CERN-LHCb/2003-124

CP sensitivity with $B^0_{(s)} \to h^+h^-$ decays at LHCb

G. Balbi, M. Bargiotti, A. Bertin, M. Bruschi, A. Carbone, L. Fabbri, P. Faccioli, D. Galli, B. Giacobbe, D. Gregori, F. Grimaldi, U. Marconi*, I. Massa, M. Piccinini, N. Semprini Cesari, R. Spighi, V. Vagnoni*, S. Vecchi, M. Villa, A. Vitale and A. Zoccoli.
(LHCb Bologna Group)

Dipartimento di Fisica dell'Università di Bologna and Ist. Naz. di Fisica Nucleare, Sezione di Bologna
**Corresponding authors:* Umberto.Marconi@bo.infn.it, Vincenzo.Vagnoni@bo.infn.it

Abstract

LHCb can collect large samples of $B^0 \to \pi^+\pi^-$, $B^0 \to K^+\pi^-$, $B^0_s \to K^+K^-$ and $B^0_s \to \pi^+K^-$ decays. The sensitivity achievable on the corresponding direct and mixing-induced CP-violating observables has been studied in detail. The combined measurement of the $B^0 \to \pi^+\pi^-$ and $B^0_s \to K^+K^-$ CP asymmetries provides a promising strategy to determine the γ angle of the unitarity triangle. The results of a study on the sensitivity on γ achievable by employing this technique are presented and discussed.

1 Introduction

LHCb is a dedicated experiment on b-quark physics currently under construction at the Large Hadron Collider (LHC) [1, 2]. It will exploit the $500\,\mu b$ beauty production cross-section in the 14 TeV proton-proton collisions of LHC, with the aim of performing precise measurements of CP-violation and rare decays of the B-mesons.

Since in the primary collisions large statistics of $B^0_s/\overline{B}^0_s$ and $B^\pm_c$ mesons, as well as other b-baryons, can be produced - differently from the case of the e^+e^- beauty factories operating at the $\Upsilon(4s)$ - LHCb provides a unique opportunity to probe Standard Model predictions in the b-quark sector beyond the $B^0/\overline{B}^0$ and $B^\pm$ mesons. By over-constraining the Cabibbo-Kobayashi-Maskawa (CKM) matrix elements, LHCb will be hopefully able to observe subtle inconsistencies with the Standard Model, therefore providing hints of New Physics.

The long awaited evidence of CP-violation in the neutral B-meson sector has been eventually established at the asymmetric e^+e^- beauty factories by the BABAR and BELLE

experiments [3, 4], which provided the first direct measurements the β angle of the unitarity triangle (UT), mainly through the "gold-plated" $B^0 \to J/\psi K^0_S$ decay. BABAR and BELLE have also provided measurements of CP violation in the $B^0 \to \pi^+\pi^-$ decay [5, 6], although with large statistical uncertainties.

Differently from the case of $B^0 \to J/\psi K^0_S$, a simple interpretation of the $B^0 \to \pi^+\pi^-$ CP-violating observables in terms of CKM phases is not possible. In fact, in addition to the $\overline{b} \to \overline{u} + W^+$ tree amplitude, sizable $\overline{b} \to \overline{d} + g\,(\gamma,\, Z^0)$ penguin amplitudes are expected to contribute to the $B^0 \to \pi^+\pi^-$ decay [7]. This penguin pollution prevents theoretically clean measurements of CKM phases.

However, recent theoretical works [8, 9] have shown that the combined measurement of the $B^0 \to \pi^+\pi^-$ and $B^0_s \to K^+K^-$ CP asymmetries, under the assumption of invariance of the strong interaction dynamics exchanging the $d \leftrightarrow s$ quarks (U-spin symmetry), provides a way to determine the γ angle of the UT.

In the following, after an introductory section where the basic formalism useful to describe CP violation in B-meson decays is presented, we will focus on the LHCb CP-reach with $B^0_{(s)} \to h^+h^-$ decays. We will describe in detail the technique adopted to extract the CP-violating observables from data, and will show the results of a simulation based on a "toy" Monte Carlo event generation, which employs as inputs the results of a full GEANT simulation study [10].

Once the sensitivity on the CP-violating observables for $B^0_{(s)} \to h^+h^-$ decays is established, we will use these information to estimate the LHCb sensitivity on the measurement of the γ angle of the UT, by using a Bayesian approach based on the strategy described in [8, 9].

All the analyses we will describe throughout this note have been performed employing the CPU power of the LHCb-Bologna computing cluster [11], hosted at the INFN-CNAF computing centre in Bologna.

2 CP violation formalism

In this section we introduce the basic formalism that will result useful to understand the contents of the following sections. The symbols B and $\overline{B}$ will be used to denote the particle and antiparticle states for both B^0 and B^0_s systems.

The decay rates for initial B and $\overline{B}$ meson, decaying into a final state f at generic proper time t, can be written as:

$$\Gamma_{B\to f}(t) = \frac{|A_f|^2}{2} e^{-\Gamma t}\left[I_+(t) + I_-(t)\right] \tag{2.1}$$

and

$$\Gamma_{\overline{B}\to f}(t) = \frac{|A_f|^2}{2}\left|\frac{p}{q}\right|^2 e^{-\Gamma t}\left[I_+(t) - I_-(t)\right]. \tag{2.2}$$

Analogously, for the CP-conjugated final state $\overline{f}$, one has:

$$\Gamma_{\overline{B}\to\overline{f}}(t) = \frac{\left|\overline{A}_{\overline{f}}\right|^2}{2} e^{-\Gamma t}\left[\overline{I}_+(t) + \overline{I}_-(t)\right] \tag{2.3}$$

and

$$\Gamma_{B\to\overline{f}}(t) = \frac{\left|\overline{A}_{\overline{f}}\right|^2}{2}\left|\frac{q}{p}\right|^2 e^{-\Gamma t}\left[\overline{I}_+(t) - \overline{I}_-(t)\right]. \tag{2.4}$$

A_f and $\overline{A}_{\overline{f}}$ are the instantaneous decay amplitudes for $B \to f$ and $\overline{B} \to \overline{f}$ respectively, and Γ is the average decay width for the two mass eigenstates $|B_L\rangle$ and $|B_H\rangle$:

$$\Gamma = \frac{\Gamma_H + \Gamma_L}{2}, \tag{2.5}$$

where the eigenstates are expressed in the $|B\rangle$, $\left|\overline{B}\right\rangle$ basis as:

$$|B_L\rangle = \frac{1}{\sqrt{|p|^2 + |q|^2}}\left(p\,|B\rangle + q\left|\overline{B}\right\rangle\right) \tag{2.6}$$

and

$$|B_H\rangle = \frac{1}{\sqrt{|p|^2 + |q|^2}}\left(p\,|B\rangle - q\left|\overline{B}\right\rangle\right). \tag{2.7}$$

The functions $I_+(t)$, $I_-(t)$, $\overline{I}_+(t)$ and $\overline{I}_-(t)$ are:

$$I_+(t) = \left(1 + |\lambda_f|^2\right)\cosh\frac{\Delta\Gamma}{2}t - 2\mathrm{Re}(\lambda_f)\sinh\frac{\Delta\Gamma}{2}t, \tag{2.8}$$

$$I_-(t) = \left(1 - |\lambda_f|^2\right)\cos\Delta m\, t - 2\mathrm{Im}(\lambda_f)\sin\Delta m\, t, \tag{2.9}$$

$$\overline{I}_+(t) = \left(1 + \left|\overline{\lambda}_{\overline{f}}\right|^2\right)\cosh\frac{\Delta\Gamma}{2}t - 2\mathrm{Re}(\overline{\lambda}_{\overline{f}})\sinh\frac{\Delta\Gamma}{2}t \tag{2.10}$$

and

$$\overline{I}_-(t) = \left(1 - \left|\overline{\lambda}_{\overline{f}}\right|^2\right)\cos\Delta m\, t - 2\mathrm{Im}(\overline{\lambda}_{\overline{f}})\sin\Delta m\, t, \tag{2.11}$$

where $\Delta\Gamma$ is the difference of the decay widths of the mass eigenstates:

$$\Delta\Gamma = \Gamma_L - \Gamma_H, \tag{2.12}$$

while Δm is the mass difference:

$$\Delta m = m_H - m_L. \tag{2.13}$$

The complex numbers λ_f and $\overline{\lambda}_{\overline{f}}$ are defined by the following equations:

$$\lambda_f = \frac{q}{p}\frac{\overline{A}_f}{A_f} \tag{2.14}$$

and

$$\bar{\lambda}_{\overline{f}} = \frac{p}{q}\frac{A_{\overline{f}}}{\overline{A}_{\overline{f}}}. \tag{2.15}$$

In the following, since the approximation $\left|\frac{q}{p}\right| \simeq 1$ is valid up to a few per mil, even in the presence of New Physics, we will assume:

$$\left|\frac{q}{p}\right| = \left|\frac{p}{q}\right| = 1 \tag{2.16}$$

2.1 Case of f CP eigenstate

If f is a CP eigenstate, one has $f = \overline{f}$, and the four decay rates reduce to two. We define the time dependent CP asymmetry as:

$$\mathcal{A}_f^{CP}(t) = \frac{\Gamma_{\overline{B}\to f}(t) - \Gamma_{B\to f}(t)}{\Gamma_{\overline{B}\to f}(t) + \Gamma_{B\to f}(t)} = -\frac{\left(1 - |\lambda_f|^2\right)\cos\Delta m\, t - 2\mathrm{Im}(\lambda_f)\sin\Delta m\, t}{\left(1 + |\lambda_f|^2\right)\cosh\frac{\Delta\Gamma}{2}t - 2\mathrm{Re}(\lambda_f)\sinh\frac{\Delta\Gamma}{2}t}. \tag{2.17}$$

Introducing the quantities:

$$\mathcal{A}_f^{dir} = -\frac{1 - |\lambda_f|^2}{1 + |\lambda_f|^2}, \tag{2.18}$$

$$\mathcal{A}_f^{mix} = \frac{2\mathrm{Im}(\lambda_f)}{1 + |\lambda_f|^2} \tag{2.19}$$

and

$$\mathcal{A}_f^{\Delta} = \frac{2\mathrm{Re}(\lambda_f)}{1 + |\lambda_f|^2}, \tag{2.20}$$

the time dependent CP asymmetry (2.17) can be re-written as:

$$\mathcal{A}_f^{CP}(t) = \frac{\mathcal{A}_f^{dir}\cos\Delta m\, t + \mathcal{A}_f^{mix}\sin\Delta m\, t}{\cosh\frac{\Delta\Gamma}{2}t - \mathcal{A}_f^{\Delta}\sinh\frac{\Delta\Gamma}{2}t}. \tag{2.21}$$

The relative decay width difference $\Delta\Gamma/\Gamma$ is expected to be of the order of 10% for the B_s^0 meson, while it is negligible for the B^0 meson. In case $\Delta\Gamma = 0$, the expression of the CP asymmetry reduces to:

$$\mathcal{A}_f^{CP}(t) = \mathcal{A}_f^{dir}\cos\Delta m\, t + \mathcal{A}_f^{mix}\sin\Delta m\, t. \tag{2.22}$$

The quantities $\mathcal{A}_f^{dir}$ and $\mathcal{A}_f^{mix}$ parametrize direct and mixing-induced CP violation respectively.

2.2 Case of f flavor specific final state

If f is a flavor specific final state, then $f \neq \overline{f}$, and one can write:

$$\lambda_f = \overline{\lambda}_{\overline{f}} = 0, \tag{2.23}$$

since only the B has instantaneous access to the decay channel f, while only the $\overline{B}$ has instantaneous access to the decay channel $\overline{f}$. In this case, the functions $I_+(t)$, $I_-(t)$, $\overline{I}_+(t)$ and $\overline{I}_-(t)$ reduce to:

$$I_+(t) = \overline{I}_+(t) = \cosh\frac{\Delta\Gamma}{2}t \tag{2.24}$$

and

$$I_-(t) = \overline{I}_-(t) = \cos\Delta m\, t. \tag{2.25}$$

By using the four decay rates, one can define the following two decay asymmetries:

$$\mathcal{A}_f(t) = \frac{\Gamma_{\overline{B}\to f}(t) - \Gamma_{B\to f}(t)}{\Gamma_{\overline{B}\to f}(t) + \Gamma_{B\to f}(t)} = -\frac{\cos\Delta m\, t}{\cosh\frac{\Delta\Gamma}{2}t} \tag{2.26}$$

and

$$\mathcal{A}_{\overline{f}}(t) = \frac{\Gamma_{\overline{B}\to \overline{f}}(t) - \Gamma_{B\to \overline{f}}(t)}{\Gamma_{\overline{B}\to \overline{f}}(t) + \Gamma_{B\to \overline{f}}(t)} = \frac{\cos\Delta m\, t}{\cosh\frac{\Delta\Gamma}{2}t} \tag{2.27}$$

It is also possible to define the following CP asymmetry:

$$\mathcal{A}^{CP}_{f,\overline{f}}(t) = \frac{\left[\Gamma_{\overline{B}\to \overline{f}}(t) + \Gamma_{B\to \overline{f}}(t)\right] - \left[\Gamma_{\overline{B}\to f}(t) + \Gamma_{B\to f}(t)\right]}{\left[\Gamma_{\overline{B}\to \overline{f}}(t) + \Gamma_{B\to \overline{f}}(t)\right] + \left[\Gamma_{\overline{B}\to f}(t) + \Gamma_{B\to f}(t)\right]}, \tag{2.28}$$

that results to be not dependent on time, and identically equal to the charge asymmetry defined as:

$$\mathcal{A}_{f,\overline{f}} = -\frac{|A_f|^2 - \left|\overline{A}_{\overline{f}}\right|^2}{|A_f|^2 + \left|\overline{A}_{\overline{f}}\right|^2} = -\frac{1 - \left|\frac{\overline{A}_{\overline{f}}}{A_f}\right|^2}{1 + \left|\frac{\overline{A}_{\overline{f}}}{A_f}\right|^2}. \tag{2.29}$$

The charge asymmetry differs from zero in the presence of direct CP violation, and parametrizes it.

3 CP violation in $B^0_{(s)} \to h^+h^-$ decays

As already mentioned in the introduction, measurements of the β angle of the UT are already available by the BABAR and BELLE experiments [3, 4]. By averaging the measurements of $\sin 2\beta$ performed by the two experiments one obtains:

$$\sin 2\beta = 0.731 \pm 0.055, \tag{3.1}$$

in very good agreement with the Standard Model prediction of β from the magnitudes of CKM elements [12].

BABAR and BELLE have also provided first measurements of direct and mixing-induced CP-violating observables in the $B^0 \to \pi^+\pi^-$ decay [5, 6]. The results of the two experiments are:

$$\mathcal{A}_{\pi\pi}^{dir} = \begin{cases} 0.30 \pm 0.25 & \text{BABAR} \\ 0.77 \pm 0.28 & \text{BELLE} \end{cases} \tag{3.2}$$

and

$$\mathcal{A}_{\pi\pi}^{mix} = \begin{cases} 0.02 \pm 0.34 & \text{BABAR} \\ -1.23 \pm 0.42 & \text{BELLE} \end{cases} . \tag{3.3}$$

As it can be seen, these results are only weakly compatible at the moment, and it is not yet possible to draw any definite conclusion.

The decay $B^0 \to \pi^+\pi^-$ is generated by the $\overline{b} \to \overline{u} + W^+$ tree diagram, as well as by $\overline{b} \to \overline{d} + g\,(\gamma,\, Z^0)$ penguin diagrams. Analogously, the $B_s^0 \to K^+K^-$ is generated by the same diagrams, if all the d $(\overline{d})$ quarks are replaced by the s $(\overline{s})$ quarks (see Figure 1). Concerning the $B^0 \to K^+\pi^-$ and $B_s^0 \to \pi^+K^-$ decays, it can be noted that they differ only in the spectator quarks from $B_s^0 \to K^+K^-$ and $B^0 \to \pi^+\pi^-$ respectively.

The direct and mixing-induced CP asymmetry terms for the $B^0 \to \pi^+\pi^-$ decay can be parametrized in the Standard Model as [8, 9]:

$$\mathcal{A}_{\pi\pi}^{dir} = \mathcal{A}_{\pi\pi}^{dir}(d,\, \vartheta,\, \gamma) = \frac{2d \sin\vartheta \sin\gamma}{1 - 2d\cos\vartheta\cos\gamma + d^2} \tag{3.4}$$

and

$$\mathcal{A}_{\pi\pi}^{mix} = \mathcal{A}_{\pi\pi}^{mix}(d,\, \vartheta,\, \gamma,\, \phi_d) = -\frac{\sin(\phi_d + 2\gamma) - 2d\cos\vartheta\sin(\phi_d + \gamma) + d^2\sin\phi_d}{1 - 2d\cos\vartheta\cos\gamma + d^2}, \tag{3.5}$$

where the dependence on the CKM angle γ is made explicit, and $\phi_d = 2\beta$ is the $B^0 - \overline{B}^0$ mixing phase. The parameters d and ϑ are real quantities defined by:

$$de^{i\vartheta} = \frac{2\,|V_{us}V_{cb}|}{2\,|V_{ub}| - |V_{ub}|\,|V_{us}|^2} \times \frac{A_{pen}^c - A_{pen}^t}{A_{cc}^u + A_{pen}^u - A_{pen}^t}, \tag{3.6}$$

where V_{ij} are CKM matrix elements, A_{pen}^j $(j \in \{u,\, c,\, t\})$ describe penguin topologies with internal j quarks, and A_{cc}^u describes current-current contributions. Sloppily speaking, d and ϑ parametrize the magnitude and the phase of the penguin-to-tree amplitude ratio respectively. Predictions of the values of the hadronic amplitudes necessary to compute

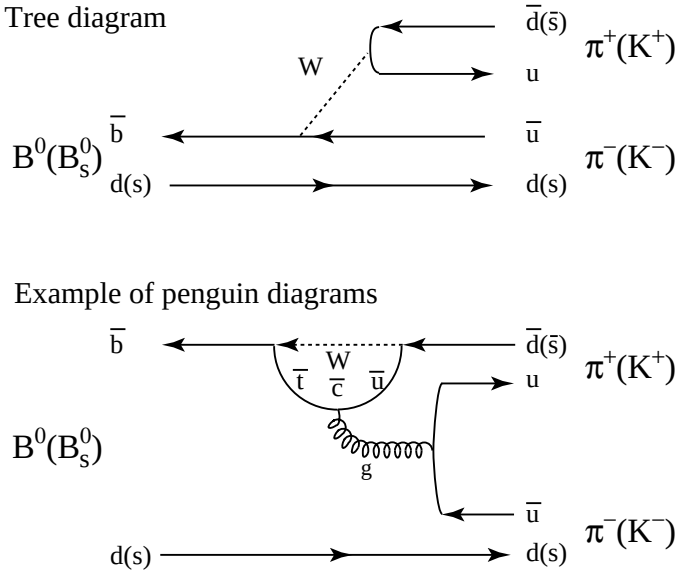

Figure 1: Tree and penguin diagrams generating $B^0 \to \pi^+\pi^-$ ($B_s^0 \to K^+K^-$) decays. By exchanging all the d ($\overline{d}$) quarks by the s ($\overline{s}$) quarks, the tree and penguin processes of the $B^0 \to \pi^+\pi^-$ generate those of the $B_s^0 \to K^+K^-$.

these parameters are very challenging at the moment, especially for what concerns the phase ϑ.

