First Observation of the Charmless Decay B to K pi0pi0 and Study of its Dalitz Plot Structure

By Eugenia Maria Teresa Irene Puccio

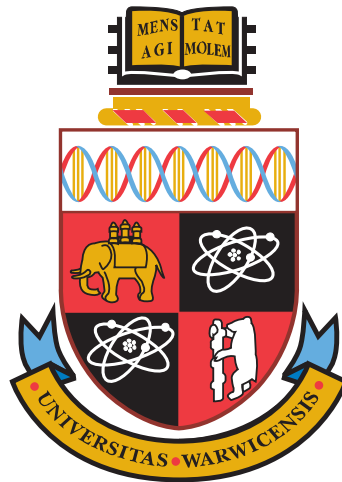# First observation of the charmless decay $B^+ \to K^+\pi^0\pi^0$ and study of the Dalitz plot structure

Eugenia Maria Teresa Irene Puccio

Thesis

Submitted to the University of Warwick for the degree of

Doctor of Philosophy

Department of Physics

University of Warwick

March 2011

Dedicata alla mia famiglia e in particolare alla memoria di mia madre.

Contents

vi

List of Figures

List of Tables

Acknowledgements

I would like to thank the Particle Physics Group at the University of Warwick for giving me the opportunity to work with them; it was a pleasure to be a part of this group and learn from such amazing people. In particular I would like to thank Dr. Tim Gershon for all his support, advice and encouragement in what turned out to be quite a tricky and challenging analysis. I also would like to thank STFC for providing the funding for my Ph.D and my long term attachment in California, which allowed me to become more involved within the Collaboration.

A particular thanks to the amazing people I met along the way both at Warwick and in California who have contributed to make these four years so enjoyable. To Andy Bennieston, Nicola McConkey, Nikola Chmel, Leigh and Mark Whitehead for always being there for me and keeping me company in the essential coffee breaks. To everyone at SLAC: Mark Tibbetts, Sudan Paramesvaran, Michael Sigamani, Jen Watson and in particular Kim Alwyn, Graham Jackson, Jasmine Hasi, Manuel Franco and Aidan Randle-Conde for making my life experience in America so fun and every Collaboration Meeting interesting! A very special thanks to Aidan and Jelena Ilic for their friendship and support in overcoming the hardest and most difficult time of my life. I would not have been able to finish this work without you! To Jean Sutherland and George Lafferty for their support while on LTA and for making those dinner parties in California so special.

A very special thanks to Tom Latham who taught me so much, not only about Particle Physics and programming, but also about life. You have been there for me and encouraged me from the start. I could not have asked for a better teacher and

partner in life.

Finally to my family, to whom I dedicate this thesis, because without your love and faith in me this thesis would have never happened. In particular to my mother who gave me all her support and encouragement even in the final moments of her life and without whom I would not have come so far.

Declaration

I declare that the work in this thesis was carried out in accordance with the Regulations of the University of Warwick. No part of this thesis has been submitted for any other academic award at this or any other university.

The data used in this analysis was recorded by the *BABAR* detector run by the *BABAR* Collaboration. The author contributed to the running of the detector by taking general shifts, being Run Quality Manager for the reprocessing of the final dataset, used in this analysis, and performing the neutral pion efficiency study in Appendix B. The event reconstruction uses code developed by the Collaboration, with packages for 3-body B meson decay reconstruction and final selection (`QnBUser` and `CharmlessFitter`) written by Dr. Fergus Wilson and Dr. Thomas Latham respectively. The software used for the likelihood fit (`Laura++`) was first developed by Prof. Paul Harrison, Dr. John Back and Dr. Thomas Latham, further extended by the author to incorporate the PDF needed for misreconstructed signal modelling. The π^0 efficiency study code in Appendix B was mostly written by Dr. Thomas Latham with significant contributions from the author and Dr. Tim Gershon. The likelihood scan (Figure 6.5) used to determine the statistical signal significance and the final calculation of the π^0 efficiency corrections (part of Section 6.3.3) was done by Dr. Thomas Latham. The remainder of the work described in this document (Chapters 4–8 and all Appendices) was carried out solely by the author with support from Dr. Tim Gershon, Dr. Thomas Latham and the Charmless 3-body Analysis Working Group from *BABAR*. All views expressed are those of the author.

 SIGNED: .... DATE: ....

Abstract

Results for the first measurement of the inclusive branching and CP asymmetry of the charmless 3-body decay $B^+ \rightarrow K^+\pi^0\pi^0$ are presented. The analysis uses a data sample with an integrated luminosity of $429.0\,\mathrm{fb}^{-1}$, recorded by the $BABAR$ detector at the PEP-II asymmetric B Factory. This sample corresponds to 470.9 ± 2.8 million $B\bar{B}$ pairs. Measurements of the branching fractions ($\mathcal{B}$) and CP asymmetries (A_{CP}) of some of the intermediate resonances in the $K^+\pi^0\pi^0$ Dalitz plot are also presented. The results are summarised here:

◇ $\mathcal{B}\left(B^+ \rightarrow K^+\pi^0\pi^0\right) = (16.2 \pm 1.2 \pm 1.4) \times 10^{-6}$

◇ $\mathcal{B}\left(B^+ \rightarrow K^*(892)^+\pi^0; K^{*0}(892) \rightarrow K^+\pi^0\right) = (2.72 \pm 0.50 \pm 0.34) \times 10^{-6}$

◇ $\mathcal{B}\left(B^+ \rightarrow f_0(980)K^+; f_0(980) \rightarrow \pi^0\pi^0\right) = (2.77 \pm 0.56 \pm 0.43) \times 10^{-6}$

◇ $\mathcal{B}\left(B^+ \rightarrow \chi_{c0}K^+; \chi_{c0} \rightarrow \pi^0\pi^0\right) = (0.51 \pm 0.22 \pm 0.09)$

◇ $A_{CP}\left(B^+ \rightarrow K^+\pi^0\pi^0\right) = (-6 \pm 6 \pm 4)\,\%$

◇ $A_{CP}\left(B^+ \rightarrow K^*(892)^+\pi^0\right) = (-4 \pm 26 \pm 4)\,\%$

◇ $A_{CP}\left(B^+ \rightarrow f_0(980)K^+\right) = (17 \pm 18 \pm 4)\,\%$

◇ $A_{CP}\left(B^+ \rightarrow \chi_{c0}K^+\right) = (-89 \pm 37 \pm 4)\,\%$

Introduction

This thesis presents the first study of the charmless 3-body decay, $B^+ \to K^+\pi^0\pi^0$. It includes the inclusive measurement of the branching fraction and CP asymmetry. Additionally a study is performed on the Dalitz plot to obtain the branching fractions and CP asymmetries of the following intermediate quasi-two-body decays: $B^+ \to K^*(892)^+(\to K^+\pi^0)\pi^0$, $B^+ \to f_0(980)(\to \pi^0\pi^0)K^+$ and $B^+ \to \chi_{c0}(\to \pi^0\pi^0)K^+$. The A_{CP} measurement of $B^+ \to K^*(892)^+\pi^0$ is particularly interesting in the light of new theoretical approaches to study the "$K\pi$" puzzle in the $K^*\pi$ system [1,2]. A previous measurement of the quasi-two-body decay of $B^+ \to K^{*+}\pi^0$ was published in Physics Review Letters by the *BABAR* Collaboration. This analysis used 232 million $B\overline{B}$ pairs and obtained a branching fraction of $\mathcal{B}\left(B^+ \to K^*(892)^+\pi^0\right) = [6.9 \pm 2.0(stat.)\pm1.3(sys.)]\times10^{-6}$, with a statistical significance of 3.6 standard deviations, and an asymmetry of $A_{CP}\left(B^+ \to K^*(892)^+\pi^0\right) = 0.04 \pm 0.29(stat.) \pm 0.05(sys.)$ [3]. The large statistical error associated to the asymmetry measurement makes it hard to extract conclusive results and a more precise measurement is needed. This thesis makes use of the full *BABAR* dataset of 470.9 million $B\overline{B}$ events collected at the $\Upsilon(4S)$ resonance, more than double the dataset available in the previous measurement. Since the asymmetries for $B^+ \to f_0(980)K^+$ and $B^+ \to \chi_{c0}K^+$ are already measured in the $K^+\pi^+\pi^-$ final state [4,5], these results serve as a cross-check of the procedure. The branching fractions of the decays $B^+ \to K^*(892)^+\pi^0$ and $B^+ \to \chi_{c0}K^+$ serve mainly as a cross-check from previous world average measurements, correcting for $K^*(892)^+ \to K^+\pi^0$ and $\chi_{c0} \to \pi^0\pi^0$ assuming isospin conservation. However it is not clear so far that dipion decays of the $f_0(980)$ do conserve isospin [6–10]. For

this mode the ratio of product branching fractions is quoted instead:

$$\frac{\mathcal{B}(B^+ \to f_0(980)K^+) \times \mathcal{B}(f_0(980) \to \pi^0\pi^0)}{\mathcal{B}(B^+ \to f_0(980)K^+) \times \mathcal{B}(f_0(980) \to \pi^+\pi^-)} = \frac{\mathcal{B}(f_0(980) \to \pi^0\pi^0)}{\mathcal{B}(f_0(980) \to \pi^+\pi^-)} \qquad (1)$$

The prediction from isospin symmetry is that the ratio of these two decays should be about 0.5.

1

Theoretical Motivations

Until 1964, the weak interaction was known to violate both charge conjugation and parity transformation individually. However in that year evidence of violation of the combination of these two transformations, known as CP, was found in the decay channel $K_L \to \pi^+\pi^-$ at Brookhaven National Laboratory [11]. CP violation is one of the three Sakharov conditions required to explain the matter-antimatter asymmetry in the Universe [12]. The primary goal of the *BABAR* experiment is the study of CP asymmetries and the measurement of CP violation parameters in the B meson system. This chapter introduces the concept of CP asymmetry and the Standard Model formalism. It also introduces the primary motivations for this measurement, *i.e.* the study of the equivalent of the so called "$K\pi$ puzzle" in $K^*\pi$ system, and the advantage of looking for this mode in the 3-body decay.

1.1 Flavour mixing and the CKM formalism

This chapter gives a brief overview of the theory behind the CKM matrix and its formulation in the Standard Model. For a more detailed derivation of the CKM formalism refer to Refs. [13–16]. In the Standard Model, all matter is made up of quarks and leptons each of which comes in six different types or "flavours". Quark flavours change under the charged current weak interaction. These flavour changing quark transitions are described by the following Lagrangian:

$$\mathcal{L}^{(CC)} = -\frac{g}{\sqrt{2}}(J^{\mu}_{(CC)}W^{(+)}_{\mu} + J^{\dagger\mu}_{(CC)}W^{(-)}_{\mu}) \tag{1.1}$$

where $W^{\pm}_{\mu}$ represents the charged vector boson field, g is the weak coupling constant. The operator $J^{\mu}_{(CC)}$ is the weak charged quark current operator and takes the form:

$$J^{\mu}_{(CC)} = \bar{u}_L\gamma^{\mu}d_L \tag{1.2}$$

where (u_L, d_L) is the notation for the left-handed SU(2) quark doublet and γ^{μ} the Dirac matrices. To represent fully the mass eigenstates that propagate through space and time, it is useful to transform the Lagrangian to the mass basis. Denoting the unitary basis transformation $\mathcal{S}$, the quark flavour basis is rotated to states of definite mass, u^m_L and d^m_L, by writing:

$$u^m_L = \mathcal{S}^u_L u_L \qquad d^m_L = \mathcal{S}^d_L d_L \tag{1.3}$$

Using the transformation in the expression for the charged weak currents coupling u_L to d_L, becomes:

$$J^{\mu}_{(CC)} = \bar{u}^m_L\gamma^{\mu}(\mathcal{S}^u_L\mathcal{S}^{d\dagger}_L)d^m_L \tag{1.4}$$

This expression shows the origin of the unitary matrix $V \equiv \mathcal{S}^u_L\mathcal{S}^{d\dagger}_L$ also known as the CKM matrix, which was named after the theorists Cabibbo, Kobayashi and Maskawa [17, 18]. Each up-type quark couples to a mixture of down-type quarks so that the CKM matrix defines the rotation from the down-type quark states produced in the strong interaction regime (d, s and b) to another set of down-type quark states

"seen" by the weak interaction (d',s' and b'):

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.5}$$

The origin of the CKM matrix is related to the fermion masses via the Yukawa interaction. The Lagrangian of the Yukawa interaction takes the form:

$$\mathcal{L}_Y = -Y_{ij}^u \bar{q}_{Li} \phi u_{Rj} - Y_{ij}^d \bar{q}_{Li} \epsilon \phi^* d_{Rj} \tag{1.6}$$

where $Y^{u,d}$ are 3×3 complex matrices, ϵ is a 2×2 antisymmetric tensor, ϕ is the Higgs field and indices run over the flavour generations. When ϕ acquires a vacuum expectation value, the Yukawa interaction yields non-diagonal mass matrices for the quarks. To determine the quark masses, Yukawa terms have to be diagonalised by different transformations for left-handed up and down quarks. As a result the charged current interactions couple to u_L^m and d_L^m states shown in Eq. 1.3 with couplings given by the elements of the CKM matrix.

1.1.1 The unitarity triangle

The CKM matrix has four independent parameters, which can be thought of as three mixing angles between the three pairs of quark generations and a complex phase. The unitarity requirement of the CKM matrix places the following constraint on its elements:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{1.7}$$

This is only one of six orthogonality constraints and each can be represented as a triangle in the complex plane. Eq. 1.7 is known by the B Factories as "the unitarity triangle" and is shown in Figure 1.1. Each angle of the triangle is given by ratios of elements of the CKM matrix as,
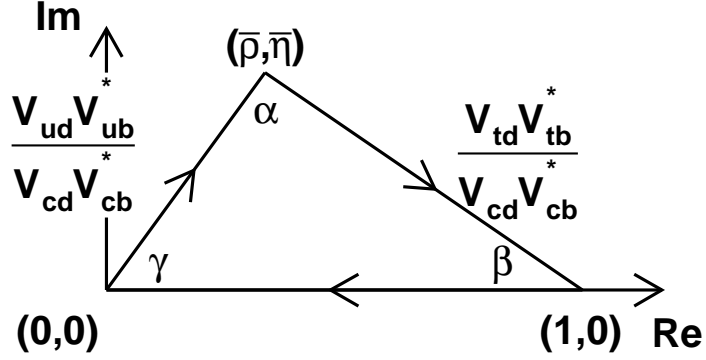
Figure 1.1: *The unitarity triangle with three mixing angles and sides as a function of the elements in the CKM matrix.*

$$\alpha = -\arg\left(\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right) \tag{1.8}$$

$$\beta = -\arg\left(\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right) \tag{1.9}$$

$$\gamma = -\arg\left(\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right) \tag{1.10}$$

and can be experimentally determined via specific decay modes. A few examples include the time dependent analysis of $B^0 \to J/\psi\, K_S^0$, also known as the golden mode for measuring β, and Cabibbo suppressed modes like $B \to D\overline{D}$. Measurements of the angles and sides have to be made through as many independent decay modes as possible to overconstrain the triangle and to be able to probe for contributions from physics beyond the Standard Model. All measurements are then combined into a fit, an example of which is shown in Figure 1.2. The area of the triangle gives a convention independent measure of the amount of *CP* violation in the Standard Model.
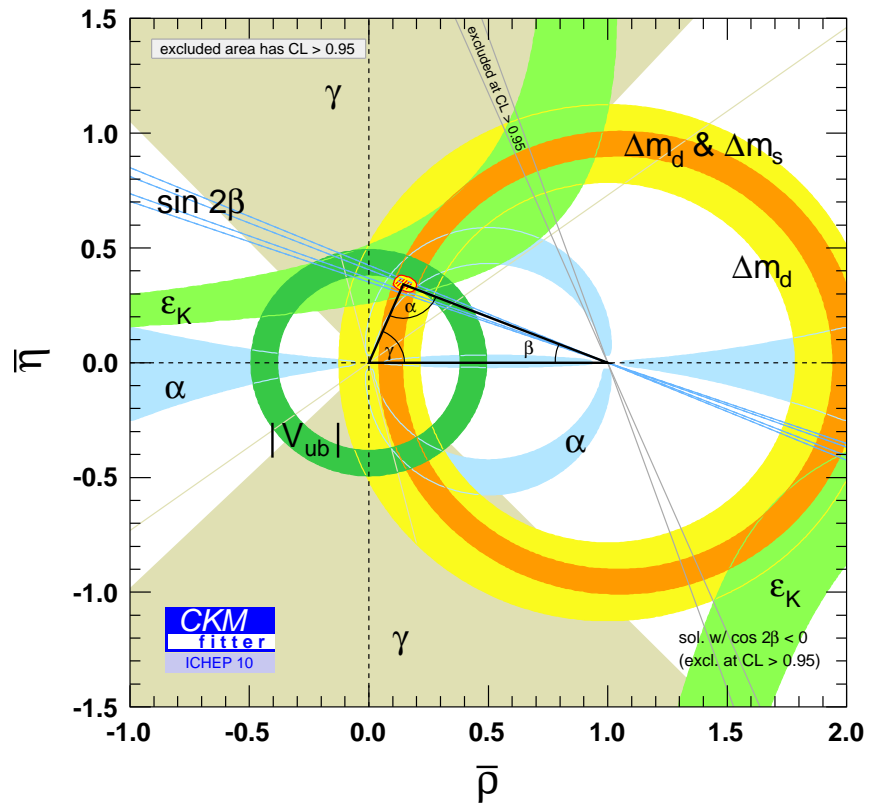
Figure 1.2: *Constraints in the $(\bar{\rho}, \bar{\eta})$ plane including recent measurements of α and γ in the global CKM fit. The red hashed region of the global combination corresponds to 68% confidence level (CL) [19].*

1.2   *CP* **violation**

This section gives a brief overview of the different types of *CP* violation observed in meson decays. For a more detailed description refer to Refs. [20, 21]. In quantum theory a transformation maps a state to another through a unitary operator U where the unitarity condition is required to maintain normalisation. This can be generalised by:

$$|\Psi\rangle \rightarrow |\Psi'\rangle = U |\Psi\rangle \tag{1.11}$$

The symmetry properties of a physical system arise by the invariance of the measured quantities when performing a transformation. A symmetry violation occurs when a symmetry, well-established in a class of physical processes, is broken under certain circumstances. The discrete transfomations of interest here in particle physics are listed below.

◇ **Charge conjugation**: this is the process associated with the exchange of particles and antiparticles under the transformation of all charges into their opposite sign by the unitary operator, C. Charge conjugation symmetry requires that for every particle, an antiparticle exists which behaves in exactly the same way except with all its internal charges reversed.

◇ **Parity**: the effect of parity transformation is defined as the inversion of spatial coordinates with respect to an origin via the unitary operator, P. The most common interpretation of this transformation is a process under which a right-handed reference system becomes a left-handed system.

◇ **Time reversal**: this is the transformation corresponding to the inversion of the time coordinate. Time reversal invariance is simply the statement that two processes related to one another by a reversal of all momenta and angular momenta have equal rates. Since momenta and angular momenta are derivatives with respect to time, reversing these quantities is mathematically equivalent to a reversal of sign of the time variable.

While there is no evidence of violation of these transformations in electromagnetic and strong interaction processes, charge conjugation and parity are found to be maximally violated by the weak interaction. The *CPT* theorem states that the product of the three transformations is a valid symmetry and is the only combination of C,P and T which is at, this time, believed to be an exact symmetry of Nature. *CP* symmetry is the discrete symmetry which attracted the most attention experimentally and theoretically, the reason being that despite its fundamental significance and connection to time reversal symmetry through *CPT*, it has been found to be violated, first in the kaon system and later on in the B meson system. In addition to this, T symmetry violation has been observed in the neutral kaon decays as expected as a corollary of *CP* violation if *CPT* is to be conserved [22]. *CP* violation can be classified in three categories which are detailed in the following subsections.

1.2.1 *CP* violation in mixing

The spontaneous oscillation of neutral mesons into their antiparticles was first observed in the kaon system [23], then in the B system [24] and most recently in the D meson system [25, 26]. A typical Feynmann diagram for $B^0\overline{B}^0$ mixing is shown in Figure 1.3.
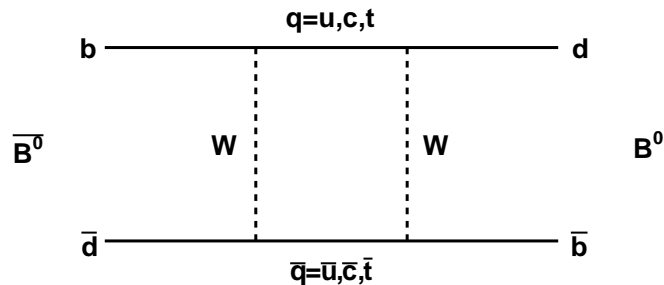
Figure 1.3: *Box diagram for $B^0\overline{B}^0$ mixing via the exchange of two W bosons.*

A generic neutral meson M^0 and its antiparticle $\bar{M}^0$ are defined by the transformations

$$CP\left|M^0\right\rangle = \eta_{CP}\left|\bar{M}^0\right\rangle \qquad CP\left|\bar{M}^0\right\rangle = \eta_{CP}^*\left|M^0\right\rangle \qquad (1.12)$$

where η_{CP} has an arbitrary phase. The mass and lifetime eigenstates of the mixing of M^0 and $\bar{M}^0$ are written in full generality as:

$$\left|M_a\right\rangle = p_a\left|M^0\right\rangle + q_a\left|\bar{M}^0\right\rangle \qquad \left|M_b\right\rangle = p_b\left|M^0\right\rangle - q_b\left|\bar{M}^0\right\rangle \qquad (1.13)$$

CPT symmetry requires the composition of each flavour eigenstate to be symmetric in terms of the two physical states hence the ratios of the complex parameters $p_{a,b}$ and $q_{a,b}$ respect the following relation $\frac{q_a}{p_a} = \frac{q_b}{p_b} = \frac{q}{p}$. Adopting a phase convention we obtain the following generic relation for mixing:

$$\left|M_a\right\rangle = p\left|M^0\right\rangle + q\left|\bar{M}^0\right\rangle \qquad \left|M_b\right\rangle = p\left|M^0\right\rangle - q\left|\bar{M}^0\right\rangle \qquad (1.14)$$

with the normalisation $|p|^2 + |q|^2 = 1$. *CP* violation in mixing occurs when the physical states do not correspond to the *CP* eigenstates, *i.e.*

$$\left|\frac{q}{p}\right| \neq 1. \qquad (1.15)$$

1.2.2 *CP* violation in decay

This type of *CP* violation occurs when the decay amplitudes of two *CP* conjugate processes into two generic final states f and $\bar{f}$ differ in modulus. If several amplitudes A_j contribute to the decay then the total amplitude A_f and its *CP* conjugate amplitude $\bar{A}_{\bar{f}}$ can be defined in terms of the weak phase term, $e^{i\phi_j}$, and strong phase term, $e^{i\delta_j}$, of the contributing amplitudes. The convention independent ratio is given by:

$$\left|\frac{\bar{A}_{\bar{f}}}{A_f}\right| = \left|\frac{\sum_j |A_j|\, e^{i(\delta_j - \phi_j)}}{\sum_j |A_j|\, e^{i(\delta_j + \phi_j)}}\right|. \qquad (1.16)$$

CP violation in decay occurs when the physical decay amplitudes for *CP* conjugate processes into final states f and $\bar{f}$ are different in modulus, $\left|\frac{\bar{A}_{\bar{f}}}{A_f}\right| \neq 1$. This requires

the presence of at least two contributing amplitudes with different weak and strong phases. *CP* violation in decay can be observed by comparing the decay rates $\Gamma(P \to f)$ and $\Gamma(\bar{P} \to \bar{f})$, where P is a generic particle decaying to a final state f. The *CP* asymmetry, A_{CP} is then defined as:

$$A_{CP} = \frac{\Gamma(P \to f) - \Gamma(\bar{P} \to \bar{f})}{\Gamma(P \to f) + \Gamma(\bar{P} \to \bar{f})} \tag{1.17}$$

or in terms of the decay amplitudes as follows:

$$A_{CP} = \frac{1 - \left|\bar{A}/A\right|^2}{1 + \left|\bar{A}/A\right|^2}. \tag{1.18}$$

This is the only type of *CP* violation that can occur in both neutral and charged mesons. Charged mesons are forbidden to mix due to conservation of charge and hence exhibit only direct *CP* violation. The measurement of direct *CP* asymmetries is particularly important for this analysis.

1.2.3 *CP* **violation in interference between mixing and decay**

This type of *CP* violation occurs from a phase mismatch between mixing and decay amplitudes for neutral mesons. When dealing with decays into a final state f that can be reached by both flavour eigenstates, a complex quantity is introduced combining the physical states from the mixing and the decay amplitudes in:

$$\lambda_f = \frac{q}{p}\frac{\bar{A}_f}{A_f} = \eta_{f_{CP}}\frac{q}{p}\frac{\bar{A}_{\bar{f}}}{A_f} \tag{1.19}$$

where $\eta_{f_{CP}}$ is the *CP* eigenstate of final state f. Some of the most interesting decays involve final states that are common to B^0 and $\bar{B}^0$. This form of *CP* violation can be observed using the time-dependent asymmetry of neutral meson decays into final *CP* eigenstates, f, given by:

$$A_f(t) = S_f \sin\left(\Delta m_d t\right) - C_f \cos\left(\Delta m_d t\right) \tag{1.20}$$

where Δm_d is the difference in the mass eigenstates of B^0 meson and

$$S_f = \frac{2\mathcal{I}m\left(\lambda_f\right)}{1 + |\lambda_f|}, \qquad C_f = \frac{1 - |\lambda_f|}{1 + |\lambda_f|} \tag{1.21}$$

From the definitions of mixing-induced and direct CP violation, it follows that the coefficient S_f is not zero when there is mixing-induced CP violation, while C_f not equal to zero indicates the presence of direct CP violation.

1.3 New Physics in CP asymmetries

Interesting results were obtained in the measurement of $B \to K\pi$ decays by the B factories where, as the data became more and more precise, phenomenological analyses could not reproduce it. The direct CP violation in the charged $B^\pm \to K^\pm \pi^0$ was observed to be different from its neutral counterpart ($B^0 \to K^+\pi^-$) contradicting current theoretical predictions. This is referred to as the "$K\pi$ puzzle" and, although these measurements are susceptible to strong interaction effects needing further clarifications [1], large deviations in CP violation between these charged and neutral B meson decays could indicate the presence of new sources of CP violation.

1.3.1 The "$K\pi$ puzzle"

The decay $B \to K\pi$ occurs via two major processes: tree and strong QCD penguin (see Figure 1.4). The interference between these two processes leads to differences in the decay amplitudes between B and $\overline{B}$ decays. In the neutral B decay, the B factories observed that the rate of $B^0 \to K^+\pi^-$ is 10% larger than the equivalent antiparticle decay [27]. It is expected that charged B mesons would produce the same asymmetry. However experimental results have shown that the decays $B^\pm \to K^\pm\pi^0$ have asymmetry of opposite sign [28, 29]. This effect is measured with a significance larger than five standard deviations showing that it is indeed real and was therefore referred to as the "$K\pi$ puzzle". Of particular interest is the difference ΔA_{CP} which is defined as:

$$\Delta A_{CP} = A_{CP}(K^+\pi^0) - A_{CP}(K^+\pi^-) \tag{1.22}$$
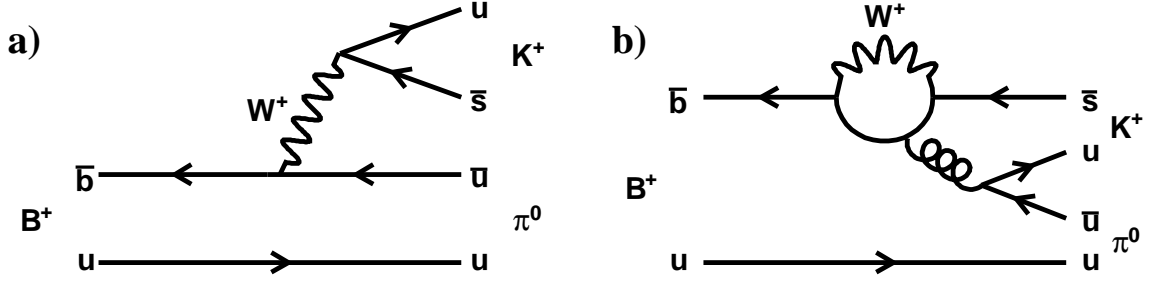
Figure 1.4: *Feynman diagrams for $B \to K\pi$: (a) tree and (b) penguin.*

where $A_{CP}(K^+\pi^0)$ and $A_{CP}(K^+\pi^-)$ are the *CP* asymmetries measured in $B^{\pm} \to K^{\pm}\pi^0$ and $B^0 \to K^+\pi^-$ respectively [30].

Decay amplitudes for $B \to K\pi$ can be described in a model independent way by using the topological contributions involved [2],

$$-A(K^+\pi^-) = V_{tb}^*V_{ts}(P_{tc} + \frac{2}{3}P_{EW}^C) + V_{ub}^*V_{us}(P_{uc} + T), \qquad (1.23)$$

$$A(K^0\pi^+) = V_{tb}^*V_{ts}(P_{tc} - \frac{1}{3}P_{EW}^C) + V_{ub}^*V_{us}(P_{uc} + A), \qquad (1.24)$$

$$-\sqrt{2}A(K^+\pi^0) = V_{tb}^*V_{ts}(P_{tc} + P_{EW} + \frac{2}{3}P_{EW}^C) + V_{ub}^*V_{us}(P_{uc} + T + C + A), \ (1.25)$$

$$\sqrt{2}A(K^0\pi^0) = V_{tb}^*V_{ts}(P_{tc} - P_{EW} - \frac{1}{3}P_{EW}^C) + V_{ub}^*V_{us}(P_{uc} + C), \qquad (1.26)$$

The notation used in the amplitude relations represent each of the following type of processes [31]:

◇ T is the "colour-favoured" tree amplitude associated with the transition $\bar{b} \to \bar{u}u\bar{s}$ where the $u\bar{s}$ forms one pseudoscalar meson and the $\bar{u}$ combines with the spectator quark to form the other,

◇ P_{tc} and P_{uc} are the QCD penguin amplitudes associated with the transitions $\bar{b} \to \bar{s}$ and are defined using the magnitudes of the CKM matrix elements which multiply them, *i.e.* $P_{qc} \equiv \left|V_{qb}^*V_{qs}\right|(\tilde{P}_q - \tilde{P}_c)$ [32],

◇ P_{EW} and P_{EW}^C are the electroweak penguin and the "colour-suppressed" electroweak penguin amplitudes respectively,

◇ C is the "colour-suppressed" tree amplitude with transition $\bar{b} \rightarrow \bar{u}u\bar{s}$ where the $u\bar{u}$ forms the π^0 meson and the $\bar{s}$ combines with the spectator quark to form the kaon,

◇ A is the annihilation process contributing only to charged B decays by means of the exchange of a W boson.

The theoretical results shown in Table 1.1 follow from a diagrammatic approach, which makes use of existing measurements and SU(3) flavour to predict decay rates and asymmetries. This method is based on the principle that a certain hierarchy between amplitudes can exist [34]. The theoretical fit is performed using theoretical input parameters, such as form factors and CKM parameters, and experimental observables. By removing one of the inputs from the fit, a prediction of the corre-

Table 1.1: *Experimental results [33] and theoretical fit predictions for the branching fractions and CP asymmetries for all $B \rightarrow K\pi$ and ΔA_{CP}, obtained using the diagrammatic approach. $C(K_S^0\pi^0)$ and $S(K_S^0\pi^0)$ are the parameters of the time-dependent amplitude in Eq. 1.20. The fit prediction of ΔA_{CP} is obtained by removing both $A_{CP}(K^+\pi^0)$ and $A_{CP}(K^+\pi^-)$ from the fit [1].*

Decay Mode	HFAG average	fit prediction
$\mathcal{B}(K^+\pi^-)$ [$\times 10^{-6}$]	19.4 ± 0.6	19.7 ± 1.0
$\mathcal{B}(K^+\pi^0)$ [$\times 10^{-6}$]	12.9 ± 0.6	12.4 ± 0.7
$\mathcal{B}(K^0\pi^+)$ [$\times 10^{-6}$]	23.1 ± 1.0	24.9 ± 1.2
$\mathcal{B}(K^0\pi^0)$ [$\times 10^{-6}$]	9.8 ± 0.6	8.7 ± 0.6
$A_{CP}(K^+\pi^-)$ [%]	-9.8 ± 1.2	3.9 ± 6.8
$A_{CP}(K^+\pi^0)$ [%]	5.0 ± 2.5	-6.2 ± 6.0
$A_{CP}(K^0\pi^+)$ [%]	0.9 ± 2.5	6.2 ± 4.5
$C(K_S^0\pi^0)$	0.01 ± 0.10	0.10 ± 0.03
$S(K_S^0\pi^0)$	0.57 ± 0.17	0.74 ± 0.04
ΔA_{CP} [%]	14.8 ± 2.8	1.7 ± 6.1

sponding experimental observable is obtained.

An alternative approach obtains the decay rates and CP asymmetries of these decays within the framework of QCD factorisation (QCDF). QCDF formalism allows to compute systematically the matrix elements of the effective weak Hamiltonian for $b \to s$ transitions and therefore extrapolate the decay amplitudes for $B \to \pi K$ and πK^* from first principle. The strength of this method is that it also allows estimates of some suppressed contributions such as the annihilation corrections. The QCDF results together with the current experimental world averages are given in Table 1.2 [35]. As expected in the SM, the CP asymmetry of $B^+ \to K^+\pi^0$ is predicted also by this method to be very close to $B^0 \to K^+\pi^-$ so that $\Delta A_{CP} \approx 1.6$. This prediction agrees with the theoretical fit value of ΔA_{CP} in Table 1.1.

