Status of the Unitarity Triangle in the Standard Model. *JHEP*, 07:028, 2005. Updated results and plots available at: http://www.utfit.org/.

[51] N. G. Deshpande, N. Sinha, and R. Sinha. Weak phase γ using isospin analysis and time dependent asymmetry in $B_d \to K_S^0\pi^+\pi^-$. *Phys. Rev. Lett.*, 90:061802, 2003.

[52] M. Gronau, D. Pirjol, A. Soni, and J. Zupan. Improved method for CKM constraints in charmless three-body B and B_s decays. *Phys. Rev.*, D75:014002, 2007.

[53] H. J. Lipkin, Y. Nir, H. R. Quinn, and A. Snyder. Penguin trapping with isospin analysis and *CP* asymmetries in B decays. *Phys. Rev.*, D44:1454–1460, 1991.

[54] S. L. Glashow, J. Iliopoulos, and L. Maiani. Weak interactions with lepton-hadron symmetry. *Phys. Rev.*, D2:1285–1292, 1970.

[55] A. Garmash et al. Dalitz analysis of the three-body charmless decays $B^+ \to K^+\pi^+\pi^-$ and $B^+ \to K^+K^+K^-$. *Phys. Rev.*, D71:092003, 2005.

[56] R. H. Dalitz. On the Analysis of τ-Meson Data and the Nature of the τ-Meson. *Phil. Mag.*, 44, 1953.

[57] G. N. Fleming. Recoupling Effects in the Isobar Model. I. General Formalism for Three-Pion Scattering. *Phys. Rev.*, 135(2B):B551–B560, Jul 1964.

[58] D. J. Herndon, P. Söding, and R. J. Cashmore. Generalized isobar model formalism. *Phys. Rev. D*, 11(11):3165–3182, Jun 1975.

[59] D. Asner. Charm Dalitz plot analysis formalism and results (expanded RPP-2004 version). 2003. Unpublished. hep-ex/0410014.

[60] D. M. Asner et al. Search for CP violation in $D^0 \to K^0_S\pi^+\pi^-$. *Phys. Rev.*, D70:091101, 2004.

[61] S. U. Chung et al. Partial wave analysis in the K-matrix formalism. *Ann. Physik.*, 4:404, 1995.

[62] J. D. Jackson. Remarks on the phenomenological analysis of resonances. *Nuov. Cim.*, 34:1644–1666, 1964.

[63] S. M. Flatté. Coupled - Channel Analysis of the $\pi\eta$ and $K\overline{K}$ systems near $K\overline{K}$ threshold. *Phys. Lett.*, B63:224, 1976.

[64] M. Ablikim et al. Resonances in J/$\psi \to \phi\pi^+ \pi^-$ and $\phi K^+ K^-$. *Phys. Lett.*, B607:243–253, 2005.

[65] D. Aston et al. A Study of $K^- \pi^+$ Scattering in the Reaction $K^-p \to K^-\pi^+n$ at 11 GeV/c. *Nucl. Phys.*, B296:493, 1988.

[66] J. Blatt and V. E. Weisskopf. *Theoretical Nuclear Physics*. J. Wiley (New York), 1952.

[67] B. Aubert et al. The *BABAR* detector. *Nucl. Instrum. Meth.*, A479:1–116, 2002.

[68] D. Lange et al. A B Flavour tagging algorithm for CP violation measurements with the *BABAR* experiment. *BABAR* Analysis Document #1025. Unpublished.

[69] B. Aubert et al. Improved measurement of CP asymmetries in $B^0 \to (c\bar{c})K^{(*)0}$ decays. *Phys. Rev. Lett.*, 94:161803, 2005.

[70] D. J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth.*, A462:152–155, 2001.

[71] T. Sjöstrand. High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4. *Comput. Phys. Commun*, 82:74, 1994.

[72] S. Agostinelli et al. GEANT4: A simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[73] P. Billoir. Track Fitting with Multiple Scattering: A New Method. *Nucl. Instr. Meth.*, A225:352, 1984.

[74] F. Martinez-Vidal et al. The *BABAR* Vertexing. *BABAR* Analysis Document #102. Unpublished.

[75] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9th printing*. Dover (New York), 1972.

[76] G. P. M. Poppe and C. M. J. Wijers. Algorithm 680: evaluation of the complex error function. *ACM Trans. Math. Softw.*, 16(1):47, 1990.

[77] W. Verkerke and D. Kirkby. The RooFit toolkit for data modeling. 2003. Unpublished. physics/0306116.

[78] W. Ford. Choice of Kinematic Variables in B Meson Reconstruction. *BABAR* Analysis Document #53. Unpublished.

[79] T. Latham et al. Time-Dependent Dalitz Plot Analysis of $B^0 \to K_S^0\pi^+\pi^-$ Decays. *BABAR* Analysis Document #1376. Unpublished.

[80] R. A. Fisher. The Use of Multiple Measurements in Taxonomic Problems. *Annals Eugen.*, 7:179–188, 1936.

[81] D. M. Asner et al. *Phys. Rev.*, D53:1039–1050, 1996.

[82] G. Cowan. *Statistical Data Analysis*. Oxford University Press, 1998.

[83] R. Tibshirani T. Hastie and J. Friedman. *The Elements of Statistical Learning*. Springer, 2001.

[84] R. Barlow. *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences*. Wiley, 1989.

[85] P. Harrison et al. LAURA++ - Likelihood Analysis Unbinned Reconstruction of Amplitudes. *BABAR* Analysis Document #806. Unpublished.

[86] F. James and M. Roos. 'MINUIT', a system for function minimization and analysis of the parameter errors and correlations. *Comput. Phys. Commun.*, 10:343–367, 1975.

[87] F. James. MINUIT, Function Minimization and Error Analysis, Reference Manual. Unpublished. wwwasdoc.web.cern.ch/wwwasdoc/minuit/minmain.html.

[88] ROOT, An Object Oriented Data Analysis Framework. http://root.cern.ch/.

[89] H. Albrecht et al. Exclusive hadronic decays of B mesons. *Z. Phys.*, C48:543–552, 1990.

[90] T. Latham et al. Dalitz Plot Analysis of $B^\pm \to K^\pm\pi^\mp\pi^\pm$ Decays (Runs 1-5 Update). *BABAR* Analysis Document #1512. Unpublished.

[91] O. Long, M. Baak, R. N. Cahn, and D. Kirkby. Impact of tag-side interference on time dependent *CP* asymmetry measurements using coherent $B^0 - \bar{B}^0$ pairs. *Phys. Rev.*, D68:034010, 2003.

[92] C. Hearty. Measurement of the Number of $\Upsilon(4S)$ Mesons Produced in Run 1 (B Counting). *BABAR* Analysis Document #134. Unpublished.

[93] E. Varnes. Measurement of the `GoodTracksLoose` efficiency using SVT tracks. *BABAR* Analysis Document #157. Unpublished.

[94] A. Telnov. $K^0_S \to \pi^+\pi^-$ Reconstruction Efficiency in *BABAR*: Release-10 Data vs. SP4 Monte Carlo. *BABAR* Analysis Document #677. Unpublished.

[95] G. J. Gounaris and J. J. Sakurai. Finite width corrections to the vector meson dominance prediction for $\rho \to e^+e^-$. *Phys. Rev. Lett.*, 21:244, 1968.

[96] M. Pivk and F. R. Le Diberder. ${}_s\mathcal{P}lot$: A statistical tool to unfold data distributions. *Nucl. Instrum. Meth.*, A555:356–369, 2005.