For the $B_s^0 \to K^+K^-$, following similar lines, one can write:

$$\mathcal{A}_{KK}^{dir} = \mathcal{A}_{KK}^{dir}(d',\, \vartheta',\, \gamma) = -\frac{2\tilde{d}' \sin\vartheta' \sin\gamma}{1 + 2\tilde{d}' \cos\vartheta' \cos\gamma + \tilde{d}'^2}, \tag{3.7}$$

$$\mathcal{A}_{KK}^{mix} = \mathcal{A}_{KK}^{mix}(d',\, \vartheta',\, \gamma,\, \phi_s) = -\frac{\sin(\phi_s + 2\gamma) + 2\tilde{d}' \cos\vartheta' \sin(\phi_s + \gamma) + \tilde{d}'^2 \sin\phi_s}{1 + 2\tilde{d}' \cos\vartheta' \cos\gamma + \tilde{d}'^2} \tag{3.8}$$

and

$$\mathcal{A}_{KK}^{\Delta} = \mathcal{A}_{KK}^{\Delta}(d',\, \vartheta',\, \gamma,\, \phi_s) = \frac{\cos(\phi_s + 2\gamma) + 2\tilde{d}' \cos\vartheta' \cos(\phi_s + \gamma) + \tilde{d}'^2 \cos\phi_s}{1 + 2\tilde{d}' \cos\vartheta' \cos\gamma + \tilde{d}'^2}, \tag{3.9}$$

where $\phi_s = -2\chi$ is the $B_s^0 - \overline{B}_s^0$ mixing phase and the parameter $\tilde{d}'$ is defined as:

$$\tilde{d}' = \frac{1 - |V_{us}|^2}{|V_{us}|^2} d'. \tag{3.10}$$

The magnitude of the CKM matrix element V_{us} is given by the sine of the Cabibbo angle, and is measured with very good precision to be $|V_{us}| = 0.2229 \pm 0.0022$ from nuclear, kaon

and hyperon decays [12]. d' and ϑ' are the analogs of d and ϑ for the $B_s^0 \to K^+K^-$ decay, similarly defined by:

$$d'e^{i\vartheta'} = \frac{2\left|V_{us}V_{cb}\right|}{2\left|V_{ub}\right| - \left|V_{ub}\right|\left|V_{us}\right|^2} \times \frac{A_{pen}^{c\prime} - A_{pen}^{t\prime}}{A_{cc}^{u\prime} + A_{pen}^{u\prime} - A_{pen}^{t\prime}}, \tag{3.11}$$

where $A_{pen}^{j\prime}$ and $A_{cc}^{u\prime}$ are the corresponding penguin and current-current contributes to the decay amplitude. The terms $\mathcal{A}_{KK}^{dir}$, $\mathcal{A}_{KK}^{mix}$ and $\mathcal{A}_{KK}^{\Delta}$ satisfy the following relation:

$$\left(\mathcal{A}_{KK}^{dir}\right)^2 + \left(\mathcal{A}_{KK}^{mix}\right)^2 + \left(\mathcal{A}_{KK}^{\Delta}\right)^2 = 1. \tag{3.12}$$

By noting that the $B^0 \to \pi^+\pi^-$ and $B_s^0 \to K^+K^-$ decays are each other related by exchanging all the d ($\overline{d}$) and the s ($\overline{s}$) quarks, in the limit of validity of the U-spin symmetry of the strong interaction dynamics, one can write:

$$d = d' \tag{3.13}$$

and

$$\vartheta = \vartheta'. \tag{3.14}$$

For a discussion on the validity of this symmetry see [8, 9] and references therein.

Finally, since the $B^0 \to K^+\pi^-$ and $B_s^0 \to \pi^+K^-$ decays differ only in the spectator quarks from $B_s^0 \to K^+K^-$ and $B^0 \to \pi^+\pi^-$ decays, respectively, by relying on the $SU(3)$ flavor symmetry and on certain dynamical assumptions [8, 9], one can write the following relations:

$$\mathcal{A}_{KK}^{dir} \simeq \mathcal{A}_{K\pi} \tag{3.15}$$

and

$$\mathcal{A}_{\pi\pi}^{dir} \simeq \mathcal{A}_{\pi K}, \tag{3.16}$$

where $\mathcal{A}_{K\pi}$ is the charge asymmetry for the $B^0 \to K^+\pi^-$ decay, and $\mathcal{A}_{\pi K}$ is the charge asymmetry for the $B_s^0 \to \pi^+K^-$ decay.

4 Experimental decay rates

In Section 2 we have introduced the theoretical expressions of the rates for B and $\overline{B}$ decays. From these expressions, it is possible to define the observed decay rates, by taking into account the tagging, the presence of background, the signal acceptance as a function of the proper time after trigger and offline-selection, and the resolution on the proper time measurement.

4.1 Case of f CP eigenstate

In case f is a CP eigenstate, one has the following two observed decay rates for events tagged as B and $\overline{B}$ respectively:

$$R_f(t) = \int \left[(1-\omega)k\Gamma_{B\to f}(\tau) + \omega k\Gamma_{\overline{B}\to f}(\tau)\right]\epsilon(\tau)G(\tau - t)d\tau + \frac{1}{2}B(t) \tag{4.1}$$

and

$$\overline{R}_f(t) = \int \left[\omega k\Gamma_{B\to f}(\tau) + (1-\omega)k\Gamma_{\overline{B}\to f}(\tau)\right]\epsilon(\tau)G(\tau - t)d\tau + \frac{1}{2}B(t), \tag{4.2}$$

where $\epsilon(t)$ is the acceptance as function of the proper time t, while $G(\Delta t)$ is a function describing the proper time resolution for the decays under study. These functions must be determined from full Monte Carlo simulations. The function $B(t)$ is an effective function describing the proper time dependence of the background rate. Its functional form for combinatorial background can be extracted from data by studying the proper time distribution in the mass sidebands spectrum. Without lack of generality, it can be written as:

$$B(t) = n_B b(t), \tag{4.3}$$

being n_B the number of tagged background events and where $b(t)$ is normalized to 1:

$$\int b(t)dt = 1. \tag{4.4}$$

The constant k is obtained by noting that the sum of the integrals of the two signal rates must give the total number of tagged signal events n_f:

$$\begin{aligned} &\int\int \left[(1-\omega)k\Gamma_{B\to f}(\tau) + \omega k\Gamma_{\overline{B}\to f}(\tau)\right]\epsilon(\tau)G(\tau - t)d\tau dt + \\ &+\int\int \left[\omega k\Gamma_{B\to f}(\tau) + (1-\omega)k\Gamma_{\overline{B}\to f}(\tau)\right]\epsilon(\tau)G(\tau - t)d\tau dt = n_f \end{aligned} \tag{4.5}$$

yielding:

$$k = \frac{n_f}{\int\int \left[\Gamma_{B\to f}(\tau) + \Gamma_{\overline{B}\to f}(\tau)\right]\epsilon(\tau)G(\tau - t)d\tau dt}. \tag{4.6}$$

By introducing the auxiliary functions:

$$h_+(t) = \int I_+(\tau)\epsilon(\tau)G(\tau - t)e^{-\Gamma\tau}d\tau \tag{4.7}$$

and

$$h_-(t) = \int I_-(\tau)\epsilon(\tau)G(\tau - t)e^{-\Gamma\tau}d\tau, \tag{4.8}$$

and their normalizations:

$$H_+ = \int h_+(t)dt \tag{4.9}$$

and

$$H_- = \int h_-(t)dt, \tag{4.10}$$

the observed decay rates assume the following simple expressions:

$$R_f(t) = \frac{h_+(t) + (1-2\omega)h_-(t)}{H_+}\frac{n_f}{2} + \frac{n_B}{2}b(t) \tag{4.11}$$

and

$$\overline{R}_f(t) = \frac{h_+(t) - (1-2\omega)h_-(t)}{H_+}\frac{n_f}{2} + \frac{n_B}{2}b(t). \tag{4.12}$$

It is straightforward to verify that, as expected, one has:

$$\int \left[R_f(t) + \overline{R}_f(t)\right] dt = n_f + n_B \tag{4.13}$$

It is then possible to write the probability density functions (p.d.f.s) for the variable t for events tagged as B and $\overline{B}$, by normalizing the decay rates, i.e.:

$$P_f(t) = \frac{R_f(t)}{\int R_f(\tau)d\tau} = \frac{\frac{h_+(t) + (1-2\omega)h_-(t)}{H_+}\frac{n_f}{2} + \frac{n_B}{2}b(t)}{\frac{H_+ + (1-2\omega)H_-}{H_+}\frac{n_f}{2} + \frac{n_B}{2}} \tag{4.14}$$

and

$$\overline{P}_f(t) = \frac{\overline{R}_f(t)}{\int \overline{R}_f(\tau)d\tau} = \frac{\frac{h_+(t) - (1-2\omega)h_-(t)}{H_+}\frac{n_f}{2} + \frac{n_B}{2}b(t)}{\frac{H_+ - (1-2\omega)H_-}{H_+}\frac{n_f}{2} + \frac{n_B}{2}} \tag{4.15}$$

Finally, by introducing the invariant mass m, which is measured for each event together with the proper time t, one can write joint p.d.f.s for t and m:

$$P_f(t,\, m) = \frac{\frac{h_+(t) + (1-2\omega)h_-(t)}{H_+}\frac{n_f\, g_S(m)}{2} + \frac{n_B\, g_B(m)}{2}b(t)}{\frac{H_+ + (1-2\omega)H_-}{H_+}\frac{n_f}{2} + \frac{n_B}{2}} \tag{4.16}$$

and

$$\overline{P}_f(t,\, m) = \frac{\dfrac{h_+(t) - (1-2\omega)h_-(t)}{H_+}\dfrac{n_f\, g_S(m)}{2} + \dfrac{n_B\, g_B(m)}{2}b(t)}{\dfrac{H_+ - (1-2\omega)H_-}{H_+}\dfrac{n_f}{2} + \dfrac{n_B}{2}}. \tag{4.17}$$

The functions $g_S(m)$ and $g_B(m)$ describe the shape of the mass spectra for signal and background events respectively, and must be normalized to 1:

$$\int g_S(m)dm = \int g_B(m)dm = 1 \tag{4.18}$$

4.2 Case of f flavor specific final state

In case f is a flavor specific final state, the previous formalism can be re-used in order to write the following four joint p.d.f.s:

$$P_f(t,\, m) = \frac{\dfrac{h_+(t) + (1-2\omega)h_-(t)}{H_+}\dfrac{n_f\, g_S(m)}{2} + \dfrac{n_B\, g_B(m)}{2}b(t)}{\dfrac{H_+ + (1-2\omega)H_-}{H_+}\dfrac{n_f}{2} + \dfrac{n_B}{2}}, \tag{4.19}$$

$$\overline{P}_f(t,\, m) = \frac{\dfrac{h_+(t) - (1-2\omega)h_-(t)}{H_+}\dfrac{n_f\, g_S(m)}{2} + \dfrac{n_B\, g_B(m)}{2}b(t)}{\dfrac{H_+ - (1-2\omega)H_-}{H_+}\dfrac{n_f}{2} + \dfrac{n_B}{2}}, \tag{4.20}$$

$$P_{\overline{f}}(t,\, m) = \frac{\dfrac{h_+(t) - (1-2\omega)h_-(t)}{H_+}\dfrac{n_{\overline{f}}\, g_S(m)}{2} + \dfrac{n_{\overline{B}}\, g_B(m)}{2}b(t)}{\dfrac{H_+ - (1-2\omega)H_-}{H_+}\dfrac{n_{\overline{f}}}{2} + \dfrac{n_{\overline{B}}}{2}} \tag{4.21}$$

and

$$\overline{P}_{\overline{f}}(t,\, m) = \frac{\dfrac{h_+(t) + (1-2\omega)h_-(t)}{H_+}\dfrac{n_{\overline{f}}\, g_S(m)}{2} + \dfrac{n_{\overline{B}}\, g_B(m)}{2}b(t)}{\dfrac{H_+ + (1-2\omega)H_-}{H_+}\dfrac{n_{\overline{f}}}{2} + \dfrac{n_{\overline{B}}}{2}} \tag{4.22}$$

In the previous expressions, n_f and $n_{\overline{f}}$ are the numbers of tagged f and $\overline{f}$ signal events respectively, while n_B and $n_{\overline{B}}$ are the numbers of tagged background events for the states f and $\overline{f}$ respectively. The functions $h_+(t)$ and $h_-(t)$ are given by:

$$\begin{aligned} h_+(t) &= \int I_+(\tau)\epsilon(\tau)G(\tau - t)e^{-\Gamma\tau}d\tau = \int \overline{I}_+(\tau)\epsilon(\tau)G(\tau - t)e^{-\Gamma\tau}d\tau = \\ &= \int \left(\cosh\frac{\Delta\Gamma}{2}\tau\right)\epsilon(\tau)G(\tau - t)e^{-\Gamma\tau}d\tau \end{aligned} \tag{4.23}$$

and

$$h_-(t) = \int I_-(\tau)\epsilon(\tau)G(\tau-t)e^{-\Gamma\tau}d\tau = \int \overline{I}_-(\tau)\epsilon(\tau)G(\tau-t)e^{-\Gamma\tau}d\tau = \\ = \int (\cos\Delta m\tau)\,\epsilon(\tau)G(\tau-t)e^{-\Gamma\tau}d\tau. \tag{4.24}$$

The normalizations H_+ and H_- can be calculated as in (4.9) and (4.10).

The charge asymmetry defined in (2.29), can be introduced in the expressions (4.19), (4.20), (4.21) and (4.22), by writing:

$$n_f = \frac{1}{2}n_{f,\overline{f}}\left(1-\mathcal{A}_{f,\overline{f}}\right) \tag{4.25}$$

and

$$n_{\overline{f}} = \frac{1}{2}n_{f,\overline{f}}\left(1+\mathcal{A}_{f,\overline{f}}\right), \tag{4.26}$$

being $n_{f,\overline{f}}$ the total number of tagged f or $\overline{f}$ signal events, i.e., defined as:

$$n_{f,\overline{f}} = n_f + n_{\overline{f}} \tag{4.27}$$

5 Construction of the likelihood function

The formalism of the previous section has been introduced in order to write the expressions of the likelihood functions, to be used to extract the physics parameters of interest from data, as it is discussed in the following.

5.1 Case of f CP eigenstate

In the case f is a CP eigenstate, the likelihood is written as:

$$\mathcal{L}_f = \prod_{i=1}^{N_f} P_f(t_i,\, m_i) \prod_{j=1}^{\overline{N}_f} \overline{P}_f(t_j,\, m_j) \tag{5.1}$$

This expression of the likelihood is formally correct, being the joint p.d.f.s for proper time and mass $P_f(t,\, m)$ and $\overline{P}_f(t,\, m)$ normalized to 1:

$$\int\int P_f(t,\, m)dtdm = \int\int \overline{P}_f(t,\, m)dtdm = 1, \tag{5.2}$$

but it has the disantvantage of using only the information on the shape of the decay rates, ignoring the additional information of the integrals of the decay rates, i.e., ignoring the information of the integrated asymmetry. This information is almost irrelevant in the case of B_s^0 decays, due to the large oscillations which lead to null integrated asymmetry. On

the other hand, it is crucial to take it into account in the case of B^0 decays, since the slow oscillations of the B^0 mesons lead to a sizable difference between the numbers of B^0 and $\overline{B}^0$ decays.

The information on the integrals of the decay rates can be naturally included by defining an extended likelihood function:

$$\mathcal{L}_f^{ext} = \frac{\left(N_f + \overline{N}_f\right)!}{N_f! \overline{N}_f!} \left(\frac{\nu}{N_f + \overline{N}_f}\right)^{N_f} \left(\frac{\overline{\nu}}{N_f + \overline{N}_f}\right)^{\overline{N}_f} \prod_{i=1}^{N_f} P_f(t_i,\, m_i) \prod_{j=1}^{\overline{N}_f} \overline{P}_f(t_j,\, m_j) \tag{5.3}$$

where ν is the average of the number of events tagged as B (N_f is a single realization of this random variable) and $\overline{\nu}$ is the average of the number of events tagged as $\overline{B}$ ($\overline{N}_f$ is a single realization). The logarithm of the extended likelihood is then:

$$\begin{aligned} \log \mathcal{L}_f^{ext} &= N_f \log \nu + \sum_{i=1}^{N_f} \log P_f(t_i,\, m_i) + \overline{N}_f \log \overline{\nu} + \sum_{j=1}^{\overline{N}_f} \log \overline{P}_f(t_j,\, m_j) + \\ &+ \log \frac{\left(N_f + \overline{N}_f\right)!}{N_f! \overline{N}_f! \left(N_f + \overline{N}_f\right)^{N_f + \overline{N}_f}} = \sum_{i=1}^{N_f} \nu \, \log P_f(t_i,\, m_i) + \sum_{j=1}^{\overline{N}_f} \overline{\nu} \log \overline{P}_f(t_j,\, m_j) + \text{const.} \end{aligned} \tag{5.4}$$

where the constant term can be dropped since irrelevant in the maximization of $\log \mathcal{L}_f^{ext}$.