Table 1.2: *Branching ratios (in units 10^{-6}) and direct CP asymmetries (in units 10^{-2}) obtained from the QCDF method [35].*

Decay Mode	BF ($\times 10^{-6}$)	A_{CP} (%)
$B^+ \to K^+\pi^0$	12.5 ± 1.6	-10.8 ± 0.8
$B^0 \to K^+\pi^-$	22.7 ± 3.3	-12.4 ± 0.7

Both of these methods contribute to show that it is very hard to accomodate a large value of ΔA_{CP} in the SM with the methods available for hadron-dynamics in B decays. An explanation for this effect is that other processes that preferentially produce u quarks rather than d quarks might affect the asymmetry, such as electroweak penguins. Alternatively the difference could be due to exotic particles entering the loop diagrams and altering the decay rates of charged B mesons [36]. All of the above amplitudes involve unknown strong phases limiting the accuracy of the prediction for the amplitudes of $B \to K\pi$ decays. Given this, New Physics (NP) contributions to $B \to K\pi$ amplitudes would be easily misidentified.

1.3.2 The "$K\pi$" puzzle in the $K^*\pi$ system

$B \to K^*\pi$ decays via identical tree and penguin processes as $B \to K\pi$ but with different weights associated to the contributing processes. The amplitudes for the $K^*\pi$ decays follow easily from Eqs. 1.23–1.26. Figure 1.5 shows all the processes contributing to $B^+ \to K^{*+}\pi^0$. The method described below uses broken flavour SU(3) and existing measurements to calculate ratios of tree (T) to penguin (P) amplitudes and infer the maximal potential *CP* asymmetry of $B \to K^*\pi$ compared to $B \to K\pi$. The values of these ratios are estimated by relating within SU(3) flavour the amplitudes of these processes to those for $B^0 \to \pi^+\pi^-$ and $B^0 \to \rho^+\pi^-$. This method is presented in detail in Ref. [2]. Table 1.3 shows the dominant terms in the amplitudes of these decay modes, with branching fraction and *CP* asymmetry averages, where the extra terms in the dominant amplitudes of $\pi^+\pi^-$ and $\rho^+\pi^-$ are:

⋄ $\tilde{\lambda} = 0.232$ – a constant term dependent on the Wolfenstein parameter, λ [37],

⋄ two ratios of meson decay constants $f_\pi/f_K = 0.84$ and $f_\rho/f_{K^*} = 0.96$ [38].

An estimate of the amplitude ratios is then given by:

$$\frac{|T_{K\pi}|}{|P_{K\pi}|} \simeq \tilde{\lambda}\left(\frac{f_K}{f_\pi}\right)\sqrt{\frac{r_\tau \mathcal{B}(\pi^+\pi^-)}{\mathcal{B}(K^0\pi^+)}} \qquad (1.27)$$

$$\frac{|T_{K^*\pi}|}{|P_{K^*\pi}|} \simeq \tilde{\lambda}\left(\frac{f_{K^*}}{f_\rho}\right)\sqrt{\frac{r_\tau \mathcal{B}(\rho^+\pi^-)}{\mathcal{B}(K^{*0}\pi^+)}} \qquad (1.28)$$

where r_τ is the ratio of the lifetimes of the charged and neutral B mesons. Including quadratic corrections in the estimates for the strong phase difference between P and T, the following bounds are obtained for the ratios in $K\pi$ and $K^*\pi$:

$$0.09 \leq \frac{|T_{K\pi}|}{|P_{K\pi}|} \leq 0.16 \qquad 0.28 \leq \frac{|T_{K^*\pi}|}{|P_{K^*\pi}|} \leq 0.35 \qquad (1.29)$$

The conclusion is that the ratio of the amplitudes in $B \to K^*\pi$ is between two to three times larger than the corresponding ratio in $B \to K\pi$. The decay $B^+ \to K^{*+}\pi^0$ of interest in this thesis involves interference of penguin and tree amplitudes

"Colour-favoured" Tree (T) "Colour-suppressed" Tree (C)

"Colour-favoured" EW penguin (P_{EW}) "Colour-suppressed" EW penguin (P_{EW}^C)

QCD penguin (P_{qq}) Annihilation (A)
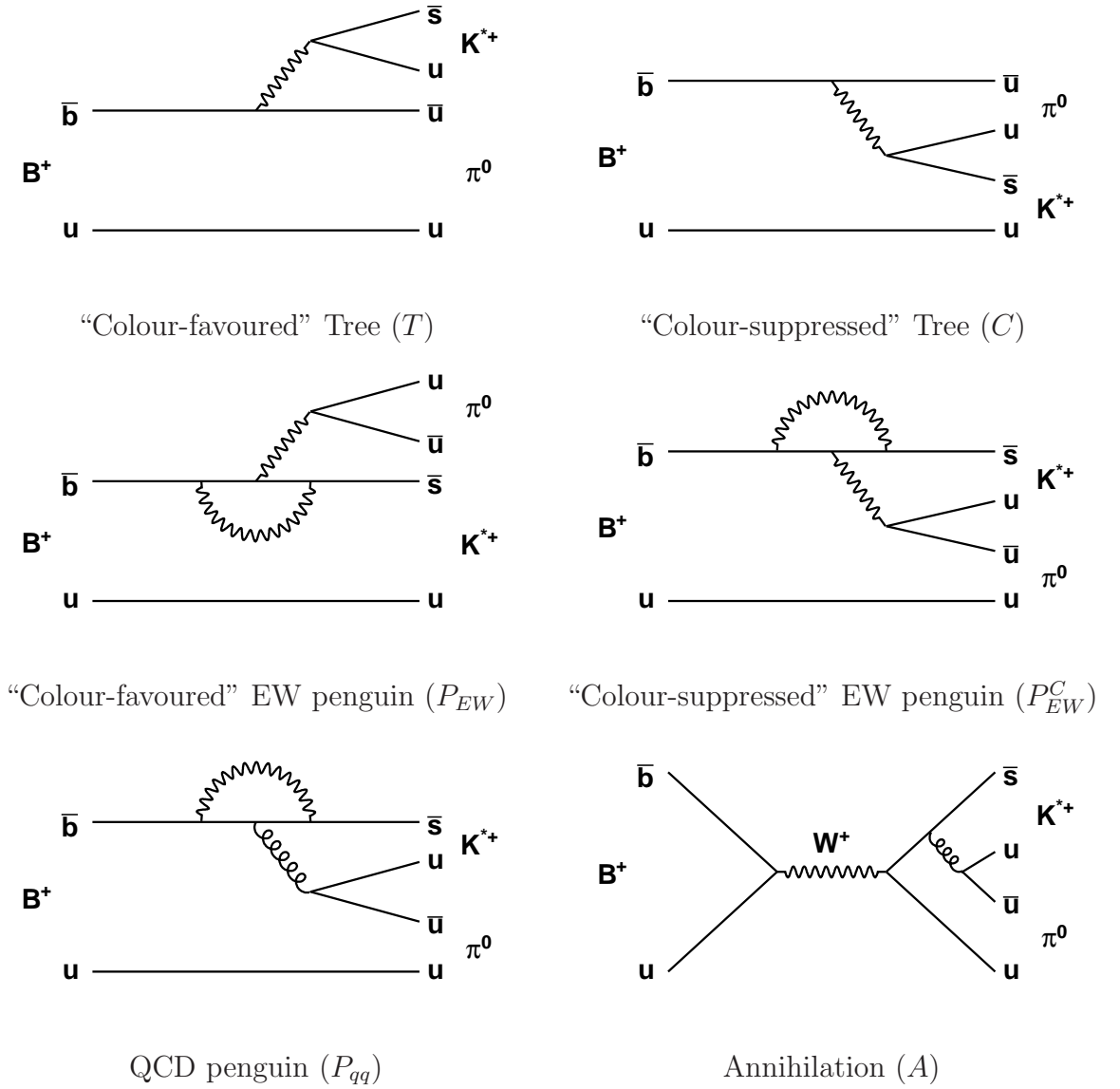
Figure 1.5: *Feynman diagrams for the decay processes in $B^+ \to K^{*+}\pi^0$*

Table 1.3: *Amplitudes, branching fractions and asymmetries for $B \to K\pi$ and $B \to K^*\pi$ modes, including $B^0 \to \pi^+\pi^-$ and $B^0 \to \rho^+\pi^-$. Branching fraction and A_{CP} averages are taken from Ref. [33].*

Mode	Amplitude	$\mathcal{B}(10^{-6})$	A_{CP}
$B^+ \to K^0\pi^+$	P	23.1 ± 1.0	0.009 ± 0.025
$B^0 \to K^+\pi^-$	$-(P+T)$	19.4 ± 0.6	$-0.098^{+0.012}_{-0.011}$
$B^+ \to \pi^+\pi^-$	$\tilde{\lambda}P - \tilde{\lambda}^{-1}T\left(\frac{f_\pi}{f_K}\right)$	5.16 ± 0.22	0.38 ± 0.06
$B^+ \to K^{*0}\pi^+$	P	$9.9^{+0.8}_{-0.9}$	-0.038 ± 0.042
$B^0 \to K^{*+}\pi^-$	$-(P+T)$	$8.6^{+0.9}_{-1.0}$	-0.18 ± 0.08
$B^+ \to \rho^+\pi^-$	$\tilde{\lambda}P - \tilde{\lambda}^{-1}T\left(\frac{f_\rho}{f_{K^*}}\right)$	15.7 ± 1.8	0.11 ± 0.06

hence the asymmetry in this mode could potentially be two to three times larger then the corresponding processes in $K\pi$. Because of the non-negligible width of the K^* resonance, these quasi-two-body modes are best studied via the analysis of the three body decay. Figure 1.6 shows the four $K^*\pi$ decays together with the corresponding three body decays where they can be studied.
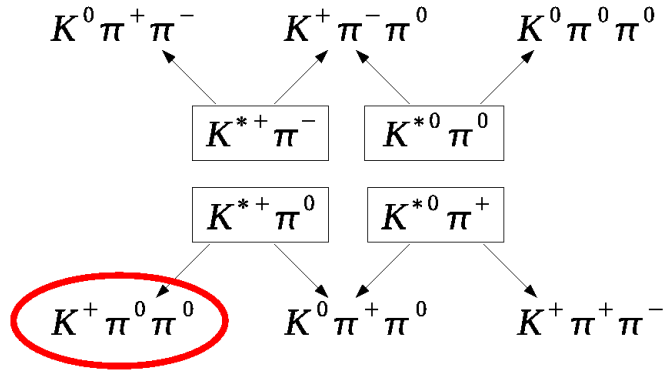
Figure 1.6: *Relation between the four $K^*\pi$ decays and the six $K\pi\pi$ decays.*

1.4 Three-body kinematics

Intermediate resonances in three-body B decays can be distinguished by looking at the phase space distribution of the decay of the three body mode. This section will outline how the phase space distribution of a three body decay is studied and how the Dalitz plot illustrates it.

1.4.1 The Dalitz Plot

The Dalitz plot [39] is a scatter plot showing the distribution of the phase space of a three body decay. Nonresonant decays which proceed only according to phase space will be uniformly distributed within the Dalitz plot (DP). Non-uniform structures in the Dalitz plot indicate the presence of a matrix element dependent on the kinematics of the decay. These are resonances, *i.e.* events that decayed to the same final state via an intermediate particle, and appear as narrow bands in the Dalitz plot distribution. Consider a B meson with mass m_B decaying at rest to three particles with masses m_1, m_2 and m_3, momenta $\vec{p_1}$, $\vec{p_2}$ and $\vec{p_3}$, and energies E_1, E_2 and E_3. If we define the four momentum combinations between daughter particles i and j as $p_{ij} = p_i + p_j$ and $m_{ij}^2 = p_{ij}^2$ then the following relation for the mass combinations with respect to daughter particle k applies:

$$m_{12}^2 + m_{13}^2 + m_{23}^2 = m_B^2 + m_1^2 + m_2^2 + m_3^2 \tag{1.30}$$

where m_{ij}^2 satisfies

$$m_{ij}^2 = (p_B - p_k)^2 = m_B^2 + m_k^2 - 2m_B E_k. \tag{1.31}$$

The orientation of the three particles with respect to one another can be derived from their energies since all three momenta lie in the same plane in the B meson rest frame. The Lorentz invariant phase space for this decay is given by:

$$dN \propto \delta^4 \left(p_B - \sum_{i=1}^{3} p_i \right) \prod_{i=1}^{3} \frac{d^3 p_i}{E_i} = \delta \left(m_B - \sum_{i=1}^{3} E_i \right) \frac{p_1^2 dp_1 p_2^2 dp_2 d\Omega_1 d\Omega_2}{E_1 E_2 E_3} \tag{1.32}$$

Since the B meson is a pseudoscalar, $\int d\Omega_1 = 4\pi$ and $\int d\Omega_2 = 2\pi d\cos\theta_{12}$ with respect to a fixed direction of the momentum $\vec{p_1}$, where θ_{12} is the angle between $\vec{p_1}$ and $\vec{p_2}$. Using $E_3 = \sqrt{p_1^2 + p_2^2 + 2p_1 p_2 \cos\theta_{12} + m_3^2}$ and simplifying using $E_i dE_i = p_i dp_i$ the expression of the decay rate is in fact proportional to the product of the mass via [40]:

$$dN \propto dE_1 dE_2 \propto dm_{12}^2 dm_{23}^2 \tag{1.33}$$

Because of the two identical π^0 in $B^+ \to K^+\pi^0\pi^0$, the only way to distinguish between the $K\pi^0$ combinations is to order them by invariant mass, the result of which is a "folded" Dalitz plot. In what follows, the Dalitz plot is drawn in terms of the variables $(m_{\pi^0\pi^0}^2)$ and $(m_{K^+\pi^0}^2)_{min}$ such as:

$$\left(m_{\pi^0\pi^0}^2\right) = \left(E_{\pi_1^0} + E_{\pi_2^0}\right)^2 - \left(p_{\pi_1^0} + p_{\pi_2^0}\right)^2 \tag{1.34}$$

$$\left(m_{K^+\pi^0}^2\right)_{min} = \left(E_{K^+} + E_{\pi^0}\right)^2 - \left(p_{K^+} + p_{\pi^0}\right)^2 \tag{1.35}$$

where π_1^0 and π_2^0 denotes the first and second π^0 in the event and $m_{K^+\pi_{min}^0}$ is the $K^+\pi^0$ combination with smaller mass. Figure 1.7 shows an example Dalitz plot for $B^+ \to K^+\pi^0\pi^0$ where the expected resonant modes in this analysis are shown with different coloured bands. The different structures of the bands are an indication of the spin of the resonance. The boundaries of the Dalitz plot at a given point along the axis occur when the momenta of the particles on the other axis are parallel or anti-parallel.

1.4.2 The square Dalitz plot

In general resonance masses are much smaller than the B mass, hence signal events tend to populate the boundaries of the conventional Dalitz plot. Due to their jet-like nature, continuum background events also accumulate at the boundaries so that the conventional Dalitz plot becomes inadequate when using histograms to describe background shapes or other experimental effects. In this analysis, the square Dalitz plot (SDP) is used instead of the conventional DP to describe the variations of

efficiency and levels of misreconstruction. The conventional DP coordinates are transformed into the SDP coordinates as follows:

$$d \left(m^2_{K^+\pi^0} \right)_{min} d \left(m^2_{\pi^0\pi^0} \right) \rightarrow |\det J| \, dm' d\theta' \tag{1.36}$$

where J is the Jacobian of the transformation,

$$m' \equiv \frac{1}{\pi} \arccos \left(2 \frac{m_{\pi^0\pi^0} - m^{min}_{\pi^0\pi^0}}{m^{max}_{\pi^0\pi^0} - m^{min}_{\pi^0\pi^0}} - 1 \right) \tag{1.37}$$

$$\theta' \equiv \frac{1}{\pi} \theta_{\pi^0\pi^0} \tag{1.38}$$

where $\theta_{\pi^0\pi^0}$ is the angle between the kaon and the one of the neutral pions in the $\pi^0\pi^0$ rest frame, $m^{max}_{\pi^0\pi^0} = m_B - m_{K^+}$ and $m^{min}_{\pi^0\pi^0} = 2m_{\pi^0}$ are the boundaries of the invariant mass $m_{\pi^0\pi^0}$. The new variables have validity between 0 and 1, except in this case the range of θ' is 0 to 0.5 as a consequence of the "folded" Dalitz plot described in Section 1.4.1 [41]. An example of a SDP is shown in Figure 1.7.
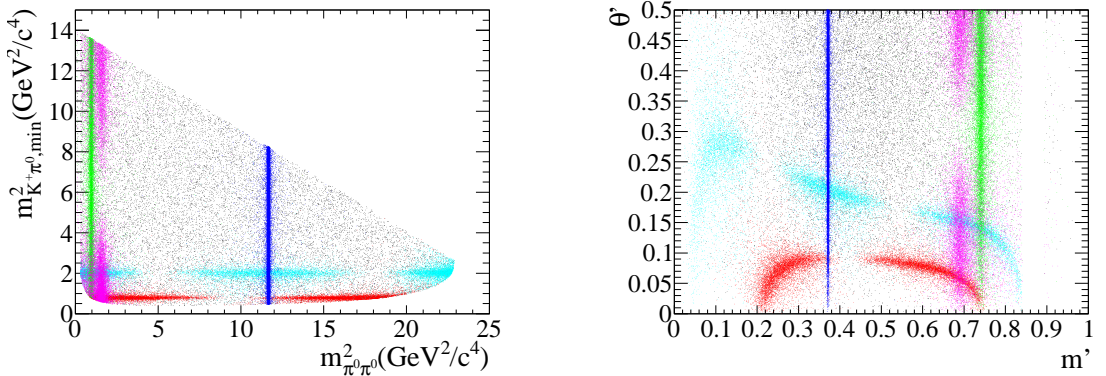
Figure 1.7: *Illustration of a sample Dalitz plot (left) and square Dalitz plot (right) obtained from toy MC events and showing $B^+ \rightarrow K^+\pi^0\pi^0$ non resonant (black) and the resonances $B^+ \rightarrow K^*(892)^+\pi^0$ (red), $B^+ \rightarrow K^{*+}_2(1430)\pi^0$ (cyan), $B^+ \rightarrow f_0(980)K^+$ (green), $B^+ \rightarrow f_2(1270)K^+$ (magenta) and $B^+ \rightarrow \chi_{c0}K^+$ (blue).*

2

The *BABAR* Experiment

The primary goal of the *BABAR* detector is the study of *CP* asymmetries in the neutral B meson system. In addition to this a number of rare B meson decays may be measured enabling constraints to be put on fundamental parameters of the Standard Model. To achieve its purpose the *BABAR* detector was designed to:

⋄ reconstruct the exclusive final states of the decay under study,

⋄ tag the flavor of the decaying particle (*i.e.* beauty or anti-beauty),

⋄ measure the time of the B^0 decay with respect to its production.

This chapter outlines the design of the PEP-II B Factory and the *BABAR* detector showing how both contribute to achieving the physics goals of the experiment.

2.1   The PEP-II accelerator

This section briefly describes the PEP-II collider. For more technical details, refer to Ref. [42]. The PEP-II B Factory is an asymmetric e^+e^- collider operating at the centre-of-mass energy of 10.58 GeV corresponding to the mass of the $\Upsilon(4S)$ resonance. This resonance decays almost exclusively to both $B^0\overline{B}^0$ and B^+B^- pairs providing an ideal laboratory for the study of B mesons. A schematic of the overall layout of the PEP-II collider is shown in Figure 2.1.
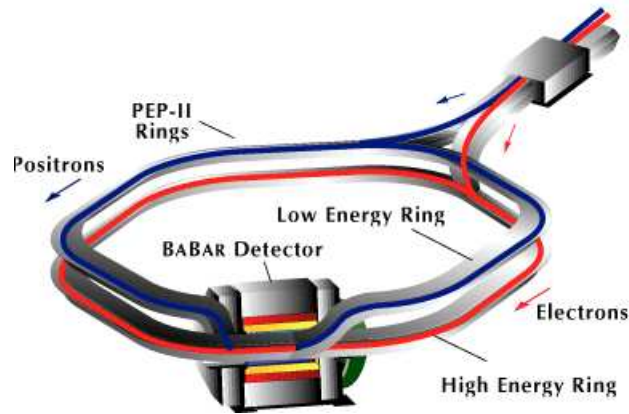
Figure 2.1: *A schematic of the PEP-II rings and the collision region. The blue ring contains the positron beam and the red ring the electron beam [43].*

PEP-II consists of two rings, one containing a 9 GeV electron beam (HER) and the other a 3.1 GeV positron beam (LER) injected into the tunnel by the SLAC linac. The two beams collide head-on in the experimental hall and the boost provided by the asymmetric beams makes it possible to reconstruct the decay vertices of the two B mesons with enough accuracy to determine the relative decay time needed for time dependent measurements.

In order to achieve the high luminosities that the physics programme requires the beams are divided into a large number of low charge bunches (~ 1500) which minimises beam-beam interference. To avoid secondary collisions, the beams are initially horizontally displaced from one another. The HER and LER are then both focused using sets of iron magnets residing just outside the detector volume. These are la-
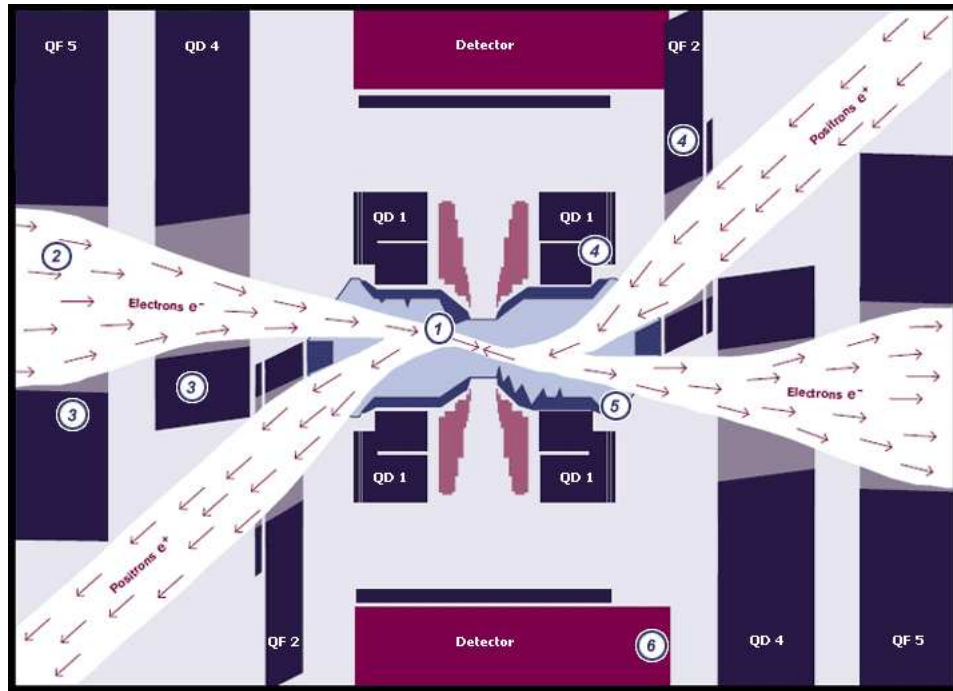
Figure 2.2: *Schematic of the PEP-II interaction region. The pink areas around the interaction point represent the dipole magnets used to bring the beams together and the regions with a Q label indicate the positions of the various quadrupole magnets. (Graphic source: SLAC Accelerator Systems Division via Ref. [44])*

belled in Figure 2.2 as QF5 and QD4 for the HER and QF2 for the LER. Finally both beams are brought together towards the interaction point (IP) by a set of permanent quadrupole magnets labelled QD1. These magnets reside just within the *BABAR* detector volume and hence affect detector acceptance and background conditions. Just after the IP, the outgoing beams are separated again using the same set-up of iron magnets on the opposite side of the IP.

PEP-II was operational from October 1999 to April 2008. Until 2004, PEP-II operated by filling both beams and continuing collisions until the instanteneous luminosity reached a certain lower limit. The detector was then ramped down whilst the beams were topped up again by injection from the SLAC linac. This method kept machine induced backgrounds due to injection low but it could only achieve
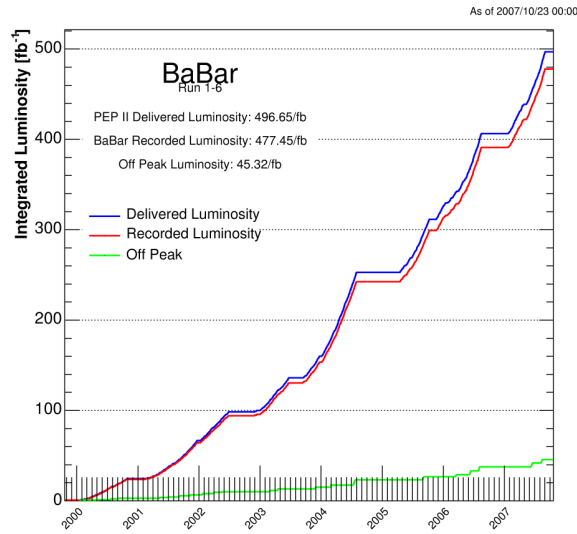
Figure 2.3: *Luminosity distribution over the experiment running period [47].*

$70 - 75\%$ of the peak luminosity. From 2004 onwards trickle injection was used (see Refs. [45,46]). This method consisted of injecting continuously without needing to ramp down the detector. This lead to greater efficiency of luminosity delivered but with the downside that machine induced backgrounds were also increased. Figure 2.3 shows the integrated luminosity of the experiment throughout its running period.

2.2 The BABAR detector

A more complete description of the BABAR detector can be found in Ref. [48]. The need for full reconstruction of final states with multiple charged particles and several π^0 mesons places stringent requirements on the detector. The most important features required to obtain good statistics in this measurement are:

⋄ a large and uniform acceptance down to small polar angles relative to the boost direction,

⋄ excellent reconstruction efficiency for charged particles and for photons,

◇ very good momentum measurement and vertex resolution to separate small signals from background,

◇ excellent energy and angular resolution for the detection of photons from π^0 decays in the range 20 MeV to 4 GeV,

◇ efficient and accurate identification of hadrons over a wide range of momenta.

To maximise the geometric acceptance for the boosted $\Upsilon(4S)$ decays, the whole detector is offset relative to the PEP-II beam-beam interaction point by 0.37 m in the direction of the lower energy beam. Figure 2.4 shows the longitudinal cross-section and end view of the *BABAR* detector around the beam line with dimensions. Since the average momentum of charged particles is less than 1 GeV/c, the precision of measured tracks is severely impacted by multiple Coulomb scattering with the material within the detector. Thus the material in the active volume of the detector was kept to a minimum to not compromise the performance of either tracking system or calorimeter. The inner detector consists of a silicon vertex tracker (SVT), a drift chamber (DCH), a ring imaging Cherenkov detector (DIRC) and a CsI calorimeter (EMC). The SVT and DCH form the main tracking system for charged particles, while hadron identification is provided by the DIRC. The EMC is designed to identify electromagnetic showers allowing detection of low energy π^0 mesons. These systems are surrounded by a superconducting solenoid designed to achieve a field of about 1.5T. Finally the instrumented steel flux return (IFR) is primarily used for the detection of muons and neutral hadrons. All these systems together contribute to achieve the requirements set by the physics goals of the experiment.

2.3 The Silicon Vertex Tracker (SVT)

The SVT has been designed to provide precise measurements of charged particle trajectories and decay vertices near the interaction region. This is achieved by precise measurements of the momenta and angles at which particles cross the detector,
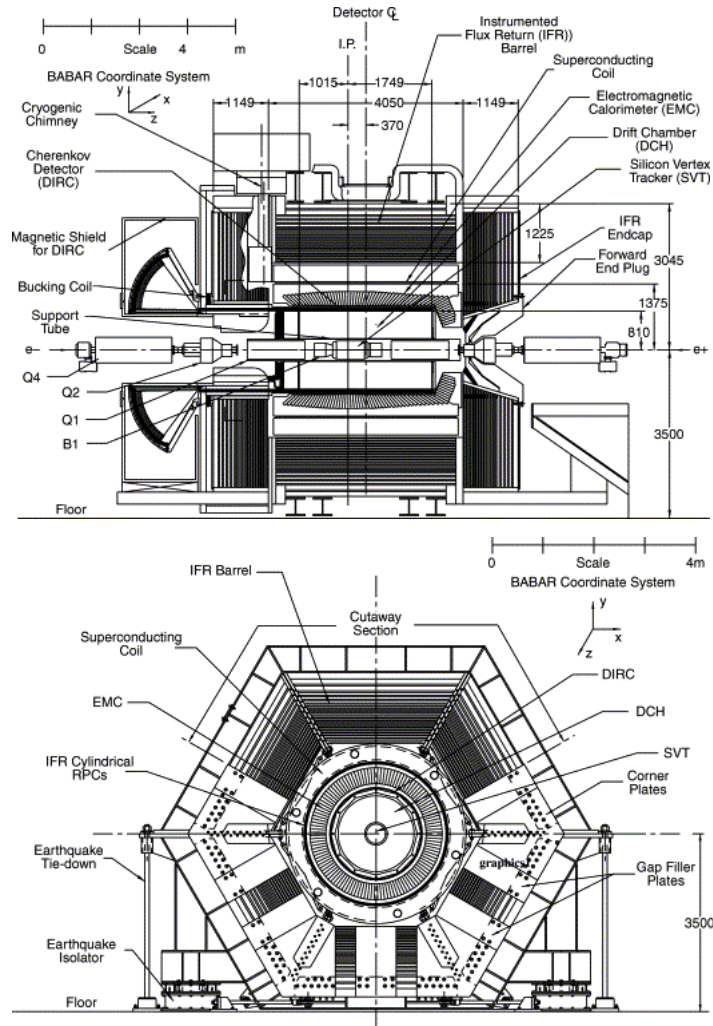
Figure 2.4: *Longitudinal section and front end view of the BaBar detector. [48]*

allowing for the reconstruction of charged B and D meson decay vertices with high resolution and low background. Due to its proximity to the interaction region, the SVT was designed to withstand 2M rad of ionising radiation. To be sensitive to the decay products of the B mesons with low transverse momentum, the SVT must provide standalone tracking for particles with transverse momentum below 120 MeV/c since this is the minimum momenta that can be reliably measured by the DCH. This feature together with the need to link the tracks reconstructed by the SVT to the tracks reconstructed by the DCH was crucial in chosing the number of silicon layers

needed for the detector.

The SVT is composed of five layers of double-sided silicon strip sensors enclosed in a carbon fibre frame with read-outs at each end. They are built on high-resistivity n-type substrates with p^+ and n^+ strips on two opposite sides. Typical depletion voltages are in the range $25 - 35V$ and inter-strip resistance is achieved at about $10V$ above the depletion voltage. Strips on opposite sides of the sensor are oriented othogonally to each other. The ϕ measuring strips run parallel to the beam whereas the z measuring strips are oriented transversely to the beam axis. Figure 2.5 shows the fully assembled SVT with visible silicon sensors of the outer layer and a transverse schematic view. The modules of the three inner layers (Layers 1,2 and 3 in Figure 2.5 (b)) are straight and slightly tilted in ϕ to allow overlapping and thus provide full azimuthal coverage. The last two outer layers are arch-shaped and designed such as to minimise the amount of silicon needed to cover the solid angle while increasing the crossing angle of particles near the edges. The outer modules cannot be tilted in ϕ like the inner modules because of their geometrical shape. To achieve a suitable overlap in the ϕ coordinate, the two outer modules were split into two sub-layers each (Layers 4a, 4b, 5a, 5b in Figure 2.5 (b)) and placed at different
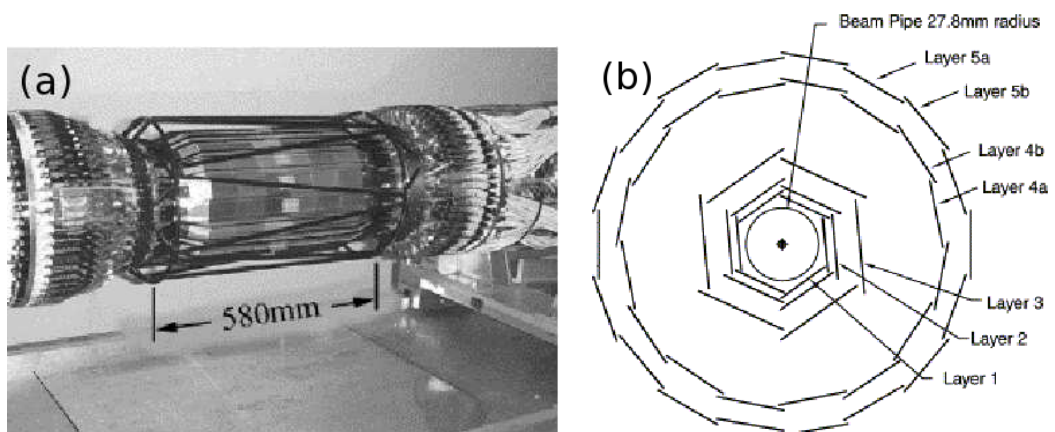
Figure 2.5: *The SVT (a) fully assembled with visible outer layers and carbon fibre frame and (b) schematic view of the transverse section with the various layers around the beam pipe [48].*

radii.