The averages ν and $\overline{\nu}$ are given exactly by the following relations:

$$\begin{aligned} \nu &= \int\int \left[\frac{h_+(t) + (1-2\omega)h_-(t)}{H_+} \frac{n_f g_S(m)}{2} + \frac{n_B g_B(m)}{2} b(t)\right] dtdm = \\ &= \frac{H_+ + (1-2\omega)H_-}{H_+} \frac{n_f}{2} + \frac{n_B}{2} \end{aligned} \tag{5.5}$$

and

$$\begin{aligned} \overline{\nu} &= \int\int \left[\frac{h_+(t) - (1-2\omega)h_-(t)}{H_+} \frac{n_f g_S(m)}{2} + \frac{n_B g_B(m)}{2} b(t)\right] dtdm = \\ &= \frac{H_+ - (1-2\omega)H_-}{H_+} \frac{n_f}{2} + \frac{n_B}{2}. \end{aligned} \tag{5.6}$$

Thus, redefining:

$$P'_f(t,\, m) = \frac{h_+(t) + (1-2\omega)h_-(t)}{H_+} \frac{n_f g_S(m)}{2} + \frac{n_B g_B(m)}{2} b(t) \tag{5.7}$$

and

$$\overline{P}'_f(t,\, m) = \frac{h_+(t) - (1-2\omega)h_-(t)}{H_+}\frac{n_f g_S(m)}{2} + \frac{n_B g_B(m)}{2} b(t), \tag{5.8}$$

one can rewrite the extended log-likelihood function obtaining the final expression:

$$\log \mathcal{L}_f^{ext} = \sum_{i=1}^{N_f} \log P'_f(t_i,\, m_i) + \sum_{j=1}^{\overline{N}_f} \log \overline{P}'_f(t_j,\, m_j). \tag{5.9}$$

It is interesting to note that the following relation for the resolution on the CP-violating observables holds:

$$\frac{1}{\sigma^2_{\mathcal{L}^{ext}}} \simeq \frac{1}{\sigma^2_{int}} + \frac{1}{\sigma^2_{\mathcal{L}}}, \tag{5.10}$$

i.e., the resolution achievable with the extended likelihood is the combination of the resolutions achievable by using the integrals of the decay rates and the ordinary likelihood. Due to the fast oscillation of the B_s^0 meson, for B_s^0 decays one has $1/\sigma^2_{int} \simeq 0$, thus the extended likelihood and ordinary likelihood methods are nearly equivalent. The same fact does not apply to B^0 decays, where the information of the integrals is important and cannot be neglected.

5.2 Case of f flavor specific final state

In case f is a flavor specific final state, the previous formalism can be easily re-used. We only write here the final expression, with obvious meaning of the symbols:

$$\log \mathcal{L}_f^{ext} = \sum_{i=1}^{N_f} \log P'_f(t_i,\, m_i) + \sum_{j=1}^{\overline{N}_f} \log \overline{P}'_f(t_j,\, m_j) + \sum_{i=1}^{N_{\overline{f}}} \log P'_{\overline{f}}(t_i,\, m_i) + \sum_{j=1}^{\overline{N}_{\overline{f}}} \log \overline{P}'_{\overline{f}}(t_j,\, m_j) \tag{5.11}$$

5.3 Comments on the numerical computation of the likelihood

The likelihood optimization, performed with the MIGRAD algorithm from the CERNLIB MINUIT library [13], is rather expensive from the point of view of CPU time. This is because of the large signal and background statistics, and because, for each event, a convolution product of the signal rates with the proper time resolution function $G(\Delta t)$ needs to be performed. These integrals in the expression of the likelihood function are in the form:

$$C_h(t_i) = \int \epsilon(\tau) e^{-\Gamma\tau} \cosh\frac{\Delta\Gamma}{2}\tau\, G(\tau - t_i) d\tau, \tag{5.12}$$

$$S_h(t_i) = \int \epsilon(\tau) e^{-\Gamma\tau} \sinh\frac{\Delta\Gamma}{2}\tau\, G(\tau - t_i) d\tau, \tag{5.13}$$

$$C(t_i) = \int \epsilon(\tau) e^{-\Gamma\tau} \cos \Delta m\tau \, G(\tau - t_i) d\tau, \tag{5.14}$$

and

$$S(t_i) = \int \epsilon(\tau) e^{-\Gamma\tau} \sin \Delta m\tau \, G(\tau - t_i) d\tau, \tag{5.15}$$

being t_i the proper time of the i-th event. A nice improvement can be achieved by noting that these integrals can be re-calculated only when the parameters Γ, $\Delta\Gamma$ and Δm are varied during the optimization. If Γ, $\Delta\Gamma$ and Δm have the same values as the previous optimization step, these integrals, for each event i, have the same values as they had in the previous step. Thus, they can be saved into arrays and re-used without being re-computed. But this is not yet very satisfactory, since while simulating hundreds of toy Monte Carlo experiments, the required total time remains very large.

A dramatic improvement can be obtained by noting that the achievable resolutions on Γ, $\Delta\Gamma$ and Δm are rather good, i.e., they are not varied too much from their nominal values during the numerical maximization procedure. Thus, an expansion in power series on Γ, $\Delta\Gamma$, and Δm to a fixed order n can be safely performed, i.e.:

$$\begin{aligned} C_h(t_i) &= \sum_{j=0}^{n} \sum_{k=0}^{j} \frac{(-1)^k}{j! 2^{j-k}} \left(\begin{array}{c} j \\ k \end{array} \right) R_h(t_i, \, j, \, \mathrm{mod}(j-k, \, 2) + 1) \times \\ &\quad \times (\Gamma - \Gamma_N)^k (\Delta\Gamma - \Delta\Gamma_N)^{j-k}, \end{aligned} \tag{5.16}$$

$$\begin{aligned} S_h(t_i) &= \sum_{j=0}^{n} \sum_{k=0}^{j} \frac{(-1)^k}{j! 2^{j-k}} \left(\begin{array}{c} j \\ k \end{array} \right) R_h(t_i, \, j, \, \mathrm{mod}(j-k+1, \, 2) + 1) \times \\ &\quad \times (\Gamma - \Gamma_N)^k (\Delta\Gamma - \Delta\Gamma_N)^{j-k}, \end{aligned} \tag{5.17}$$

$$\begin{aligned} C(t_i) &= \sum_{j=0}^{n} \sum_{k=0}^{j} \frac{(-1)^k (-1)^{\mathrm{int}\left(\frac{j-k+1}{2}\right)}}{j!} \left(\begin{array}{c} j \\ k \end{array} \right) R(t_i, \, j, \, \mathrm{mod}\,(j-k, \, 2) + 1) \times \\ &\quad \times (\Gamma - \Gamma_N)^k (\Delta m - \Delta m_N)^{j-k} \end{aligned} \tag{5.18}$$

and

$$\begin{aligned} S(t_i) &= \sum_{j=0}^{n} \sum_{k=0}^{j} \frac{(-1)^k (-1)^{\mathrm{int}\left(\frac{j-k}{2}\right)}}{j!} \left(\begin{array}{c} j \\ k \end{array} \right) R(t_i, \, j, \, \mathrm{mod}(j-k+1, \, 2) + 1) \times \\ &\quad \times (\Gamma - \Gamma_N)^k (\Delta m - \Delta m_N)^{j-k}, \end{aligned} \tag{5.19}$$

where Γ_N, $\Delta\Gamma_N$ and Δm_N are nominal values for Γ, $\Delta\Gamma$, and Δm (i.e., they can be chosen to be not too far from the true values), the operator int$(\cdot)$ stands for truncated integral part of its argument, and the operator mod$(n,\, m)$ stands for n module m. The functions R_h and R are given by:

$$R_h(t_i,\, l,\, 1) = \int \tau^l \epsilon(\tau) e^{-\Gamma_N \tau} \cosh \frac{\Delta\Gamma_N}{2} \tau \, G(\tau - t_i) d\tau, \tag{5.20}$$

$$R_h(t_i,\, l,\, 2) = \int \tau^l \epsilon(\tau) e^{-\Gamma_N \tau} \sinh \frac{\Delta\Gamma_N}{2} \tau \, G(\tau - t_i) d\tau, \tag{5.21}$$

$$R(t_i,\, l,\, 1) = \int \tau^l \epsilon(\tau) e^{-\Gamma_N \tau} \cos \Delta m_N \tau \, G(\tau - t_i) d\tau \tag{5.22}$$

and

$$R(t_i,\, l,\, 2) = \int \tau^l \epsilon(\tau) e^{-\Gamma_N \tau} \sin \Delta m_N \tau \, G(\tau - t_i) d\tau. \tag{5.23}$$

These integrals, which are constants for each event i because they no longer depend on the current parameters Γ, $\Delta\Gamma$, and Δm, can be calculated only at the very first step of the maximization procedure, their values saved into arrays and re-used for all subsequent optimization steps until the fit convergence. An expansion of order $n = 5$ has been checked to be more than sufficient, with the fit yielding the same results if compared to the full non-approximated time-consuming calculation. This approximation reduces the completion time of a fit from the order of several hours to the order of ten minutes on a powerful modern CPU.

6 Combined likelihood fit

By looking at the experimental CP asymmetry for the $B^0 \to \pi^+\pi^-$ decay that, in the case of no background and perfect proper time resolution for simplicity, can be written as:

$$\mathcal{A}^{CP}_{\pi\pi}(t) = (1 - 2\omega) \left(\mathcal{A}^{dir}_{\pi\pi} \cos \Delta m_d \, t + \mathcal{A}^{mix}_{\pi\pi} \sin \Delta m_d \, t \right), \tag{6.1}$$

one can easily convince him/herself that it is not possible to extract ω, $\mathcal{A}^{dir}_{\pi\pi}$ and $\mathcal{A}^{mix}_{\pi\pi}$ simultaneously from the $B^0 \to \pi^+\pi^-$ data only, since the problem is undetermined.

Similarly, for the $B^0_s \to K^+K^-$ decay, one can write the experimental CP asymmetry as:

$$\mathcal{A}^{CP}_{KK}(t) = (1 - 2\omega) \frac{\mathcal{A}^{dir}_{KK} \cos \Delta m_s \, t + \mathcal{A}^{mix}_{KK} \sin \Delta m_s \, t}{\cosh \frac{\Delta\Gamma_s}{2} t - \mathcal{A}^{\Delta}_{KK} \sinh \frac{\Delta\Gamma_s}{2} t}, \tag{6.2}$$

where, in the Standard Model, the relation (3.12) between $\mathcal{A}^{dir}_{KK}$, $\mathcal{A}^{mix}_{KK}$ and $\mathcal{A}^{\Delta}_{KK}$ holds.

In this case, the indetermination of the problem is in principle removed by the presence of the denominator in (6.2), but, since the expected value of $\Delta\Gamma_s$ is small ($\Delta\Gamma_s/\Gamma_s \simeq 10\%$), i.e., the denominator does not differ dramatically from unity for every interesting proper

time t, useful information on ω could be extracted from data with good precision only with very large statistics.

A way out consists in making use of the $B^0 \to K^+\pi^-$ ($B_s^0 \to \pi^+K^-$) decay, which is topologically similar to the $B^0 \to \pi^+\pi^-$ ($B_s^0 \to K^+K^-$) decay. This similarity, apart from small systematics that can be taken into account, leads to same tagging efficiency and mistag probability.

For the $B^0 \to K^+\pi^-$ decay, two experimental decay asymmetries can be written, given by:

$$\mathcal{A}_{K^+\pi^-}(t) = -(1-2\omega)\cos\Delta m_d\, t \tag{6.3}$$

and

$$\mathcal{A}_{K^-\pi^+}(t) = (1-2\omega)\cos\Delta m_d\, t. \tag{6.4}$$

Analogously, for the $B_s^0 \to \pi^+K^-$ one has:

$$\mathcal{A}_{\pi^+K^-}(t) = -(1-2\omega)\frac{\cos\Delta m_s\, t}{\cosh\frac{\Delta\Gamma_s}{2}t} \tag{6.5}$$

and

$$\mathcal{A}_{\pi^-K^+}(t) = (1-2\omega)\frac{\cos\Delta m_s\, t}{\cosh\frac{\Delta\Gamma_s}{2}t}. \tag{6.6}$$

The expressions of the asymmetries (6.3), (6.4), (6.5) and (6.6) clearly state that for the $B^0 \to K^+\pi^-$ and $B_s^0 \to \pi^+K^-$ channels it is possible to extract the mistag probability ω from data.

One can conclude that, in order to use the maximal content of information and determine ω, the most general approach consists in making a combined likelihood fit of $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$ ($B_s^0 \to K^+K^-$ and $B_s^0 \to \pi^+K^-$) data[1]. This is discussed in the following.

6.1 Case of $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$

A combined fit of $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$ data can be realized by writing a combined extended likelilood function in the following way:

$$\log\mathcal{L}^{ext}_{\pi\pi,\,K\pi} = \sum_{i=1}^{N_{\pi\pi}}\log P'_{\pi\pi}(t_i,\,m_i) + \sum_{j=1}^{\overline{N}_{\pi\pi}}\log\overline{P}'_{\pi\pi}(t_j,\,m_j) + \sum_{i=1}^{N_{K^+\pi^-}}\log P'_{K^+\pi^-}(t_i,\,m_i) +$$

[1]The mistag probability can be also determined from Monte Carlo simulations and used in the likelihood fit, once there is enough confidence that the Monte Carlo reproduces the data with sufficient accuracy. In this work we ignore this possibility. In this sense, this can be considered as a conservative choice.

$$+ \sum_{j=1}^{\overline{N}_{K^+\pi^-}} \log \overline{P}'_{K^+\pi^-}(t_j,\, m_j) + \sum_{i=1}^{N_{K^-\pi^+}} \log P'_{K^-\pi^+}(t_i,\, m_i) + \sum_{j=1}^{\overline{N}_{K^-\pi^+}} \log \overline{P}'_{K^-\pi^+}(t_j,\, m_j). \tag{6.7}$$

In order to improve the sensitivity on the CP parameters, prior information, such as the known resolution and values of Δm_d and Γ_d, can be naturally included in the fit by multiplying the likelihood by Gaussian functions, i.e.:

$$\log \mathcal{L}^{ext'}_{\pi\pi,\, K\pi} = \log \mathcal{L}^{ext}_{\pi\pi,\, K\pi} - \frac{(\Delta m_d - \overline{\Delta m}_d)^2}{2\sigma^2_{\Delta m_d}} - \frac{(\Gamma_d - \overline{\Gamma}_d)^2}{2\sigma^2_{\Gamma_d}}. \tag{6.8}$$

For $\sigma_{\Delta m_d}$ and σ_{Γ_d}, their current knowledge has been used [14]. On the contrary, no prior on $\Delta\Gamma_d$ has been used, since the current experimental knowledge on this parameter consists only in a large upper limit.

In the following, we will define all the terms that are present in the expression of the likelihood. For more details on how the values of the relevant quantities have been estimated, see [10].

The signal efficiency $\epsilon(t)$ as function of the proper time has been studied for each $B^0_{(s)} \to h^+h^-$ channel. It can be parametrized as:

$$\epsilon(t) = \frac{\alpha}{1 + \exp\left[(t \times ps^{-1})^{\beta}\right]}. \tag{6.9}$$

In the expression of the likelihood, the constant α is cancelled (e.g., it can be set to 1), as it can be argued by looking at (5.7) and (5.8), where it can be factorized in the expressions of $h_+(t)$, $h_-(t)$ and H_+, and thus eliminated in the ratio. The parameter β has been fixed to the central value obtained from the full simulation.

The proper time resolution function $G(\Delta t)$ has also been studied. It turns out to be well described by a double Gaussian[2]:

$$G(\Delta t) = \frac{f}{\sqrt{2\pi}\sigma_{\Delta t_1}} \exp\left(-\frac{\Delta t^2}{2\sigma^2_{\Delta t_1}}\right) + \frac{(1-f)}{\sqrt{2\pi}\sigma_{\Delta t_2}} \exp\left(-\frac{\Delta t^2}{2\sigma^2_{\Delta t_2}}\right), \tag{6.10}$$

where $\Delta t = t_{rec} - t_{true}$, and t_{rec} is the reconstructed proper time while t_{true} is the true proper time. A detailed study of the dependence of Δt on the true proper time has been performed. Only a moderate dependence is observed, and it can be ignored at first order without dramatic loss of information. In this study this dependence is considered

[2]In the studies described in [10], the presence of a small bias in the proper time resolution is observed. It is due to the fact that B decay products are included in the primary vertex finding, so biasing the primary vertex position, and in turn biasing the distance of flight and thus the proper time measurement [2]. This bias can be eliminated by removing the B decay products once the final state has been reconstructed, and then re-fitting the primary vertex. For the studies presented in this note, we assume that this bias is not present. This is why the expression (6.10) does not contain it, differently from what shown in [10].

negligible and it is ignored[3]. The parameters of the resolution function are fixed in the fit as determined from the full simulation.

The normalized rate $b(t)$ for combinatorial background events has been guessed to be the same as the untagged signal one, and written as:

$$b(t) = \frac{e^{-\eta t}}{1+\exp\left[(t \times ps^{-1})^{\delta}\right]} \Big/ \int \frac{e^{-\eta \tau}}{1+\exp\left[(\tau \times ps^{-1})^{\delta}\right]} d\tau. \tag{6.11}$$

The values of δ and η have been determined specifically for each decay under study, by fitting this expression to triggered $b\bar{b}$ data, selected with slightly reduced offline-selection cuts in order to have a sizable statistics. These parameters are left free during the likelihood maximization and their best values are determined from the fit.

The signal mass distribution from the full GEANT simulation has been found to be consistent with a single Gaussian:

$$g_S(m) = \frac{1}{\sqrt{2\pi}\sigma_m} \exp\left(-\frac{(m-\overline{m})^2}{2\sigma_m^2}\right). \tag{6.12}$$

The two parameters $\overline{m}$ and σ_m are left free during the likelihood maximization, and their best values are determined from the fit.

The combinatorial background mass distribution has been guessed to have an exponential shape:

$$g_B(m) = \frac{\mu \exp(-\mu m)}{\exp(-\mu m_{min}) - \exp(-\mu m_{max})}, \tag{6.13}$$

being $m_{min} = 4.9\,GeV/c^2$ and $m_{max} = 5.7\,GeV/c^2$ the minimum and maximum mass values accepted by the trigger. This mass range is assumed to be available for the $B^0_{(s)} \to h^+h^-$ offline analyses. The value of the variable μ has been determined by using the full simulation, fitting the previous expression to triggered $b\bar{b}$ data, with slightly reduced offline-selection cuts. Also these parameters are left free to be determined from the likelihood fit.

All values of the parameters present in the previous expressions can be found in [10]. For the reader's convenience, they will be summarized in Table 1.

As it can be argued from what said above, only combinatorial background is included in this study. Indeed, specific decay backgrounds are expected to contribute to the total decay rate, even if they are estimated to be kept relatively small by the offline-selection, as discussed in [10]. These backgrounds can be particularly dangerous, since they may exhibit their own CP asymmetries, thus biasing the CP measurements of the decays under consideration. For the moment they have been ignored, but in future developments of this work they must be included.

[3]In principle, in real data analysis the knowledge of the event-by-event estimated error on t_{rec} can be used as an additional information to be included in the fit. This has not been done in the studies presented in this note. This improvement can be included in future developments of this work.

The complete list of parameters left free during the likelihood maximization (17 in total) is:

- $\mathrm{Re}(\lambda_{\pi\pi})$ and $\mathrm{Im}(\lambda_{\pi\pi})$;
- the charge asymmetry for the $B^0 \to K^+\pi^-$ decay, $\mathcal{A}_{K\pi}$;
- the number of tagged $B^0 \to \pi^+\pi^-$ decays $n_{\pi\pi}$;
- the number of tagged $B^0 \to K^+\pi^-$ decays $n_{K\pi} = n_{K^+\pi^-} + n_{K^-\pi^+}$;
- two parameters for the $B^0 \to \pi^+\pi^-$ background rate, $\delta(B^0 \to \pi^+\pi^-)$ and $\eta(B^0 \to \pi^+\pi^-)$;
- two parameters for the $B^0 \to K^+\pi^-$ background rate, $\delta(B^0 \to K^+\pi^-)$ and $\eta(B^0 \to K^+\pi^-)$;
- one parameter for the $B^0 \to \pi^+\pi^-$ background mass distribution, $\mu(B^0 \to \pi^+\pi^-)$;
- one parameter for the $B^0 \to K^+\pi^-$ background mass distribution, $\mu(B^0 \to K^+\pi^-)$;
- two parameters for the B^0 mass distribution, $\overline{m}$ and σ_m;
- the wrong tagging fraction ω;
- the difference of the decay widths of the B^0 mass eigenstates $\Delta\Gamma_d$;
- the mass difference of the B^0 mass eigenstates Δm_d;
- the average decay width of the B^0 mass eigenstates Γ_d.

Note that the number of tagged background events for each channel is not varied as a free parameter, since it can be obtained as the difference between the total number of tagged events, which is observed and thus a constant, and the number of tagged signal events, which is left free in the fit.

6.2 Case of $B_s^0 \to K^+K^-$ and $B_s^0 \to \pi^+K^-$

The combined fit of $B_s^0 \to K^+K^-$ and $B_s^0 \to \pi^+K^-$ follows the same lines as $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$. The combined extended likelihood function takes the form:

$$\log \mathcal{L}^{ext}_{KK,\pi K} = \sum_{i=1}^{N_{KK}} \log P'_{KK}(t_i,\, m_i) + \sum_{j=1}^{\overline{N}_{KK}} \log \overline{P}'_{KK}(t_j,\, m_j) + \sum_{i=1}^{N_{\pi^+K^-}} \log P'_{\pi^+K^-}(t_i,\, m_i) + \\ + \sum_{j=1}^{\overline{N}_{\pi^+K^-}} \log \overline{P}'_{\pi^+K^-}(t_j,\, m_j) + \sum_{i=1}^{N_{\pi^-K^+}} \log P'_{\pi^-K^+}(t_i,\, m_i) + \sum_{j=1}^{\overline{N}_{\pi^-K^+}} \log \overline{P}'_{\pi^-K^+}(t_j,\, m_j). \tag{6.14}$$

Prior information on $\Delta\Gamma_s$, Δm_s and Γ_s are included in the fit by multiplying the likelihood by Gaussian functions:

$$\log \mathcal{L}^{ext\,\prime}_{KK,\pi K} = \log \mathcal{L}^{ext}_{KK,\pi K} - \frac{(\Delta\Gamma_s - \overline{\Delta\Gamma}_s)^2}{2\sigma^2_{\Delta\Gamma_s}} - \frac{(\Delta m_s - \overline{\Delta m}_s)^2}{2\sigma^2_{\Delta m_s}} - \frac{(\Gamma_s - \overline{\Gamma}_s)^2}{2\sigma^2_{\Gamma_s}}. \qquad (6.15)$$

For $\sigma_{\Delta\Gamma_s}$ and σ_{Γ_s}, the results of the LHCb $B^0_s \to J/\psi\phi$ analysis have been used [15], while for $\sigma_{\Delta m_s}$ the results of the LHCb $B^0_s \to D_s h$ analysis have been used [16].

The central values of the parameters determined from the full GEANT simulation for the $B^0_s \to K^+K^-$ and $B^0_s \to \pi^+K^-$ [10], analogous to the $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$ ones, are summarized in Table 1.

The complete list of parameters left free during the likelihood maximization (always 17 in total) is:

- $\mathrm{Re}(\lambda_{KK})$ and $\mathrm{Im}(\lambda_{KK})$;
- the charge asymmetry for the $B^0_s \to \pi^+K^-$ decay, $\mathcal{A}_{\pi K}$;
- the number of tagged $B^0_s \to K^+K^-$ decays n_{KK};
- the number of tagged $B^0_s \to \pi^+K^-$ decays $n_{\pi K} = n_{\pi^+K^-} + n_{\pi^-K^+}$;
- two parameters for the $B^0_s \to K^+K^-$ background rate, $\delta(B^0_s \to K^+K^-)$ and $\eta(B^0_s \to K^+K^-)$;
- two parameters for the $B^0_s \to \pi^+K^-$ background rate, $\delta(B^0_s \to \pi^+K^-)$ and $\eta(B^0_s \to \pi^+K^-)$;
- one parameter for the $B^0_s \to K^+K^-$ background mass distribution, $\mu(B^0_s \to K^+K^-)$;
- one parameter for the $B^0_s \to \pi^+K^-$ background mass distribution, $\mu(B^0_s \to \pi^+K^-)$;
- two parameters for the B^0_s mass distribution, $\overline{m}$ and σ_m;
- the wrong tagging fraction ω;
- the difference of the decay widths of the B^0_s mass eigenstates $\Delta\Gamma_s$;
- the mass difference of the B^0_s mass eigenstates Δm_s;
- the average decay width of the B^0_s mass eigenstates Γ_s.

	$B^0 \to \pi^+\pi^-$	$B^0 \to K^+\pi^-$	$B_s^0 \to K^+K^-$	$B_s^0 \to \pi^+K^-$
Annual Signal Yield	$26\,k$	$135\,k$	$37\,k$	$5.3\,k$
$b\bar{b}$ B/S ratio	0.42	0.16	0.31	0.67
ϵ_{tag} [in %]	41.8	41.8	49.8	49.8
ω [in %]	34.9	34.9	33.0	33.0
β	-2.35	-2.28	-1.76	-2.52
f [in %]	86	88	85	80
$\sigma_{\Delta t_1}$ [fs]	32.6	32.8	34.4	29.8
$\sigma_{\Delta t_2}$ [fs]	59.9	60.6	67.3	54.1
δ	-3.09	-2.16	-1.03	-2.19
η [ps^{-1}]	1.02	1.15	0.89	0.87
$\overline{m}$ [GeV/c^2]	5.279	5.279	5.369	5.369
σ_m [MeV/c^2]	17.5	17.5	17.5	17.5
μ [$(GeV/c^2)^{-1}$]	1.71	0.54	1.21	0.44
Γ [ps^{-1}]	1/1.54	1/1.54	1/1.46	1/1.46

Table 1: Values of the parameters used as inputs to the fast Monte Carlo simulation. They are obtained from the respective $B^0_{(s)} \to h^+h^-$ analyses with the full GEANT simulation [10]. The $b\bar{b}$ B/S ratios quoted in [10] are the 90% CL upper limits calculated before the trigger, while the values reported in this table and used in the fast simulation are the respective central values. ϵ_{tag} and ω are the signal tagging efficiency and the mistag probability respectively. They are assumed to be identical for $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$ ($B_s^0 \to K^+K^-$ and $B_s^0 \to \pi^+K^-$) decays (see text). The meaning of the other parameters is described in the text.

7 Fast Monte Carlo simulation

In order to generate the data samples required to study the sensitivity on CP-violating observables for the $B^0_{(s)} \to h^+h^-$ decay channels, a fast Monte Carlo simulation program has been written. The program first generates proper time and mass values for triggered and offline-selected untagged signal and combinatorial background events, according to the annual signal yields and background-to-signal (B/S) ratios, as well as proper time and mass distributions, obtained from the full GEANT simulation.

In the signal event generation, the effect of CP violation is obviously included, while no CP violation is assumed for the combinatorial background. The program also simulates the effect of the trigger and offline-selection acceptance as function of the proper time, the effect of the proper time resolution (smearing of the proper time by means of a double Gaussian), and the tagging procedure.

All the input parameters used in the fast Monte Carlo generation are reported in Table 1, and are taken from [10]. Signal yields and B/S ratios reported in the table are calculated for events with a mass value inside a window $\pm\delta m$ around the nominal $B^0_{(s)}$ mass, for each

$B^0_{(s)} \to h^+h^-$ analysis, as described in [10]. In the fast Monte Carlo simulation, the mass is generated from a value $m_{min} = 4.9\,GeV/c^2$ up to a value $m_{max} = 5.7\,GeV/c^2$, in order to include the mass sideband spectrum and simulate the extraction of the combinatorial background properties from data, as discussed in the previous section. Then, the generation of signal and background events is performed by reproducing the signal yields and B/S ratios inside the respective $\pm\delta m$ mass windows.

The tagging efficiencies and mistag probabilities for $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$ ($B^0_s \to K^+K^-$ and $B^0_s \to \pi^+K^-$) decays are assumed to be indentical. From the full GEANT simulation, they result to be compatible within statistical uncertainties [2, 10, 17]. The assumed common values are the ones from $B^0 \to \pi^+\pi^-$ ($B^0_s \to K^+K^-$), which have smaller statistical errors due to the larger Monte Carlo statistics available in the full simulation. For the tagging of combinatorial background events, the same tagging efficiency of the signal is conservatively assumed, while the mistag probability is assumed to be 50% (i.e., one combinatorial background event can be randomly tagged as B or $\overline{B}$ with same probability).

Table 1 does not include the values of $\Delta\Gamma$, Δm, $\mathcal{A}^{dir}$, $\mathcal{A}^{mix}$ and $\mathcal{A}_{K\pi}$ ($\mathcal{A}_{\pi K}$). Due to the fact that these values are unknown (except $\Delta\Gamma_d$ which has been set to zero in the generation), a scan over reasonable ranges (i.e., around the Standard Model expectations) has been performed for them, by running different sets of fast Monte Carlo generations, and then making likelihood fits on each independent data set (see next section).

Instead of scanning directly the values of $\mathcal{A}^{dir}$, $\mathcal{A}^{mix}$ and $\mathcal{A}_{K\pi}$ ($\mathcal{A}_{\pi K}$), we perform scans on the parameters d, ϑ, γ and ϕ_s defined in Section 3 ($\phi_d = 2\beta$ has been set to the central value of the current world average, and not varied). The corresponding values of $\mathcal{A}^{dir}$, $\mathcal{A}^{mix}$ and $\mathcal{A}_{K\pi}$ ($\mathcal{A}_{\pi K}$) are then calculated from equations (3.4), (3.5), (3.7), (3.8), (3.15) and (3.16), assuming the validity of U-spin symmetry.

Table 2 summarizes the scanned values of $\Delta\Gamma_s/\Gamma_s$, Δm_s, d, ϑ, γ and ϕ_s. The parameters are varied one at a time, with the other ones kept to their assumed nominal values, which are also indicated in Table 2. This results in a total of 13 independent sets of values for the $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$ scan, and 21 for the $B^0_s \to K^+K^-$ and $B^0_s \to \pi^+K^-$ scan.

For each set of values, 375 fast Monte Carlo experiments and subsequent likelihood fits have been performed. Repeated experiments are simulated in order to assess the average values of the sensitivities on the physics parameters of interest, and also to produce pull distributions for the fitted parameters to demonstrate the reliability and consistency of the event generation and fit procedure, as it will be shown in the next section.

8 Likelihood fit results

In this section, we present the results of the likelihood fits used to extract the physics parameters from data generated by the fast Monte Carlo simulation.

$\Delta\Gamma_s/\Gamma_s$	0	(0.1)	0.2			
$\Delta m_s\,[ps^{-1}]$	15	(20)	25	30		
d	0.1	0.2	(0.3)	0.4		
ϑ	120^o	140^o	(160^o)	180^o	200^o	
γ	55^o	(65^o)	75^o	85^o	95^o	105^o
ϕ_s [rad]	0	(-0.04)	-0.1	-0.2		

Table 2: Scanned values of the unknown physics parameters. The physics parameters are varied one at a time, with the other ones kept to their assumed nominal values. The nominal values are those enclosed in parentheses.

8.1 Propagation from Re(λ) and Im(λ) to $\mathcal{A}^{dir}$ and $\mathcal{A}^{mix}$

As discussed in Section 6, the CP-violating observables returned by the fit are Re(λ), Im(λ) for the $B^0 \to \pi^+\pi^-$ ($B_s^0 \to K^+K^-$) and the charge symmetry for $B^0 \to K^+\pi^-$ ($B_s^0 \to \pi^+K^-$). In order to obtain the best estimates of the observables $\mathcal{A}^{dir}$ and $\mathcal{A}^{mix}$, with their uncertainties and correlation, a propagation from Re(λ) and Im(λ) to $\mathcal{A}^{dir}$ and $\mathcal{A}^{mix}$, each other related by equations (2.18) and (2.19), is necessary.

The analytic procedure consists in calculating a new 17×17 covariance matrix (17 is the number of parameters in the fit), where Re(λ) and Im(λ) are substituted with $\mathcal{A}^{dir}$ and $\mathcal{A}^{mix}$. However, an easier way to do it, that is the one we follow, is to perform a Monte Carlo propagation.

The Monte Carlo propagation is realized by generating a sample of 17-tuples of random numbers from a multi-variate Gaussian p.d.f., using the best estimates of the parameters and the covariance matrix returned by the fit. Then, the values of Re(λ) and Im(λ) in each 17-tuple of random numbers are used to compute $\mathcal{A}^{dir}$ and $\mathcal{A}^{mix}$ through (2.18) and (2.19), and the uncertainties and correlation of $\mathcal{A}^{dir}$ and $\mathcal{A}^{mix}$ are estimated from the full sample. In order not to introduce additional uncertainties in the propagation, a large sample of Monte Carlo generations has been performed ($n = 10^6$)[4]. The distortion from gaussianity of the p.d.f. for $\mathcal{A}^{dir}$ and $\mathcal{A}^{mix}$ has been checked to be completely negligible.

8.2 Sensitivity on CP-violating observables

The sensitivity on $\mathcal{A}^{dir}_{\pi\pi}$, $\mathcal{A}^{mix}_{\pi\pi}$ and charge asymmetry $\mathcal{A}_{K\pi}$ from $B^0 \to K^+\pi^-$ is explored for different values of the physics parameters d, ϑ and γ. For $\mathcal{A}^{dir}_{KK}$, $\mathcal{A}^{mix}_{KK}$ and charge asymmetry $\mathcal{A}_{\pi K}$ from $B_s^0 \to \pi^+K^-$ an additional scan on the values of $\Delta\Gamma_s/\Gamma_s$, Δm_s and ϕ_s has been done. These parameters are varied one at a time, with the other ones kept to their assumed nominal values. The results are summarized in the Tables 3 and 4.

Figure 2 shows the pull distributions from 375 experiments simulated with nominal

[4]The Monte Carlo generation of the multi-variate Gaussian distributed random numbers has been realized by employing the function CORGEN from the CERNLIB library [19]. The generation of 10^6 17-tuples requires just few seconds of (a modern CPU) time.

d	0.1	0.2	(0.3)	0.4
$\sigma(\mathcal{A}_{\pi\pi}^{dir})$	0.063	0.063	0.064	0.064
$\sigma(\mathcal{A}_{\pi\pi}^{mix})$	0.053	0.054	0.055	0.056
$\rho(\mathcal{A}_{\pi\pi}^{dir}, \mathcal{A}_{\pi\pi}^{mix})$	-0.43	-0.43	-0.43	-0.43
$\sigma(\mathcal{A}_{K\pi})$	0.0035	0.0035	0.0035	0.0035

ϑ	120^o	140^o	(160^o)	180^o	200^o
$\sigma(\mathcal{A}_{\pi\pi}^{dir})$	0.065	0.064	0.064	0.063	0.064
$\sigma(\mathcal{A}_{\pi\pi}^{mix})$	0.055	0.055	0.055	0.054	0.054
$\rho(\mathcal{A}_{\pi\pi}^{dir}, \mathcal{A}_{\pi\pi}^{mix})$	-0.45	-0.45	-0.43	-0.42	-0.41
$\sigma(\mathcal{A}_{K\pi})$	0.0035	0.0035	0.0035	0.0035	0.0035

γ	55^o	(65^o)	75^o	85^o	95^o	105^o
$\sigma(\mathcal{A}_{\pi\pi}^{dir})$	0.064	0.064	0.064	0.064	0.064	0.064
$\sigma(\mathcal{A}_{\pi\pi}^{mix})$	0.057	0.055	0.053	0.052	0.053	0.056
$\rho(\mathcal{A}_{\pi\pi}^{dir}, \mathcal{A}_{\pi\pi}^{mix})$	-0.41	-0.43	-0.44	-0.42	-0.40	-0.37
$\sigma(\mathcal{A}_{K\pi})$	0.0035	0.0035	0.0035	0.0035	0.0035	0.0035

Table 3: Average statistical uncertainties and correlation on $\mathcal{A}_{\pi\pi}^{dir}$ and $\mathcal{A}_{\pi\pi}^{mix}$ from $B^0 \to \pi^+\pi^-$, and average uncertainty on the charge asymmetry $\mathcal{A}_{K\pi}$ from $B^0 \to K^+\pi^-$, corresponding to one year of data, depending on the true values of the physics parameters d, ϑ and γ. The physics parameters are varied one at a time, with the other ones kept to their assumed nominal values. The nominal values are those enclosed in parentheses. The correlation between $\mathcal{A}_{K\pi}$ and $\mathcal{A}_{\pi\pi}^{dir}$ ($\mathcal{A}_{\pi\pi}^{mix}$) is negligible, and is not reported in the table.