The alignment of the SVT is performed in two stages: the first step consists in determining the relative position of all the silicon sensors and the second aligns the SVT as a whole with respect to the global coordinate system of the DCH. The SVT has a combined hit reconstruction efficiency of about 97%. The geometrical acceptance is 90% of the solid angle in the centre of mass frame, of which 86% is used in charged particle tracking. The SVT is also used to measure energy losses, dE/dx, deposited in the sensors by the passing particles to provide an extra measurement of particle identification. The double sided sensor provides up to ten measurements of dE/dx per track. The average dE/dx gives a 2σ separation between kaons and pions up to momentum of $500\,\mathrm{MeV}/c$ and between kaons and protons up to $1\,\mathrm{GeV}/c$.

2.4 Drift Chamber (DCH)

The DCH was also designed to measure charged particles' momenta and angular distributions. It complements the measurements of parameters and directions of charged tracks obtained by the SVT near the interaction region but provides the sole information for all long lived particles that decay outside the SVT volume. At lower momenta the DCH measurements dominate the errors for the extrapolation of tracks to the DIRC, EMC and IFR. The DCH is able to measure not only transverse momenta and positions but also the longitudinal position of tracks. The longitudinal cross-section of the DCH with dimensions can be found in Figure 2.6.

The DCH is made up of low-mass aluminium field wires and a helium-based gas mixture that minimises multiple scattering within the detector. This gas mixture has a radiation length that is five times larger than commonly used argon-based gases. The wires are disposed such as to form 40 layers of small hexagonal cells providing up to 40 spatial and ionisation loss measurements for charged particles with transverse momentum greater than $180\,\mathrm{MeV}/c$. These cells form circular layers around the
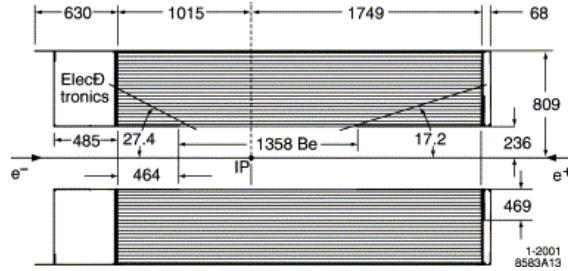
Figure 2.6: *Longitudinal cross section of the DCH with the principal dimensions in* mm *and offset with respect to interaction point (IP) [48].*

axis of the drift chamber and a group of four layers form so called "superlayers". Longitudinal position is measured by placing the wires in six superlayers at small angles with respect to the z axis. The superlayers alternate from axial (A) to positive and negative stereo (U,V) in the following order: AUVAUVAUVA. The angle of each stereo layer increases from 45 mrad in the innermost layer to 76 mrad in the outermost layer. The inner cylindrical wall was kept thin to facilitate the matching of tracks from the SVT to the DCH and reduce the background from photon conversions and interactions. Material in the outer wall was also minimised to avoid degrading the performance of the DIRC and EMC. The DCH is bound by a support tube at its inner radius and by the DIRC at its outer radius. The readout electronics, at the endplate of the chamber, is designed to provide measurements of the drift time and integrated charge for every wire with a signal. The total charge deposited in the drift cells provides a measurement of the energy loss dE/dx like the SVT. Figure 2.7 shows the corrected dE/dx measurements as a function of track momenta with superimposed the Bethe-Bloch predictions for particles of different masses as determined from selected control samples.
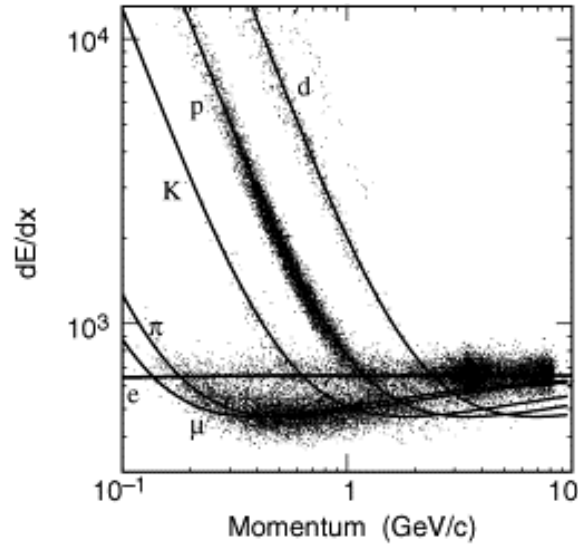
Figure 2.7: *Measurement of the average* dE/dx *as a function of track momenta from the DCH. The curves superimposed to the data show the Bethe-Bloch predictions for energy loss of a sample of particles of different masses [48].*

2.5 The Detector of Internally Reflected Cerenkov Light (DIRC)

Kaon and pion separation is essential for reconstructing B decays, like $B^+ \rightarrow K^+\pi^0\pi^0$ and for tagging particle flavour. The lab frame momenta can range between 1.7 GeV/c and 4.2 GeV/c with strong momentum-polar angle correlations, hence the DIRC is designed to provide particle identification (PID) and π/K separation of at least 4σ for all tracks from the pion Cherenkov threshold up to 4.2 GeV/c. Its principle is based on the magnitudes of angles being preserved after reflection upon a flat surface. The variable θ_c represents the Cherenkov angle and is calculated from the following relation:

$$\cos\theta_c = \frac{c}{nv} \qquad (2.1)$$

where v is the velocity of the particle, c the speed of light and n the refractive index of fused silica.
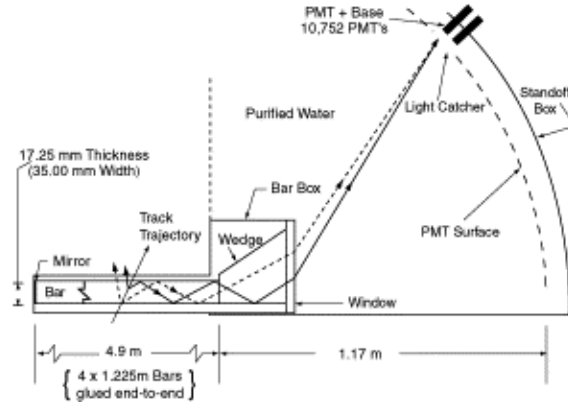
Figure 2.8: *Schematics of the DIRC fused silica radiator bar [48].*

The DIRC is a three-dimensional imaging device, that makes use of the position and arrival time of the signal at an array of densely packed photomultiplier tubes (PMTs). It consists of a radiator material made of synthetic, fused silica bars that serve both as radiators and as light pipes for the light trapped in the radiator by total internal reflection. A schematic of the DIRC geometry illustrating the principle of light production, transport and imaging is shown in Figure 2.8. Photons generated by particles above the Cherenkov threshold are trapped inside the bars and emerge into a water-filled expansion region, called a *standoff box*. A fused silica wedge is used to reflect photons at large angles to reduce the size of the required detection surface and hence recover those photons that would be lost due to internal reflection at the fused silica and water interface. The photons are then detected by the PMTs. In order to associate the photon signal with the track that emitted it, the vector pointing from the centre of the bar end to the centre of each PMTs is measured in terms of the photon propagation angles. Using the track positions and angles measured by the tracking system described in Section 2.3 and Section 2.4, the photon propagation angles can be used to determine the Cherenkov angles θ_c. These are then used to identify the type of particle. Figure 2.9 shows the K/π separation as a function of track momentum. A reduction in separation power is observed as the momentum of a track increases. This is due to a reduction in the average separation in the Cherenkov angle for the different particles hypothesis, increasing
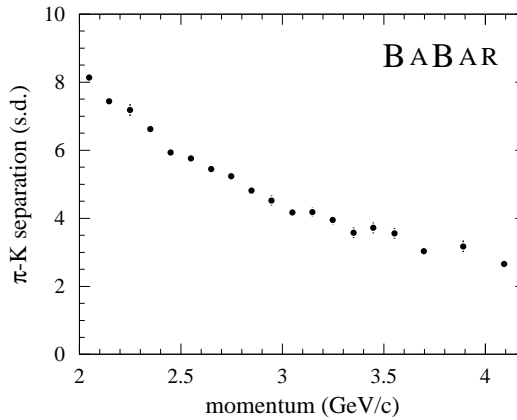
Figure 2.9: *Expected K/π separation as a function of track momentum [48].*

the ambiguity in K/π separation [49].

2.6 Electromagnetic Calorimeter (EMC)

The EMC is designed to measure electromagnetic showers, energy and angular distributions over the energy range from 20 MeV to 9 GeV. This allows for the detection of π^0 mesons, which decay 99% of the time to two photons (see Ref. [38]), as well as electromagnetic and radiative processes. Most of the photons produced by neutral pion decays have energies below 200 MeV hence the lower energy bound is set to allow reconstruction of B decays containing multiple π^0 mesons, like the decay studied in this analysis, with high efficiency. The higher bound is necessary for calibration and luminosity monitoring of photons produced via the process $e^+e^- \rightarrow e^+e^-\gamma$. The EMC is also used to identify electrons which allows the study of rare B, D mesons and τ lepton decays.

The EMC consists of a cylindrical barrel and a conical forward endcap. This fully covers the azimuthal range and extends from 15.8^o to 141.8^o in polar angle, which corresponds to a solid angle coverage of 90% in the centre of mass frame. The barrel contains 5760 thallium doped caesium iodide crystals (CsI(Tl)) arranged in
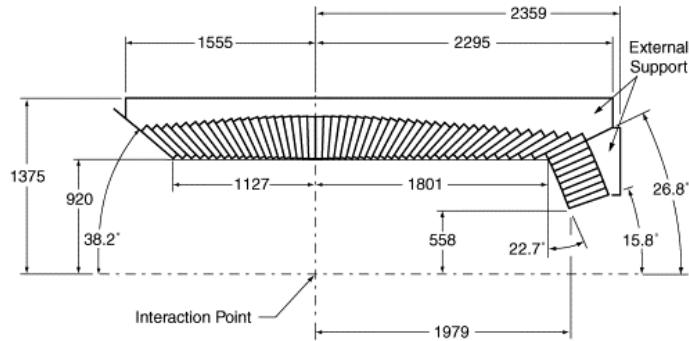
Figure 2.10: *A longitudinal cross-section of the EMC indicating the arrangements of the 56 crystal rings. All dimensions are given in* mm *[48].*

48 rings with 120 identical crystals each. The endcap is made up of 8 rings holding 820 crystals. Figure 2.10 shows the longitudinal cross section of the EMC with the arrangement of the 56 crystal rings. Reducing material in front of the EMC is essential to minimise pre-showering. To achieve this, the crystals are supported at their outer radius, with only a thin gas seal at the front. In addition the DIRC was designed in such a way as to minimise the material in front of the calorimeter. The crystals are read out with silicon photodiodes that are matched to the spectrum of scintillation light. The photodiodes have a quantum efficiency of 85% for the CsI(Tl) scintillation light. Each photodiode is connected to a low-noise preamplifier to shape and filter the signal and remove high and low frequency noise components.

A typical electromagnetic shower spreads over many adjacent crystals, forming a cluster of energy deposits. Pattern recognition algorithms are used to differentiate between single clusters with one energy maximum and merged clusters with more than one local energy maximum. These energy maximums are referred to as bumps. Signal clusters are required to contain at least one crystal with energy above 10 MeV and surrounding crystals are considered as part of the cluster if their energy exceeds a threshold of 1 MeV. An iterative algorithm is used to determine the energy of the bumps by giving each crystal a weight, w_i, which is dependent on the number of bumps in the cluster. The bump energy is then defined as $E_{bump} = \sum_i w_i E_i$ where the sum runs over all crystals in the cluster. A bump is then associated with a

charged particle if the angle and momentum of the track projected onto the inner face of the calorimeter is consistent with the distance between the track impact point and the bump centre. If no track is found to be associated to a bump then it is associated with a neutral particle. The transverse shower shape is used to identify the type of particle at the origin of the energy deposition. The shape can be quantified using the lateral moment (LAT) of the shower cluster produced by the crystal hits in the EMC. This is defined as the ratio of the sum of all the energies excluding the two most energetic crystals weighted by the square of the distance to the centre of the cluster, r, and the sum of all energies including the two most energetic crystals weighted by the square of the length scale of a crystal, l, as shown in Eq. 2.2.

$$\texttt{LAT} = \frac{\sum_{i=3}^{N} E_i r_i^2}{\sum_{i=1}^{N} E_i l^2} \tag{2.2}$$

Electromagnetic showers tend to deposit a large proportion of energy in one or two crystals and therefore have LAT values close to zero whereas hadronic showers tend to be more spread out and hence have LAT values close to one [50].

Calibration of the EMC is performed using the energy resolution of the detector measured at different energy thresholds. Below 2 GeV, the mass resolution of π^0 mesons decaying into two photons of approximately equal energies is used to measure the EMC energy and angular resolution.

2.7 Instrumented Flux Return (IFR)

The IFR was designed to identify muons and detect neutral hadrons with high efficiency and over a wide range of momenta and angles. Muons are important to tag the flavor of neutral B mesons and for the study of rare decays involving leptons. To achieve these goals, the IFR needs to have a large solid angle coverage and high background rejection for muons down to momenta below 1 GeV/c.

The IFR uses the steel flux return of the magnet as a muon filter and hadron
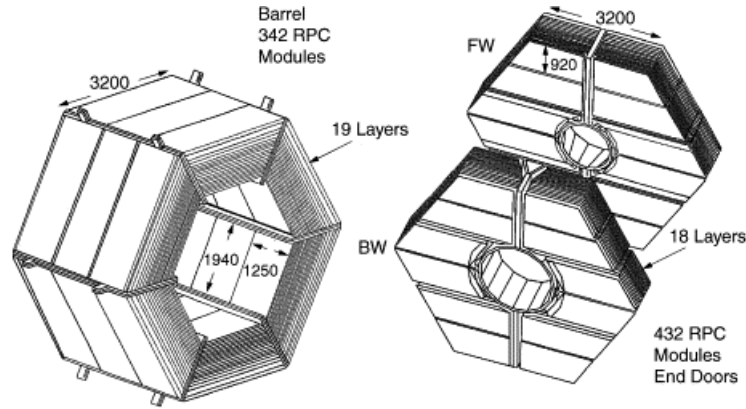
Figure 2.11: *Overview of the IFR showing barrel sectors and forward and backward end doors. The shape and dimensions of the RPC modules are also shown. [48]*

absorber. It consists of 19 Resistive Plate Chambers (RPC) layers in the barrel, 18 in a forward and backward endcaps and 2 between the EMC and the magnet cryostat to detect particles exiting the EMC. The IFR barrel and endcaps structures are shown in Figure 2.11. The RPCs are installed in the gaps of the finely segmented steel of the end doors of the flux return and are used to detect left over charge from ionising particles via capacitive readout strips. Each barrel module contains 32 strips running perpendicular to the z coordinate and 96 strips in the orthogonal direction to measure ϕ. The endcaps modules have strips running in the x and y directions providing full three-dimensional position information.

Over the first running period, it was seen that the muon efficiency was degrading substantially in many RPCs. The RPCs were showing signs of overheating but after additional cooling was installed, the RPCs continued to deteriorate until eventually some of them were "dead", *i.e.* less than 10% efficient. These were later replaced partially by brass and the remainder by Limited Streamer Tubes (LST) which were shown, after extensive testing, to be more practical and reliable than RPCs. The LSTs are made of a cathode ray tube with an anode wire strung at both ends and are coated in a resistive layer of graphite. Figure 2.12 shows cross-sections for an RPC and LST. The signal for the measurement of one coordinate is then either
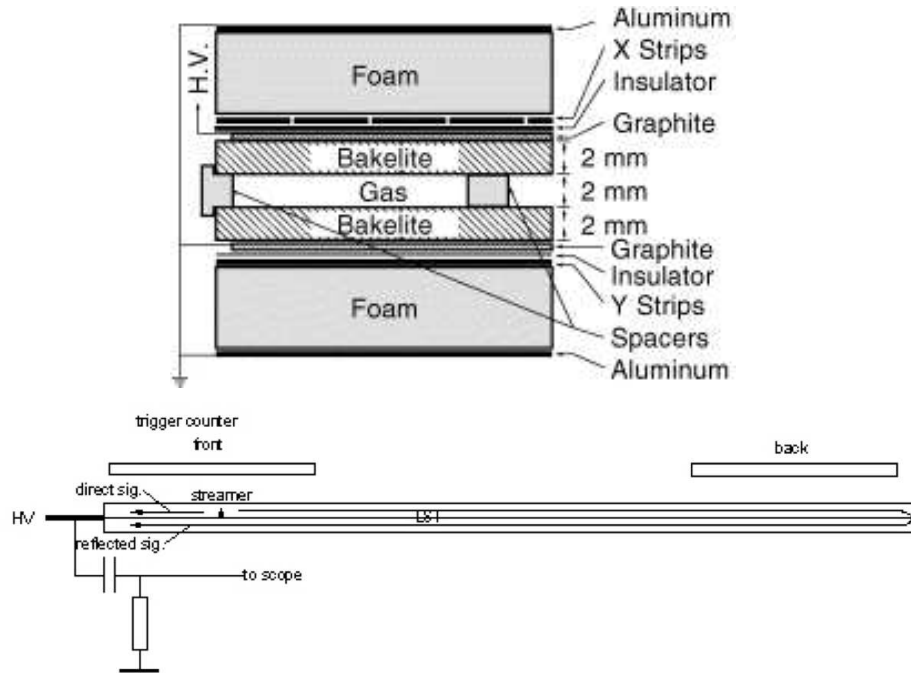
Figure 2.12: *Schematic view of an RPC (top) and LST (bottom) cathode tube [51].*

read directly from the wires or by external strips attached to both sides. For more information on the LST detector upgrade see Ref. [51].

The performance of muon identification was determined from control samples of muons and pions gathered from data. Neutral hadrons interact with the steel of the IFR and are identified as clusters not associated with a charged track. Since a significant fraction of hadrons interact before reaching the IFR, the information from the EMC, detailed in Section 2.6, and the information from the cylindrical RPC layers between the EMC and the magnet is combined with the IFR cluster information. Neutral showers in the EMC are associated with IFR clusters by matching the production angles.

2.8 Trigger and data acquisition

The basic requirement of the trigger system is the selection of events of interest with high, stable and well-understood efficiency while rejecting beam-induced backgrounds. The total trigger efficiency is required to be greater than 99% for all $B\overline{B}$ events and at least 95% for continuum events. Less stringent requirements apply to other event types, such as $\tau^+\tau^-$ events, which only need between $90-95\%$ efficiency depending on specific channels.
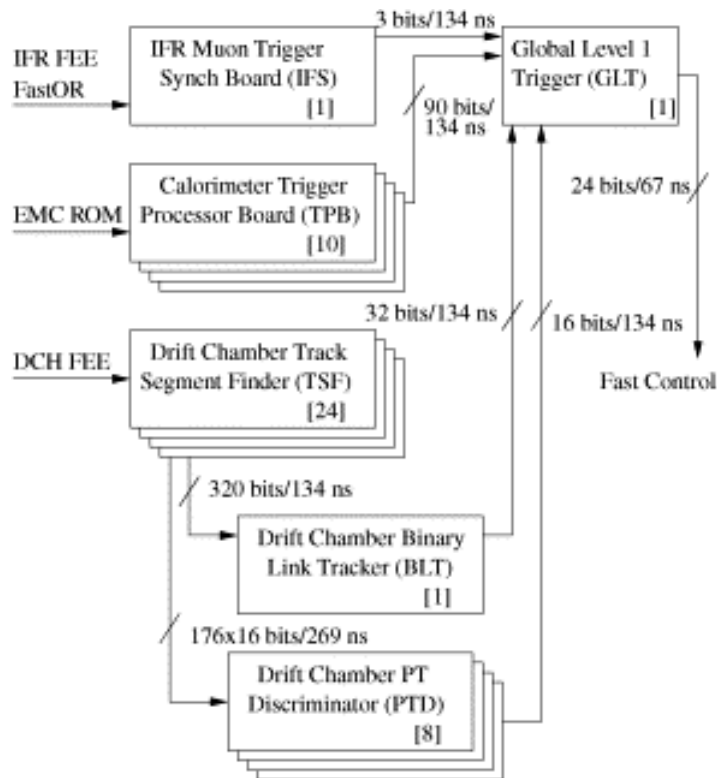
Figure 2.13: *Simplified L1 trigger schematic. Indicated in the figure are the number of components and the transmission rates between them in terms of signal bits. [48]*

The trigger system is implemented in two stages: the Level 1 (L1) trigger in hardware followed by the Level 3 (L3) in software. The L1 trigger data is processed by three different hardware structures: the DCH trigger (DCT), the EMC trigger (EMT) and the IFR trigger (IFT). Each L1 trigger receives information indicating the presence

of a track, calorimeter energy deposit or IFR cluster from their parent system and produces a set of trigger variables which summarise the data in terms of position and energy of the particles. The DCT identifies tracks down to a transverse momentum of 120 MeV/c and the EMT works with energy deposits for each crystal above a threshold of 20 MeV. The IFT only requires single clusters or back-to-back events to detect cosmic rays for calibration and $\mu^+\mu^-$ events. As shown in Figure 2.13, these trigger variables are then sent to the global trigger (GLT). If an event is accepted, a L1 *Accept* is issued to initiate the event readout from all subsystem buffers. The L3 trigger software includes event reconstruction and classification. The software has access to the complete event data information including the output of the L1 trigger processors. All events passing L3 are then transferred to the Offline Prompt Reconstruction (OPR) farms. Information on detector conditions such as temperature, voltages, gas supply and humidity are extracted from the Online Detector Control systems (ODC). This information is placed in the condition database for calibration and later use in the event reconstruction by the OPR farms.

3

Analysis Techniques

This chapter describes the various tools and techniques used in this analysis. The high luminosity provided by PEP-II and the complex features of the *BABAR* detector mean that the volume of recorded data is immense. MC simulations are used to construct a model for the decay in study and take into account not only the decay models but all experimental issues such as machine backgrounds and calibration effects. Charmless B decays typically have fairly small branching fractions. In addition to this, this analysis has the inherent difficulty of having to identify two π^0 mesons in the final state. Neutral particles are particularly easy to misidentify. Hence a small number of signal events are expected and need to be separated from a potentially large number of background events. A set of discriminating variables is used to isolate these background events.

3.1 Data sample and Monte Carlo simulation

This analysis makes use of a bias-free method, commonly known as "blind" analysis. In a blind analysis the physics result, *i.e.* the signal region, is hidden until the fit model is validated. The major advantage of a blind analysis is to minimise the potential for experimenter's bias in the result [52]. In a blind analysis, the selection and fitting model are constructed and tested using Monte Carlo (MC) simulation to simulate the signal region and background contributions from other $B\overline{B}$ decays. Data taken below the $\Upsilon(4S)$ resonance, also known as "offpeak" data, is used to estimate the continuum background, *i.e.* background from other $q\overline{q}$ events. The MC simulation process is split in two stages: the event generation and the detector response to the passage of particles through the sub-detectors, with subsequent particle decays and electronic response.

3.1.1 Event Generators

Two different event generators are commonly used in *BABAR* simulations. The `EvtGen` package simulates the decays of B mesons, other particles and resonances. The `Jetset7.4` package is used to simulate generic continuum events and B decays for which `EvtGen` does not have an implementation. A typical B decay simulation contains approximately 50% exclusive final states while the other 50% are produced by `Jetset7.4`.

3.1.2 Full Detector Simulation

This is handled by a framework making extensive use of the `GEANT`4 package. It provides tools to construct the detector geometry and to simulate the full variety of interactions and decays from charged and neutral particles passing through the detector. Each subsystem-specific package consists of a standard set of routines called by the framework package at various stages of the simulation. These routines

contain information on materials, shapes, positions of sensitive and insensitive areas of the subdetector components. Each particle emits so called `GHits` when interacting with the active components of the detector. A `GHit` contains all information needed to simulate the detector response to the passage of the interacting particle such as position, time of flight, angle of track or cluster, energy lost and Monte Carlo track number together with other subsystem-specific quantities. At the end of the simulated event, all subsystem `GHits` are written to an output file.

The next stage is a simulation of the processing of the detector signals through the front-end electronics and dataflow through the data acquisition system. The information on the response of the trigger system is also included allowing determination of when an event would be triggered on and stored. During data taking the trigger takes samples between each bunch crossings where no physics event is taking place, so that they represent a good sample of the background conditions in the detector. These samples are overlaid on the simulated data to create the full simulated event [50, 53].

An accurate MC description of the inclusive signal decay is achieved by looking at nonresonant together with a number of resonant MC decaying to the final state $K^+\pi^0\pi^0$ via an intermediate particle. These MC samples, listed in Table 3.1, have been produced with large quantities of events. The final stage of the MC simulation is the reconstruction which shares the same reconstruction tools as data. Except for the additional truth matching information, MC events are treated in the same way as data events by the analysis package.

3.1.3 Data sample

The data sample in this analysis comprises the full *BABAR* dataset ("R24a3-v03"), *i.e.* Run 1 to 6, and contains a total of 429.0 fb^{-1} of onpeak data, *i.e.* data taken at the $\Upsilon(4S)$ resonance energy. As mentioned before offpeak data is also used in the blind stage of the analysis. Table 3.2 shows onpeak and offpeak luminosities as

Table 3.1: *List of nonresonant and resonant MC modes. The "SP" followed by the mode number is a unique identifier for each signal decay.*

Decay Mode	SP mode number	Generated events (10^3)
$K^+\pi^0\pi^0$(NR)	1934	860
$K^*(892)^+\pi^0$	1941	860
$K_2^{*+}(1430)\pi^0$	5061	217
$K^{*+}(1680)\pi^0$	5065	217
$K^{*+}(1410)\pi^0$	5066	217
$f_2(1270)K^+$	5067	217
$f_0(980)K^+$	5068	217
$K_0^{*+}(1430)\pi^0$	5069	217
$\chi_{c0}K^+$	9504	217

recorded for each running period.

The number of $B\bar{B}$ pairs in the dataset is essential to calculate branching fractions and needs to be determined as accurately as possible. In *BABAR*, a method known as "B Counting" is used for this purpose. The method is a weighted subtraction of the number of multi-hadronic events (N_{MH}) recorded offpeak from the number recorded onpeak. Since the offpeak CM energy is below the $B\bar{B}$ threshold, the difference must be entirely due to $B\bar{B}$ production once the energy dependence of the continuum cross-section is taken into account. As shown in Table 3.2, the onpeak and offpeak samples have different integrated luminosities and this has to be taken into account in the final determination. The ratio of luminosities is taken to be equal to the ratio of the number of $\mu^+\mu^-$ pairs ($N_{\mu^+\mu^-}$). So $N_{B\bar{B}}$ is given by:

$$N_{B\bar{B}} = \frac{1}{\epsilon_{B\bar{B}}}\left(N_{MH}^{\text{on}} - N_{MH}^{\text{off}}\kappa\frac{N_{\mu^+\mu^-}^{\text{on}}}{N_{\mu^+\mu^-}^{\text{off}}}\right) \tag{3.1}$$

where $\epsilon_{B\bar{B}}$ is the efficiency with which $B\bar{B}$ events pass the multi-hadronic selection (determined from MC simulation) and $\kappa \approx 1$ is a constant representing the energy dependence of the continuum cross-section. Applying this procedure, the number of

Table 3.2: *The "R24a3-v03" dataset*

Sample	$\mathcal{L}_{\mathrm{on}}(\,\mathrm{fb}^{-1})$	$N_{B\bar{B}}(10^6)$	$\mathcal{L}_{\mathrm{off}}(\,\mathrm{fb}^{-1})$
Run 1	20.6	22.6	2.6
Run 2	62.1	68.4	7.0
Run 3	32.7	35.7	2.5
Run 4	100.8	111.4	10.2
Run 5	133.9	147.6	14.5
Run 6	79.0	85.2	7.9
Total	429.0	470.9	44.8

$B\bar{B}$ pairs in the full dataset is found to be 470.9 ± 2.8 million [54].

3.2 Discriminating Variables

Discriminating variables are used to distinguish signal events from the large background. Most *BABAR* analyses, including this analysis, use two types of discriminating variables: kinematic and topological. Kinematic variables are obtained from the reconstructed B meson momentum and energy and the known beam energy. Topological variables describe the distribution of an event and are often combined in a Neural Network (NN) where the NN output ($\mathrm{NN_{out}}$) is used as a discriminating variable. All these variables can be used in two ways to increase the signal to background ratio:

◇ if the signal distribution of a variable falls within a specific range, a cut is applied to the variable distribution to eliminate as much background as possible and keep most of the signal events;

◇ if the distributions for signal and background have different shapes but overlap in the same region then a model of the distribution is made for both signal and background and is used in a fit, *e.g.* a Maximum Likelihood fit (see

Section 3.3).

3.2.1   Kinematic Variables

Correctly reconstructed candidates have a mass distribution centred around the B mass, $5.279 \, \text{GeV}/c^2$. The distribution is generally quite broad due to detector resolution, especially when multiple neutral clusters are involved. The invariant mass of the B is defined as:

$$m_B = \sqrt{E_B^2 - \vec{p}_B^2} \tag{3.2}$$

Since the *BABAR* detector operates at the $\Upsilon(4S)$ energy, the four momentum of the B mesons are quite well constrained by the beam energy, which is itself well measured. From that we can construct two sets of almost uncorrelated discriminating variables, the beam energy substituted mass (m_{ES}) and the difference between the reconstructed and expected B meson energy (ΔE), see Ref. [55]. These are defined in the centre of mass (CM) frame as:

$$\Delta E = E_B^* - E_{\text{beam}}^* \tag{3.3}$$

$$m_{\text{ES}} = \sqrt{E_{\text{beam}}^{*2} - \vec{p}_B^{*2}} \tag{3.4}$$

where $(E_B^*, \vec{p}_B^*)$ is the reconstructed B meson four-momentum and E_{beam}^* is the beam energy in the CM frame.

3.2.2   Topological Variables

The other way to discriminate between signal and background is to look at the topology of an event. Since the mass difference between the $B\bar{B}$ pair and the $\Upsilon(4S)$ is quite small, the $B\bar{B}$ pairs are produced almost at rest in the CM frame. This implies that the distribution of their decay products should be isotropic. Continuum events however are produced with large kinetic energy resulting in the decay

products forming highly collimated jets around the momentum vector of the original quark and anti-quark. Several variables can be used to distinguish between the two geometries of decays. These variables are constructed by making a distinction between the final state particles that reconstruct a signal B candidate and those that make up the "rest of the event" (ROE), *i.e.* the all other decay products in the event not associated to the B candidate. Each variable does not necessarily provide much discrimination, but combined into a NN, they help to eliminate most of the continuum background with minimal loss in signal.

3.2.3 Neural Networks

An Artificial Neural Network denotes any simulated collection of inter-connected neurons, with each neuron producing a certain response to a given set of input signals. These neurons are a simplified simulation of neural functions of the biological nervous system. They receive an input signal from the input neurons and respond by transmitting a weighted response to one or several output neurons. There are various types of NN implementation. The one used in this analysis is commonly called a `MultiLayer Perceptron` (MLP).

An MLP, described in detail in [56], simplifies the problem of n neurons with n^2 directional connections by organising the neurons in layers and allowing the directional connection only between one layer to the immediate next layer. The first layer to a multilayer perceptron is the input layer, the last one is the output layer and all layers in between are the hidden layers. For a classification problem containing n_{var} input variables and 2 output classes (signal or background), the input layer will consist of n_{var} neurons holding input values $x_1, ..., x_{n_{var}}$ and one neuron in the output layer holding the output variable, *i.e.* y_{NN}. The neuron response function, *e.g.* linear, sigmoid, an hyperbolic tangent or radial, maps the neuron input into the output. Each directional connection between the output of one neuron and the input of another holds an associated weight by which the output of a neuron is mul-
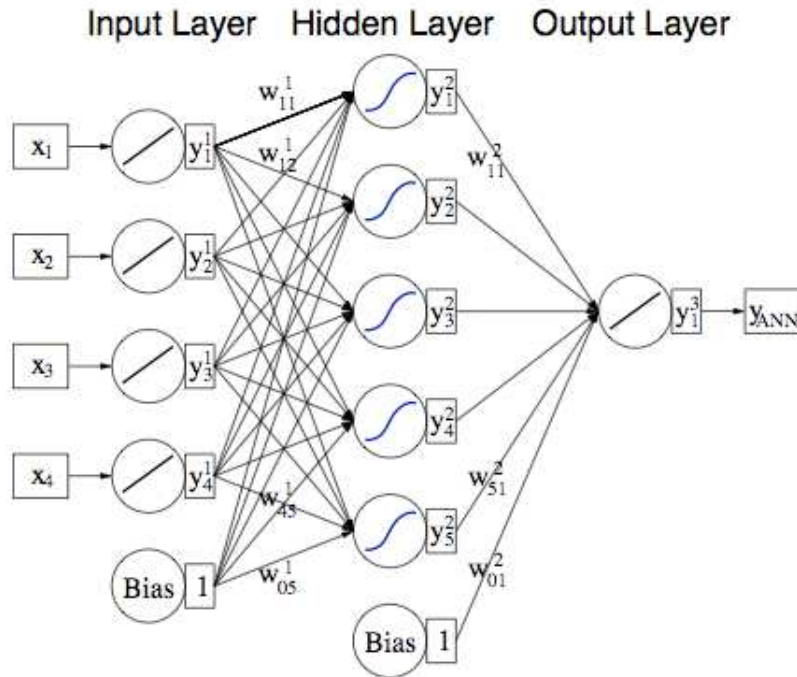
Figure 3.1: *Layout of a Multilayer Perceptron with one hidden layer [56].*

tiplied and passed as input value to the next. The weights are adjusted by a process of back propagation. This is a learning process by which the network is supplied with N training events. For each of these training events the neural network output, y_{NN}, is computed and compared to the known output $\hat{y} \in 1, 0$ where 1 is signal and 0 is background. An error function measuring the rate of agreement of the network response with the known class of the training event is formed and the set of weights are inferred by the minimisation of the error function, provided that the neuron response function is differentiable with respect to the input weights. Figure 3.1 shows a layout of a typical MLP configuration with one hidden layer.