$\Delta\Gamma_s/\Gamma_s$	0	(0.1)	0.2
$\sigma(\mathcal{A}_{KK}^{dir})$	0.054	0.051	0.047
$\sigma(\mathcal{A}_{KK}^{mix})$	0.070	0.067	0.061
$\rho(\mathcal{A}_{KK}^{dir}, \mathcal{A}_{KK}^{mix})$	-0.20	-0.22	-0.21
$\sigma(\mathcal{A}_{\pi K})$	0.023	0.024	0.024

Δm_s $[ps^{-1}]$	15	(20)	25	30
$\sigma(\mathcal{A}_{KK}^{dir})$	0.043	0.051	0.060	0.076
$\sigma(\mathcal{A}_{KK}^{mix})$	0.058	0.067	0.077	0.093
$\rho(\mathcal{A}_{KK}^{dir}, \mathcal{A}_{KK}^{mix})$	-0.24	-0.22	-0.20	-0.17
$\sigma(\mathcal{A}_{\pi K})$	0.024	0.024	0.024	0.024

d	0.1	0.2	(0.3)	0.4
$\sigma(\mathcal{A}_{KK}^{dir})$	0.060	0.056	0.051	0.049
$\sigma(\mathcal{A}_{KK}^{mix})$	0.065	0.079	0.067	0.061
$\rho(\mathcal{A}_{KK}^{dir}, \mathcal{A}_{KK}^{mix})$	-0.18	-0.33	-0.22	-0.15
$\sigma(\mathcal{A}_{\pi K})$	0.024	0.024	0.024	0.024

ϑ	120^o	140^o	(160^o)	180^o	200^o
$\sigma(\mathcal{A}_{KK}^{dir})$	0.061	0.055	0.051	0.049	0.050
$\sigma(\mathcal{A}_{KK}^{mix})$	0.053	0.059	0.067	0.070	0.066
$\rho(\mathcal{A}_{KK}^{dir}, \mathcal{A}_{KK}^{mix})$	-0.27	-0.30	-0.22	0	0.22
$\sigma(\mathcal{A}_{\pi K})$	0.023	0.023	0.024	0.025	0.024

γ	55^o	(65^o)	75^o	85^o	95^o	105^o
$\sigma(\mathcal{A}_{KK}^{dir})$	0.050	0.051	0.051	0.050	0.050	0.050
$\sigma(\mathcal{A}_{KK}^{mix})$	0.065	0.067	0.067	0.067	0.067	0.066
$\rho(\mathcal{A}_{KK}^{dir}, \mathcal{A}_{KK}^{mix})$	-0.20	-0.22	-0.22	-0.20	-0.21	-0.18
$\sigma(\mathcal{A}_{\pi K})$	0.023	0.024	0.023	0.023	0.023	0.023

ϕ_s [rad]	0	(-0.04)	-0.1	-0.2
$\sigma(\mathcal{A}_{KK}^{dir})$	0.051	0.051	0.051	0.052
$\sigma(\mathcal{A}_{KK}^{mix})$	0.064	0.067	0.072	0.079
$\rho(\mathcal{A}_{KK}^{dir}, \mathcal{A}_{KK}^{mix})$	-0.21	-0.22	-0.22	-0.23
$\sigma(\mathcal{A}_{\pi K})$	0.023	0.024	0.023	0.024

Table 4: Average statistical uncertainties and correlation on $\mathcal{A}_{KK}^{dir}$ and $\mathcal{A}_{KK}^{mix}$ from $B_s^0 \to K^+K^-$, and average uncertainty on the charge asymmetry $\mathcal{A}_{\pi K}$ from $B_s^0 \to \pi^+K^-$, corresponding to one year of data, depending on the true values of the physics parameters $\Delta\Gamma_s/\Gamma_s$, Δm_s, d, ϑ, γ and ϕ_s. The physics parameters are varied one at a time, with the other ones kept to their assumed nominal values. The nominal values are those enclosed in parentheses. The correlation between $\mathcal{A}_{\pi K}$ and $\mathcal{A}_{KK}^{dir}$ ($\mathcal{A}_{KK}^{mix}$) is negligible, and is not reported in the table.

values of the input physics parameters, for all the 17 parameters varied in the $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$ combined fit (Figure 3 for the $B_s^0 \to K^+K^-$ and $B_s^0 \to \pi^+K^-$ combined fit). The pulls are compatible with normal distributions centered in zero with unitary variance. This condition has been checked to be fulfilled for all the sets of $(13+21) \times 375$ experiments simulated in the scan.

As an example, Figure 4 shows the mass distributions generated by the fast Monte Carlo simulation, with a projection of the likelihood fit superimposed.

8.3 Discussion

The resulting resolutions on ($\mathcal{A}_{\pi\pi}^{dir}$, $\mathcal{A}_{\pi\pi}^{mix}$) and ($\mathcal{A}_{KK}^{dir}$, $\mathcal{A}_{KK}^{mix}$) are quite similar, even if the $B_s^0 \to K^+K^-$ decay is characterized by larger signal yield, larger tagging efficiency and smaller mistag probability, as well as smaller background, with respect to the $B^0 \to \pi^+\pi^-$ decay. This is due to the fact that the mistag probability is measured with smaller precision for the $B_s^0 \to K^+K^-$, due to the much smaller yield of $B_s^0 \to \pi^+K^-$ with respect to the yield of $B^0 \to K^+\pi^-$. The mistag probability is measured with an average relative resolution of 1.0% for the $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$, while the average relative resolution for $B_s^0 \to K^+K^-$ and $B_s^0 \to \pi^+K^-$ goes from 6.1% to 9.1%, depending on the true value of Δm_s.

Within the considered parameter ranges, the dependence of the resolutions and correlation of $\mathcal{A}_{\pi\pi}^{dir}$ and $\mathcal{A}_{\pi\pi}^{mix}$ on their true values (i.e., on d, ϑ and γ to which they are related) is negligible. The same thing does not hold for $\mathcal{A}_{KK}^{dir}$ and $\mathcal{A}_{KK}^{mix}$. This is still due to the different resolution on the mistag probability. A simplified semi-quantitative explanation is given in the following.

The experimental CP asymmetry for $B^0 \to \pi^+\pi^-$ and $B_s^0 \to K^+K^-$, in the case of no background, perfect proper time resolution and negligible $\Delta\Gamma$ for simplicity, is:

$$\mathcal{A}^{CP}(t) = (1-2\omega)\left(\mathcal{A}^{dir}\cos\Delta m\, t + \mathcal{A}^{mix}\sin\Delta m\, t\right). \tag{8.1}$$

Let us suppose to have infinite statistics of $B^0 \to \pi^+\pi^-$ and $B_s^0 \to K^+K^-$ data at our disposal. Even in this case, the problem of extracting ω, $\mathcal{A}^{dir}$ and $\mathcal{A}^{mix}$ simultaneously from $B^0 \to \pi^+\pi^-$ and $B_s^0 \to K^+K^-$ data only, is undetermined. By combining $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$ ($B_s^0 \to K^+K^-$ and $B_s^0 \to \pi^+K^-$) data, as already discussed, ω can be extracted from data. Now suppose that the measurement of ω for one experiment gives the best estimate:

$$\hat{\omega} = \omega_{true} + \delta\omega, \tag{8.2}$$

where ω_{true} is the true value of the mistag probability and $\delta\omega$ is the difference between the measured and true values of ω. Since we have assumed to have infinite statistics for $B^0 \to \pi^+\pi^-$ and $B_s^0 \to K^+K^-$, we can write:

$$(1-2\hat{\omega})\hat{\mathcal{A}}^{dir} = (1-2\omega_{true})\mathcal{A}_{true}^{dir} \tag{8.3}$$

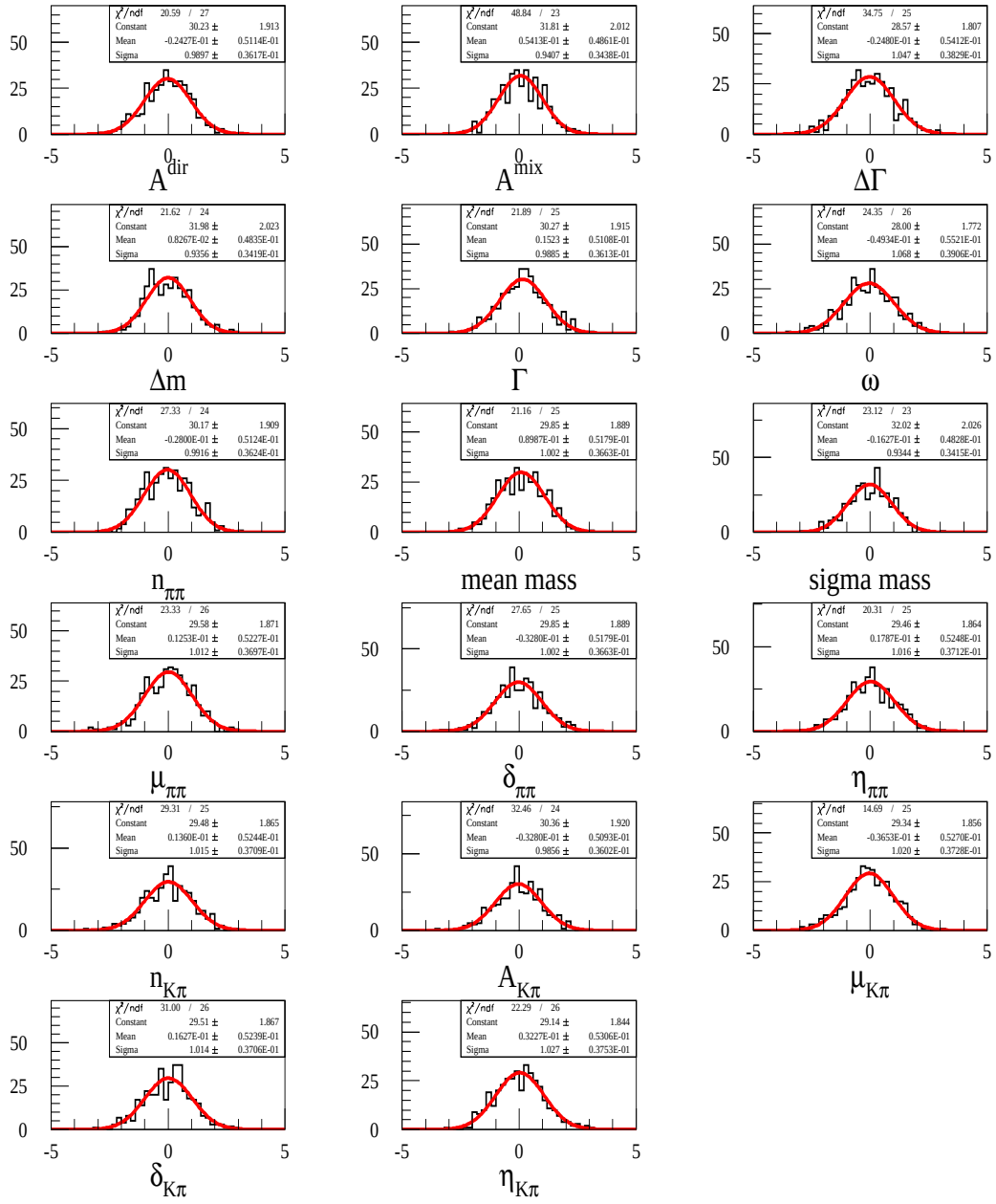

Figure 2: Pull distributions from 375 experiments simulated with nominal values of the input physics parameters, for all the 17 parameters varied in the $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$ combined fit.

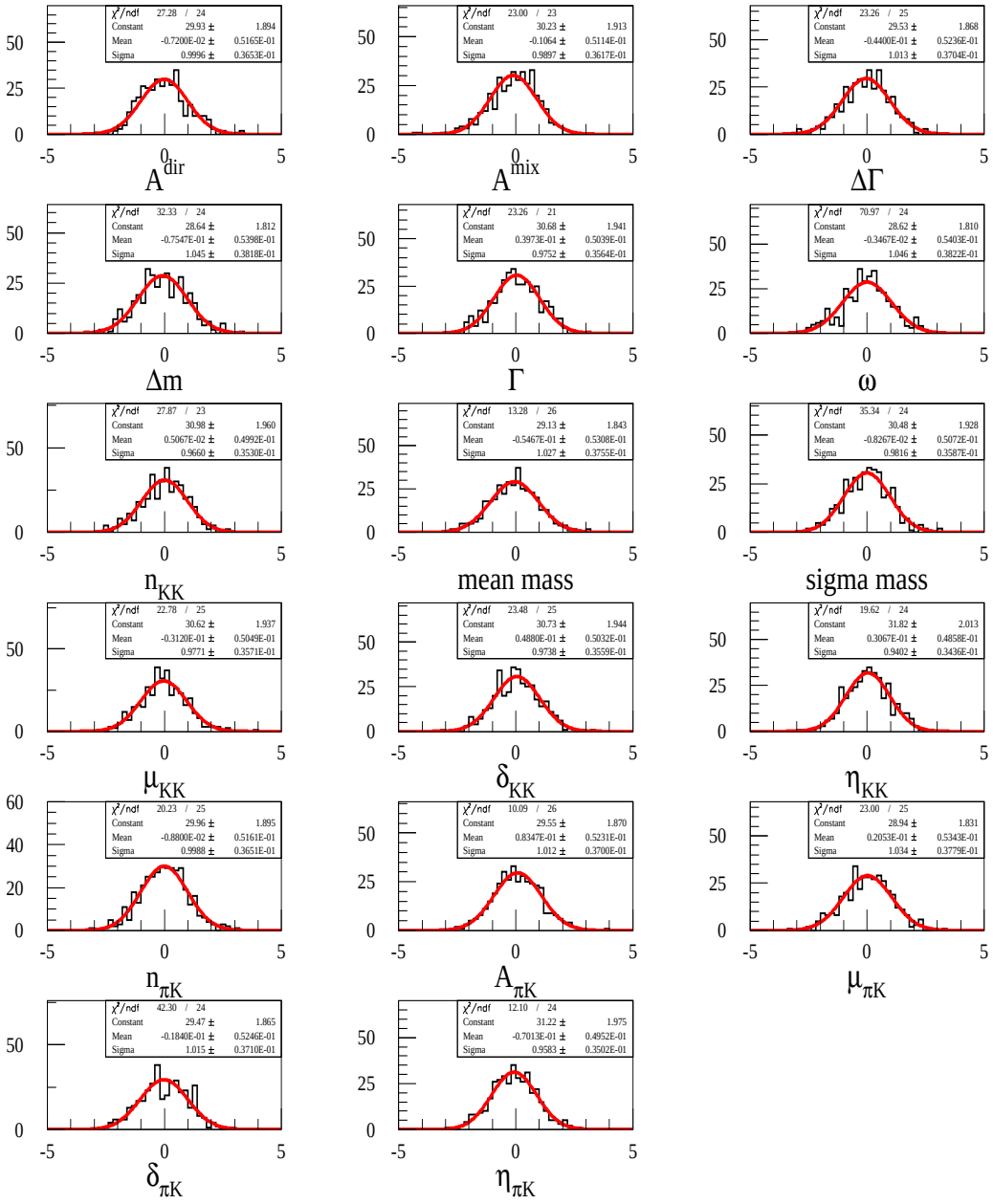

Figure 3: Pull distributions from 375 experiments simulated with nominal values of the input physics parameters, for all the 17 parameters varied in the $B_s^0 \to K^+K^-$ and $B_s^0 \to \pi^+K^-$ combined fit.