3.3 Extended Maximum Likelihood Fit

A Maximum Likelihood fit is a statistical method used for fitting models to data in order to determine the model's parameters. It has many advantages over other

methods but its most important features are that it uses the Probability Density Function (PDF) shape information to infer the probability of each event category and that it treats each event individually, hence the data does not require binning. Both of these features maximise the signal sensitivity, allowing the study of very rare decays.

The Extended Maximum Likelihood (EML) differs from the standard method through the relaxation of the normalisation condition, in the cases where this depends on an event yield. The use of probability distributions with free normalisation is particularly useful to solve problems where the number of observed signal events is not fixed and unknown a priori, being subject to random Poisson fluctuations. The expected value may depend on the parameters to be estimated. EML is explained in detail in Ref. [57]. If the distribution of a given variable x has a particular shape, such as Gaussian, then it is possible to express this as a normalised PDF with parameters $\vec{\alpha}$ denoted $P(x, \vec{\alpha})$. The likelihood for a particular data set $x_1, x_2...x_N$ must incorporate not only the information that events were observed at $x_1, x_2...x_n$ but also the total number of events N. The extended likelihood is thus a combined set of probabilities where the normalisation depends on event yield distributed according to a Poisson distribution of mean ν written as:

$$\mathcal{L}(\nu, \vec{\alpha}) = \frac{e^{-\nu}\nu^N}{N!} \prod_i^N P(x_i, \vec{\alpha}) \tag{3.5}$$

where $\mathcal{P}(x_i)$ is the probability density and N the total number of observed events. Thus the extended log likelihood is given by:

$$\ln \mathcal{L} = \sum_i^N \ln \left[\nu \mathcal{P}(x_i, \vec{\alpha}) \right] - \nu \tag{3.6}$$

Increasing the normalisation increases the log likelihood through the first term but at the same time decreases it through the second term. Maximising the log likelihood involves finding the correct balance between the two terms.

3.3.1 Error in Likelihood Estimates

The error on a parameter can be calculated in two different ways during the fitting procedure. The first involves inverting the covariance matrix (H), obtained from the fit, to yield the error matrix:

$$E_{ij} = H_{ij}^{-1} = \left(\frac{\partial^2 \ln \mathcal{L}}{\partial \alpha_i \partial \alpha_j} \right)^{-1} , \tag{3.7}$$

where α_i and α_j are parameters of the fit and $\mathcal{L}$ is the likelihood. The vector of errors can then be extracted from $\vec{\sigma}_\alpha = E\vec{\alpha}$. This is the method used in this analysis to determine the statistical errors, unless stated otherwise. The second method approximates the Taylor expansion of the log-likelihood about its maximum, $l_{\max}$. This method is commonly used to determine asymmetric errors.

3.3.2 Fitting packages

Maximum likelihood fitting is a method widely used in particle physics analysis and a few software packages exist that allow performing such fits. `Minuit` [58] is an example of such a package capable of minimising a user-defined function (*e.g.* $-\ln \mathcal{L}$) and returning parameters and their errors at the minimum value. Several routines are included in `Minuit`. MIGRAD is the most commonly used minimisation routine which finds a function minimum and makes an initial approximation of the errors. HESSE is designed to make a more precise error calculation using the matrix inversion method and MINOS makes even more precise error calculations including asymmetric errors using the Taylor expansion technique both of which are mentioned in Section 3.3.1.

One of the software packages used in this analysis, called `Laura++` [59], was developed primarily for performing Dalitz amplitude fits and as such contains definitions of the likelihood functions needed to reconstruct the Dalitz plot variables required by this analysis. This package makes use of all the available routines mentioned in the previous paragraph to minimise the log likelihood and calculate errors. It was used

to do the main fit for the inclusive branching fraction and *CP* fit. The other package called `RooFit` [60] was developed to provide an interface to `Minuit` via the `Root` [61] software. The normalisation of the likelihood function is performed directly by the `RooFit` framework before being passed to `Minuit`. `RooFit` contains large number of possible PDFs and these can be combined by addition, multiplication or convolution. This software package was used to fit the resonant mass distributions in the quasi-two body mass regions.

3.4 The $_s\mathcal{P}lot$ Technique

$_s\mathcal{P}lot$ is a method for extrapolating unknown distributions using the results obtained from the EML fit, *i.e.* the yields for signal and background and the covariance matrix, and the PDFs of the discriminating variables. This is particularly useful in this analysis since the Dalitz plot distribution is unknown and essential to calculate the inclusive branching fraction and study the invariant mass distributions. The variables describing the dataset can be split in the maximum likelihood fit in two categories: the first component is a set of discriminating variables with distributions included in the fit for all event sources, the second component is a set of variables for which the distributions of some sources of events are either unknown or regarded as unknown since these are not part of the fit. This last set of variables are referred to as *control* variables. For more details on $_s\mathcal{P}lot$, see Refs. [62, 63].

3.4.1 The $_s\mathcal{P}lot$ Formalism

An essential assumption to use this technique is that the control variables are uncorrelated with the discriminating variables. When performing the maximum likelihood fit relying only on a variable y, no a priori knowledge of a control variable x is used. In this case the appropriate weight for event e is given by the covariance-weighted

quantity, called $_s\mathcal{W}eights$ and defined by:

$$_s\mathcal{P}_n(y_e) = \frac{\sum_{j=1}^{N_s} \mathbf{V}_{nj} f_j(y_e)}{\sum_{k=1}^{N_s} N_k f_k(y_e)} \qquad (3.8)$$

where:

◇ $\mathbf{V}_{ij}$ is the covariance matrix obtained from the fit,

◇ $f_i(y_e)$ are the PDFs of the discriminating variables,

◇ N_i are the yields of all the species.

With the $_s\mathcal{W}eight$, the distribution of the control sample variable x for a given species n, denoted by $\mathbf{M}_n(x)$, can be obtained from the $_s\mathcal{P}lot$ histogram:

$$N_n \bar{M}_n(\bar{x}) \delta x \equiv \sum_{e \subset \delta x} {}_s\mathcal{P}_n(y_e) \qquad (3.9)$$

The sum runs over the $N_{\delta x}$ events for which the value of x lies in the x-bin centered on $\bar{x}$ and of total width δx. $N_n \bar{M}_n(\bar{x})$ is then the x-distribution obtained by histogramming events using the weights in Eq. 3.8.

$_s\mathcal{P}lots$ bear the following properties useful in statistical analysis:

◇ the distribution $\bar{M}_n$ defined above is definitely normalised to unity;

◇ the sum over all events $\sum_{e=1}^{N} {}_s\mathcal{P}_n(y_e) = N_n$;

◇ in each bin, the sum over all species of the expected number of events equals the number of events actually observed *i.e.* for any event $\sum_{e=1}^{N_s} {}_s\mathcal{P}_l(y_e) = 1$.

Therefore an $_s\mathcal{P}lot$ is a consistent representation of how all events from various species are distributed in the control variable x. Summing over the yield of all species, $_s\mathcal{P}lots$ recover the data sample distribution in x and summing over the number of events in an $_s\mathcal{P}lot$ for a given species recovers the yield of the species, as determined by the fit.
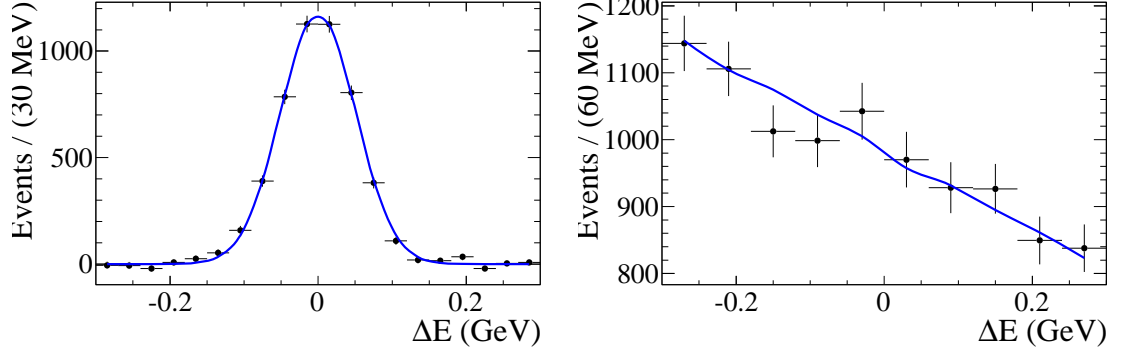
Figure 3.2: *An example ΔE $_s\mathcal{P}lot$ (data points) for signal (left) and background (right) obtained from a maximum likelihood fit to the sample m_{ES} distribution only. Overlayed is the distribution generated from the Gaussian and linear PDFs (blue line).*

A set of toy experiments for two species, signal and background, were generated from PDFs of two discriminating variables, m_{ES} and ΔE, to illustrate the power of this technique. In this simplified likelihood fit, the signal PDFs of both variables are Gaussians and for the backgrounds, an ARGUS for m_{ES} and linear PDF for ΔE. Figure 3.2 shows that by fitting only the m_{ES} distribution it is possible to reconstruct the distribution of the other variable.

3.4.2 Extended $_s\mathcal{P}lots$ for fixed yields

In this analysis the yields of some of the species are fixed to a value and not derived from the data sample at hand. These species are there after denoted with a '0' so that the number of expected events for this species, N_0, is a known value and held fixed in the fit. The $_s\mathcal{P}lot$ technique can be extended to deal with this issue assuming the distributions, $\mathbf{M}_0(x)$ for the fixed species are known. Now the Eq. 3.9 can be modified to include the fixed yields:

$$N_n \bar{M}_n(\bar{x})\delta x = c_n \mathbf{M}_0(x)\delta x + \sum_{e \subset \delta x} {}_s\mathcal{P}_n \tag{3.10}$$

where $_s\mathcal{P}_n$ is the $_s\mathcal{W}eight$ from Eq. 3.8 except that the sum now runs over all floating species only and c_n is the species dependent coefficient. This coefficient is simply the sum of the covariance matrix elements subtracted to the number of events for the unknown species defined below:

$$c_n = N_n - \sum_j \mathbf{V}_{nj} \tag{3.11}$$

The above extended $_s\mathcal{P}lots$ share the same properties described in Section 3.4.1 *i.e.* they reproduce the true marginal distribution and are properly normalised.

4

Event Selection

The first step towards measuring the branching fraction is to reduce the background levels in the data sample. This is achieved via cuts on selected variables of the final state particles, use of standard tracking, neutral particle lists and a NN (see Section 3.2.3). The effects of the selection criteria applied are tested on nonresonant signal MC, $B\overline{B}$ background MC and offpeak data, the latter being a good sample of continuum background. Each selection was carefully tested and analysed to minimise the background contributions at the fitting stage but to keep as many signal events as possible.

4.1 Particle identifications

This section lists the selections applied to variables in the event to reduce misidentification of the final state particles. The neutral particles are selected by combining information from a predefined list containing specific π^0 energy and mass criteria to a tighter set of selections calibrated using signal MC. Charged kaons are selected by applying one of the *BABAR* standard particle identification lists. These make use of neural networks composed from specific discriminating variables to associate charged tracks to the most likely contributing charged particle.

4.1.1 Neutral selections

The π^0 candidates are selected from the standard *BABAR* `PiOAllLoose` list. This list is a combination of two different π^0 lists, `PiOLooseMass` and `MergedPiOLoose`. The `PiOLooseMass` list is used to select composite π^0, *i.e.* π^0 candidates for which the pair of photons daughters contribute to one bump each in the EMC (for a complete description of bumps and energy deposits refer to Section 2.6). The `MergedPiOLoose` list is used to select merged π^0, *i.e.* π^0 candidates for which photon daughters contribute to a single bump in the EMC. To be considered a π^0 candidate by the neutral lists a bump must have a polar angle between 0.42 and 2.4 and have energy greater than 30 MeV [64]. Each list selects candidates by applying additional sets of cuts on the photon energy distributions and the π^0 candidate mass. The `PiOLooseMass` list requires:

⋄ minimum uncalibrated energy of photon needs to be 0.030 GeV,

⋄ maximum LAT (see Eq. 2.2) of photon shower needs to be 0.8,

⋄ minimum π^0 energy of 0.2 GeV,

⋄ a π^0 mass between 0.1-0.16 GeV/c^2 with an additional mass constrained fit.

The `MergedPi0Loose` list requires the single bump to be consistent with a π^0 candidate with a significance greater than 1%. This is decided from an algorithm looking at the characteristics of the shape of the bump and the energy distribution within it. This algorithm is shown to retain approximately 97% of neutral pion candidates found in a data control samples [65]. See Refs. [66,67] for details on the various lists and selections. The loose lists were chosen to avoid eliminating too much potential signal. Additional selections on the photons and π^0 variables have been calibrated from the distributions observed in nonresonant signal MC and were used to eliminate most background events. Events were selected if the π^0 candidate passed the following criteria:

⋄ the lateral moment of each photon needs to be between 0.01 and 0.6,

⋄ the energy of all decay photon is required to be greater than 0.05 GeV,

⋄ the helicity angle of the π^0, defined as the angle between one of the decay photons and the π^0 direction of flight in the CM frame of the π^0, is less than 0.9,

⋄ the π^0 mass is further constrained to $0.115 \, \mathrm{GeV}/c^2 < m_{\pi^0} < 0.150 \, \mathrm{GeV}/c^2$.

4.1.2 Kaon PID

By using all the tracking information available on charged tracks, it is possible to discriminate between different hypotheses on the identity of the charged particle passing through the detector. There are two types of kaon PID in *BABAR*, each with different degrees of rigidity in the identification requirements:

⋄ Kaon KM (KKM) selector: The KM selector combines multiple binary classifiers all trained in different ways to form a multiclass classifier. Each classifier gives an output between -1 and 1 according to its definition of a signal track. The results obtained correspond to a specific particle hypothesis. To allow

for different levels of tightness that might be required for various analysis, the minimum selection is written as the comparison of the ratio of the various resultant variables and a constant. This constant is then varied according to the level of tightness of the selector [68].

◇ Kaon BDT selector: Whereas the kaon KM selector allows K identification from most other charged particle species *i.e.* pions, protons and electrons, the BDT selector specialises in K/π separation only. By knowing only the distinguishable factors between kaons and pions, the KBDT selector achieves higher efficiencies and lower pion-as-kaon mis-ID [69].

In order to determine the best setting for the PID selectors, a study was performed where all the different choice of selectors and tightness were applied to nonresonant signal MC, offpeak data and generic $B\overline{B}$ MC. The number of events were used to calculate the so-called Punzi Figure Of Merit (FOM) given by [70]:

$$\text{FOM} = \frac{\varepsilon}{\frac{a}{2} + \sqrt{B}}\,. \tag{4.1}$$

where ε is the signal efficiency, B is the total expected background and $a = 3$, expressed in terms of standard deviations, is the statistical significance of the expected signal. This FOM was chosen since it does not make any assumption on the effective presence of a signal. The best signal to background ratio is achieved by the PID which maximises the FOM. Results shown in Table 4.1.2 favour the KKMTight selector with FOM $= 4.87 \times 10^{-4}$. These results were obtained before applying any vetoes and before making any requirements on a continuum fighting MVA.

4.2 Continuum background rejection

Background from $q\bar{q}$ events is one of the major backgrounds found in this analysis. As described in Section 3.2.2, $q\bar{q}$ decays have a very distinct topology compared to $B\overline{B}$ events. The combination of the following three topological variables (ratio of

Table 4.1: *PID selector performance for $B^+ \to K^+ \pi^0 \pi^0$.*

Kaon PID Selector	Signal Efficiency	$N_{q\bar{q}}$ Expected	$N_{B^+B^-}$ Expected	$N_{B^0\bar{B}^0}$ Expected	FOM
KKMSuperLoose	0.21688	241569	1787.5	908.6	0.000438
KKMVeryLoose	0.20676	189178	1368.4	733.8	0.000471
KKMLoose	0.19062	154151	1127.0	620.2	0.000481
KKMTight	0.18064	135177	985.4	557.1	0.000487
KKMVeryTight	0.18054	135072	982.4	556.1	0.000486
KKMSuperTight	0.17125	124025	905.8	512.9	0.000481
KBDTNotPion	0.21866	225841	1752.6	910.2	0.000456
KBDTVeryLoose	0.21138	225841	1425.1	766.9	0.000441
KBDTLoose	0.20046	186832	1214.5	660.5	0.000460
KBDTTight	0.19550	173038	1125.6	619.2	0.000466
KBDTVeryTight	0.19027	160852	1043.1	582.6	0.000470

Legendre polynomials, angular distributions of the B momentum and thrust) [71] combined with the absolute value of flavour tagging algorithm are identified in this analysis as providing the best discrimination power when combined in a Neural Network (NN). Furthermore these variables show almost no dependences with the Dalitz plot position. This is important since it was observed that any Dalitz plot dependance in the input variables of the NN can lead to non negligible fit biases (See "Punzi biases" in Appendix A).

4.2.1   Ratio of Legendre polynomials

The 0th and 2nd order Legendre polynomial are momentum weighted sums over the ROE (see Section 3.2.2). Both these variables can act as topological discriminants, however they are highly correlated with one another. The ratio of these variables is used instead and is defined as:

$$\frac{L_2}{L_0} = \frac{\frac{1}{2} \sum\limits_{i \in \text{ROE}} p_i \left(3 \cos^2 \theta_i - 1\right)}{\sum\limits_{i \in \text{ROE}} p_i} \tag{4.2}$$

where p_i is the momentum and θ_i is the angle between each track or neutral cluster and the thrust axis of the B candidate, which is the direction maximising the sum of

the longitudinal momenta of the B candidate daughters [72]. From this definition, a track belonging to a jet-like continuum event would be approximately colinear with the thrust axis of the decay candidates. Therefore the ratio in Eq. 4.2 would have values approaching 1. On the other hand tracks originating from isotropic B decays would have an approximately uniform distribution of $\cos(\theta)$ and therefore the ratio would tend to have smaller, even negative, values. This is illustrated in Figure 4.1, which shows the distribution of the ratio for nonresonant signal MC and offpeak data.
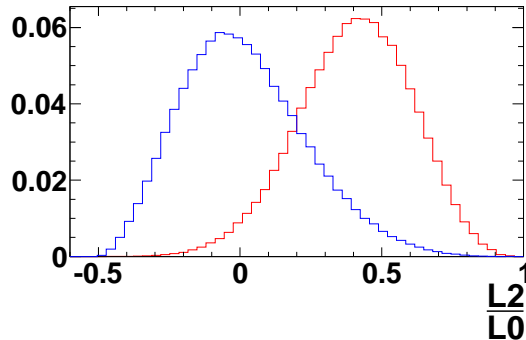
Figure 4.1: *Distributions for signal MC (blue line) and offpeak data (red line) of the ratio of Legendre polynomials, $\frac{L_2}{L_0}$. Both distributions were normalised to unity.*

4.2.2   Angular Variables of B Decay

Two additional topological variables are used to define the angular distribution of the decay of the B meson. The first contains information on the absolute value of the cosine of the angle ($\theta_{B_{\mathrm{mom}}}$) between the direction of the B momentum and the beam axis. In a true B event, the angular distribution of the decay of the spin one $\Upsilon(4S)$ to two spin zero B mesons is proportional to $\sin^2(\theta_{B_{\mathrm{mom}}})$, while in a $q\bar{q}$ event the distribution is more or less uniform. The second is the absolute value of the cosine of the angle ($\theta_{B_{\mathrm{thr}}}$) between the direction of the B thrust axis and the beam axis. In this case, for true B events the distribution should be uniform due to the

spherical nature of B decays, while for $q\bar{q}$ events the distribution is proportional to $1 + \cos^2\theta_{B_{\text{thr}}}$. Distributions for these two variables for signal MC and offpeak data are shown in Figure 4.2. The distributions for $\cos\theta_{B_{\text{thr}}}$ differ quite dramatically from expectations due to the effects of detector acceptance and selection requirements.
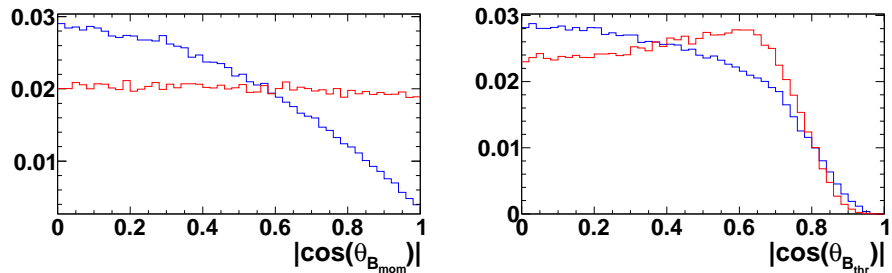
Figure 4.2: *Absolute value of the cosine of the B direction on the left and the B thrust to the right with respect to the beam axis for signal MC (blue line) and offpeak data (red line). Both distributions were normalised to unity.*

4.2.3 Flavour tagging

The last variable used in the NN is the absolute value of the output of the flavour tagging algorithm. While one B meson is fully reconstructed in the final state being studied, the other is used to determine the flavour of the other B meson. Flavour tagging at BABAR is achieved via another neural network. This neural network is provided primary information to help determine the flavour of the tagging B meson, such as the presence in the decay products of primary and secondary leptons or kaons and fast charged tracks [73]. If the B meson in the event is tagged as a B^+ or B^0 then the output is positive and if it is a B^- or $\overline{B}^0$ then the output is negative. The output of the NN is closer to zero if the outcome of the NN is more uncertain. The absolute value of the output is shown in Figure 4.3 for signal MC and offpeak data. As expected offpeak is shown to have more events with absolute value closer to zero and signal MC has slightly more events closer to unity.
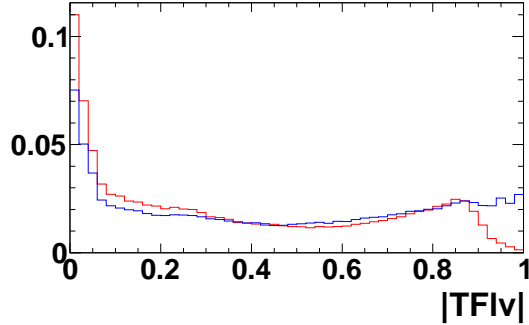
Figure 4.3: *Absolute value of the output of the flavour tagging NN for signal MC (blue line) and offpeak data (red line)*

4.2.4 Neural Network training and output selection

The variables listed above were combined to form a NN, Fisher and Boosted Decision Tree (BDT) discriminant trained on half the events in signal MC and offpeak data. The output was used to compare the performance of these methods and the results of the test is shown in Figure 4.4. The MLP exhibits marginally higher signal efficiency to background rejection. From the distribution of the input variables, the most powerful at discriminating between signal and continuum is the ratio of Legendre polynomials in Figure 4.1. The distributions of these variables is seen to be
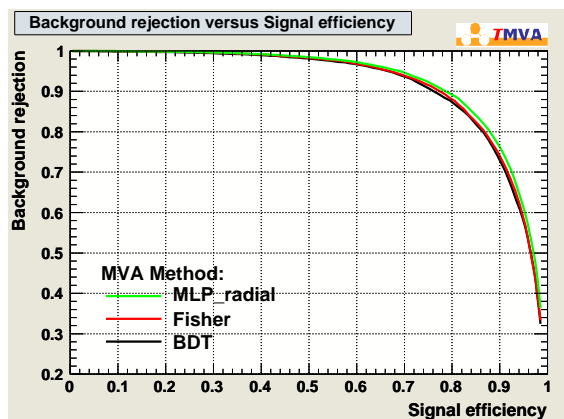
Figure 4.4: *Comparison of the performance of three MVAs using the same variables. The MLP Neural Network gives the best performance for this mode.*
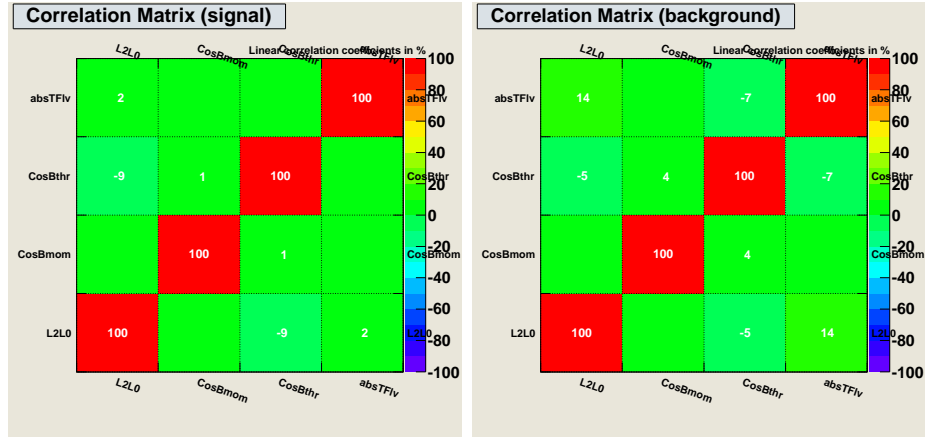
Figure 4.5: *Correlation matrices between the five event shape variables for signal MC and offpeak data.*

almost Gaussian. This fact together with the correlations between variables, shown in Figure 4.5, being almost negligeable allow a linear discriminant like the Fisher to perform as well as the non-linear approaches of the NN and BDT. The output of the MLP is used to apply the selection and as a discriminating variable later in the fit (see Section 3.2.3 for a more detailed description on MLP). The following NN architecture was chosen to obtain good performance:

- ◇ neuron function: radial;

- ◇ number of cycles: 200;

- ◇ size of the training samples: 18658 signal MC events and 18658 offpeak data events.

The selection in the MLP at $NN_{out} > 0.35$ (see green arrow in Figure 4.6) is chosen to keep approximately 90% of signal and reject 76% of the background.
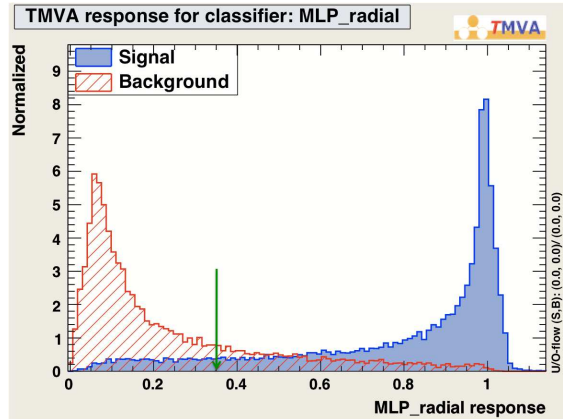
Figure 4.6: *Neural Network distribution for signal Monte Carlo and offpeak data. The green arrow indicates the position of the selection applied on the* NN$_{out}$ *variable.*

4.3 Selection Optimisation

The selection of events is further optimised by carefully chosing candidates within an event and applying vetos of candidates corresponding to signal-like B background decays. The variation in signal efficiency over the DP is then studied and later used to correct the fitted yield.

4.3.1 Final Candidate Selection

In the full Run 1–6 onpeak dataset, following the application of the previously described selection criteria, 20.9% of events with at least one candidate have more than one candidate. Of these events, 13.2% have two candidates, 4.1% have three, 1.8% have four, and the remaining 1.8% have five or more candidates. The average number of B candidates found per selected event is equal to 1.3. To chose the best candidate in a multiple candidate event, the χ^2 is formed from the two π^0 masses. The best candidate is selected to have the smallest χ^2 formed from the sum of the χ^2 values of the two π^0 candidates. This selection was found to have a rate of success of 91.2% to select the best reconstructed candidate.

4.3.2 Vetoed Regions

In order to reduce contamination from B-background modes, vetoes are used against certain regions of the Dalitz plot. If any candidate in the event falls in the vetoed region then the entire event is rejected. In this analysis only one veto is applied for the decay mode $B^+ \to K_S^0 K^+$, $K_S^0 \to \pi^0\pi^0$. The $\pi^0\pi^0$ invariant mass distribution in this mode is shown in Figure 4.7. The veto is applied to the region $0.4\,\mathrm{GeV}/c^2 < m_{\pi^0\pi^0} < 0.55\,\mathrm{GeV}/c^2$. Relative efficiencies of applying the K_S^0 veto for each signal MC mode are shown in Table 4.3 at the end of this chapter.
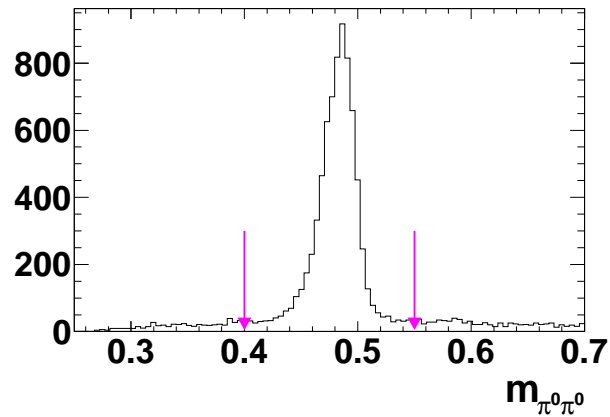
Figure 4.7: *$\pi^0\pi^0$ invariant mass distribution in $B^+ \to K_S^0 K^+$, $K_S^0 \to \pi^0\pi^0$ Monte Carlo before the veto is applied. The arrows indicate the vetoed region.*

4.3.3 Signal Efficiency

A summary of the cuts and their efficiencies is given in Table 4.2 for the nonresonant and $K^*(892)^+$ resonance MC. It is observed that the main differences in efficiency occur for the selections on the neutral pions' angular distributions and masses. These differences in efficiency are expected to be even larger with resonances like the $f_0(980)$, which decay to two neutral pions.

Nonresonant MC events are used to study the variation in signal efficiency over the Dalitz plot. The MC events were generated with a phase space distribution and re-

Table 4.2: *Selection cut summary and efficiencies for the NR mode $K^+\pi^0\pi^0$ and the resonant $K^*(892)^+\pi^0$.*

Selection Cut	Relative efficiency (%)	
	Phase space MC	$K^*(892)^+\pi^0$ MC
Reconstruction and preselection	45.5	52.1
Kaon/pion PID requirements	78.7	78.5
$E_{\text{tot}} < 20\,\text{GeV}$	99.9	99.9
$\left\|\frac{\Delta t}{\sigma^2_{\Delta t}}\right\|$	98.8	99.0
$5.20\,\text{GeV}/c^2 < m_{ES} < 5.28\,\text{GeV}/c^2$	98.2	98.8
$-0.2\,\text{GeV} < \Delta E < 0.3\,\text{GeV}$	98.0	98.0
$\text{NN}_{\text{out}} > 0.35$	90.1	89.9
$0.01 < LAT_{\pi_1^0,\gamma_1} < 0.6$	98.8	98.5
$0.01 < LAT_{\pi_1^0,\gamma_2} < 0.6$	97.6	96.7
$0.01 < LAT_{\pi_2^0,\gamma_1} < 0.6$	99.3	99.4
$0.01 < LAT_{\pi_2^0,\gamma_2} < 0.6$	98.8	99.0
$E_{\pi_1^0,\gamma_1} > 0.05\,\text{GeV}$	99.6	99.5
$E_{\pi_1^0,\gamma_2} > 0.05\,\text{GeV}$	96.3	95.8
$E_{\pi_2^0,\gamma_1} > 0.05\,\text{GeV}$	99.7	99.7
$E_{\pi_2^0,\gamma_2} > 0.05\,\text{GeV}$	98.3	98.7
$\|\cos\theta^{\gamma_1}_{\text{helicity}}\| < 0.9$	97.6	98.3
$\|\cos\theta^{\gamma_2}_{\text{helicity}}\| < 0.9$	94.6	94.1
$0.115\,\text{GeV}/c^2 < m_{\pi_1^0} < 0.150\,\text{GeV}/c^2$	93.7	94.4
$0.115\,\text{GeV}/c^2 < m_{\pi_2^0} < 0.150\,\text{GeV}/c^2$	94.1	93.3
Veto K_S^0	99.5	97.9
Total efficiency	16.1	14.7
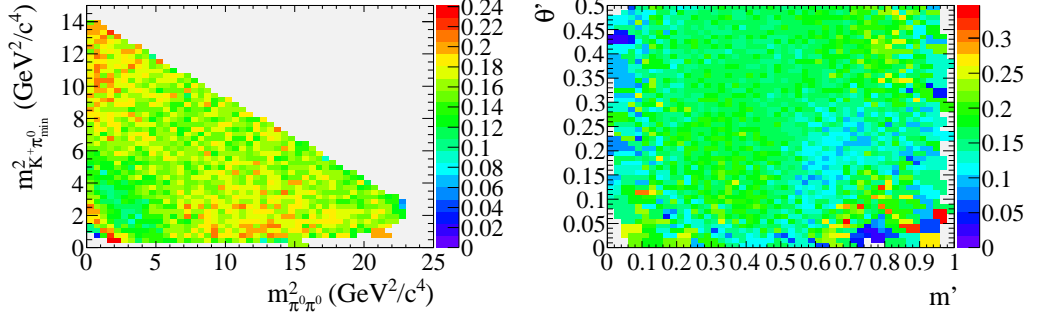
Figure 4.8: *Variation of signal efficiency over the conventional Dalitz plot (left) and square Dalitz plot (right).*

constructed using the criteria detailed in Table 4.2 together with their corresponding efficiencies. The variation in efficiency is then studied by dividing the distribution of all reconstructed events plotted at their generated position by the distribution of all the generated events. The resultant distribution for both the conventional and square Dalitz plot is shown in Figure 4.8. The square Dalitz plot histogram is later used to correct the $_s\mathcal{P}lot$ Dalitz distribution for efficiency (see Section 6.3.2). The lower left-hand corner of the Dalitz plot contains events with a low momentum π^0 meson while the right-hand corner contains events with a low momentum charged kaon. Low momentum kaons have smaller probability of passing the PID requirements since they do not reach the DIRC, hence the efficiency in the corresponding corner of the Dalitz plot is lower. Conversely, the lower left-hand corner shows higher efficiency because the chances of forming a low momentum π^0 are relatively high due to the high multiplicity of low momentum photons. However the probability of this being correctly reconstructed is correspondingly small, see Section 4.4. Moving away from the very corner the π^0 momentum increases and the efficiency drops since the energy of the photons, required to make up the π^0 meson, is higher than expected from background photons. See Table 4.3 at the end of the chapter for a list of average signal efficiency for all signal MC samples.

4.4 Classification of Misreconstructed Events

"Self cross feed" (SCF) can occur in non-negligible fractions due to low momentum particles being exchanged with particles from the decay of the other B meson in the event. These are signal events that have been misreconstructed. Correctly reconstructed events are denoted "truth matched" events (TM). Due to the presence of two π^0 mesons in the final state, SCF is expected to be a non-negligible effect in this analysis. The effect of a significant SCF fraction of events is that the reconstructed Dalitz plot positions can significantly differ from the true (generated) position. The distributions for m_{ES} and ΔE are also affected.

4.4.1 SCF Definition

SCF events are distinguished by looking at the generator level information and in particular by using the following quantity:

$$p_{\mathrm{pull}} = \frac{p_{\mathrm{gen}} - p_{\mathrm{rec}}}{\sigma_{p_{\mathrm{rec}}}} \tag{4.3}$$

where p_{gen} is the generated momentum, p_{rec} the reconstructed momentum and $\sigma_{p_{\mathrm{rec}}}$ the error on the reconstructed momentum. This quantity is then plotted against p_{rec} in Figure 4.9 for the three final state particles in nonresonant signal MC. TM events
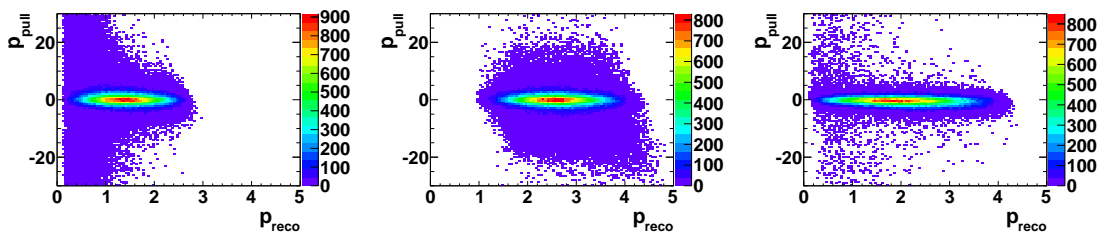
Figure 4.9: *Difference between generated and reconstructed momenta divided by reconstructed momentum error plotted against reconstructed momenta for lowest momentum π^0 candidate (left), highest momentum π^0 candidate (centre), kaon candidate (right).*

seem to cluster within a momentum pull ± 10.0. To determine the best value of Eq. 4.3 that helps distinguish between TM and SCF, the effect of varying selection ranges of this p_{pull} on m_{ES} and ΔE was examined within the limit determined from Figure 4.9. This was done to determine when the distributions become predominantly SCF. These distributions are plotted at regular intervals of x_i, see Figure 4.10, where x_i is defined to be the largest of the absolute value of the quantity defined in Eq. 4.3 for the three final state particles, K^+ and the two π^0. It is clear that events with a momentum pull within $0.0 < x_i < 2.0$ have very different distributions to those with larger ranges of x_i. The question is at which value the change occurs, the distributions seem to incorporate mainly SCF events over the range $4.0 < x_i < 6.0$.

The boundary value was therefore chosen to be:

$$TM : p_{\mathrm{pull}} < 5.0 \qquad\qquad SCF : p_{\mathrm{pull}} > 5.0 \qquad\qquad (4.4)$$

If any of the three particles have $p_{\mathrm{pull}} > 5$ than the event is classified as SCF. It is observed that the ΔE distribution is asymmetric towards low energy even for small values of x_i. This suggests that even at low x_i values some events are retained in which one of the particles, presumably a photon, is misreconstructed due to energy losses in the detector. Final distributions for m_{ES} and ΔE of TM and SCF events from nonresonant MC, obtained using the classification scheme in Eq. 4.4, are shown in Figure 4.11.

4.4.2 SCF fraction

The average SCF fraction, defined as:

$$f_{\mathrm{SCF}} = \frac{N_{\mathrm{SCF}}}{N_{\mathrm{SCF}} + N_{\mathrm{TM}}} \qquad\qquad (4.5)$$

where N_{SCF} is the number of SCF events and N_{TM} the number of truth matched events, is plotted in Figure 4.12 across the Dalitz plot. The SCF fraction is observed to be strongly dependent on Dalitz plot position. Low momentum π^0 are very hard
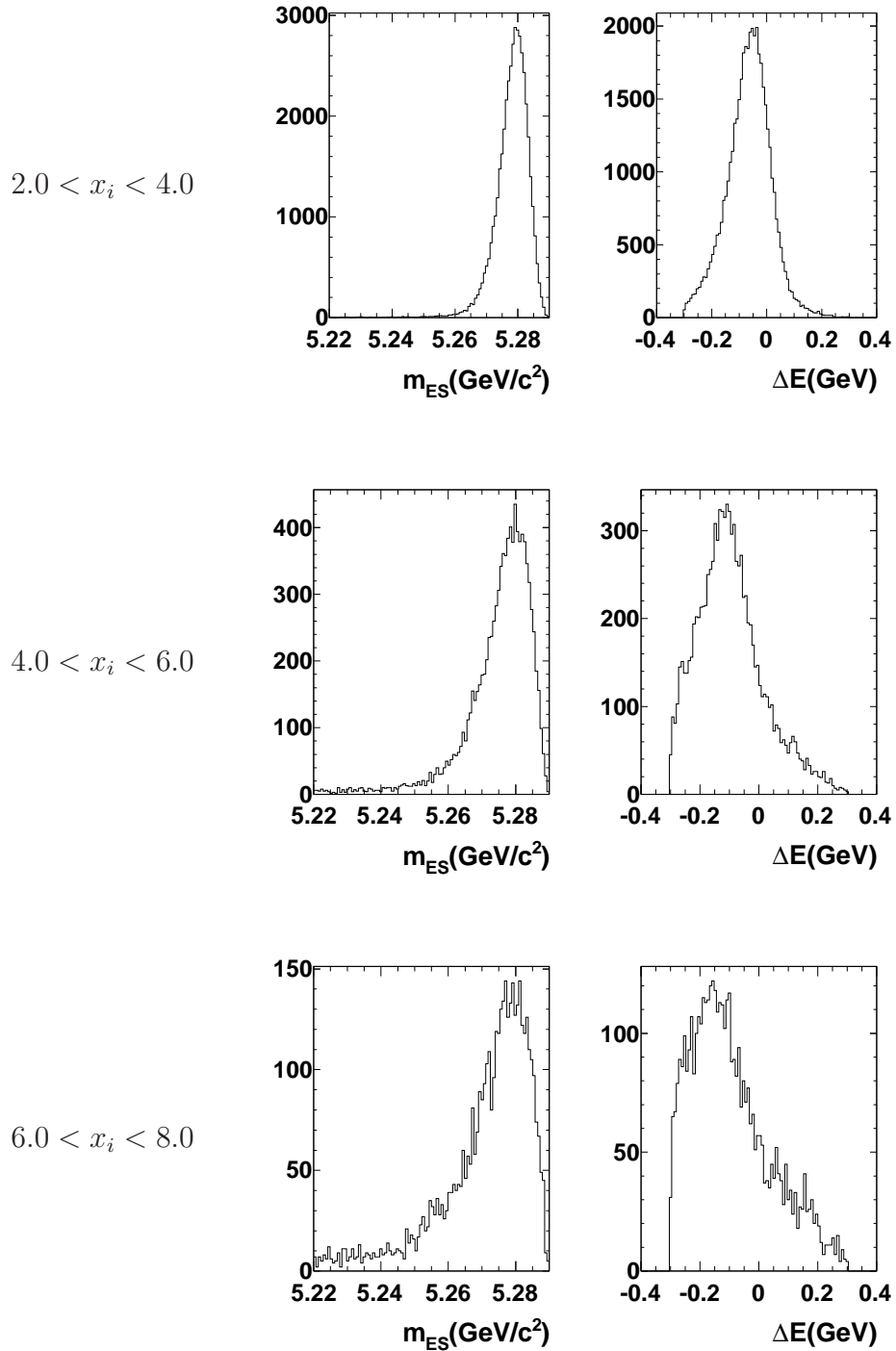
Figure 4.10: m_{ES} and ΔE distributions for different ranges of x_i, where x_i is p_{pull} for all three final state particles.
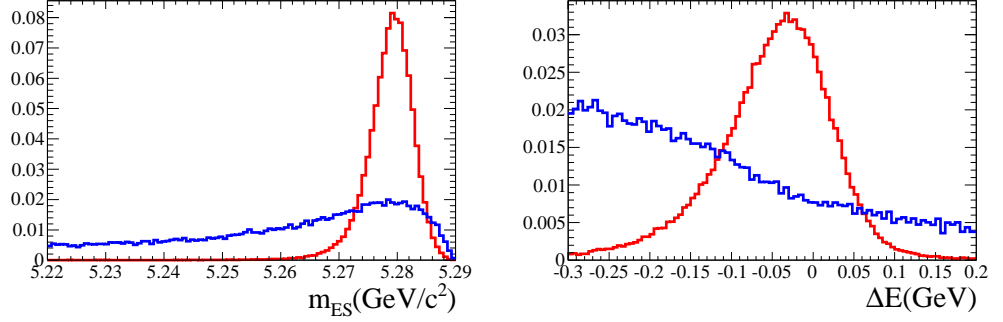
Figure 4.11: m_{ES} *(left) and ΔE (right) distributions for TM (red histogram) and SCF (blue histogram) events based upon a definition of SCF from $p_{\mathrm{pull}} >$ 5.0. Both TM and SCF histograms have been normalised.*

to correctly reconstruct therefore the largest fraction of SCF events is found in the corner of the Dalitz plot where the momentum of one of the π^0 mesons is very low. In order to avoid biases from Punzi effect, the Dalitz plot distribution of the SCF fraction is not directly included in the fit (see Appendix A). The method used instead determines the SCF fraction in data by an iterative procedure involving $_s\mathcal{P}lots$ (see Section 3.4). This procedure is described in detail in the next chapter in Section 5.4 and uses the square DP in Figure 4.12 to calculate the SCF fraction. A summary of average SCF fractions in MC is listed in Table 4.3 and the large variation is evident (higher for low mass K^* and broad $\pi^0\pi^0$ resonances).
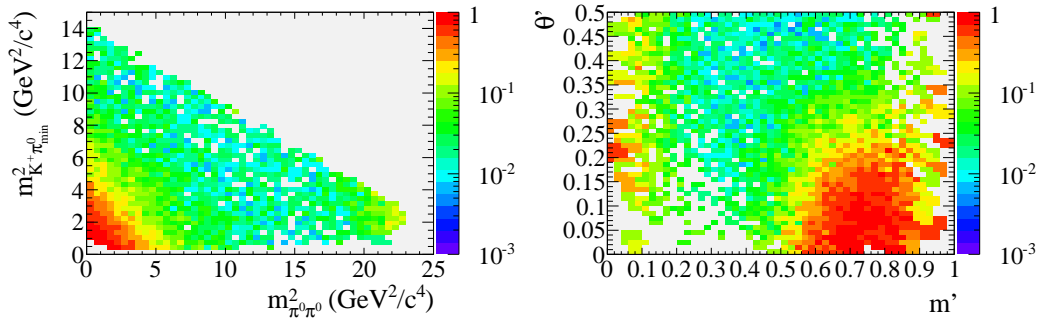
Figure 4.12: *Fraction of self cross feed events as a function of Dalitz plot position in conventional (left) and square (right) Dalitz plot form.*

Table 4.3: *Summary of veto efficiency, average efficiency and SCF fraction for all nonresonant and resonant signal modes.*

Decay mode	Veto efficiency (%)	Average efficiency (%)	SCF fraction (%)
$B^+ \to K^+\pi^0\pi^0$	99.5	16.1	5.3
$B^+ \to K^*(892)^+\pi^0$, $K^*(892)^+ \to K^+\pi^0$	97.9	14.7	21.2
$B^+ \to K_2^{*+}(1430)\pi^0$, $K_2^{*+}(1430) \to K^+\pi^0$	97.3	15.4	16.9
$B^+ \to K^{*+}(1680)\pi^0$, $K^{*+}(1680) \to K^+\pi^0$	97.1	14.3	14.5
$B^+ \to K^{*+}(1410)\pi^0$, $K^{*+}(1410) \to K^+\pi^0$	97.5	14.6	18.1
$B^+ \to f_2(1270)K^+$, $f_2(1270) \to \pi^0\pi^0$	96.5	15.3	28.4
$B^+ \to f_0(980)K^+$, $f_0(980) \to \pi^0\pi^0$	96.5	16.1	14.6
$B^+ \to K_0^{*+}(1430)\pi^0$, $K_0^{*+}(1430) \to K^+\pi^0$	98.8	16.1	8.9
$B^+ \to \chi_{c0}K^+$; $\chi_{c0} \to \pi^0\pi^0$	100	16.8	1.7

5

The Fitting Model

This chapter includes the analysis of the PDF distributions of the discriminating variables to include in the fit to obtain the branching fraction and CP asymmetry of the inclusive decay. The previous chapter showed how the presence of two neutral pions in the final state increased the likelihood for particle misreconstruction. The signal PDF therefore includes a SCF and a TM PDF component. Background PDFs on the other hand account for continuum and four $B\overline{B}$ categories, grouped depending on the shape of the m_{ES} and ΔE distributions. A summary of the PDFs used in this analysis is provided in Table 5.1. The chapter also includes the optimisation of the fitting techniques, to avoid biases due to Dalitz plot dependences, and the details of the method used to obtain the branching fractions of the intermediate resonances.

Table 5.1: *An overview of the fitting model giving a description of the PDFs including if the parameters are fixed or floated in the fit.*

Event category	Fit Components			
	m_{ES}		NN_{out}	
Signal TM	Cruijff	all parameters fixed	Histogram	-
Signal SCF	3rd order Chebychev	all parameters fixed	Histogram	-
$q\bar{q}$	Argus	endpoint m_0 fixed c parameter floated	Step Function	all bins floated
$B\bar{B}$ background	Histogram	-	Histogram	-
Signal yield	Floating			
$q\bar{q}$ yield	Floating			
$B\bar{B}$ yield	Fixed			
Signal A_{CP} (CP fit only)	Floating			
$q\bar{q}$ A_{CP} (CP fit only)	Floating			
$B\bar{B}$ A_{CP} (CP fit only)	Fixed			

5.1 Fitting Regions

The fitting regions are defined using the kinematic variables m_{ES} and ΔE (see Section 3.2.1 for a better description of these variables). These were optimised to take into account possible correlations with the Dalitz plot and avoid fit biases at a later stage. Correlations of the discriminating variables and the NN_{out} with Dalitz plot coordinates are listed in Table 5.2. ΔE exhibits the largest correlation in particular with respect to the $\pi^0\pi^0$ invariant mass. This is further enhanced in Figure 5.1, which shows the distributions of the values corresponding to the mean and rms of m_{ES}, ΔE and NN_{out} distributions over the Dalitz plot . Since ΔE is correlated with the Dalitz plot, it is not used in the fit and a tight cut on the distribution is applied instead (see Appendix A). The mean and width of the m_{ES} and NN_{out} distributions show no substantial correlation with the DP and can be used in the maximum likelihood fit. The signal region of these two distributions is therefore kept loose.
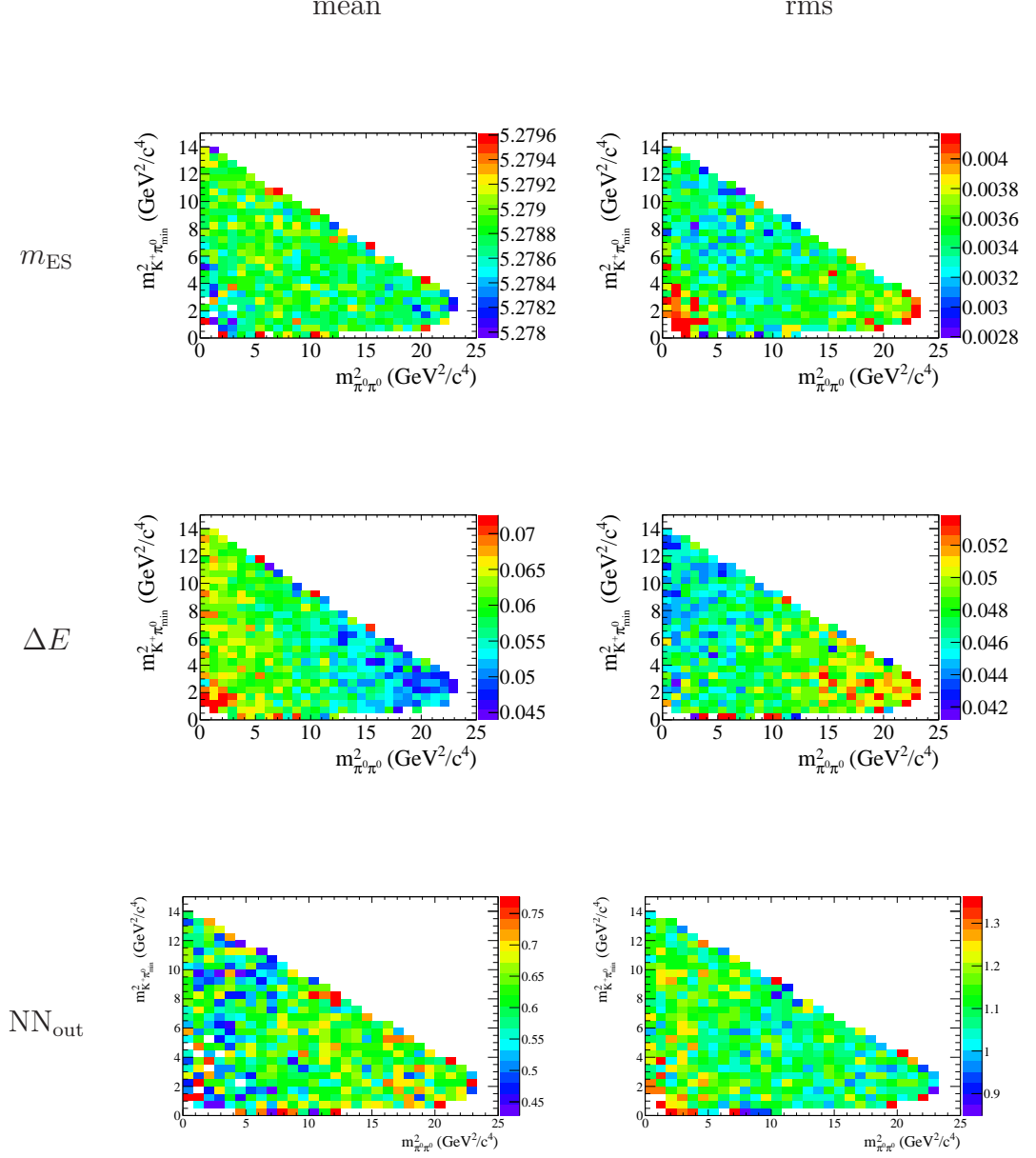
Figure 5.1: *Variation of the signal m_{ES}, ΔE and NN_{out} distributions over the Dalitz plot in terms of the mean and rms of the distributions. These Dalitz plots were constructed from nonresonant signal MC events that lie in the signal region of m_{ES} and ΔE (see Section 5.1.2). The events were selected as described in Chapter 4 except that the $K_S^0 K^+$ veto was not applied.*

Table 5.2: *Correlations with Dalitz plot coordinates of m_{ES}, ΔE and $\mathrm{NN}_{\mathrm{out}}$ distributions.*

Discriminating variable	correlation with $m_{K^+\pi^0_{\min}}$ (%)	correlation with $m_{\pi^0\pi^0}$ (%)
m_{ES}	2.0	-1.1
ΔE	3.5	-8.0
$\mathrm{NN}_{\mathrm{out}}$	-1.1	2.3

5.1.1 ΔE signal region optimisation

A tighter cut must be applied on the ΔE distribution since this kinematic variable is no longer included in the fit. The ΔE signal region is optimised using the Punzi figure of merit described in Eq. 4.1 and results are given by in Table 5.3. The optimal ΔE signal region is the region that maximises the figure of merit and is found to be $-0.15\,\mathrm{GeV} < \Delta E < 0.05\,\mathrm{GeV}$.

5.1.2 Definitions of fitting and sideband regions

In the initial selection, events are selected with kinematic variable within the range:

⋄ $5.22\,\mathrm{GeV}/c^2 < m_{\mathrm{ES}} < 5.29\,\mathrm{GeV}/c^2$,

⋄ $-0.2\,\mathrm{GeV} < \Delta E < 0.3\,\mathrm{GeV}$.

The signal region with tight ΔE cut is given by:

⋄ $5.260\,\mathrm{GeV}/c^2 < m_{\mathrm{ES}} < 5.286\,\mathrm{GeV}/c^2$,

⋄ $-0.15\,\mathrm{GeV} < \Delta E < 0.05\,\mathrm{GeV}$.

The extreme endpoint of the m_{ES} distribution is excluded from the fitting region, since this has been shown in previous analyses [74] to reduce the sensitivity to fit

Table 5.3: *Optimisation of the ΔE cut. For each set of cut values, the total signal efficiency, expected number of background events, the Punzi FOM and the ΔE cut efficiency based on nonresonant MC. The coloured row indicates the signal selection used.*

ΔE_{min}	ΔE_{max}	Signal Efficiency	N_{bkgd}	FOM ($\times 10^{-4}$)	Cut efficiency
-0.3	0.3	0.1998	77841	7.12	0.91
-0.3	0.25	0.1993	73064	7.33	0.908
-0.3	0.2	0.1986	68120	7.57	0.905
-0.3	0.15	0.1974	62965	7.82	0.899
-0.3	0.1	0.1949	57662	8.06	0.888
-0.3	0.05	0.1857	52301	8.07	0.846
-0.25	0.3	0.1949	67369	7.46	0.888
-0.25	0.25	0.1944	62593	7.73	0.886
-0.25	0.2	0.1937	57649	8.02	0.882
-0.25	0.15	0.1924	52494	8.35	0.877
-0.25	0.1	0.19	47191	8.68	0.865
-0.25	0.05	0.1808	41789	8.78	0.824
-0.2	0.3	0.1878	58054	7.75	0.856
-0.2	0.25	0.1874	53278	8.07	0.854
-0.2	0.2	0.1867	48334	8.43	0.85
-0.2	0.15	0.1854	43179	8.86	0.845
-0.2	0.1	0.1829	37876	9.33	0.833
-0.2	0.05	0.1738	32514	9.56	0.792
-0.15	0.3	0.1755	49798	7.81	0.8
-0.15	0.25	0.1751	45021	8.2	0.798
-0.15	0.2	0.1744	40077	8.65	0.794
-0.15	0.15	0.1731	34922	9.19	0.789
-0.15	0.1	0.1706	29619	9.83	0.778
-0.15	**0.05**	**0.1615**	**24258**	**10.3**	**0.736**
-0.1	0.3	0.1512	43170	7.23	0.689
-0.1	0.25	0.1508	38393	7.64	0.687
-0.1	0.2	0.1501	33449	8.14	0.684
-0.1	0.15	0.1488	28295	8.77	0.678
-0.1	0.1	0.1463	22993	9.56	0.667
-0.1	0.05	0.1372	17630	10.2	0.625
-0.05	0.3	0.1059	37001	5.46	0.482
-0.05	0.25	0.1055	32225	5.83	0.48
-0.05	0.2	0.1047	27281	6.28	0.477
-0.05	0.15	0.1035	22126	6.89	0.471
-0.05	0.1	0.101	16823	7.7	0.46
-0.05	0.05	0.0918	11461	8.46	0.418
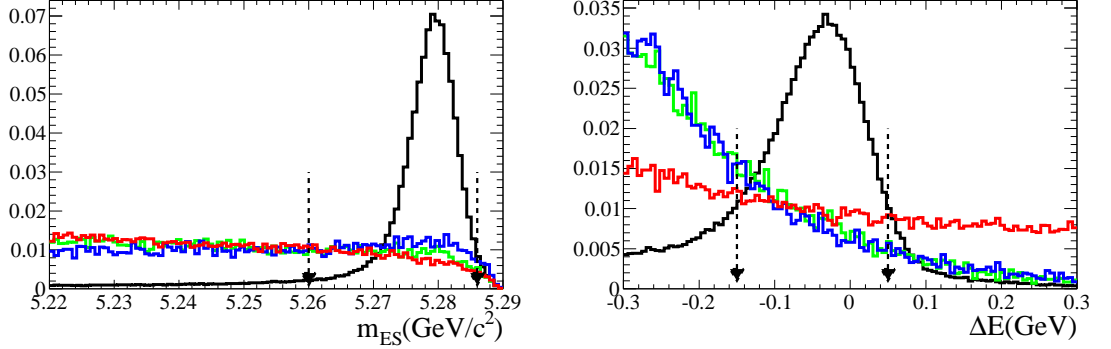
Figure 5.2: *m_{ES} and ΔE distributions for signal MC (black line), continuum background (red line), generic B^+B^- MC (green line) and generic $B^0\overline{B}^0$ MC (blue histogram). The black dashed arrows indicate the signal region. All histograms have been normalised to unity.*

instabilities and biases that can occur due to the fact that the true endpoint can vary during data taking. Figure 5.2 shows distributions for m_{ES} and ΔE taken from $B^+ \to K^+\pi^0\pi^0$ signal MC and various types of backgrounds (continuum and generic $B\overline{B}$) with arrows indicating the fitting region. The ΔE distribution is observed to be broader than in the case of decays with only charged particles in the final state and with a larger tail component. This is due to the large combined uncertainty for the energy measurement of the two π^0 mesons.

Two sideband regions are also defined. These are used only to obtain the initial values of the continuum background PDF parameters in onpeak data (see Section 5.2.2). The Upper Side Band (USB) is defined within the initial selection as:

◇ $m_{ES} > 5.22 \, \text{GeV}/c^2$,

◇ $\Delta E > 0.2 \, \text{GeV}$.

This region was chosen to minimise contamination from signal and $B\overline{B}$ backgrounds. As can be seen from Figure 5.2 the corresponding lower sideband ($\Delta E < -0.2 \, \text{GeV}$) is heavily contaminated with these events and so is not considered. The USB contains enough statistics to obtain an initial estimate of these floated PDF parameters.

The Grand Side Band (GSB) is defined as:

$\diamond$ $m_{\mathrm{ES}} < 5.26\,\mathrm{GeV}/c^2$,

$\diamond$ $-0.15\,\mathrm{GeV} < \Delta E < 0.05\,\mathrm{GeV}$.

5.2 Signal and background PDFs

This analysis makes use of the Extended Maximum Likelihood method (see Section 3.3) to discriminate between each class of events, signal and background. Therefore an important step of this analysis is the correct parametrisation of each category of events. The first step to constructing the fit is determining the PDF that accurately models the discriminating variables for each event category and extracting the parameters to include in the final fit to data.

5.2.1 Signal PDF

As discussed in Section 4.4 there are high levels of SCF in this analysis and they have strong dependence on the Dalitz-plot position. It is essential that these effects are correctly modelled in the likelihood function. As such, separate PDFs are created for TM and SCF events and are combined together as follows to create the complete signal PDF:

$$\mathcal{P}_{\mathrm{sig}} = (1 - f_{\mathrm{SCF}})\,\mathcal{P}_{\mathrm{TM}} + f_{\mathrm{SCF}}\mathcal{P}_{\mathrm{SCF}}\,, \tag{5.1}$$

where f_{SCF} is the SCF fraction, $\mathcal{P}_{\mathrm{TM}}$ and $\mathcal{P}_{\mathrm{SCF}}$ are the products of the PDFs for each discriminating variable, m_{ES} and $\mathrm{NN}_{\mathrm{out}}$.

5.2.1.1   TM signal

The m_{ES} distribution for the TM signal is modelled using the following function, known as a "Cruijff" [75]:

$$f_{\mathrm{cruijff}}(x) = \exp\left(-\frac{(x-m)^2}{2\sigma_{\pm}^2 + \alpha_{\pm}(x-m)^2}\right) \tag{5.2}$$

where m is the mean, the $+/-$ corresponds to $x > m$ or $x < m$ respectively, $\sigma_{\pm}$ is the width and $\alpha_{\pm}$ corresponds to the tail parameter of the distribution. The Cruijff function is essentially an asymmetric width Gaussian with a tail component. The PDF takes the form $f_{\mathrm{cruijff}}/\int_{x=5.260}^{x=5.286} f_{\mathrm{cruijff}}dx$. The NN$_{\mathrm{out}}$ PDF is modelled by a one dimensional histogram, obtained from nonresonant MC. The TM signal m_{ES} PDF parameters are determined from fitting nonresonant MC as shown in Figure 5.3. Note that, due to the way `Laura++` (see Section 3.3.2) is designed, the PDF in Figure 5.3 is obtained by generating a 100 toy experiments and drawing the distributions. PDFs are therefore not perfectly smooth due to finite statistics. Any PDF shown in Chaps. 5-7 have been obtained using this method, except for Section 6.4 where `RooFit` was used.
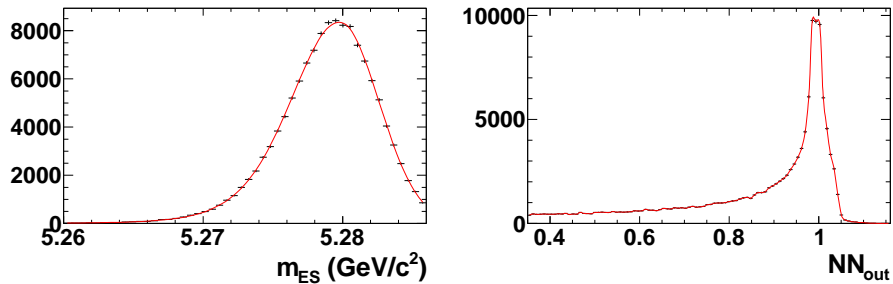
Figure 5.3: *Signal PDF distributions (red line) overlaid on nonresonant MC (black data points) for TM m_{ES} (left) and NN$_{\mathrm{out}}$ histogram (right).*

5.2.1.2   Calibration of TM signal parameters

The TM parameters obtained from the fit to the nonresonant MC in Section 5.2.1.1 need to be corrected for data/MC differences using a control sample. A control

sample of $B^+ \rightarrow \bar{D}^0 \rho^+ \rightarrow (K^+\pi^-\pi^0)(\pi^+\pi^0)$ was chosen because of the following characteristics:

◇ it provides high statistics *i.e.* has a high branching fraction $((1.87 \pm 0.28) \times 10^{-3}$ [38]);

◇ it is topologically similar to the signal decay (it has two π^0 mesons and a charged kaon in the final state);

◇ it is a well measured mode.

Events are selected using same requirements to those for the signal decay except for two additional selections for the D and ρ masses of:

◇ D mass cut: $1.84\,\text{GeV}/c^2 < m_{K^+\pi^-\pi^0} < 1.88\,\text{GeV}/c^2$,

◇ ρ mass cut: $0.65\,\text{GeV}/c^2 < m_{\pi^+\pi^0} < 0.85\,\text{GeV}/c^2$.

The selection efficiency for the control sample was found to be 3.16%. The selection was also applied to the generic $B\bar{B}$ background samples after removing events from the "signal" control sample decay channel. The major sources of $B\bar{B}$ background belong to $\bar{D}^0\pi^+$, $\bar{D}^{*0}\rho^+$, $\bar{D}^{*0}\pi^+$, $D^{*-}\rho^+$ and $D^{*-}\pi^+$.
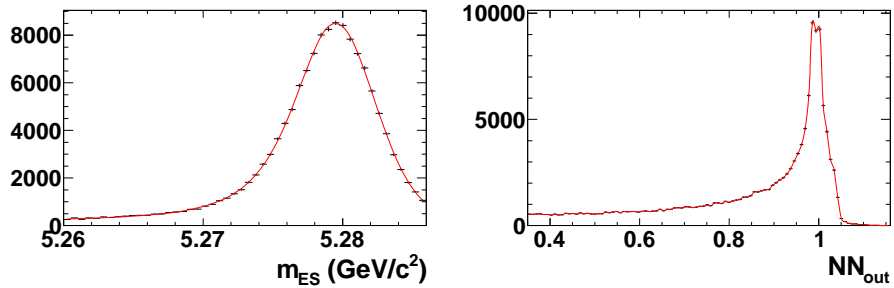
Figure 5.4: *Control channel, $B^+ \rightarrow \bar{D}^0 \rho^+ \rightarrow (K^+\pi^-\pi^0)(\pi^+\pi^0)$, PDF distributions (red line) overlaid on MC (black data points) for m_{ES} (left) and NN_{out} histogram (right).*

The fit model for the control sample is kept as close as possible to the one for the inclusive mode. The signal PDFs include a Cruijff PDF for m_{ES} and histogram for NN_{out} as shown in Figure 5.4. The $B\overline{B}$ and $q\overline{q}$ PDFs are also the same PDFs as the one listed in Table 5.1. All calibration mode PDF parameters are determined from signal MC and allowed to float in the fit to onpeak data except for the tail parameters of the Cruijff function which need to be kept fixed in order for the fit to converge. In the fit to the control sample, the $B\overline{B}$ background yields are fixed to the MC-based expectation to remain consistent with the signal fit procedure.