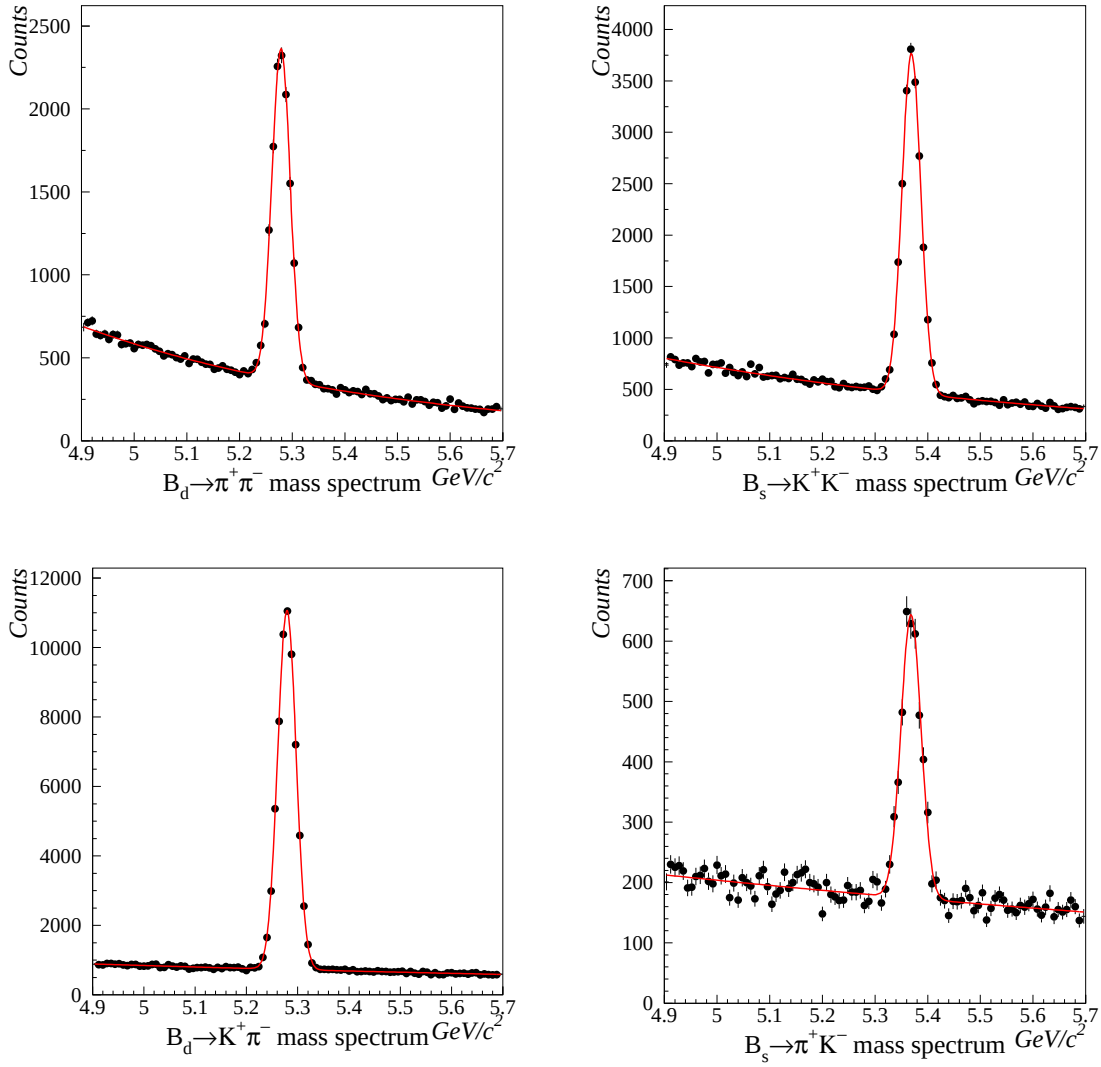

Figure 4: Mass distributions for each $B^0_{(s)} \to h^+h^-$ generated by the fast Monte Carlo simulation for one year of data, with a projection of the likelihood fit superimposed.

and

$$(1-2\hat{\omega})\hat{\mathcal{A}}^{mix} = (1-2\omega_{true})\mathcal{A}^{mix}_{true}, \tag{8.4}$$

where $\hat{\mathcal{A}}^{dir}$ and $\hat{\mathcal{A}}^{mix}$ are the best estimates of $\mathcal{A}^{dir}$ and $\mathcal{A}^{mix}$, while $\mathcal{A}^{dir}_{true}$ and $\mathcal{A}^{mix}_{true}$ are their true values. By inserting (8.2) in (8.3) and (8.4), we have:

$$\hat{\mathcal{A}}^{dir} = \frac{1-2\omega_{true}}{1-2\omega_{true}-2\delta\omega}\mathcal{A}^{dir}_{true} \simeq \left(1 + \frac{2\delta\omega}{1-2\omega_{true}} + \mathcal{O}\left[\left(\frac{2\delta\omega}{1-2\omega_{true}}\right)^2\right]\right)\mathcal{A}^{dir}_{true} \tag{8.5}$$

and

$$\hat{\mathcal{A}}^{mix} = \frac{(1-2\omega_{true})}{(1-2\omega_{true}-2\delta\omega)}\mathcal{A}^{mix}_{true} \simeq \left(1 + \frac{2\delta\omega}{1-2\omega_{true}} + \mathcal{O}\left[\left(\frac{2\delta\omega}{1-2\omega_{true}}\right)^2\right]\right)\mathcal{A}^{mix}_{true}, \tag{8.6}$$

Hence, for small $\delta\omega$, we finally obtain:

$$\delta\mathcal{A}^{dir} = \hat{\mathcal{A}}^{dir} - \mathcal{A}^{dir}_{true} \simeq \frac{2\delta\omega}{1-2\omega_{true}}\mathcal{A}^{dir}_{true} \tag{8.7}$$

and

$$\delta\mathcal{A}^{mix} = \hat{\mathcal{A}}^{mix} - \mathcal{A}^{mix}_{true} \simeq \frac{2\delta\omega}{1-2\omega_{true}}\mathcal{A}^{mix}_{true}, \tag{8.8}$$

i.e., the absolute errors on $\mathcal{A}^{dir}$ and $\mathcal{A}^{mix}$ depend on their true values, and are proportional to $\delta\omega$. In the limit of $\delta\omega$ tending to zero, this dependence on the true values is not present.

One can note that:

$$\delta\mathcal{A}^{dir}\delta\mathcal{A}^{mix} \simeq \left(\frac{2\delta\omega}{1-2\omega_{true}}\right)^2 \mathcal{A}^{dir}_{true}\mathcal{A}^{mix}_{true}, \tag{8.9}$$

suggesting that the covariance depends on the product $\mathcal{A}^{dir}_{true}$ and $\mathcal{A}^{mix}_{true}$, and should be negative if $\mathcal{A}^{dir}_{true}$ and $\mathcal{A}^{mix}_{true}$ have opposite signs, negligible if $\mathcal{A}^{dir}_{true}$ or $\mathcal{A}^{mix}_{true}$ are close to zero, and positive if $\mathcal{A}^{dir}_{true}$ and $\mathcal{A}^{mix}_{true}$ have same signs. In fact, this is exactly what observed in $B^0_s \to K^+K^-$ and $B^0_s \to \pi^+K^-$ fits. In Table 4, in the row corresponding to the scan on ϑ, one can see that the correlation is negative for $\vartheta < 180^o$, null for $\vartheta = 180^o$ and positive for $\vartheta > 180^o$. It is possible to verify by using (3.7) and (3.8), with $d = d'$ and $\vartheta = \vartheta'$, that $\mathcal{A}^{dir}$ and $\mathcal{A}^{mix}$ have opposite signs in the first case, $\mathcal{A}^{dir}$ is zero in the second case, while in the third case $\mathcal{A}^{dir}$ and $\mathcal{A}^{mix}$ have same signs. The same thing does not apply to $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$ fits, where the main contribute to the correlation between $\mathcal{A}^{dir}$ and $\mathcal{A}^{mix}$ is not introduced through ω. This will be clear in the following.

In order to validate our initial assertion it is now sufficient to demonstrate that if ω is perfectly known (i.e., in case of infinite statistics for $B^0 \to K^+\pi^-$ and $B^0_s \to \pi^+K^-$),

the uncertainties and correlation on $\mathcal{A}^{dir}$ and $\mathcal{A}^{mix}$ are not dependent on their true values. This is true since the asymmetry is linear in $\mathcal{A}^{dir}$ and $\mathcal{A}^{mix}$. We demonstrate it in the simplified but conceptually equivalent case of a χ^2 fit of the binned asymmetry. The χ^2 is given by:

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{(1-2\omega_{true})(\mathcal{A}^{dir}\cos\Delta m\, t_i + \mathcal{A}^{mix}\sin\Delta m\, t_i) - \mathcal{A}_i^{CP}}{\sigma(\mathcal{A}_i^{CP})} \right)^2, \tag{8.10}$$

where the sum runs over all the N bins, t_i is the proper time of the i-th bin, $\mathcal{A}_i^{CP}$ is the observed asymmetry in the i-th bin, and $\sigma(\mathcal{A}_i^{CP})$ is its error.

The second order derivatives are:

$$\frac{\partial^2\chi^2}{\partial\mathcal{A}^{dir\,2}} = 2\sum_{i=1}^{N} \left(\frac{(1-2\omega_{true})\cos\Delta m\, t_i}{\sigma(\mathcal{A}_i^{CP})} \right)^2, \tag{8.11}$$

$$\frac{\partial^2\chi^2}{\partial\mathcal{A}^{mix\,2}} = 2\sum_{i=1}^{N} \left(\frac{(1-2\omega_{true})\sin\Delta m\, t_i}{\sigma(\mathcal{A}_i^{CP})} \right)^2 \tag{8.12}$$

and

$$\frac{\partial^2\chi^2}{\partial\mathcal{A}^{dir}\partial\mathcal{A}^{mix}} = 2\sum_{i=1}^{N} \left(\frac{1-2\omega_{true}}{\sigma(\mathcal{A}_i^{CP})} \right)^2 \cos\Delta m\, t_i \sin\Delta m\, t_i. \tag{8.13}$$

Hence, the covariance matrix, that is given by the inverse of the second derivative matrix, does not depend on the true values of $\mathcal{A}^{dir}$ and $\mathcal{A}^{mix}$, if no such dependence is present for $\sigma(\mathcal{A}_i^{CP})$. But this is true, since $\sigma(\mathcal{A}_i^{CP})$ depends essentially only on the statistics in each bin.

Furthermore, in the case of the $B_s^0 \to K^+K^-$, the second mixed derivative (8.13) cancels, since within a period of the fast B_s^0 oscillations (e.g., $2\pi/20\,ps$ for $\Delta m_s = 20\,ps^{-1}$), there is no sizable variation of statistics, and hence of $\sigma(\mathcal{A}_i^{CP})$. This leads to a diagonal covariance matrix, i.e., to null correlation. This means that a correlation between $\mathcal{A}^{dir}$ and $\mathcal{A}^{mix}$ for the $B_s^0 \to K^+K^-$ can be introduced only through ω, as previously discussed. The same fact does not hold for the $B^0 \to \pi^+\pi^-$, where the oscillations are slow (the period is $2\pi/0.5\,ps$), and so there is a sizable variation of statistics in one period, and then of $\sigma(\mathcal{A}_i^{CP})$. In this case, the main contribute to the correlation is not due to ω.

One further comment concerns the resolution on the charge asymmetry. First of all, the uncertainty on the charge asymmetry is much larger for the $B_s^0 \to \pi^+K^-$ with respect to the $B^0 \to K^+\pi^-$, due to the large difference of the yields. Secondly, since the charge asymmetry is given by:

$$\mathcal{A}_{f,\overline{f}} = \frac{n_{\overline{f}} - n_f}{n_{\overline{f}} + n_f}, \tag{8.14}$$

where $n_{\overline{f}}$ and n_f are the yields of $K^-\pi^+$ and $K^+\pi^-$ decays from $B^0 \to K^+\pi^-$ (π^-K^+ and π^+K^- decays from $B_s^0 \to \pi^+K^-$) respectively, its uncertainty depends only on the yields, and in particular its measurement in not correlated with those of $\mathcal{A}^{dir}$ and $\mathcal{A}^{mix}$.

Finally, since it will be useful in the following, we note that due to the presence of correlation the joint p.d.f. for $\mathcal{A}_{\pi\pi}^{dir}$ and $\mathcal{A}_{\pi\pi}^{mix}$ is best described by a bi-variate Gaussian $G_\pi(\mathcal{A}_{\pi\pi}^{dir}, \mathcal{A}_{\pi\pi}^{mix})$, and similarly for $\mathcal{A}_{KK}^{dir}$ and $\mathcal{A}_{KK}^{mix}$ by $G_K(\mathcal{A}_{KK}^{dir}, \mathcal{A}_{KK}^{mix})$.

9 Extraction of the CKM angle γ

Once the best estimates of the CP-violating observables $\mathcal{A}_{\pi\pi}^{dir}$, $\mathcal{A}_{\pi\pi}^{mix}$, $\mathcal{A}_{KK}^{dir}$ and $\mathcal{A}_{KK}^{mix}$ have been determined from the likelihood fit, together with their uncertainties and correlations, it is possible to determine the value of the CKM angle γ by assuming the validity of the U-spin symmetry (see Section 3).

In fact, if (3.13) and (3.14) hold, equations (3.4), (3.5), (3.7) and (3.8) relate $\mathcal{A}_{\pi\pi}^{dir}$, $\mathcal{A}_{\pi\pi}^{mix}$, $\mathcal{A}_{KK}^{dir}$ and $\mathcal{A}_{KK}^{mix}$ to d, ϑ, γ, ϕ_d and ϕ_s. If the values of ϕ_d and ϕ_s are determined elsewhere, e.g., from the $B^0 \to J/\psi K_S^0$ and $B_s^0 \to J/\psi\phi$ decays respectively, we have a system of four equations and three unknowns (d, ϑ and γ), which is then over-determined and can be solved to extract the values of d, ϑ and γ. This problem consists in propagating the experimental joint p.d.f. for $\mathcal{A}_{\pi\pi}^{dir}$, $\mathcal{A}_{\pi\pi}^{mix}$, $\mathcal{A}_{KK}^{dir}$, $\mathcal{A}_{KK}^{mix}$, ϕ_d and ϕ_s, to a joint p.d.f. for d, ϑ and γ, by making use of the theoretical relations between these parameters (in the following, these theoretical relations will be called "constraints"). For this purpose, we adopt a Bayesian approach.

9.1 Bayesian determination of γ

By using the Bayes' theorem one can write:

$$f\left(d,\,\vartheta,\,\gamma,\,\{\mathcal{O}_{CP}\},\,\phi_d,\,\phi_s \mid \{\hat{\mathcal{O}}_{CP}\}\right) \propto \; f\left(\{\hat{\mathcal{O}}_{CP}\} \mid d,\,\vartheta,\,\gamma,\,\{\mathcal{O}_{CP}\},\,\phi_d,\,\phi_s\right) \times \\ \times f\left(d,\,\vartheta,\,\gamma,\,\{\mathcal{O}_{CP}\},\,\phi_d,\,\phi_s\right) \qquad (9.1)$$

where the generic notation $f(\cdot)$ is used to indicate p.d.f.s. The symbol $\{\mathcal{O}_{CP}\}$ stands for the 4-tuple of CP-violating observables ($\mathcal{A}_{\pi\pi}^{dir}$, $\mathcal{A}_{\pi\pi}^{mix}$, $\mathcal{A}_{KK}^{dir}$, $\mathcal{A}_{KK}^{mix}$), while $\{\hat{\mathcal{O}}_{CP}\}$ stands for their respective experimental best estimates ($\hat{\mathcal{A}}_{\pi\pi}^{dir}$, $\hat{\mathcal{A}}_{\pi\pi}^{mix}$, $\hat{\mathcal{A}}_{KK}^{dir}$, $\hat{\mathcal{A}}_{KK}^{mix}$). Since $\{\hat{\mathcal{O}}_{CP}\}$ depends on d, ϑ, γ, ϕ_d and ϕ_s only via $\{\mathcal{O}_{CP}\}$, the following relation holds:

$$f\left(\{\hat{\mathcal{O}}_{CP}\} \mid d,\,\vartheta,\,\gamma,\,\{\mathcal{O}_{CP}\},\,\phi_d,\,\phi_s\right) = f\left(\{\hat{\mathcal{O}}_{CP}\} \mid \{\mathcal{O}_{CP}\}\right). \qquad (9.2)$$

By using simple rules of probability, we can also write:

$$f\left(d,\,\vartheta,\,\gamma,\,\{\mathcal{O}_{CP}\},\,\phi_d,\,\phi_s\right) = f\left(\{\mathcal{O}_{CP}\} \mid d,\,\vartheta,\,\gamma,\,\phi_d,\,\phi_s\right) \times f\left(d,\,\vartheta,\,\gamma,\,\phi_d,\,\phi_s\right). \qquad (9.3)$$

In the Standard Model and under the assumption of U-spin symmetry (i.e., $d = d'$, $\vartheta = \vartheta'$), the CP-violating observables $\{\mathcal{O}_{CP}\}$ are exactly related to d, ϑ, γ, ϕ_d and ϕ_s by the constraints (3.4), (3.5), (3.7) and (3.8). Hence, one can write:

$$
\begin{aligned}
f\left(\{\mathcal{O}_{CP}\} \mid d,\, \vartheta,\, \gamma,\, \phi_d,\, \phi_s\right) = & \\
= \delta\left(\mathcal{A}^{dir}_{\pi\pi} - \mathcal{A}^{dir}_{\pi\pi}(d,\, \vartheta,\, \gamma)\right) \times & \quad \delta\left(\mathcal{A}^{mix}_{\pi\pi} - \mathcal{A}^{mix}_{\pi\pi}(d,\, \vartheta,\, \gamma,\, \phi_d)\right) \times \\
\times\ \delta\left(\mathcal{A}^{dir}_{KK} - \mathcal{A}^{dir}_{KK}(d,\, \vartheta,\, \gamma)\right) \times & \quad \delta\left(\mathcal{A}^{mix}_{KK} - \mathcal{A}^{mix}_{KK}(d,\, \vartheta,\, \gamma,\, \phi_s)\right), \qquad (9.4)
\end{aligned}
$$

where the symbol $\delta(\cdot)$ stands for the generalized Dirac delta function.