The expected number of signal events for the calibration sample is about 27891 and the fit obtains a yield of 27490 ± 355 which is consistent with expectation. Figure 5.5 shows the projections of the results of the fit to the calibration sample onto m_{ES} and NN_{out}, together with the distributions of each event category. The NN_{out} distribution for the continuum shows an unexpected peak, corresponding to several hundred events, in the signal region. The most likely origin of this is thought to be mismodelling of the $B\overline{B}$ backgrounds, *i.e.* $B\overline{B}$ background events being misidentified to continuum background events. Attempting to float the $B\overline{B}$ yield produced a similar peak in the continuum NN_{out}, most likely since without
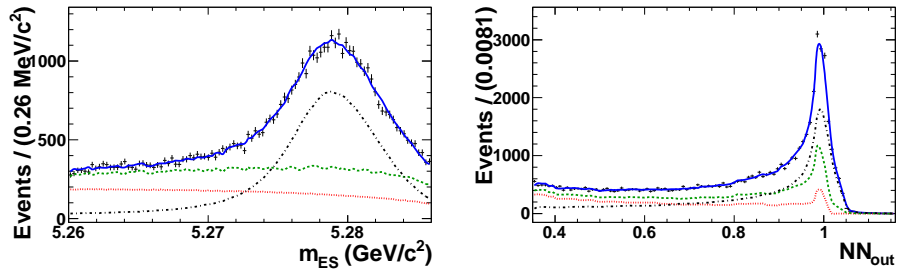
Figure 5.5: *(Left) m_{ES} and (right) NN_{out} projection distributions from the fit to the control channel $B^+ \to \overline{D}^0 \rho^+ \to (K^+\pi^-\pi^0)(\pi^+\pi^0)$. Black markers are the data points with fit overlaid (blue line), green dashed lines are the $B\overline{B}$ background, red dotted lines the continuum background and black dashed lines the signal.*
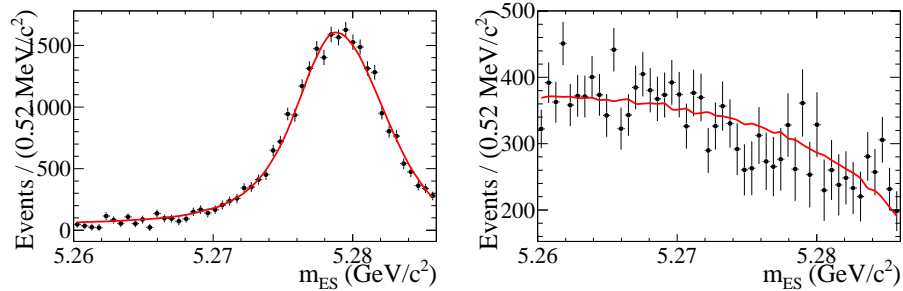
Figure 5.6: *Signal (left) and continuum (right) $_s\mathcal{P}lot$ distributions for m_{ES} obtained from a fit to the control channel $B^+ \to \overline{D}^0\rho^+ \to (K^+\pi^-\pi^0)(\pi^+\pi^0)$. Black dots show the $_s\mathcal{P}lot$ distributions and the red lines show the fit results.*

ΔE the discriminating power is insufficient. The main purpose of the fit to the control sample is to calibrate the signal m_{ES} PDF, and since the projection of the fit onto m_{ES} appears excellent both for signal and continuum, further supported by the $_s\mathcal{P}lot$ distributions shown in Figure 5.6 for signal and for continuum, the results obtained are considered reliable for the intended purpose.

The result for the m_{ES} fit parameters obtained from the calibration sample, together with the calibration factors, are given in Table 5.4. The calibration factor for the mean, μ, is applied additively, while those for the right and left rms, $\sigma_{R,L}$, are applied multiplicatively. Finally Table 5.5 gives the uncorrected and corrected values for the TM signal m_{ES} distribution parameters. All the parameters are kept fixed to their corrected values in the fit to the inclusive $B^+ \to K^+\pi^0\pi^0$ decay in onpeak data.

5.2.1.3 SCF signal

The m_{ES} distribution for SCF events is modelled using a 3rd order Chebychev polynomial with parameters obtained from nonresonant MC. The recurrence relation for this Chebychev polynomial is of the form [76]:

$$T_0(x) = 1 \tag{5.3}$$

$$T_1(x) = x \tag{5.4}$$

Table 5.4: *Signal $m_{\rm ES}$ PDF parameters obtained from fit to control sample in MC and data, together with the obtained data/MC calibration factors.*

Parameter	MC(GeV/c^2)	Data(GeV/c^2)	Correction factor
μ	5.279481 ± 0.000036	5.278779 ± 0.000077	-0.000703 ± 0.000085
σ_R	0.002804 ± 0.000047	0.003370 ± 0.000072	1.202 ± 0.033
σ_L	0.00277 ± 0.000028	0.00267 ± 0.00007	0.966 ± 0.027
α_R	0.074 ± 0.013	0.074	Not applicable
α_L	0.2528 ± 0.0019	0.2528	Not applicable

Table 5.5: *The uncorrected values of the parameters for the TM signal $m_{\rm ES}$ Cruijff, obtained from a fit to $B^+ \to K^+\pi^0\pi^0$ nonresonant MC with errors, together with the values calibrated using the data/MC correction factors and errors obtained from the control sample Table 5.4.*

Parameter	Uncorrected (MeV/c^2)	After calibration (MeV/c^2)
μ	5279.77 ± 0.03	5279.07 ± 0.09
σ_R	2.82 ± 0.03	3.39 ± 0.23
σ_L	3.42 ± 0.02	3.30 ± 0.16
α_R	-0.005 ± 0.011	-0.005
α_L	0.101 ± 0.002	0.101

$$T_{n+1}(x) \;\; = \;\; 2xT_n(x) - T_{n-1}(x) \,. \tag{5.5}$$

The Chebychev polynomial PDF is then constructed as:

$$C_n(x) = 1 + \sum_{i=1,n} a_i T_i(x) \tag{5.6}$$

where a_i are the fit parameters. The NN$_{\rm out}$ PDF is modelled from the one dimensional histogram of the distribution in nonresonant signal MC. Figure 5.7 shows the fit to the $m_{\rm ES}$ distribution and the NN$_{\rm out}$ histogram. Table 5.6 lists the parameters obtained from the fit to the MC $m_{\rm ES}$ distribution. These are all fixed in the final fit to data.
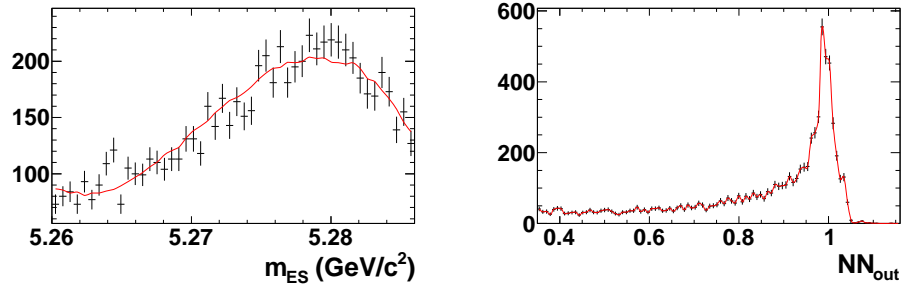
Figure 5.7: *Signal PDF distributions (red line) overlaid on nonresonant MC (black data points) for SCF signal m_{ES} (left) and NN_{out} (right).*

Table 5.6: *Parameters of the signal SCF m_{ES} PDF (a 3rd order Chebychev polynomial). Values and their uncertainties are obtained from a fit to the nonresonant MC distributions. All of these parameters are fixed in the fit to data.*

Parameter	Value	Units
a_1	344.6 ± 22.1	$(\text{MeV}/c^2)^{-1}$
a_2	-205.5 ± 21.5	$(\text{MeV}/c^2)^{-2}$
a_3	-194.7 ± 19.7	$(\text{MeV}/c^2)^{-3}$

5.2.2 Continuum background PDFs

Continuum is the dominant background in rare B decays. Both m_{ES} and the NN_{out} in the fit provide discrimination from signal. The m_{ES} distribution for continuum is modelled by an ARGUS distribution, written as follows [77]:

$$f(x) = x\sqrt{\left(1 - \left(\frac{x}{m_0}\right)^2\right)} \exp\left[-c\left(1 - \left(\frac{x}{m_0}\right)^2\right)\right] \qquad (5.7)$$

where c is the shape parameter and m_0 the endpoint $(= \sqrt{s}/2)$. The initial value of the x parameter (floated in the fit) is determined from a fit to offpeak data, and is given in Table 5.7.

The NN_{out} distribution is modelled using a parametric step function with 20 bins. The initial values of the parameters are obtained from a fit to upper and grand side-

Table 5.7: *Parameters for the $q\bar{q}$ m_{ES} PDF. The initial value given is obtained from the fit to offpeak data.*

Parameter	m_{ES} $q\bar{q}$ (ARGUS)	
c	12.76 ± 0.94	Floating
m_0	5289 MeV/c^2	Fixed

band in onpeak data to eliminate any signal. To eliminate most $B\bar{B}$ background so that the onpeak sample becomes a good representation of the continuum distribution, expected number of $B\bar{B}$ background events were calculated in the sideband regions. Then histograms for the NN_{out} were formed by scaling the B^+B^- and $B^0\bar{B}^0$ generic MC by the number of expected events in each sideband and finally subtracted from the distribution in onpeak data. All parameters of the step function are listed in Table 5.8. Both PDFs together with the original distributions for m_{ES} and NN_{out} are shown in Figure 5.8. All the parameters for the continuum are kept floating in the fit to data except for the endpoint of the ARGUS function.
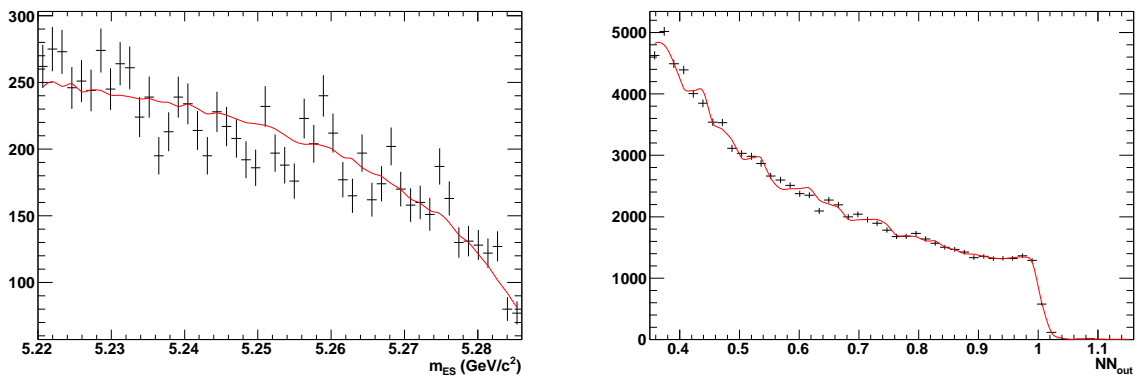
Figure 5.8: *$q\bar{q}$ PDF distributions (red line) overlaid on: offpeak data for m_{ES} (left) and USB and GSB onpeak data with $B\bar{B}$ background subtracted for NN_{out} (right).*

Table 5.8: *Initial values of the parameters for the $q\bar{q}$ $\mathrm{NN_{out}}$ PDF obtained from a fit to the sideband regions (USB and GSB) in onpeak with $B\overline{B}$ backgrounds subtracted.*

Parameter	$\mathrm{NN_{out}}$ $q\bar{q}$ (Parametric step function)		
	Bin lower limit		
Bin 0	0.35	–	–
Bin 1	0.38	2.855	Floating
Bin 2	0.4	2.516	Floating
Bin 3	0.45	2.098	Floating
Bin 4	0.49	1.836	Floating
Bin 5	0.55	1.641	Floating
Bin 6	0.63	1.435	Floating
Bin 7	0.67	1.258	Floating
Bin 8	0.75	1.077	Floating
Bin 9	0.8	1.107	Floating
Bin 10	0.84	0.948	Floating
Bin 11	0.87	0.780	Floating
Bin 12	0.9	0.857	Floating
Bin 13	0.92	0.837	Floating
Bin 14	0.96	0.928	Floating
Bin 15	0.98	0.787	Floating
Bin 16	1.0	0.363	Floating
Bin 17	1.01	0.202	Floating
Bin 18	1.02	0.040	Floating
Bin 19	1.04	0.001	Floating
Upper Limit	1.16		

5.2.3 $B\overline{B}$ Background PDFs

This type of background includes decays of $B\overline{B}$ pairs to modes other than the signal modes. It also includes misreconstructed decays via intermediate charm, charmonium states or other charmless decays. In order to look for these specific $B\overline{B}$ decays in the data set, the same selection process as applied to offpeak data and signal MC is applied to generic $B\overline{B}$ MC. For modes that are identified as potential sources of significant background, further studies are carried out using dedicated MC samples. The remaining $B\overline{B}$ background modes are modelled using the generic samples where all background modes explicitly studied were removed. Modes with similar ΔE distributions are combined into the following categories:

1. 2-body modes (mainly $B^+ \rightarrow K^+\pi^0$);

2. 3-body or 2-body modes with $\pi - K$ misidentification (mainly $B^+ \rightarrow K^{*+}\gamma$ and $B^+ \rightarrow \pi^+\pi^0\pi^0$);

3. missing tracks (mainly generic and B to charm);

4. 4-body modes (mainly $B^+ \rightarrow K^+\pi^0\pi^0\pi^0$ with or without intermediate states (including charm)).

As shown in Figure 5.9, the only significant difference between the distributions in signal and in the $B\overline{B}$ backgrounds is in the shape of the ΔE distribution. Whereas ΔE peaks around zero in signal, the $B\overline{B}$ backgrounds do not peak in ΔE since these events had either misreconstructed particles or missing particles and therefore the sum of the energies do not correspond to the energy of the B meson. Since ΔE is not included in the fit, there is not enough discrimination power to distinguish between signal and $B\overline{B}$ background therefore $B\overline{B}$ background yields are fixed to the expected values determined from MC to avoid under or over estimating the signal. To compensate for this, systematic uncertainties are added to the final results (see Section 8.1.3). The m_{ES} and $\mathrm{NN}_{\mathrm{out}}$ PDFs for each $B\overline{B}$ background category are
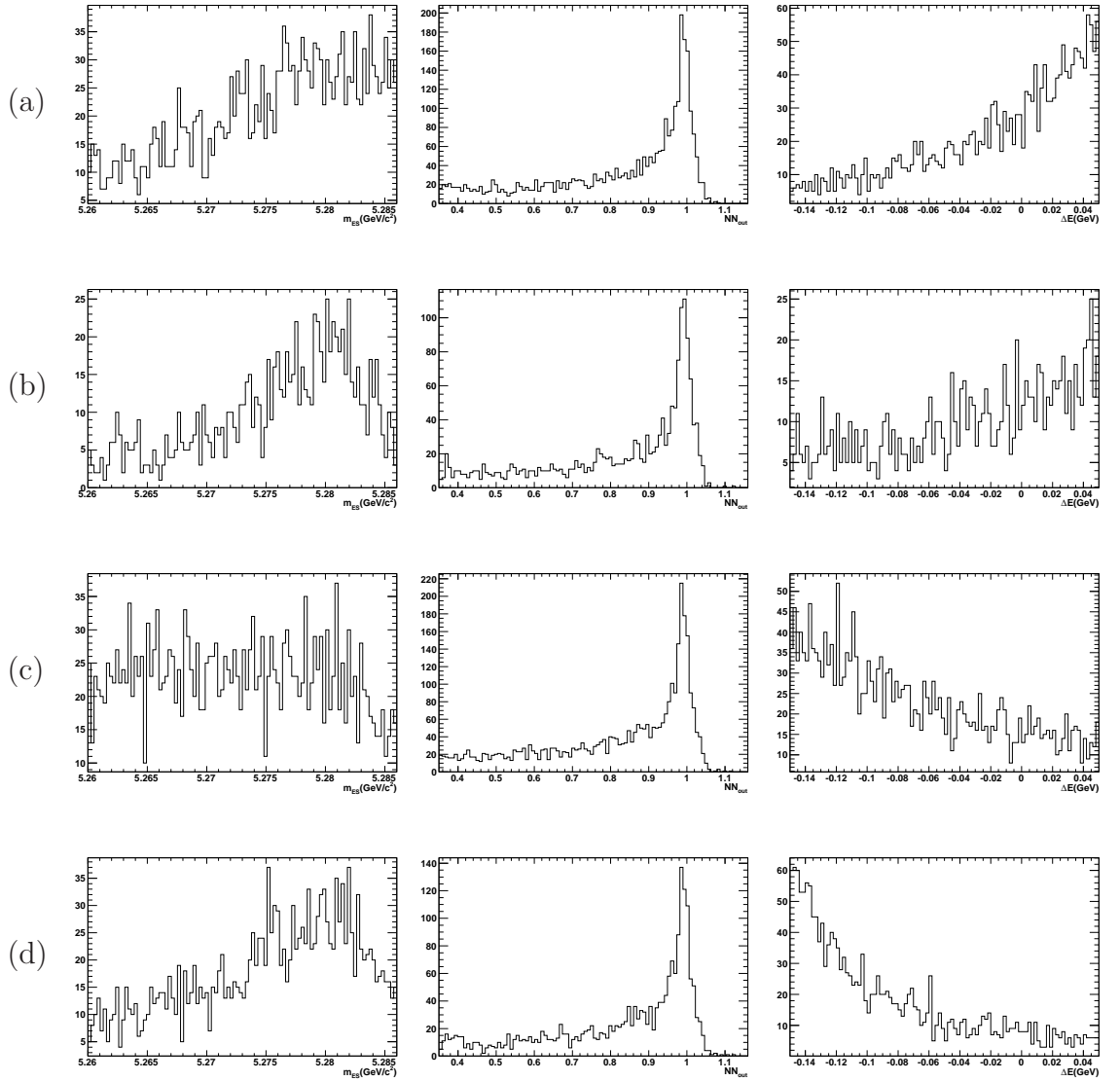
Figure 5.9: *Distributions of m_{ES} (left), NN_{out} (centre) and ΔE (right) for: (a) category 1 $B\overline{B}$ backgrounds, (b) category 2 $B\overline{B}$ backgrounds, (c) category 3 $B\overline{B}$ backgrounds, (d) category 4 $B\overline{B}$ backgrounds.*

formed from the one dimensional histograms scaled by the number of expected events for each of these background decays in data. The histograms can be seen in Figure 5.9.

5.3 Expected yields

It is useful to have a rough estimate of how many events it is expected to have in signal and each type of background. The number of expected yields is used to test the fit model in toy experiments (see Section 6.1) or validate methods later on in this chapter. The number of expected events in data for a specific decay mode follows from the definition of the branching fraction as:

$$N_{\text{exp}} = N_{B\overline{B}} \times \mathcal{B} \times \epsilon_{\text{eff}} \tag{5.8}$$

where N_{exp} is the number of expected events of that specific decay in the onpeak dataset, $N_{B\overline{B}}$ is the total number of $B\overline{B}$ pairs in the data sample (see Section 3.1.3), $\mathcal{B}$ is the branching fraction of that mode and $\epsilon_{\text{eff}} = \frac{N_{\text{sig}}}{N_{\text{gen}}}$ is the efficiency and is measured as the number of events in the signal region over the total generated events. Sometimes the branching fraction of the full decay or part of the decay is unknown, like for $B^+ \to K^+\pi^0\pi^0$. In these cases the branching fraction is estimated using known measurements and isospin rules.

5.3.1 Expected signal yield

The expected number of signal events was estimated from yields quoted in the analysis of $B^0 \to K^0_S\pi^0\pi^0$ [78]. This analysis did not include a branching fraction measurement but the efficiency from their measurement was used to estimate the branching fraction of $B^0 \to K^0_S\pi^0\pi^0$ which was found to be around 3.44×10^{-6}. By SU(3) flavour symmetry the branching fraction for $B^+ \to K^+\pi^0\pi^0$ is expected to be equal to the branching fraction of $B^+ \to K^0\pi^0\pi^0$. Since there is a 50% chance that

the K^0 is a K^0_S or a K^0_L, the branching fraction for $K^+\pi^0\pi^0$ was estimated to be roughly twice that of $B^0 \to K^0_S\pi^0\pi^0$:

$$\mathcal{B}\left(B^+ \to K^+\pi^0\pi^0\right) = 2\mathcal{B}\left(B^0 \to K^0_S\pi^0\pi^0\right) = 6.87 \times 10^{-6} \tag{5.9}$$

Using Eq. 5.8, 520 number of signal events are expected in the full dataset.

5.3.2 Background yields

The number of expected events in the continuum sample is estimated by scaling the number of events found in the signal region in offpeak data, N_{off}, by the ratio of onpeak and offpeak luminosities, $\mathcal{L}_{\text{on}}$ and $\mathcal{L}_{\text{off}}$, as follows:

$$N^{q\bar{q}}_{\text{exp}} = N^{q\bar{q}}_{\text{off}}\frac{\mathcal{L}_{\text{on}}}{\mathcal{L}_{\text{off}}} \tag{5.10}$$

The luminosities in the full dataset are listed in Table 3.2. Using these numbers, 28785 continuum events are expected to be present in the onpeak data.

For the $B\bar{B}$ backgrounds, the number of expected events in each $B\bar{B}$ category is the sum of the expected number of events of the decays composing it. All of the contributions are listed in Table 5.10. Expected number of events for exclusive decay were found using Eq. 5.8 and the branching fractions obtained from PDG [38] or HFAG [33]. In some cases, either the full branching fraction or part of the product branching has not yet been measured (decays with † in Table 5.10). In those cases techniques similar to the one used to estimate the signal yield were used. The expected number of events in the generic $B\bar{B}$ sample were simply calculated by multiplying the efficiency by the total number of $B\bar{B}$ events in the full dataset, $N_{B\bar{B}}$:

$$N^{B\bar{B}}_{\text{exp}} = N_{B\bar{B}}\frac{N^{B\bar{B}}_{\text{sig}}}{N^{B\bar{B}}_{\text{gen}}} \tag{5.11}$$

In total 69.8 ± 8.7 events are expected in the first category, 39.4 ± 18.1 events in the second, 1092.5 ± 44.6 events in the third and finally 166.7 ± 33.8 events in the last category. A summary of expected yields for signal and background and which yields were allowed to float in the fit can be found in Table 5.9.

Table 5.9: *Expected numbers of events used in the generation process in each signal and background category and their status in the fit. Uncertainties on the measured values are given for the fixed yields.*

Event category	Generated events	Fit status
Signal	520	floating
$q\bar{q}$	28785	floating
$B\bar{B}$ 1	70 ± 9	fixed
$B\bar{B}$ 2	39 ± 18	fixed
$B\bar{B}$ 3	1092 ± 45	fixed
$B\bar{B}$ 4	168 ± 34	fixed

5.4 Determination of SCF fraction

As mentioned in Section 4.4.2, the Dalitz plot distribution of the SCF fraction cannot be used since Dalitz plot parameters are not included in the fit to data. f_{SCF} also cannot be floated since the fit cannot determine the proportions of SCF and TM signal from the PDFs. The SCF fraction must therefore be fixed. To solve this problem, an iterative procedure was adopted where the fit is repeated until both SCF fraction and signal yield have converged. The procedure uses the $_s\mathcal{P}lots$ information (see Section 3.4 on the $_s\mathcal{P}lots$ technique) to retrieve the Dalitz plot distribution and calculate the SCF fraction as follows:

◇ fit with the value of f_{SCF} fixed to half way between highest and lowest SCF fraction measured in the MC signal modes – 17.5% (Table 4.3);

◇ determine the signal Dalitz plot distribution using $_s\mathcal{W}eights$ corrected for the fixed $B\bar{B}$ background (see Section 3.4.2);

◇ calculate the average f_{SCF} from

$$f_{\mathrm{SCF}} = \frac{\sum_{i\in\mathrm{DP}} \left(s\mathcal{W} \times f_{\mathrm{SCF}}^{DP} \right)}{\sum_{i\in\mathrm{DP}} s\mathcal{W}} \tag{5.12}$$

Table 5.10: *Table of branching fractions and CP asymmetry (if known) for each B background mode along with the expected number of events in the signal region. The "DP" next to the mode description indicates that the Dalitz plot model is used and therefore MC includes nonresonant and resonant contributions. The values listed are found using the world averages taken from HFAG [33] and PDG [38]. A decay with † indicates that the full or part of the branching fraction was estimated using isospin relations.*

Mode #	Mode + CC	BF($\times 10^{-5}$)	A_{CP}(%)	Number of Expected Events in Signal Box	Background Category
1587	$B^+ \to K^+\pi^0$	(1.29 ± 0.06)	(5 ± 2)	48.1 ± 2.4	1
1940	$B^+ \to \rho^+\pi^0$	(1.09 ± 0.02)	(2 ± 11)	21.3 ± 0.8	1
1048	$B^+ \to \pi^+\pi^0$	(0.56 ± 0.04)	(6 ± 5)	0.3 ± 0.1	1
1713	$B^+ \to K^{*+}\gamma;\ K^{*+} \to K^+\pi^0$	(1.33 ± 0.26)	(18 ± 29)	12.3 ± 2.4	2
1938	$B^+ \to \pi^+\pi^0\pi^0$	(1.09 ± 1.09)	–	16.7 ± 16.7	2
6948	$B^0 \to \pi^+\pi^-\pi^0$	(2.5 ± 0.2)	–	6.3 ± 0.6	2
1765	$B^+ \to K_2^{*+}(1430)\gamma$	(1.4 ± 0.4)	–	2.5 ± 0.7	2
1972	$B^+ \to K^{*+}(1680)\gamma^\dagger$	(1.4 ± 1.4)	–	1.6 ± 1.6	2
2436	$B^+ \to \overline{D}^0\rho^+;\ \overline{D}^0 \to K^+\pi^-$	(52.1 ± 7.0)	0	122.7 ± 17.0	3
2441	$B^+ \to \overline{D}^0\rho^+;\ \overline{D}^0 \to K^+\pi^-\pi^0$	(186 ± 10)	0	81.4 ± 6.0	3
8523	$B^0 \to K^+\pi^-\pi^0$(DP)	(3.6 ± 0.3)	(0 ± 10)	149.9 ± 12.4	3
3585	$B^+ \to \rho^+\pi^0\pi^0$	(6.0 ± 0.5)	–	21.6 ± 2.1	3
4957	$B^+ \to a_1^+\pi^0;\ a_1^+ \to \rho^+\pi^{0\dagger}$	(1.3 ± 1.3)	–	13.0 ± 13.0	3
4874	$B^+ \to a_1^0 K^+;\ a_1^0 \to \rho^\mp\pi^\pm$	(0.7 ± 0.3)	(12 ± 11)	6.1 ± 3.0	3
9611	$B^+ \to K^+ K_S^0$	(0.068 ± 0.027)	(12^{+17}_{-18})	1.3 ± 0.5	3
9624	$B^+ \to f_0(980)\rho^+;\ f_0(980) \to \pi^0\pi^{0\dagger}$	(0.05 ± 0.05)	(5 ± 5)	0.5 ± 0.5	3
2191	$B^0 \to \overline{D}^0\pi^0;\ \overline{D}^0 \to K^+\pi^-\pi^0$	(3.63 ± 0.16)	0	48.5 ± 2.9	4
9595	$B^+ \to K^+\pi^0\pi^0\pi^{0\dagger}$	(3.77 ± 3.77)	–	23.0 ± 23.0	4
4148	$B^0 \to K^{*0}\pi^0\pi^0;\ K^{*0} \to K^+\pi^{-\dagger}$	(1.81 ± 0.26)	–	19.9 ± 2.9	4
9597	$B^+ \to K^{*+}\pi^0\pi^0;\ K^{*+} \to K^+\pi^{0\dagger}$	(1.2 ± 0.5)	–	17.2 ± 7.0	4
4960	$B^0 \to a_1^+ K^-;\ a_1^+ \to \rho^+\pi^{0\dagger}$	(0.81 ± 0.81)	(-16 ± 12)	15.7 ± 16.1	4
9596	$B^+ \to \eta'K^+;\ \eta' \to \eta\pi^0\pi^0$	(1.45 ± 0.09)	(-26 ± 27)	15.6 ± 1.1	4
7615/6	$B^0 \to \rho^- K^{*+\dagger}$	(0.4 ± 0.4)	–	11.1 ± 10.7	4
9598	$B^+ \to \overline{D}^0 K^+;\ \overline{D}^0 \to K_S^0\pi^0$	(0.50 ± 0.04)	-10 ± 8	9.7 ± 0.8	4
9623	$B^+ \to f_0(980)K^{*+};\ f_0(980) \to \pi^0\pi^{0\dagger}$	(0.09 ± 0.02)	(-34 ± 21)	3.5 ± 0.9	4
9612	$B^+ \to \eta K^+;\ \eta \to \pi^0\pi^0\pi^0$	(0.09 ± 0.03)	(-37 ± 9)	2.4 ± 0.8	4
1235	Generic B^+B^-	—	—	401.7 ± 11.8	3
1237	Generic $B^0\overline{B}^0$	—	—	294.2 ± 10.0	3
	Total B backgrounds	—	—	1368 ± 47	

where the sum is over the bins in the Dalitz plot, $s\mathcal{W}$ is the $_s\mathcal{W}eights$ value obtained after correction for fixed $B\bar{B}$ backgrounds, f_{SCF}^{DP} is the SCF value in the relevant square Dalitz plot bin (refer to Figure 4.12);

◇ refit using the new value of f_{SCF};

◇ repeat until the results for the total signal yield and the SCF fraction are stable.

This method was tested by generating a single toy experiments for each nonresonant and four other resonant signal MC for which the SCF fractions are known from Table 4.3. In each toy, 520 signal events were randomly selected from the corresponding MC samples. This method is known as "embedded toy experiments" and is described in detail later on in Section 6.1.2. The generated MC experiment is then fitted multiple times and for each iterations f_{SCF} and yields are calculated until both quantities have converged. The number of signal yields and the SCF fractions at which these toy experiments converge to, reported in Table 5.11, seem to indicate

Table 5.11: *Table of values of f_{SCF}, calculated using $_s\mathcal{W}eights$, and signal yields after each iteration, up to convergence, of the fit to single toy experiments generated using MC for each nonresonant (NR), $K^*(892)^+\pi^0$, $K_2^{*+}(1430)\pi^0$, $K^{*+}(1410)\pi^0$ and $f_2(1270)K^+$.*

Mode	MC f_{SCF} (%)	Initial f_{SCF} (%)	Calculated f_{SCF} (%)	N_{SIG}
NR	5.3	17.5	7.25	696 ± 80
		7.25	7.22	644 ± 75
		7.22	7.21	642 ± 75
$K_2^{*+}(1430)\pi^0$	16.9	17.5	15.0	604 ± 79
		15.0	14.9	591 ± 78
$f_2(1270)K^+$	28.4	17.5	23.8	691 ± 80
$K^*(892)^+\pi^0$	21.2	17.5	17.5	658 ± 79
$K^{*+}(1410)\pi^0$	18.1	17.5	16.5	609 ± 79
		16.5	16.4	604 ± 79

the presence of small biases when compared to their expected values.

To check if these biases are symptomatic of an underlying problem or just lack of statistics, another test was performed using successful fits to 100 toy MC experiments instead of just 1. These toy experiments were generated using the nonresonant and $K^*(892)^+\pi^0$ MC. Out of 100 pseudo-experiments, 76 (79) fits succeed to complete all iterations in the nonresonant ($K^*(892)^+$) ensembles. Figure 5.10 shows how the signal yields and SCF fractions are distributed. Gaussian fits give mean signal yields of 552 ± 10 for nonresonant and 550 ± 9 for $K^{*+}(892)\pi^0$, corresponding to biases of 32 and 31 events, respectively. A Gaussian fit to the distribution of f_{SCF} shows that the iterative procedure converges to values close to the correct f_{SCF} but
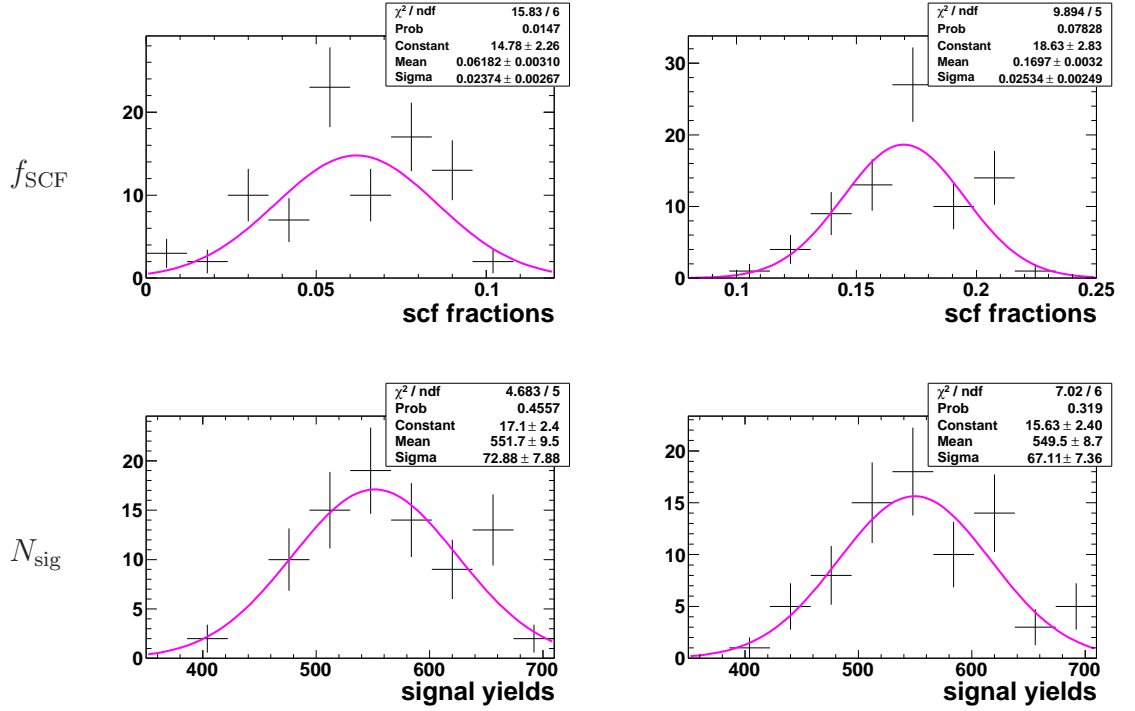
Figure 5.10: *Gaussian fits to the distributions of fitted SCF fractions and signal yields. The toy experiments are generated from: nonresonant MC – expected SCF fraction 5.3 % (left); $K^*(892)^+\pi^0$ MC – expected SCF fraction 21.2 % (right).*

with some bias. This is not surprising since the value of f_{SCF} is obtained from the reconstructed Dalitz plot distribution obtained using $_s\mathcal{W}eights$ rather than the true distribution. However, the results are sufficient to allow the extraction of a conservative range within which to vary f_{SCF} and estimate associated systematic uncertainties (see Section 8.1.3). This range will be given as a linear function of the fitted f_{SCF}, varying from 0.026 for fitted f_{SCF} of 6% to 0.049 for fitted f_{SCF} of 17%. The statistical uncertainty on the value of f_{SCF} as a result of this procedure is obtained by the width of the f_{SCF} distribution of these toy experiments.

In order to test whether the DP distribution reconstructed from $_s\mathcal{W}eights$ after iterations converge gives a good description of the true underlying distribution, a high-statistics MC test was carried out. This test was performed using 100 times the number of expected signal events. Figure 5.11 shows the square Dalitz plot distributions obtained from nonresonant signal MC and from the $_s\mathcal{W}eights$. The two appear consistent.
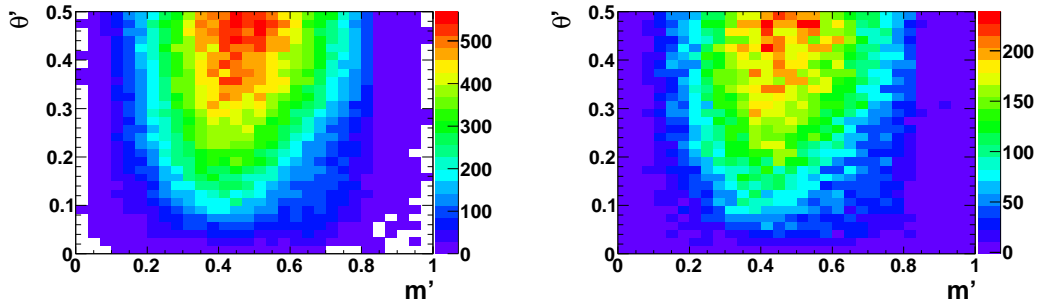
Figure 5.11: *Signal Dalitz plot distributions calculated from the signal MC (right) and from the $_s\mathcal{W}eights$ (left) for the nonresonant mode. The Dalitz plot distribution to the left is obtained by running a toy experiment of 100 times the number of expected signal events.*

5.5 Determining the resonant branching ratios

One of the aims of this analysis is to use the Dalitz plot of $B^+ \to K^+\pi^0\pi^0$ to obtain information on the quasi-two-body contributions from $B^+ \to K^*(892)^+\pi^0$, $B^+ \to f_0(980)K^+$ and $B^+ \to \chi_{c0}K^+$ (see Section 1.3). Since the fit returns the $_s\mathcal{P}lot$ of the signal decay, the quasi-two-body branching fractions can be measured using the following approach, elements of which have been used in previous analyses [74, 79]:

◇ select the resonance band by placing a tight cut in the appropriate mass region;

◇ sum the efficiency corrected $_s\mathcal{W}eights$;

◇ correct for the inefficiency of the selection evaluated from MC;

◇ divide by the number of $B\overline{B}$ pairs and hence extract the branching fraction.

The approach is limited by the fact that each resonance is not the only contribution to the Dalitz plot. The presence of nonresonant and other quasi-two-body decays can bias the result. Interference effects are assumed to be negligeable and although these effects between the vector $K^*(892)^+$ and scalar nonresonant contributions cancel in the mass projection, it is not the case for the $f_0(980)$ and χ_{c0}. For this assumption to be valid, the resonance needs to be reasonably narrow. The signal region was chosen as to incorporate most of the signal and at the same time avoid including too many events from nearby resonances. The mass region was defined to be $\pm 2\sigma$ around the characteristic mean of the resonance. This bias is corrected to first order by applying the same procedure to the sidebands on either side of and 1σ away from the signal region. The signal region and sideband definition for each resonance can be found in Table 5.12. The extra contribution under the signal peak is then calculated as the average yield from these sidebands:

$$N_{\mathrm{NR}} = \frac{\omega_{\mathrm{sig}}}{2} \left(\frac{N_{\mathrm{left}}}{\omega_{\mathrm{left}}} + \frac{N_{\mathrm{right}}}{\omega_{\mathrm{right}}} \right) \tag{5.13}$$

where ω_i is the width of the signal, left and right sideband and $N_{\text{left(right)}}$ the efficiency corrected yields in the sidebands. N_{NR} is the estimated yield within the resonance signal for all other $B^+ \to K^+\pi^0\pi^0$ processes which I will refer to as "nonresonant" in this context and which need subtracting from the signal yield.

Table 5.12: *Selection requirements used to isolate resonances in the relevant MC samples. Units of* GeV/c^2 *have been suppressed.*

Resonance	Signal region	Sideband region
$K^{*+}(892)$	$0.795 < m_{K^+\pi^0_{\min}} < 0.995$	$0.645 < m_{K^+\pi^0_{\min}} < 0.745$
		$1.045 < m_{K^+\pi^0_{\min}} < 1.145$
$f_0(980)$	$0.873 < m_{\pi^0\pi^0} < 1.073$	$0.723 < m_{\pi^0\pi^0} < 0.823$
		$1.123 < m_{\pi^0\pi^0} < 1.223$
χ_{c0}	$3.373 < m_{\pi^0\pi^0} < 3.453$	$3.313 < m_{\pi^0\pi^0} < 3.353$
		$3.473 < m_{\pi^0\pi^0} < 3.513$

Additionally an extended binned maximum likelihood fit is performed on the efficiency corrected invariant mass $_s\mathcal{P}lot$ distribution. The fit includes a signal component, corresponding to the resonance of interest in $B^+ \to K^+\pi^0\pi^0$, and a second component for the "nonresonant" contribution. Note that by "nonresonant" it is in fact referred to not only nonresonant but also events from the tail of nearby resonances and therefore this background component is fairly difficult to model accurately. The branching fraction extracted from this method was therefore deemed unreliable, further supported by evidence at the validation stage that the background component was being undermodelled in some of the toy experiments in Section 6.4.1. The fit is therefore only used as a comparison and to make sure results are consistent between the two methods.

The fitting region is defined to be $\pm 5\sigma$ around the characteristic mean of the resonance. This in fact includes signal region, sidebands and the gap between them

Table 5.13: *Parameters obtained from an unbinned fit to the signal MC mass distributions. These are all fixed in the fit to the mass $_s\mathcal{P}lot$.*

Resonance	Type	Parameter (GeV/c^2)	Error (GeV/c^2)
	Gaussian 1 μ	0.93218	0.00075
	Gaussian 1 σ	0.11786	0.00079
$K^*(892)^+\pi^0$	Gaussian 2 μ	0.89126	0.00014
	Gaussian 2 σ	0.02614	0.00017
	fraction	0.433	0.0033
	Gaussian 1 μ	0.9733	0.0012
	Gaussian 1 σ	0.1215	0.0017
$f_0(980)K^+$	Gaussian 2 μ	0.97182	0.00024
	Gaussian 2 σ	0.02624	0.00028
	fraction	0.3798	0.0060
	Gaussian 1 μ	3.41468	0.00049
	Gaussian 1 σ	0.03896	0.00073
$\chi_{c0}K^+$	Gaussian 2 μ	3.41281	0.00012
	Gaussian 2 σ	0.01516	0.00018
	fraction	0.272	0.0110

defined in the previous method. Each signal is fitted using a double Gaussian with parameters fixed to the values obtained in an unbinned fit to the respective MC and shown in Table 5.13. The fit to the nonresonant contribution is tested by fitting the nonresonant MC in the signal region of each resonance. For the $K^*(892)^+$ the nonresonant background was found to be best fitted by a third order polynomial due to the proximity to the mass threshold whereas the nonresonant contributions in the $f_0(980)$ and χ_{c0} regions are found to be approximately linear. The parameters for the background are allowed to float from the values obtained in the MC and listed in Table 5.14. Figure 5.12 shows the resonant and nonresonant MC distributions with the fit superimposed (blue line). The red and green arrows around the signal

Table 5.14: *Fit parameters obtained from the nonresonant MC in the reso-
nance mass region. These parameters are allowed to float in the fit to the mass
${}_s\mathcal{P}lot$.*

Mass region in NR MC	Type	Parameters (GeV/c^2)	Error (GeV/c^2)	χ^2/dof
	p_1	-43.8	1.4	
$K^*(892)^+$	p_2	99.6	7.9	1.4
$p_1 m + p_2 m^2 + p_3 m^3$	p_3	-47.3	1.7	
$f_0(980)$	p_1	3.0	1.0	0.75
$p_1 m$				
χ_{c0}	p_1	0.03	0.22	0.81
$p_1 m$				

peak indicate the signal region and sidebands respectively.

The same method is also used to measure the CP asymmetry of the resonances,
except that the distinction is made between B^+ and B^- yields and no efficiency
corrections are applied since these cancel in the ratio. The A_{CP} is calculated as:

$$A_{CP} = \frac{N_{B^-} - N_{B^+}}{N_{B^-} + N_{B^+}} \tag{5.14}$$

where $N_{B^\pm}$ are the yields after sideband subtraction of the positive and negative B
decays.
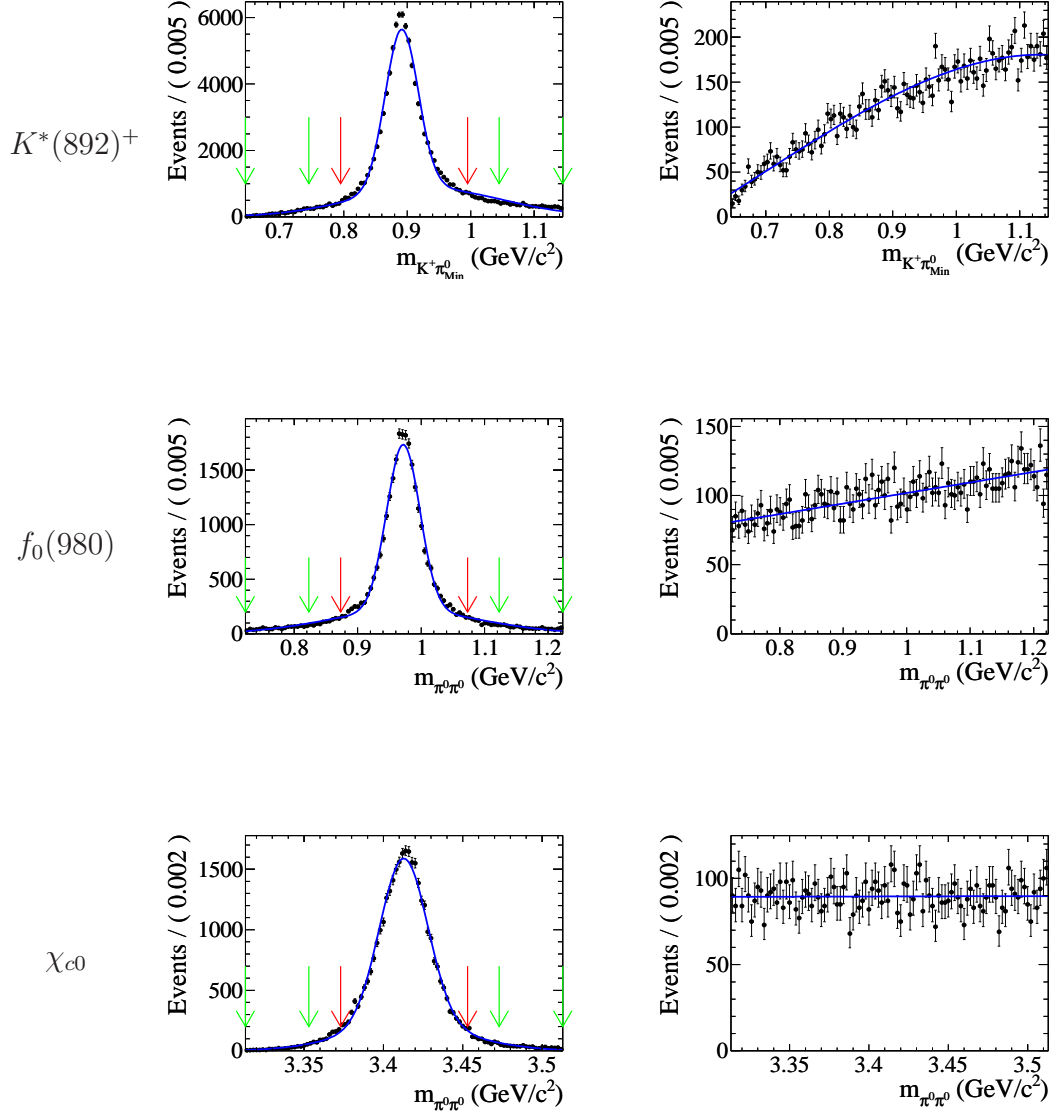
Figure 5.12: *(Top) $m_{K^+\pi^0_{min}}$ distribution in the $K^{*+}(892)$ region from $K^*(892)^+\pi^0$ (left) and nonresonant (right) MC; (Middle) $m_{\pi^0\pi^0}$ distribution in the $f_0(980)$ region, from $f_0(980)K^+$ (left) and nonresonant (right) MC; (Bottom) $m_{\pi^0\pi^0}$ distribution in the χ_{c0} region, from $\chi_{c0}K^+$ (left) and nonresonant (right) MC. The arrows in the left plots indicate the selection requirements of signal (red) and sidebands (green) used for the sideband subtraction. The blue line represents the fit to the MC data.*

6

Results: Branching Fractions

In the previous chapter, the building blocks of the maximum likelihood model were selected and optimised separately to fit the MC and offpeak data set. PDFs were established from these fits and are now combined to form the fit model. Before this model can be used on onpeak data, it needs to go through various stages of testing. These involve generating a data set and fitting the data to look for biases in the model. Finally the model is used to fit the onpeak data sample. This chapter will outline the procedure used to obtain the inclusive branching fraction of $B^+ \to K^+ \pi^0 \pi^0$ and of the three intermediate resonant contributions $B^+ \to K^*(892)^+ \pi^0$, $B^+ \to f_0(980)K^+$ and $B^+ \to \chi_{c0}K^+$. It will include the testing of the Extended Maximum Likelihood (EML) fit and the final results.

6.1 Toy tests for inclusive branching fraction

Various tests are used to validate the fit and check for biases in the model. Pure toy tests generate data in each category (signal, continuum, $B\overline{B}$) from the PDFs and then fit this data to look for biases in the pull distributions. Embedded toys on the other hand generate data by picking random events in the MC samples (signal and signal+$B\overline{B}$) and fit the new data samples to retrieve the event distribution. The mean of the signal yield distribution should be within error of the number of the expected signal events in the onpeak data sample. A toy test consists of multiple experiments, each toy experiment is generated with numbers of events based on the estimated numbers of events in the onpeak sample (see Table 5.9). All parameters are either fixed or floated according to Table 5.1. For the purpose of these validation studies the SCF fraction is fixed to the values obtained in the MC (refer to Table 4.3).

6.1.1 Pure toys

Pure toy tests are performed specifically to check if there is undercoverage or overcoverage of errors. They are also intended to check for fit instabilities and intrinsic biases. This is done by looking at the distribution of the pull on the number of signal events returned by the fit. This is calculated as:

$$\text{pull}_{signal} = \frac{n_{sig}^{fit} - n_{sig}^{gen}}{\sigma_{n_{sig}^{fit}}} \tag{6.1}$$

where n_{sig}^{fit} is the number of signal events returned by the fit, n_{sig}^{gen} is the number of signal events generated from the PDFs and $\sigma_{n_{sig}}$ is the error on the number of events returned by the fit. For each experiment, signal and background datasets are sampled using the PDFs shown in Section 5.2. The number of events generated is varied around a mean value according to the Poisson distribution. The mean value is set to be the number of expected events listed in Table 5.9. The EML fit is then performed to extract the signal yield. A Gaussian is fitted to the pull distribution. For an unbiased fit model with correct error coverage, the Gaussian for the pull

distribution should be centered on zero with unit width ($\mu = 0$ and $\sigma = 1$). If the mean shows large variations from expectations then the model is biased, *i.e.* there are large correlations between PDF parameters or floated parameters cannot be determined properly by the fit. The width of the pull distribution tests errors are correctly evaluated, *i.e.* the average signal yield uncertainty is more or less equal to the standard deviation of the difference between the mean of the signal yield distribution and the expected number of events, also known as the signal yield residual.

Table 6.1: *Mean and width of the pull distributions together with the mean of the signal yield uncertainties and their errors obtained from pure toy experiments for $B^+ \to K^+\pi^0\pi^0$ nonresonant, $B^+ \to K^*(892)^+\pi^0$ and $B^+ \to f_0(980)K^+$.*

Decay type	Mean pull	σ pull	Signal sensitivity
NR	0.13 ± 0.05	0.96 ± 0.04	76.67 ± 0.07
$K^*(892)^+\pi^0$	0.05 ± 0.05	1.02 ± 0.04	87.09 ± 0.10
$f_0(980)K^+$	0.15 ± 0.05	1.00 ± 0.05	83.11 ± 0.09

Table 6.1 lists the mean and width of the pull with the signal sensitivity obtained from the fit to the distributions of the pull and signal yield uncertainties in Figure 6.1 for nonresonant (NR), $K^*(892)^+$ and $f_0(980)$ models. The signal sensitivity is the average statistical uncertainty on the signal yield. All pull distributions are approximately centered on zero with negligeable biases and with unit width. The sensitivity of the model to the signal yield is expected to be about 7σ.

6.1.2 Embedded signal toys

The same study is repeated but instead of generating signal events from the PDFs, these are randomly selected from nonresonant and resonant MC events, see Table 3.1. This study is used to test for problems in the modelling of the distributions which

Decay type	Signal yield pull	Signal yield uncertainties
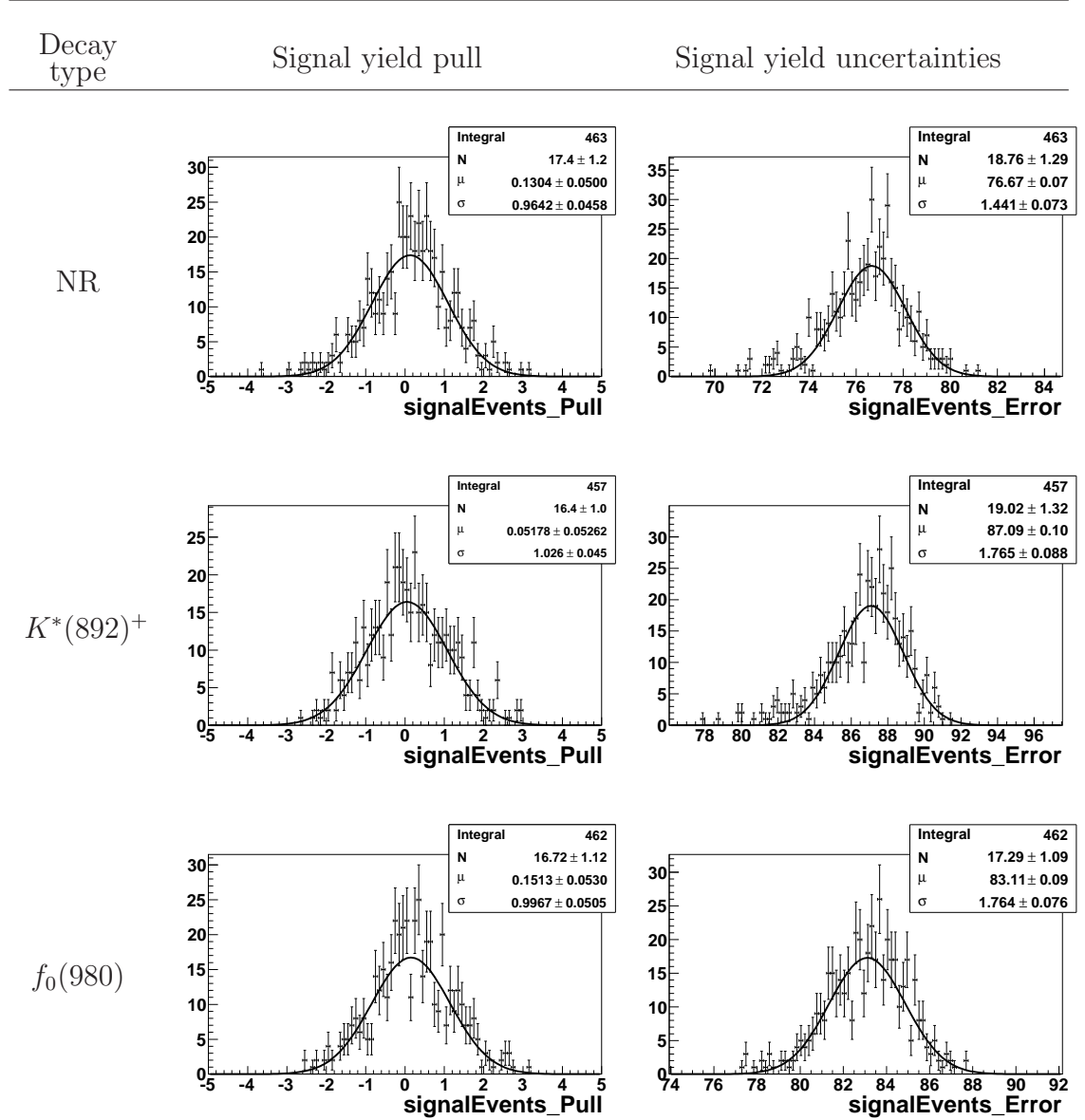
Figure 6.1: *Pull and fitted uncertainties distributions for $B^+ \to K^+\pi^0\pi^0$ nonresonant, $B^+ \to K^*(892)^+\pi^0$ and $B^+ \to f_0(980)K^+$ obtained from 500 pure toy experiments.*

could lead to biases in the fit, *e.g.* dependences of PDFs with other event variables. The Punzi effect, for example, described in Appendix A became apparent when performing this study. Poisson smearing of the number of generated events is not used in these toys. A bias in the average is inferred by looking at the distribution of the fitted number of signal events with respect to the signal yield generated by the toy experiments.

The signal yields and residuals are fitted with a Gaussian. The mean values of the Gaussian fits are listed in Table 6.2 showing the average signal yield and the biases for each of resonant and nonresonant decay modes. All average yields show good agreement with the expected yield. Slightly larger biases are visible in modes with resonance type $X \to \pi^0\pi^0$ but also those are within error. These biases will be accounted for in the systematic uncertainties discussed in Section 8.1.

Table 6.2: *List of signal yields and biases for the embedded toy experiments using nonresonant and all resonant signal MC.*

Mode	Signal yield	Signal bias
NR	532 ± 69	11 ± 4
$K^*(892)^+\pi^0$	530 ± 82	18 ± 4
$K_2^{*+}(1430)\pi^0$	530 ± 83	9 ± 4
$K^{*+}(1680)\pi^0$	521 ± 78	-1 ± 4
$K^{*+}(1410)\pi^0$	529 ± 79	8 ± 4
$K_0^{*+}(1430)\pi^0$	526 ± 77	6 ± 4
$f_0(980)K^+$	546 ± 77	29 ± 4
$f_2(1270)K^+$	550 ± 89	25 ± 4
$\chi_{c0}K^+$	537 ± 74	20 ± 3

6.1.3 Embedded signal and $B\bar{B}$ toys

The study above is repeated but this time also the $B\bar{B}$ backgrounds are embedded in the generation stage. This performs a more complete check for problems in the modelling of the distributions including backgrounds. Due to the low number of events in the second category of $B\bar{B}$ backgrounds, generating 500 experiments with the $B\bar{B}$ events would randomly chose the same event too many times and reduce the statistical significance of the test. Instead a single toy experiment was performed with three times the number of expected signal and $B\bar{B}$ events. Table 6.3 shows the results on the extracted signal yield, which is consistent with the generated value in all cases.

Table 6.3: *Results for the signal yield in one toy experiment embedding signal and $B\bar{B}$ backgrounds with $3\times$ the statistics for nonresonant, $K^*(892)^+$ and $f_0(980)$.*

Decay type	Yield from fit	Error	Generated value
NR	1606.96	99.83	1559.7
$K^*(892)^+\pi^0$	1624.86	110.15	1559.7
$f_0(980)K^+$	1612.75	105.28	1559.7

6.2 Testing the fit to data

In the previous section the fit was tested by generating events from the MC samples and PDFs. The fit model is now tested using offpeak data in the signal region and a blind fit to onpeak data. The latter is called a "blind" fit because only the background distributions are extracted from the fit. These further tests to the model are done to establish that no signal yield is returned when there is no signal and backgrounds in data are correctly modelled.

The first test was performed on 3007 offpeak events in the fitting region with a signal

component and a SCF fraction of 10%. The $B\bar{B}$ background yields were all fixed to zero. Table 6.4 gives the yields obtained for both continuum and signal. The signal yield is zero within error, as expected.

Table 6.4: *Yield results for the fit to offpeak data allowing for a signal component including the errors.*

Event type	Fitted yield	Error on yield
Signal	-14.9	21.1
Continuum	3022	59

In the blind fit to onpeak data, only the continuum background $_s\mathcal{P}lots$ are plotted with the fit overlaid. This fit also includes $B\bar{B}$ backgrounds and their yields are fixed again to the values in Table 5.9. For this fit, parameters for the TM signal m_{ES} PDF have been calibrated using the corrections obtained from the control sample, see Section 5.2.1.2. The projection distributions in Figure 6.2 show the distribution for m_{ES} and NN_{out} of the events classified by the fit as continuum background. The distributions are in good agreement with the expected distribution for continuum in Figure 5.8.

6.3 Inclusive branching fraction of $B^+ \to K^+\pi^0\pi^0$

To obtain the branching fraction the following steps have to be taken:

◇ apply fit model to the onpeak data and extract the signal yield,

◇ obtain the $_s\mathcal{W}eights$ Dalitz plot and correct for efficiency,

◇ correct yield for fit biases and vetos.
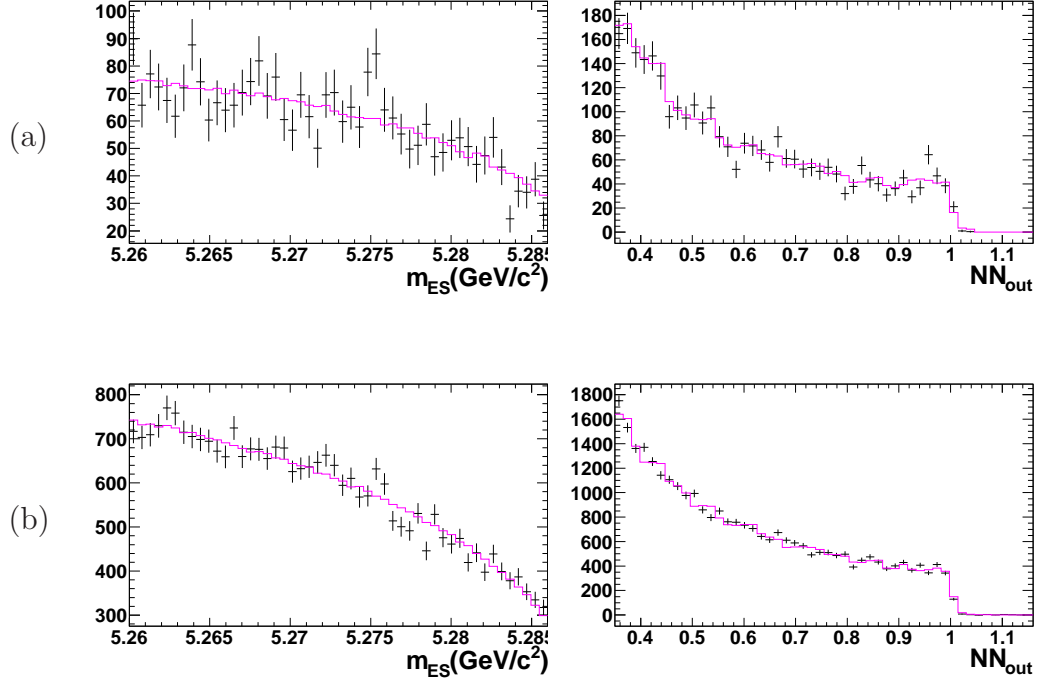
Figure 6.2: *Projections of m_{ES} (left) and NN_{out} (right) for the fit to: (a) offpeak data and (b) blind fit to onpeak. The data points (black) show the $_s\mathcal{P}lot$ distribution for the continuum background and the line (magenta) is the background PDF generated from the fit.*

The inclusive branching fraction and its statistical error are then obtained from the following relations:

$$\mathcal{B}(B^+ \to K^+\pi^0\pi^0) = \frac{N_{\text{corr}}}{N_{B\bar{B}}} \tag{6.2}$$

$$\sigma_{\mathcal{B}}^{stat} = \frac{\sigma_{n_{\text{sig}}}}{N_{B\bar{B}} \times \epsilon_{\text{avg}}} \tag{6.3}$$

where N_{corr} is the signal yield corrected for efficiency, biases and vetos, $\epsilon_{\text{avg}} = \frac{N_{\text{sig}}}{N_{\text{corr}}}$ is the average efficiency as a ratio of uncorrected and corrected signal yield, $N_{B\bar{B}} = (470.9 \pm 2.8) \times 10^6$ is the total number of $B\bar{B}$ in the onpeak data set and $\sigma_{n_{sig}}$ is the error on the yield returned by the fit.

Table 6.5: *Results at each iteration for f_{SCF} and signal yield of the fit to data up to convergence.*

Initial f_{SCF} used in fit (%)	Calculated f_{SCF} after fit (%)	N_{sig}
17.5	9.939	1291
9.939	9.713	1222
9.713	9.707	1220

6.3.1 Fit results

The fit is applied to the onpeak data sample containing $31,673$ events after all selection requirements. The iteration procedure is applied to determine the SCF fraction and, as indicated in Table 6.5, the fit is found to converge after three iterations to $f_{\mathrm{SCF}} = 9.7\%$ and a signal yield of 1220 ± 85.

Figure 6.3 shows the projection of the fit results onto the discriminating variables.
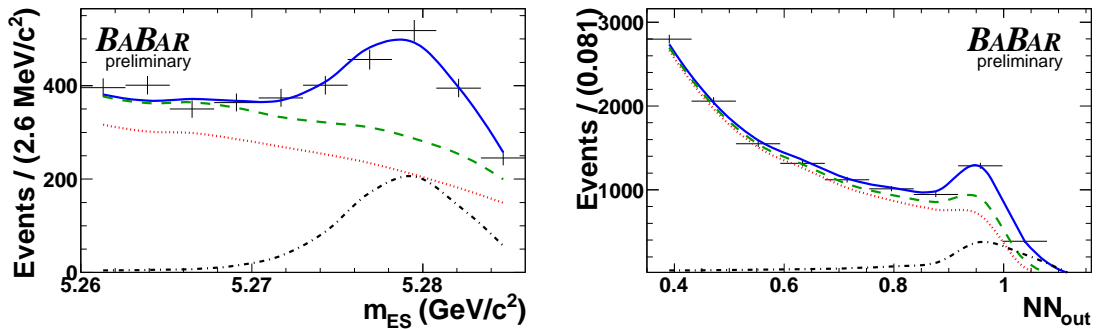
Figure 6.3: *Projection distributions for m_{ES} (left) and $\mathrm{NN}_{\mathrm{out}}$ (right) after implementing the additional requirements on the other fit variable to enhance the signal visibility. Points with error bars show the data, the solid (blue) line represent the total fit result, the dashed (green) curves show the total background contribution and the dotted (red) curve is the $q\bar{q}$ component of the background. The dash-dotted curve represents the signal contribution.*

In each case an additional cut on the other variable was applied to enhance the signal component in the following way:

◇ the m_{ES} projection plot has a cut at $\mathrm{NN}_{\mathrm{out}} > 0.9$;

◇ the $\mathrm{NN}_{\mathrm{out}}$ projection plot has a cut at $m_{\mathrm{ES}} > 5.274\,\mathrm{GeV}/c^2$.