Equation (9.1) can then be re-written as:

$$
\begin{aligned}
f\left(d,\, \vartheta,\, \gamma,\, \{\mathcal{O}_{CP}\},\, \phi_d,\, \phi_s \mid \{\hat{\mathcal{O}}_{CP}\}\right) & \propto f\left(\{\hat{\mathcal{O}}_{CP}\} \mid \{\mathcal{O}_{CP}\}\right) \times \\
\times\delta\left(\mathcal{A}^{mix}_{\pi\pi} - \mathcal{A}^{mix}_{\pi\pi}(d,\, \vartheta,\, \gamma,\, \phi_d)\right) & \times \delta\left(\mathcal{A}^{dir}_{\pi\pi} - \mathcal{A}^{dir}_{\pi\pi}(d,\, \vartheta,\, \gamma)\right) \\
\times\delta\left(\mathcal{A}^{dir}_{KK} - \mathcal{A}^{dir}_{KK}(d,\, \vartheta,\, \gamma)\right) & \times \delta\left(\mathcal{A}^{mix}_{KK} - \mathcal{A}^{mix}_{KK}(d,\, \vartheta,\, \gamma,\, \phi_s)\right) \times \\
& \times f\left(d,\, \vartheta,\, \gamma,\, \phi_d,\, \phi_s\right). \qquad (9.5)
\end{aligned}
$$

$f\left(\{\hat{\mathcal{O}}_{CP}\} \mid \{\mathcal{O}_{CP}\}\right)$ is the joint experimental p.d.f. for $\mathcal{A}^{dir}_{\pi\pi}$, $\mathcal{A}^{mix}_{\pi\pi}$, $\mathcal{A}^{dir}_{KK}$ and $\mathcal{A}^{mix}_{KK}$. Following what discussed in Section 8.3, it can be written as the product of two bi-variate Gaussian p.d.f.s:

$$
f\left(\{\hat{\mathcal{O}}_{CP}\} \mid \{\mathcal{O}_{CP}\}\right) = G_\pi(\mathcal{A}^{dir}_{\pi\pi},\, \mathcal{A}^{mix}_{\pi\pi}) \times G_K(\mathcal{A}^{dir}_{KK},\, \mathcal{A}^{mix}_{KK}). \qquad (9.6)
$$

The p.d.f. $f\left(d,\, \vartheta,\, \gamma,\, \phi_d,\, \phi_s\right)$ describes our *a priori* knowledge on d, ϑ, γ, ϕ_d and ϕ_s. Since we have independent measurements of ϕ_d and ϕ_s, but we ignore any knowledge on d, ϑ and γ, we can write:

$$
f\left(d,\, \vartheta,\, \gamma,\, \phi_d,\, \phi_s\right) = F_0(d,\, \vartheta,\, \gamma) \times g^d_0(\phi_d) \times g^s_0(\phi_s), \qquad (9.7)
$$

where $F_0(d,\, \vartheta,\, \gamma)$ is a *prior* p.d.f., which we assume to be uniform, while $g^d_0(\phi_d)$ and $g^s_0(\phi_s)$ are *prior* p.d.f.s for ϕ_d and ϕ_s. We assume $g^d_0(\phi_d)$ and $g^s_0(\phi_s)$ to be Gaussians, with resolutions obtained from the LHCb $B^0 \to J/\psi K^0_S$ and $B^0_s \to J/\psi\phi$ analyses respectively [15, 18].

By introducing all these ingredients in equation (9.5), we end up with the final expression:

$$
\begin{aligned}
f\left(d,\, \vartheta,\, \gamma,\, \{\mathcal{O}_{CP}\},\, \phi_d,\, \phi_s \mid \{\hat{\mathcal{O}}_{CP}\}\right) & \propto G_\pi(\mathcal{A}^{dir}_{\pi\pi},\, \mathcal{A}^{mix}_{\pi\pi}) \times G_K(\mathcal{A}^{dir}_{KK},\, \mathcal{A}^{mix}_{KK}) \times \\
\times\delta\left(\mathcal{A}^{dir}_{\pi\pi} - \mathcal{A}^{dir}_{\pi\pi}(d,\, \vartheta,\, \gamma)\right) & \times \delta\left(\mathcal{A}^{mix}_{\pi\pi} - \mathcal{A}^{mix}_{\pi\pi}(d,\, \vartheta,\, \gamma,\, \phi_d)\right) \times \\
\times\delta\left(\mathcal{A}^{dir}_{KK} - \mathcal{A}^{dir}_{KK}(d,\, \vartheta,\, \gamma)\right) & \times \delta\left(\mathcal{A}^{mix}_{KK} - \mathcal{A}^{mix}_{KK}(d,\, \vartheta,\, \gamma,\, \phi_s)\right) \times \\
& \times F_0(d,\, \vartheta,\, \gamma) \times g^d_0(\phi_d) \times g^s_0(\phi_s). \qquad (9.8)
\end{aligned}
$$

Finally, this p.d.f. can be integrated over the variables $\mathcal{A}^{dir}_{\pi\pi}$, $\mathcal{A}^{mix}_{\pi\pi}$, $\mathcal{A}^{dir}_{KK}$, $\mathcal{A}^{mix}_{KK}$, ϕ_d and ϕ_s to get the (unnormalized) joint p.d.f. for d, ϑ and γ that we are looking for:

$$
\begin{aligned}
F(d,\,\vartheta,\,\gamma) \;&\propto\; \int f\left(d,\,\vartheta,\,\gamma,\,\{\mathcal{O}_{CP}\},\,\phi_d,\,\phi_s \mid \{\hat{\mathcal{O}}_{CP}\}\right) \mathrm{d}\mathcal{A}^{dir}_{\pi\pi}\mathrm{d}\mathcal{A}^{mix}_{\pi\pi}\mathrm{d}\mathcal{A}^{dir}_{KK}\mathrm{d}\mathcal{A}^{mix}_{KK}\mathrm{d}\phi_d\mathrm{d}\phi_s \\
&= \int G_\pi(\mathcal{A}^{dir}_{\pi\pi}(d,\,\vartheta,\,\gamma),\,\mathcal{A}^{mix}_{\pi\pi}(d,\,\vartheta,\,\gamma,\,\phi_d)) \times \\
&\times\; G_K(\mathcal{A}^{dir}_{KK}(d,\,\vartheta,\,\gamma),\,\mathcal{A}^{mix}_{KK}(d,\,\vartheta,\,\gamma,\,\phi_s)) \times g_0^d(\phi_d) \times g_0^s(\phi_s)\mathrm{d}\phi_d\mathrm{d}\phi_s, \qquad (9.9)
\end{aligned}
$$

where the flat *prior* $F_0(d,\,\vartheta,\,\gamma)$ has been dropped, since it does not carry information and would be in any case cancelled by normalizing $F(d,\,\vartheta,\,\gamma)$ to 1.

9.2 The $(d,\,\gamma)$ plane

Equation (9.9) provides a conceptually simple way to compute a joint p.d.f. for d, ϑ and γ. Since this p.d.f. is a function of three variables, trivial graphical representations are not possible. However, useful plots can be realized in two variables, for example for the joint p.d.f. of d and γ:

$$
F(d,\,\gamma) = \int F(d,\,\vartheta,\,\gamma)\mathrm{d}\vartheta. \qquad (9.10)
$$

The $(d,\,\gamma)$ plane was also chosen by R. Fleischer in his original work [8] to represent the analytical solution of the system of constraints, in case of perfect knowledge of the CP-violating observables and of the weak phases ϕ_d and ϕ_s. This is discussed in Section 9.2.1.

Figure 5*a* shows the 68% and 95% confidence regions for d and γ in the $(d,\,\gamma)$ plane, for one year of data and nominal values of the unknown physics parameters (see Table 2), while Figures 5*b* and 5*c* show the one-dimensional p.d.f.s for γ and d respectively, calculated as:

$$
F(d) = \int F(d,\,\vartheta,\,\gamma)\mathrm{d}\vartheta\mathrm{d}\gamma \qquad (9.11)
$$

and

$$
F(\gamma) = \int F(d,\,\vartheta,\,\gamma)\mathrm{d}\vartheta\mathrm{d}d. \qquad (9.12)
$$

The confidence regions for d and γ are numerically computed as highest posterior density regions[5]. All the multiple integrals in expressions (9.9), (9.10), (9.11) and (9.12) have been computed by means of a Monte Carlo integration method [20].

Figure 5*a* shows also two distinct 95% confidence regions for d and γ (continuous bands), obtained by combining the two constraints (3.4), (3.5) and the two constraints (3.7), (3.8) independently. A discussion on how they are obtained is reported later in Section 9.2.2, after the introduction of additional useful formalism in the following section.

[5]Highest posterior density confidence regions (intervals) of a given confidence level α are defined as regions (intervals) which enclose a probability α, with the p.d.f. having larger values everywhere inside the region (interval), with respect to the outside. Different definitions of confidence regions exist in the literature. However, they are equivalent if the p.d.f. is symmetric, and almost equivalent in case the p.d.f. does not spread dramatically far from the mode.

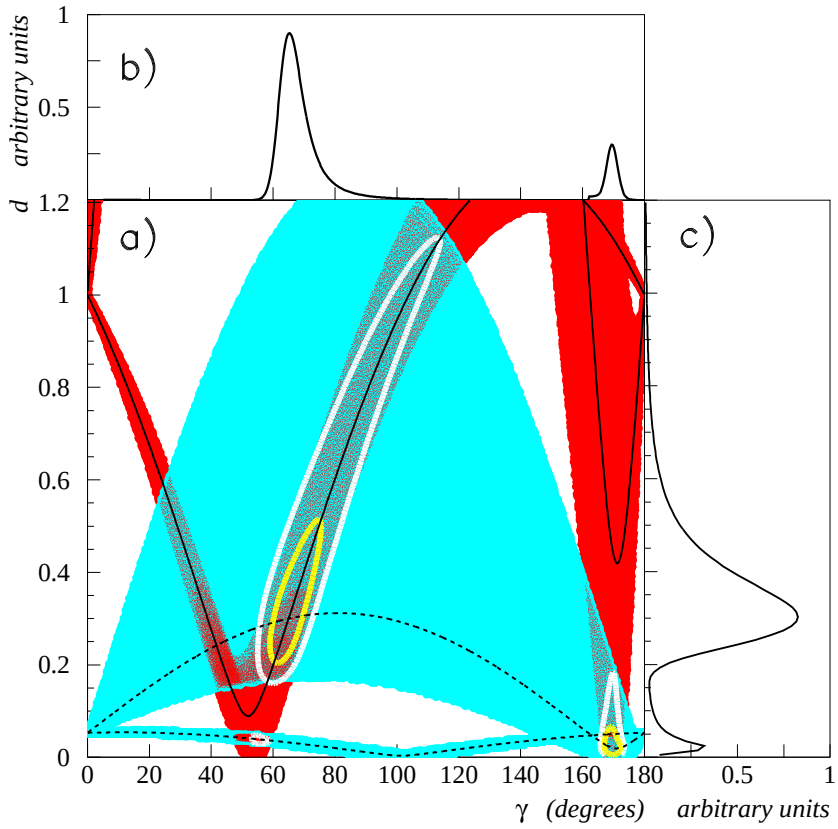

Figure 5: a) Confidence regions in the (d, γ) plane for one year of data, generated with nominal values of the unknown physics parameters (see Table 2); the darker (red) bands delimit the 95% confidence regions obtained by combining the two constraints (3.4) and (3.5) from $B^0 \to \pi^+\pi^-$; the lighter (cyan) bands delimit the 95% confidence regions obtained by combining the two constraints (3.7) and (3.8) from $B^0_s \to K^+K^-$; the black solid and dashed curves represent analytical relations between d and γ obtained with perfect knowlegde of the CP-violating observables and of the weak phases ϕ_d and ϕ_s (see Section 9.2.1); the light (yellow and white) contours enclose the 68% and 95% confidence regions obtained by using all the four constraints, as described in the text. b) p.d.f. for γ. c) p.d.f. for d.

9.2.1 Analytical solution of the system of constraints

The black solid and dashed curves in Figure 5*a*, represent analytical relations between d and γ obtained by combining the two constraints (3.4), (3.5) and the two constraints (3.7), (3.8) independently [8].

In fact, by combining (3.4) and (3.5), it is possible to eliminate ϑ (e.g., obtain ϑ as a function of $\mathcal{A}_{\pi\pi}^{dir}$, d and γ from the first constraint, and insert it into the second). After some algebra, one arrives to the following equation:

$$d = \sqrt{\frac{1}{k}\left[l \pm \sqrt{l^2 - hk}\right]}, \tag{9.13}$$

where h, k and l are given by:

$$h = u^2 + D(1 - u\cos\gamma)^2, \tag{9.14}$$

$$k = v^2 + D(1 - v\cos\gamma)^2 \tag{9.15}$$

and

$$l = 2 - uv - D(1 - u\cos\gamma)(1 - v\cos\gamma), \tag{9.16}$$

with

$$u = \frac{\mathcal{A}_{\pi\pi}^{mix} + \sin(\phi_d + 2\gamma)}{\mathcal{A}_{\pi\pi}^{mix}\cos\gamma + \sin(\phi_d + \gamma)}, \tag{9.17}$$

$$v = \frac{\mathcal{A}_{\pi\pi}^{mix} + \sin\phi_d}{\mathcal{A}_{\pi\pi}^{mix}\cos\gamma + \sin(\phi_d + \gamma)} \tag{9.18}$$

and

$$D = \left(\frac{\mathcal{A}_{\pi\pi}^{dir}}{\sin\gamma}\right)^2. \tag{9.19}$$

If the values of $\mathcal{A}_{\pi\pi}^{dir}$, $\mathcal{A}_{\pi\pi}^{mix}$ and ϕ_d are known, by means of equation (9.13) one can plot two curves of d as functions of γ, depending on the choice of the plus or minus sign in (9.13).

By following similar lines, combining (3.7) and (3.8), it is possible to eliminate ϑ', and after some algebra one obtains the following equation:

$$\tilde{d}' = \sqrt{\frac{1}{k'}\left[l' \pm \sqrt{l'^2 - h'k'}\right]}, \tag{9.20}$$

where h', k' and l' are given by:

$$h' = u'^2 + D'(1 - u'\cos\gamma)^2, \tag{9.21}$$

$$k' = v'^2 + D'(1 - v'\cos\gamma)^2 \tag{9.22}$$

and

$$l' = 2 - u'v' - D'(1 - u'\cos\gamma)(1 - v'\cos\gamma), \tag{9.23}$$

with

$$u' = \frac{\mathcal{A}_{KK}^{mix} + \sin(\phi_s + 2\gamma)}{\mathcal{A}_{KK}^{mix}\cos\gamma + \sin(\phi_s + \gamma)}, \tag{9.24}$$

$$v' = \frac{\mathcal{A}_{KK}^{mix} + \sin\phi_s}{\mathcal{A}_{KK}^{mix}\cos\gamma + \sin(\phi_s + \gamma)} \tag{9.25}$$

and

$$D' = \left(\frac{\mathcal{A}_{KK}^{dir}}{\sin\gamma}\right)^2. \tag{9.26}$$

If the values of $\mathcal{A}_{KK}^{dir}$, $\mathcal{A}_{KK}^{mix}$ and ϕ_s are known, by means of equation (9.20), and using in addition the relation between $\tilde{d}'$ and d' given in (3.10), one can plot two curves of d' as functions of γ [6].

If U-spin symmetry holds, then $d = d'$, and all the four curves from (9.13) and (9.20) can be plotted in the same (d, γ) plane. In case of perfect knowledge of the true values of the CP-violating observables and of the weak phases ϕ_d and ϕ_s, the curves must intersect at the true value of d and γ, as it is shown Figure 5*a* (black solid curves for $B^0 \to \pi^+\pi^-$ and black dashed curves for $B_s^0 \to K^+K^-$). However, note that in this way we have reduced the original system of four constraints to only two constraints, i.e., we are not using the entire available information. Even if the system of four contraints has unambiguous solution, there is no guarantee that no fake solutions are obtained so doing. In fact, this is exactly what can be observed in Figure 5*a*, where the curves intersect also at different points than the right solution (d, γ) = (0.3, 65^o).

The fake solutions can be eliminated by, for instance, re-inserting the obtained (d, γ) solutions in the original equations, then solving for ϑ. If the employed (d, γ) solution is fake, the solution in ϑ cannot be found. Another mathematical trick consists in obtaining a further relation between d and γ, and plot the corresponding curve on the (d, γ) plane, as it is done in [8].

It is worth to note that the adopted Bayesian approach described in the Section 9.1, which without lack of generality uses the four original constraints independently, i.e., it makes use of the full available information, automatically finds the true (d, γ) solution, in

[6] To compute d' from $\tilde{d}'$ by using (3.10), also the value of the sine of the Cabibbo angle needs to be known. Its value is nowadays measured with a small relative error of the order of 1% [12], i.e., much smaller than the errors on $\mathcal{A}_{KK}^{dir}$, $\mathcal{A}_{KK}^{mix}$ and ϕ_s for one year of data. Hence, for the sake of simplicity, in the following we have considered the uncertainty on this parameter as negligible, or, in other words, the parameter as perfectly known.

the limit to Dirac delta functions of the p.d.f.s of the four CP-violating observables and the two weak phases ϕ_d and ϕ_s. However, due to finite experimental resolutions on $\mathcal{A}_{\pi\pi}^{dir}$, $\mathcal{A}_{\pi\pi}^{mix}$, $\mathcal{A}_{KK}^{dir}$, $\mathcal{A}_{KK}^{mix}$, ϕ_d and ϕ_s, fake solutions, i.e., fake "peaks" in the $F(d,\, \vartheta,\, \gamma)$ p.d.f., may still occur, but, as it can be argued, for a different reason than that discussed before. For instance, these wrong peaks are visible in Figure 5. How to deal with them, is discussed in Section 9.4.

9.2.2 Construction of 95% confidence regions

In this section, we describe how the 95% confidence regions for d and γ obtained by combining the two constraints (3.4), (3.5) and the two constraints (3.7), (3.8) independently, are calculated. These confidence regions are shown in Figure 5*a* as continuous bands. As it can be argued, this is related to the formalism introduced to calculate the analytical solution discussed in the previous section.

In practice, these 95% confidence regions are obtained by projecting the 95% confidence regions from the space of the observables (i.e., $\mathcal{A}_{\pi\pi}^{dir}$, $\mathcal{A}_{\pi\pi}^{mix}$, $\mathcal{A}_{KK}^{dir}$, $\mathcal{A}_{KK}^{mix}$, ϕ_d and ϕ_s) to the corresponding space in the $(d,\, \gamma)$ plane. This can be done since the probability must conserve in this transformation[7].

We have shown in the previous section that combining (3.4) and (3.5), it is possible to obtain from (9.13) a relation between d and γ, if the values of $\mathcal{A}_{\pi\pi}^{dir}$, $\mathcal{A}_{\pi\pi}^{mix}$ and ϕ_d are known (i.e., from an experimental point of view, if they are measured). Assuming the gaussianity of the experimental p.d.f.s for $\mathcal{A}_{\pi\pi}^{dir}$, $\mathcal{A}_{\pi\pi}^{mix}$ and ϕ_d, the 95% confidence region in the ($\mathcal{A}_{\pi\pi}^{dir}$, $\mathcal{A}_{\pi\pi}^{mix}$, ϕ_d) observable space is an ellipsoid. By taking in mind that the measurements of $\mathcal{A}_{\pi\pi}^{dir}$ and $\mathcal{A}_{\pi\pi}^{mix}$ are each other correlated (see Section 8), the equation of the surface of this 95% confidence ellipsoid is:

$$\begin{aligned} &\frac{1}{2\left[1-\rho\left(\mathcal{A}_{\pi\pi}^{dir},\, \mathcal{A}_{\pi\pi}^{mix}\right)\right]}\left[\left(\frac{\mathcal{A}_{\pi\pi}^{dir}-\hat{\mathcal{A}}_{\pi\pi}^{dir}}{\sigma\left(\mathcal{A}_{\pi\pi}^{dir}\right)}\right)^2+\left(\frac{\mathcal{A}_{\pi\pi}^{mix}-\hat{\mathcal{A}}_{\pi\pi}^{mix}}{\sigma\left(\mathcal{A}_{\pi\pi}^{mix}\right)}\right)^2+\right. \\ &\left.-2\rho\left(\mathcal{A}_{\pi\pi}^{dir},\, \mathcal{A}_{\pi\pi}^{mix}\right)\left(\frac{\mathcal{A}_{\pi\pi}^{dir}-\hat{\mathcal{A}}_{\pi\pi}^{dir}}{\sigma\left(\mathcal{A}_{\pi\pi}^{dir}\right)}\right)\left(\frac{\mathcal{A}_{\pi\pi}^{mix}-\hat{\mathcal{A}}_{\pi\pi}^{mix}}{\sigma\left(\mathcal{A}_{\pi\pi}^{mix}\right)}\right)\right]+\frac{1}{2}\left(\frac{\phi_d-\hat{\phi}_d}{\sigma\left(\phi_d\right)}\right)^2 \;=\; f, \end{aligned} \tag{9.27}$$

where $\hat{\mathcal{A}}_{\pi\pi}^{dir}$, $\hat{\mathcal{A}}_{\pi\pi}^{mix}$ and $\hat{\phi}_d$ are the best estimates of the corresponding observables, $\sigma\left(\mathcal{A}_{\pi\pi}^{dir}\right)$, $\sigma\left(\mathcal{A}_{\pi\pi}^{mix}\right)$ and $\sigma\left(\phi_d\right)$ are their estimated uncertainties, and $\rho\left(\mathcal{A}_{\pi\pi}^{dir},\, \mathcal{A}_{\pi\pi}^{mix}\right)$ is the correlation between the measurements of $\mathcal{A}_{\pi\pi}^{dir}$ and $\mathcal{A}_{\pi\pi}^{mix}$. The value of the parameter f corresponding to 95% C.L. depends on the number of random variables (in this case 3). It can be found for instance in [14], and it is:

[7] In principle, these 95% confidence regions could be determined by obtaining p.d.f.s for d and γ with a Bayesian approach, following the same lines as in the discussion of Section 9.1, but employing as constraints equations (9.13) and (9.20). This way was first tried. The so obtained p.d.f.s resulted to have regular but very complex shapes. For this reason, it was not easy to give a reasonable definition of confidence region, and instead of choosing an arbitrary definition, difficult to be justified, we decided to follow the more intuitive approach described in the text.

$$f \simeq \frac{7.82}{2} = 3.91. \tag{9.28}$$

By uniformly scanning the volume of this ellipsoid, one can explore the 95% confidence region of the ($\mathcal{A}_{\pi\pi}^{dir}$, $\mathcal{A}_{\pi\pi}^{mix}$, ϕ_d) observable space. By uniformly scanning γ in the range [0^o, 180^o], and for each value of γ making a uniform scan in the 95% confidence region of the ($\mathcal{A}_{\pi\pi}^{dir}$, $\mathcal{A}_{\pi\pi}^{mix}$, ϕ_d) observable space, one obtains sets of 4-tuples ($\mathcal{A}_{\pi\pi}^{dir}$, $\mathcal{A}_{\pi\pi}^{mix}$, ϕ_d, γ). With the values of each 4-tuple, by means of (9.13), the corresponding value of d can be computed (if it exists). Each pair of values for d and γ is then used to populate a bi-dimensional histogram. By suitable choice of the steps used while scanning the volume of the ellipsoid, it is possible to obtain uniform bands in the (d, γ) plane, as it can be seen in Figure 5*a*.

One may object that with this technique, one indeed obtains the union of 95% confidence intervals for the random variable d conditioned by the value of the variable γ (interpretation A), obtained for all the allowed values of γ, instead of the 95% confidence region for the random variables d and γ (interpretation B). In fact, it is simply true. But since we are now using only one constraint, given by equation (9.13), that can be generically re-written (ignoring for simplicity the presence of two solutions corresponding to the plus or minus sign) as:

$$d = d(\mathcal{A}_{\pi\pi}^{dir}, \mathcal{A}_{\pi\pi}^{mix}, \phi_d, \gamma), \tag{9.29}$$

the problem is undetermined (two unknowns, d and γ, and one constraint). To better understand it, suppose that $\mathcal{A}_{\pi\pi}^{dir}$, $\mathcal{A}_{\pi\pi}^{mix}$ and ϕ_d are very well known (i.e., their p.d.f.s are very narrow, and at limit tend to Dirac delta functions). In this very favourable case, all we obtain from (9.29) is an (almost) exact dependence of d on γ, i.e., a function. All we can say is that, given a value for γ, we can compute a value for d (and vice versa, inverting the function). Thus, the only thing that we can be sure of on d (or γ), is its dependence on γ (or d) and its allowed region.

At this point one might be tempted to say that the individual p.d.f.s $f(d)$ and $f(\gamma)$ must be uniform inside the respective allowed regions for d and γ. But, by looking at (9.29), if $f(\gamma)$ is uniform inside the allowed region for γ, it is obvious that for an arbitrary dependence of d on γ, $f(d)$ cannot be uniform too (and vice versa). Then, one can conclude that stating the perfect ignorance on γ by assuming the uniformity of its p.d.f., the p.d.f. for d is automatically fixed by the constraint (9.29). In this case one can write:

$$f(d, \gamma) = f(d \mid \gamma) \times f(\gamma) \propto f(d \mid \gamma), \tag{9.30}$$

since $f(\gamma)$ is flat. Hence, interpretations A and B are equivalent. But one may also have followed the opposite path, by choosing $f(d)$ to be uniform inside the allowed region for d, so obtaining:

$$f(d, \gamma) = f(\gamma \mid d) \times f(d) \propto f(\gamma \mid d). \tag{9.31}$$

The two relations are not simultaneously true, and in this scheme, the choice of which p.d.f. assume as uniform is arbitrary. The arbitrariness reflects the fact that perfect ignorance on a random variable cannot be translated so blindly into a uniform p.d.f., as it can be argued for example by noting that the p.d.f. of the inverse of a random variable distributed uniformly is not uniform, and this cannot mean that we know something on the inverse of the variable, if we don't know anything on the variable itself.

In conclusion, we need more information to unambiguously solve such a problem, otherwise we must cohabit with this ambiguity. Hopefully, this ambiguity does not lead to dramatic difference of the results. The connection to Bayesian reasoning and the usage of flat *priors* seems evident.

However, if the previous discussion holds in the very favourable case of perfect knowledge of the observables $\mathcal{A}_{\pi\pi}^{dir}$, $\mathcal{A}_{\pi\pi}^{mix}$ and ϕ_d, there is no need of any mathematical proof to convince anyone that no additional information can be obtained in the real case of sizable uncertainties on $\mathcal{A}_{\pi\pi}^{dir}$, $\mathcal{A}_{\pi\pi}^{mix}$ and ϕ_d. What can only happen is a further enlargement of the allowed regions for d and γ. The conclusions remain unchanged, and interpretations A and B are still equivalent.

The case of the combination of the constraints (3.7) and (3.8) is analogous, with the equation of the 95% confidence ellipsoid given by:

$$\frac{1}{2\left[1-\rho\left(\mathcal{A}_{KK}^{dir},\,\mathcal{A}_{KK}^{mix}\right)\right]}\left[\left(\frac{\mathcal{A}_{KK}^{dir}-\hat{\mathcal{A}}_{KK}^{dir}}{\sigma\left(\mathcal{A}_{KK}^{dir}\right)}\right)^2+\left(\frac{\mathcal{A}_{KK}^{mix}-\hat{\mathcal{A}}_{KK}^{mix}}{\sigma\left(\mathcal{A}_{KK}^{mix}\right)}\right)^2+\right. \tag{9.32}$$
$$\left.-2\rho\left(\mathcal{A}_{KK}^{dir},\,\mathcal{A}_{KK}^{mix}\right)\left(\frac{\mathcal{A}_{KK}^{dir}-\hat{\mathcal{A}}_{KK}^{dir}}{\sigma\left(\mathcal{A}_{KK}^{dir}\right)}\right)\left(\frac{\mathcal{A}_{KK}^{mix}-\hat{\mathcal{A}}_{KK}^{mix}}{\sigma\left(\mathcal{A}_{KK}^{mix}\right)}\right)\right]+\frac{1}{2}\left(\frac{\phi_s-\hat{\phi}_s}{\sigma\left(\phi_s\right)}\right)^2 \;=\; f,$$

where the meaning of the symbols is now obvious. The only difference with the case discussed previously, is that one now has to use equation (9.20), and not forget to pass from $\tilde{d}$ to d through (3.10).

9.3 Sensitivity on γ

The sensitivity on γ is explored for different values of the physics parameters $\Delta\Gamma_s/\Gamma_s$, Δm_s, d, ϑ, γ and ϕ_s. These parameters are varied one at a time, with the other ones kept to their assumed nominal values. The results are summarized in Table 5.

As it is visible for instance in Figure 5*b*, the p.d.f. for γ (9.12) can be asymmetric for some of the input physics parameters. For this reason, the sensitivities $\sigma(\gamma)$ shown in Table 5 are calculated as half the range covered by the 68% highest posterior density confidence intervals. Furthermore, since secondary fake solutions can be present (see discussion in the next Section 9.4), these are ignored in the construction of the confidence intervals.

$\Delta\Gamma_s/\Gamma_s$	0	(0.1)	0.2			
$\sigma(\gamma)$	5.2^o	4.9^o	4.5^o			
Δm_s $[ps^{-1}]$	15	(20)	25	30		
$\sigma(\gamma)$	4.0^o	4.9^o	5.9^o	8.5^o		
d	0.1	0.2	(0.3)	0.4		
$\sigma(\gamma)$	1.8^o	2.7^o	4.9^o	9.0^o		
ϑ	120^o	140^o	(160^o)	180^o	200^o	
$\sigma(\gamma)$	3.8^o	3.8^o	4.9^o	6.7^o	5.2^o	
γ	55^o	(65^o)	75^o	85^o	95^o	105^o
$\sigma(\gamma)$	5.8^o	4.9^o	4.3^o	4.7^o	4.7^o	4.7^o
ϕ_s [rad]	0	(-0.04)	-0.1	-0.2		
$\sigma(\gamma)$	4.9^o	4.9^o	4.9^o	5.4^o		

Table 5: Statistical uncertainty on γ for one year of data, depending on the true values of the physics parameters $\Delta\Gamma_s/\Gamma_s$, Δm_s, d, ϑ, γ and ϕ_s. The physics parameters are varied one at a time, with the other ones kept to their assumed nominal values. The nominal values are those enclosed in parentheses.

9.4 Discussion

At this point, it is worth pointing out some remarks. As it can be noted in Figure 5*a*, the two 95% confidence regions obtained from the combination of (3.4), (3.5) and (3.7), (3.8) cross each other not only in the right solution region, but also in other regions where the p.d.f. calculated by means of the full Bayesian approach described in Section 9.1 does not extend. This is due to the same reason mentioned at the end of Section 9.2.1, i.e., the two 95% confidence regions are obtained using partial information only.

However, what is really important for the determination of γ is the approach described in Section 9.1. The two 95% confidence regions and the analytical relations from the combination of (3.4), (3.5) and (3.7), (3.8) are reported in Figure 5*a* with the sole aim of providing means to understand the consistency of the technique adopted to extract γ. They are of no relevance and not used for the evaluation of the sensitivity on γ, as presented in the previous Section 9.3.

By looking at the 68% and 95% confidence regions for d and γ in Figure 5*a*, as well as at the p.d.f.s for γ and d in the Figures 5*b* and 5*c*, it is possible to see that a fake solution at large values of γ ($\gamma \simeq 170^o$) and small values of d ($d \simeq 0.05$) is obtained (we remind that all the plots in Figure 5 are obtained for one year statistics and nominal values of the unknown physics parameters reported in Table 2). Within the parameter ranges used in the sensitivity scan, these wrong components to the p.d.f. $F(d, \gamma)$ are always observed at large values of γ ($\gamma > 160^0$) and small values of d ($d < 0.15$), mainly depending on the true values of d, γ and ϑ. Furthermore, they are not observed in all the sensitivity scans.

This effect is not due to the method itself, but to the finite resolution on the CP-

violating observables. It can be eliminated with larger statistics (in the sense of more years of data taking), or by combining the measurement of d and γ with information from other independent analyses (i.e., start from non-uniform prior $F_0(d,\, \vartheta,\, \gamma)$ in equation (9.8)). For this reason, the occasional presence of a wrong solution is ignored in the determination of the resolutions reported in Table 5.

One final remark is on the assumed validity of U-spin symmetry, which is required not to be dramatically broken in order to have reliable and useful results from this measurement. A starting point for a theoretical discussion on the validity of this symmetry, useful for this study, can be found in [8]. From an experimental point of view, this symmetry can be checked to some extent from data itself. For example, since the problem is over-constrained (3 unknowns, d, ϑ and γ, and 4 constraints) as discussed at the beginning of Section 9, one can release an additional parameter, such as no longer impose $d = d'$ (or alternatively $\vartheta = \vartheta'$), and study the difference $\Delta d = d' - d$ (or $\Delta\vartheta = \vartheta' - \vartheta$).

10 Conclusions

The LHCb sensitivity in measuring the CP-violating observables in $B^0_{(s)} \to h^+h^-$ decays has been estimated. Due to the large statistics collected per year of data taking [10], it has been found that LHCb can realize very precise measurements of CP-violation with these decays.

The event generation has been realized by means of a dedicated fast Monte Carlo simulation, while the extraction of the physics parameters from simulated data has been achieved by employing an extended maximum likelihood approach.

The combination of $B^0 \to \pi^+\pi^-$ and $B^0_s \to K^+K^-$ CP measurements can be used to extract the CKM phase γ [8], relying on the validity of the U-spin symmetry of the strong interaction dynamics. An implementation of this method, based on a Bayesian approach, has been realized, and the resulting sensitivity on γ has been estimated.

The relevance of this measurement can be made immediately visible considering that, since both $\overline{b} \to \overline{d} + g\,(\gamma,\, Z^0)$ and $\overline{b} \to \overline{s} + g\,(\gamma,\, Z^0)$ penguin processes are involved in the $B^0 \to \pi^+\pi^-$ and $B^0_s \to K^+K^-$ decay amplitudes, the value of γ determined in this way can be affected by sizable contributions from New Physics. This would lead to discrepancies with the values of γ predicted by the Standard Model through indirect CKM fits, as well as with those determined by using other B decays, such as $B^0_s \to D^\mp_s K^\pm$ [16], which is generated by pure $\overline{b} \to \overline{u} + W^+$ and $\overline{b} \to \overline{c} + W^+$ tree processes.

Aknowledgements

The authors are grateful to T. Nakada and O. Schneider for stimulating discussions and suggestions, as well as for their constant presence and attention on this work. M. Merk and G. Raven are kindly acknowledged for useful suggestions, and the former also for his careful review of the manuscript. Many thanks go to F. Ruggieri and the entire INFN-CNAF staff for their efficiency in keeping the LHCb-Bologna computing cluster, massively employed for this analysis, fully operational.

References

[1] LHCb Collaboration, LHCb Technical Proposal, CERN-LHCC/98-004.

[2] LHCb Collaboration, LHCb Technical Design Report, CERN-LHCC/2003-030.

[3] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **89** (2002) 201802.

[4] BELLE Collaboration, K. Abe *et al.*, Phys. Rev. D **66** (2002) 071102.

[5] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **89** (2002) 281802.

[6] BELLE Collaboration, K. Abe *et al.*, Phys. Rev. D **68** (2003) 012001.

[7] R. Fleischer and T. Mannel, Phys. Lett. B **397** (1997) 269.

[8] R. Fleischer, Phys. Lett. B **459** (1999) 306.

[9] R. Fleischer and J. Matias, Phys. Rev. D **66** (2002) 054009.

[10] V. Vagnoni *et al.*, Selection of $B^0_{(s)} \to h^+h^-$ decays at LHCb, CERN-LHCb/2003-123.

[11] G. Avoni *et al.*, A Beowulf-class computing cluster for the Monte Carlo production of the LHCb experiment, Comput. Phys. Commun. **150** (2003) 129.

[12] M. Battaglia *et al.*, The CKM matrix and the unitarity triangle, hep-ph/0304132.

[13] F. James, Minuit Reference Manual, `http://wwwinfo.cern.ch/asdoc/minuit/minmain.html`.

[14] Particle Data Group, K. Hagiwara *et al.*, Phys. Rev. D **66** (2002) 010001, and 2003 off-year partial update for the 2004 edition available on the PDG web site: `http://pdg.lbl.gov/`.

[15] G. Raven, Sensitivity studies of χ and $\Delta\Gamma$ with $B^0_s \to J/\psi(\mu^+\mu^-)\phi(K^+K^-)$, CERN-LHCb/2003-119.

[16] R. Hierck, J. van Hunen and M. Merk, The sensitivity for Δm_s and $\gamma + \phi_s$ from $B^0_s \to D^-_s\pi^+$ and $B^0_s \to D^\mp_s K^\pm$, CERN-LHCb/2003-103.

[17] M. Calvi, O. Dormond and M. Musy, LHCb flavor tagging performance, CERN-LHCb/2003-115.

[18] S. Amato, J.R.T. de Mello Neto and C. Nunes, The LHCb sensitivity to $\sin 2\beta$ from $B^0 \to J/\psi(\mu\mu)K^0_S$, CERN-LHCb/2003-107.

[19] CORGEN - Correlated Gaussian-distributed Random Numbers, `http://wwwasdoc.web.cern.ch/wwwasdoc/shortwrupsdir/v122/top.html`.

[20] DIVON4 - Multidimensional Integration or Random Number Generation, `http://wwwasdoc.web.cern.ch/wwwasdoc/Welcome.html`.