Figure 6.4 shows the $_s\mathcal{P}lot$ distributions for m_{ES} and $\mathrm{NN}_{\mathrm{out}}$ with the PDF overlaid. Additionally the $_s\mathcal{P}lot$ distribution for ΔE is showing similar features to the expected distribution in the MC, *i.e.* broad tail and peaking at around zero.
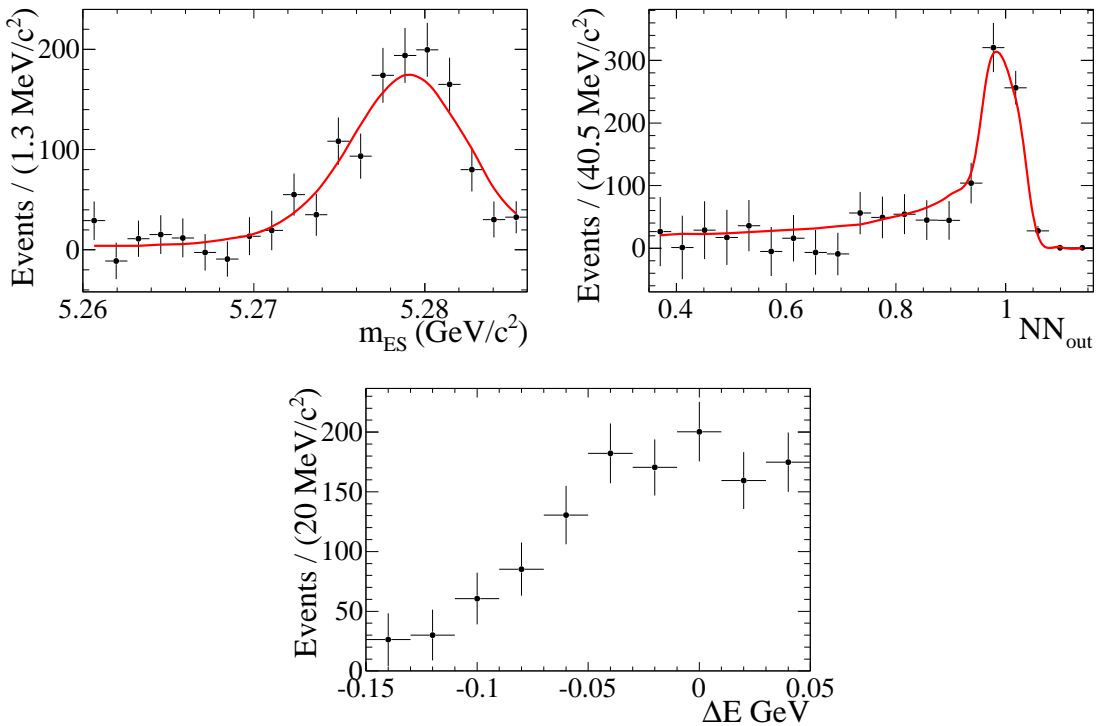
Figure 6.4: *Signal $_s\mathcal{P}lot$ distributions (black data points) with PDF (red line) overlaid, where appropriate, for m_{ES} (top left), $\mathrm{NN}_{\mathrm{out}}$ (top right), ΔE (bottom center).*

The statistical significance of the result is extrapolated from Figure 6.5, which shows the distribution of negative log likelihoods obtained by fixing the signal yield to various values in the range $-200 < N_{\mathrm{sig}} < 3000$. The statistical significance is given
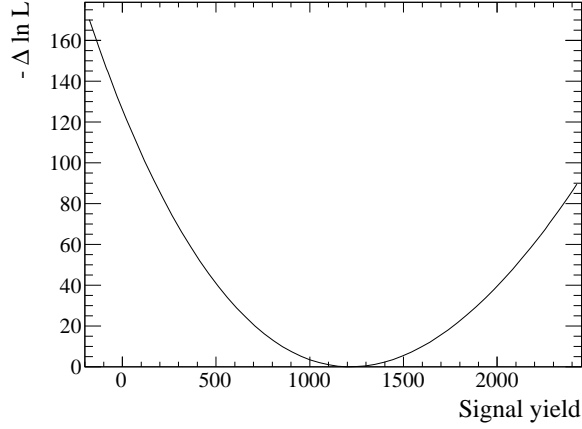
Figure 6.5: *Negative log likelihood distribution versus signal yield.*

from:

$$\sqrt{2(\ln \mathcal{L}(1220) - \ln \mathcal{L}(0))} = 15\sigma \tag{6.4}$$

The next step to obtaining the inclusive branching fraction is the extraction of the signal Dalitz plot from the $_s\mathcal{W}eights$.

6.3.2 The signal Dalitz plot

The $_s\mathcal{W}eights$ are used to obtain the signal Dalitz plot distribution. Using the event-by-event $_s\mathcal{W}eights$, it is possible to apply an efficiency correction to the Dalitz plot distribution using the signal efficiency Dalitz plot obtained from the MC and shown in Figure 4.8. The Dalitz plots before and after correction are shown in Figure 6.6. The efficiency corrected signal yield is calculated as the integral of the resultant Dalitz plot distribution and is found to be 7427 ± 518 giving an average efficiency of 16.4%.

6.3.3 Determination of the inclusive branching fraction

Before using the efficiency corrected signal yield, the following corrections have to be applied:
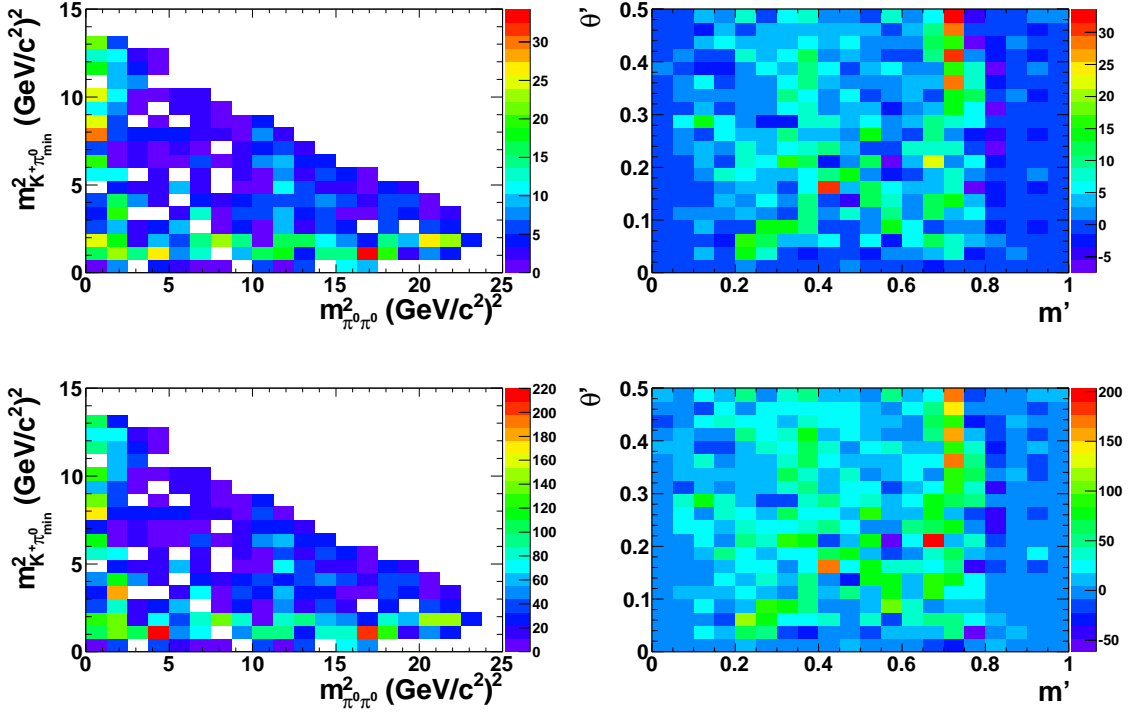
Figure 6.6: *Signal Dalitz plot distribution, obtained using $_s\mathcal{W}eights$, for conventional (left) and square (right) Dalitz plots for uncorrected (top) and corrected for efficiency using the signal MC efficiency from Figure 4.8 (bottom).*

◇ Correcting for fit biases: to estimate a correction for the biases new embedded toys are run using a signal MC cocktail based on the mode-by-mode f_{SCF} fractions given in Table 4.3 to reproduce the average f_{SCF} in data. The cocktail composition is given in Table 6.6. The result of the ensemble of 500 embedded toy experiments gives a signal yield of 1264 ± 4 events and a bias of 44 events. Using the average efficiency obtained in Section 6.3.2, the correction to the efficiency corrected yield is -268 events;

◇ Correcting for the K_S^0 veto: as shown in Table 4.3, the K_S^0 veto removes a fraction of the signal events that depends on the true Dalitz plot distribution and is estimated to be between 0 and 4%. The correction applied to the signal yield is of $(98 \pm 2)\%$;

Table 6.6: *Number of events for each signal MC used to reproduce the total yield and preserve the overall SCF fraction. This mixture also reproduces the broad features of the Dalitz plot.*

Decay mode	Number of events
NR	500
$B^+ \to K^*(892)^+\pi^0$	150
$B^+ \to K_0^{*+}(1430)\pi^0$	350
$B^+ \to f_0(980)K^+$	200
$B^+ \to \chi_{c0}K^+$	20

$\diamond$ Correcting for π^0 efficiency: as mentioned in Appendix B, a correction needs to be applied for the π^0 efficiency. This correction depends on the π^0 momentum distribution and is calculated using the signal MC cocktail described above. A weighted average correction for each π^0 is determined using the momentum distribution of each π^0 in the event and the fit to the double ratio, shown in Figure B.3, as follows:

$$\epsilon_{corr} = \frac{\sum_i^n f(p_i)N_i}{\sum_i^n N_i} \tag{6.5}$$

where the sum runs over all bins in the momentum distribution, $f(p_i)$ is the result of the double ratio fit function (given in Eq. B.8) in bin i and N_i is the number of events in that bin. The error on the correction is determined using the covariance matrix of the linear fit to the double ratio. The two corrections are multiplied together to give a combined correction for both neutral pions of $(95.8 \pm 1.1)\%$.

The final signal yield after efficiency and corrections is found to be 7305 ± 529. After dividing by the total number of $B\overline{B}$ events in the data sample, the branching fraction obtained is:

$$\mathcal{B}(B^+ \to K^+\pi^0\pi^0) = (16.2 \pm 1.2) \times 10^{-6} \tag{6.6}$$

where the error quoted is statistical only. Note that the daughter branching fraction of $\pi^0 \to \gamma\gamma$ is not required in the inclusive branching fraction calculation since the π^0 meson is not forced to decay this way in the MC.

6.4 The branching ratio of the resonant decays

The next step involves measuring the branching fractions for the resonant decays: $B^+ \to K^*(892)^+(\to K^+\pi^0)\pi^0$, $B^+ \to f_0(980)(\to \pi^0\pi^0)K^+$ and $B^+ \to \chi_{c0}(\to \pi^0\pi^0)K^+$. Figure 6.7 and Figure 6.8 show the $_s\mathcal{P}lot$ distributions projected onto the Dalitz plot axes to check for peaks in the mass combination $K\pi^0_{min}$ and $\pi^0\pi^0$. Excesses of events due to the decays $B^+ \to K^*(892)^+\pi^0$ and $B^+ \to f_0(980)K^+$ are clearly seen in Figure 6.7 (b) and Figure 6.8 (b). A slight peak is also observed in the mass range of the decay $B^+ \to \chi_{c0}K^+$ in Figure 6.8 (c).
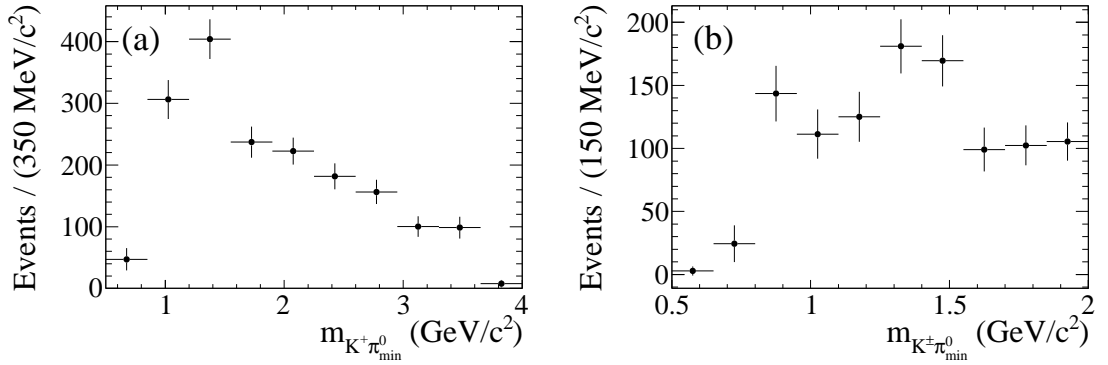
Figure 6.7: *Signal $_s\mathcal{P}lot$ distributions for $m_{K\pi^0_{min}}$ (a) over all mass range, (b) zoomed into mass range $0.5 < m_{K\pi^0_{min}} < 2.0$ GeV/c^2.*

The branching fractions are measured using the method described in Section 5.5 which needs to be first validated to establish if it returns unbiased results. To this effect embedded toy experiments are performed on three different "cocktails" made by combining different number of events from nonresonant, $K^*(892)^+$, $f_0(980)$ and χ_{c0} MC. The numbers of events were decided based on the proportions of expected events from each resonant sample. Expected events have been estimated using
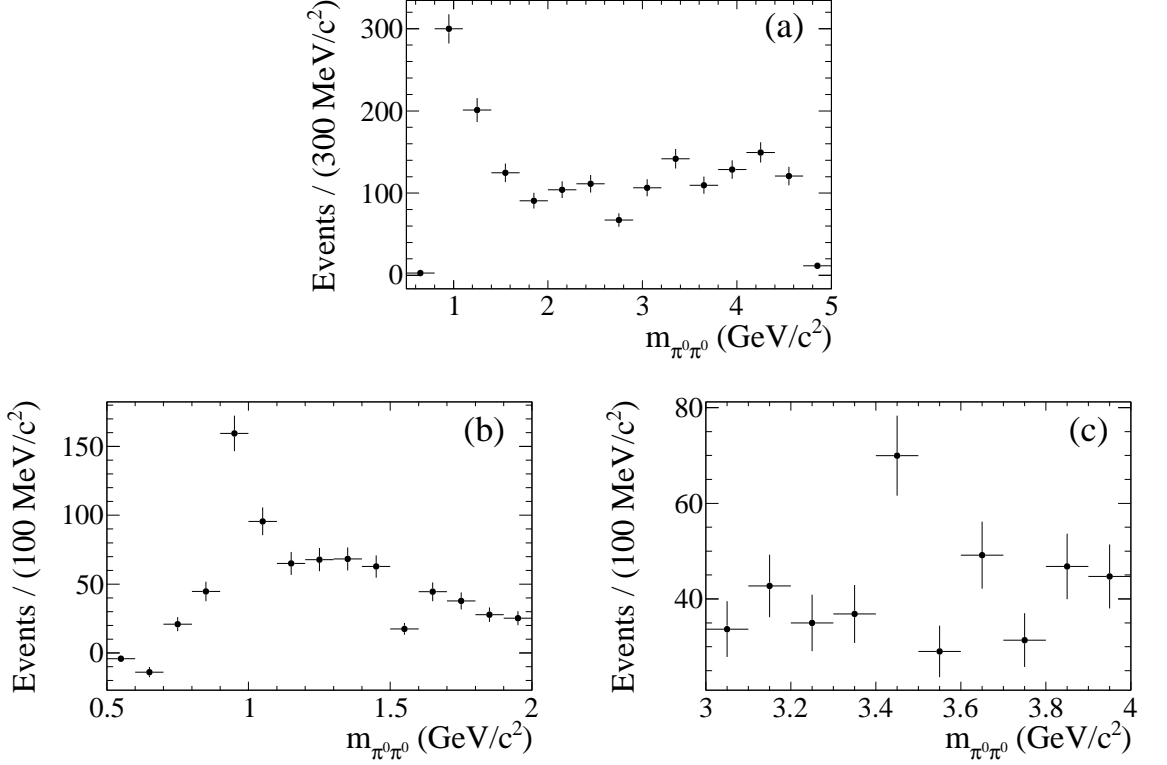
Figure 6.8: *Signal $_s\mathcal{P}$lot distributions for $m_{\pi^0\pi^0}$ (a) over all mass range, (b) zoomed into mass range $0.5 < m_{\pi^0\pi^0} < 2.0$ GeV/c², (c) zoomed into mass range $3.0 < m_{\pi^0\pi^0} < 4.0$ GeV/c².*

branching fractions calculated from PDG [38], HFAG [33] and isospin. In each cocktail shown in Table 6.7, one resonant mode was omitted on purpose to cross-check that the distribution of branching fractions returned by these resonances in those cocktail samples is zero within error. This section details the process of extracting the resonant branching ratios from validation to the measurement on onpeak data.

6.4.1 Validation of the method

For each cocktail, 500 experiments were run of which 339 for cocktail 1, 336 for cocktail 2 and 337 for cocktail 3 were successful. The large number of failed experiments is believed to be due to a large difference between the SCF fraction of the generated

Table 6.7: *Composition of the cocktail Monte Carlos. Events are drawn from large samples – the numbers quoted are the average numbers of events from each SP mode in each cocktail experiment. The corresponding branching fractions (in units of 10^{-6}) for the Q2B decays are also given.*

	Cocktail 1		Cocktail 2		Cocktail 3	
	# events	BF	# events	BF	# events	BF
$K^{*+}(892)$	158	2.28	0	0	158	2.28
$f_0(980)$	356	4.70	356	4.70	0	0
χ_{c0}	0	0	40	0.51	40	0.51
NR	665	15.5	665	15.5	665	15.5

data set and the one used in the first iteration of the maximum likelihood fit. The successful fits provide enough statistics to validate the method. An example, taken from one of the successful embedded toy experiments, for each MC cocktails of the $_s\mathcal{P}lots$ distributions for the invariant mass regions of the $K^*(892)^+\pi^0$, $f_0(980)K^+$ and $\chi_{c0}K^+$ is shown in Figure 6.9. Red (green) arrows indicate the signal (sidebands) region. The resonance peaks are clearly visible within the red arrow when the resonance is present and absent when the resonance is missing from that specific cocktail mixture.

All corrections used are the same as used in the inclusive branching fraction measurement (see Section 6.3.3) except no K_S^0 veto correction is needed for the $f_0(980)$ and χ_{c0} measurement and an additional correction is included to account for the efficiency of the tight cut in the resonance invariant mass in MC, listed in Table 6.8. For the subtraction method, this correction is the efficiency of the signal region requirement. For the fit method, this is the efficiency of a loose cut in the mass region that includes the sidebands.

The branching fraction distributions are then fitted with a Gaussian, as shown in Figure 6.10, to obtain the mean and rms of the distribution, which is used as the

Cocktail 1

Cocktail 2
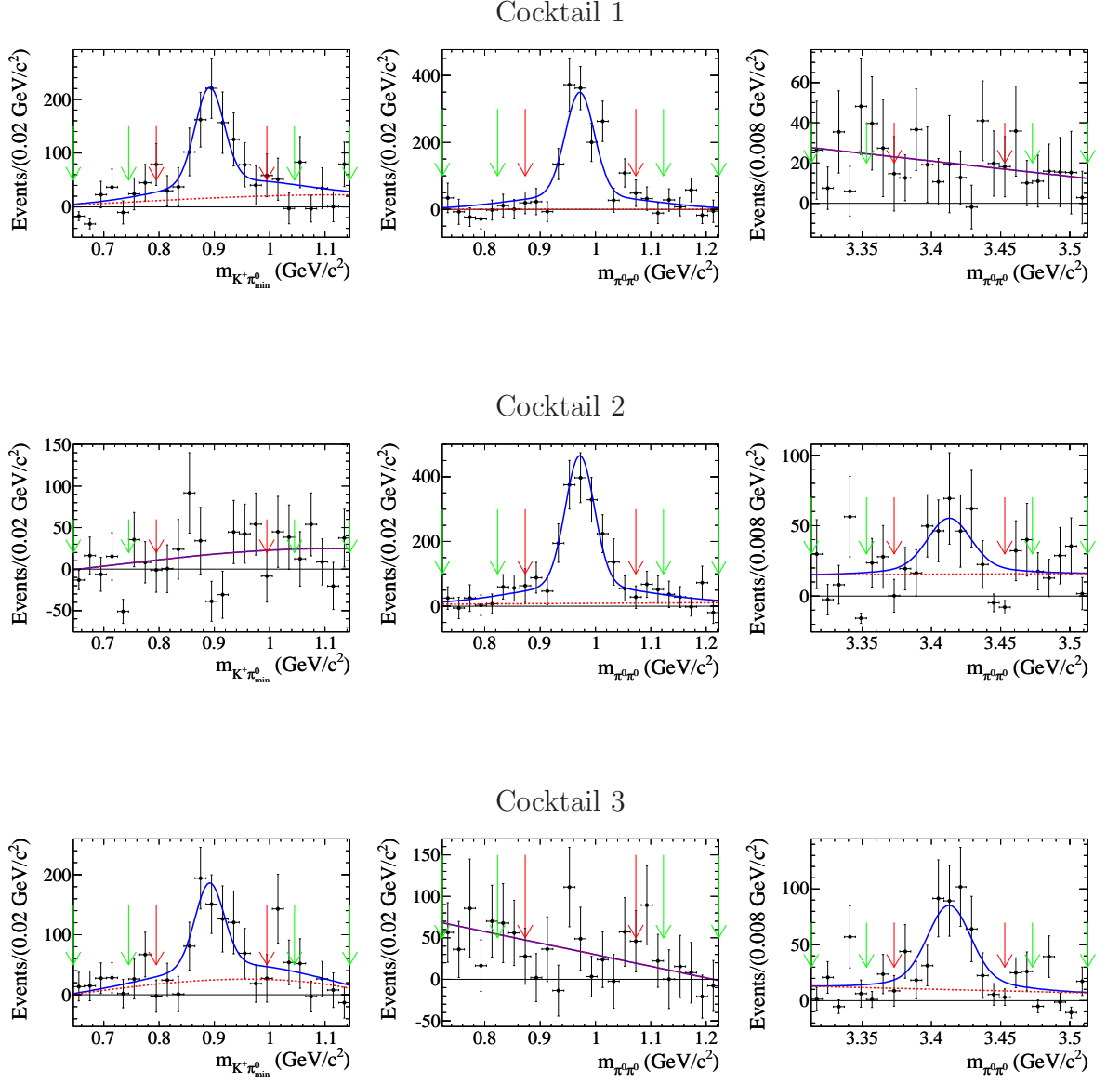
Cocktail 3

Figure 6.9: *Typical $_s\mathcal{P}$lot distributions obtained from one experiment in each cocktail mixtures for: (left to right) $m_{K^+\pi^0_{\min}}$ distribution in the $K^{*+}(892)$ region, $m_{\pi^0\pi^0}$ distribution in the $f_0(980)$ region; and $m_{\pi^0\pi^0}$ distribution in the χ_{c0} region. χ_{c0}, $K^*(892)^+$ and $f_0(980)$ are missing from one cocktail in this order. Red (green) arrows indicate the signal (sidebands) region.*

Table 6.8: *Efficiencies of mass cut selections of the signal region and (signal+sidebands) regions obtained from the MC samples of each resonance.*

Resonance	Signal region cut efficiency(%)	Mass region cut efficiency(%)
$K^*(892)^+\pi^0$	73.9	88.0
$f_0(980)K^+$	79.9	92.5
$\chi_{c0}K^+$	90.4	99.0

error on the measured branching fraction. The mean branching fraction and its error are given in Table 6.9 and show that results obtained from the fit and subtraction method are comparable with minimal biases.

Table 6.9: *Validation of the method to determine quasi-two-body branching fractions. The values given are the measured branching fractions (in units of 10^{-6}) for each cocktail and are taken from Gaussian fits to the BF distributions in Figure 6.10. The uncertainty is the width of the Gaussian, which we use as an estimate of the statistical uncertainty of the result of one experiment.*

	$K^{*+}(892)$	$f_0(980)$	χ_{c0}
	Expected Branching Fractions (when non-zero) ($\times 10^{-6}$)		
	2.28	4.70	0.51
	With sideband subtraction method		
Cocktail 1	2.03 ± 0.47	4.45 ± 0.58	0.005 ± 0.154
Cocktail 2	-0.009 ± 0.383	4.38 ± 0.56	0.52 ± 0.22
Cocktail 3	2.01 ± 0.47	0.13 ± 0.43	0.53 ± 0.21
	With fit method		
Cocktail 1	1.99 ± 0.43	4.80 ± 0.56	-0.002 ± 0.125
Cocktail 2	-0.055 ± 0.303	4.70 ± 0.52	0.55 ± 0.20
Cocktail 3	2.02 ± 0.47	0.03 ± 0.35	0.53 ± 0.19
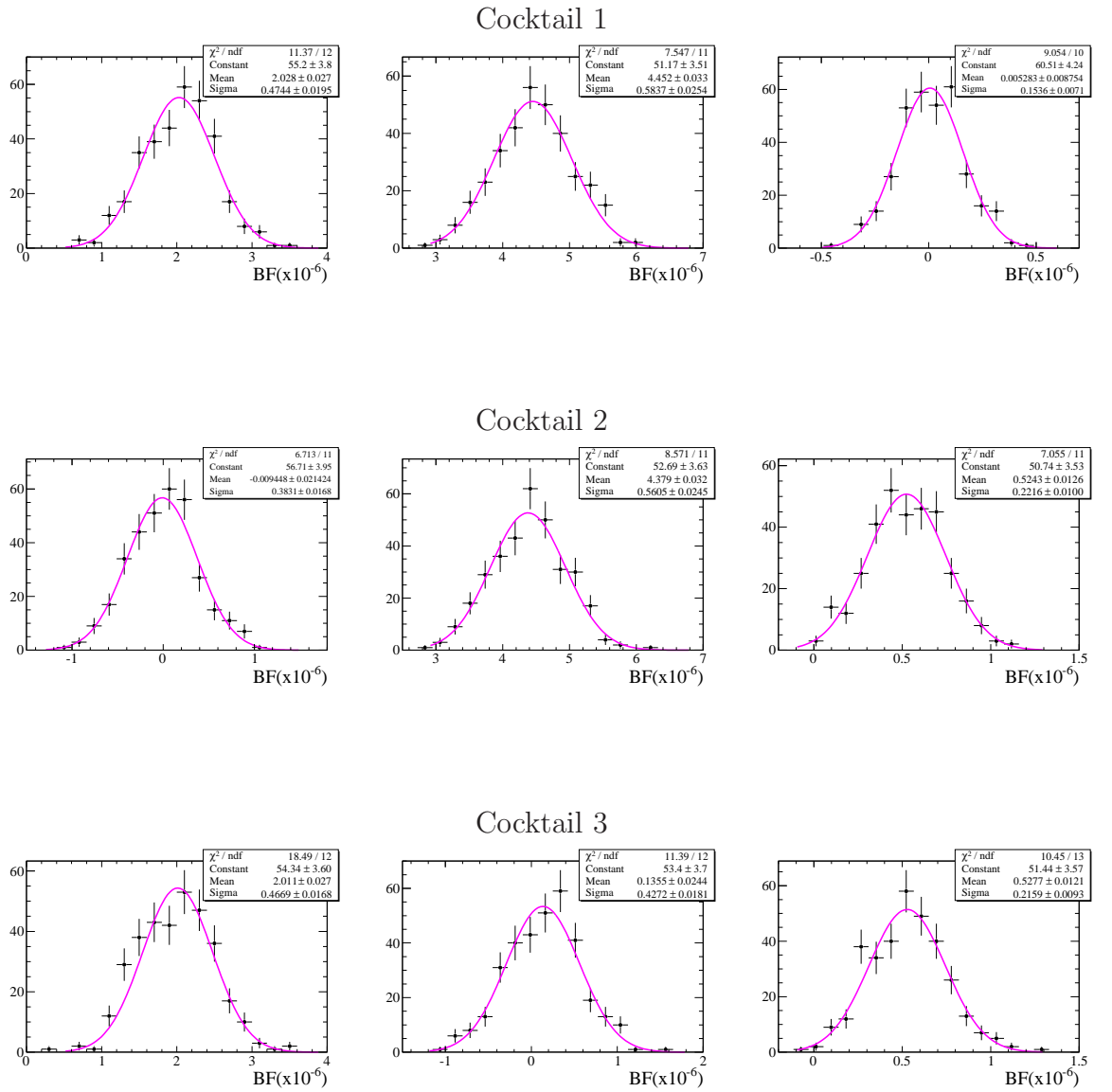
Figure 6.10: *Distribution of branching fractions from each cocktail for: (from left to right) $K^*(892)^+\pi^0$, $f_0(980)K^+$ and $\chi_{c0}K^+$. The distributions were all fitted with Gaussians.*

6.4.2   Results from onpeak data

The subtraction and fit methods were applied to the $_s\mathcal{P}lot$ distributions obtained from the fit to data in Section 6.3. Results obtained from the distributions in Figure 6.11 are given in Table 6.10. The results obtained using both methods are shown to be again comparable.

Table 6.10: *List of results for fit and background subtraction method.*

Resonant decay	Method	Corrected yield	BF($\times 10^{-6}$)
$B^+ \to K^*(892)^+\pi^0$	Sub	1078	2.29 ± 0.48
	Fit	872	1.85 ± 0.41
$B^+ \to f_0(980)K^+$	Sub	1186	2.52 ± 0.53
	Fit	1166	2.48 ± 0.46
$B^+ \to \chi_{c0}K^+$	Sub	245	0.52 ± 0.21
	Fit	262	0.56 ± 0.19

The statistical errors are obtained from embedded toy studys of a cocktail mixture reflecting the yields found in data. Nonresonant MC is added to complete the inclusive yield. The embedded study is also used to determine the bias correction on the measurements. The composition of this cocktail and measured bias are given in Table 6.11. A π^0 efficiency correction is also determined from the same MC

Table 6.11: *Mean branching fractions and bias obtained for each resonance from a cocktail reflecting the number of events observed in data. These results were obtained using the subtraction method.*

Resonance	Number of events	BF($\times 10^{-6}$)	Bias ($\times 10^{-6}$)
Nonresonant	829	-	-
$K^*(892)^+\pi^0$	159	$1.95 \pm 0.03(\pm 0.48)$	0.34
$f_0(980)K^+$	191	$2.43 \pm 0.03(\pm 0.53)$	0.09
$\chi_{c0}K^+$	41	$0.55 \pm 0.01(\pm 0.21)$	0.03
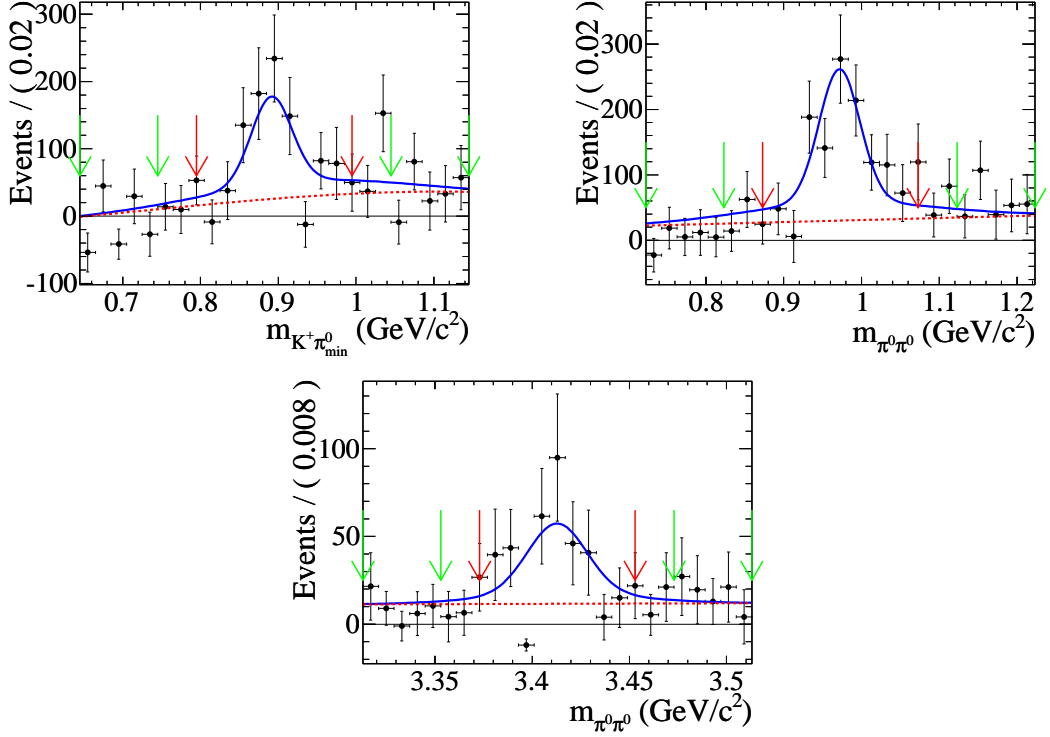
Figure 6.11: *Mass region distributions for $K^*(892)^+$ (top left), $f_0(980)$ (top right) and χ_{c0} (bottom) from data with fit result overlaid. The black data point show the $_s\mathcal{P}lot$ data, the blue continuous line is the overall fit and the red dashed the nonresonant contribution. The red and green arrows indicate the signal and sideband regions used in the subtraction method.*

cocktail in the same way as for the inclusive mode with the difference that only the momentum of each π^0 in the signal region of the resonance is considered. The π^0 efficiency corrections to apply to the branching fractions are as follows: $(96.4 \pm 1.2)\%$ for $K^*(892)^+\pi^0$, $(94.0 \pm 0.9)\%$ for $f_0(980)K^+$ and $(96.6 \pm 1.0)\%$ for $\chi_{c0}K^+$.

The final results of the product branching fractions quoting statistical uncertainties only and correcting for the fit bias and π^0 efficiencies are:

$$\mathcal{B}(B^+ \to K^{*+}(\to K^+\pi^0)\pi^0) = (2.72 \pm 0.50) \times 10^{-6} \tag{6.7}$$

$$\mathcal{B}(B^+ \to f_0(980)(\to \pi^0\pi^0)K^+) = (2.77 \pm 0.56) \times 10^{-6} \tag{6.8}$$

$$\mathcal{B}(B^+ \to \chi_{c0}(\to \pi^0\pi^0)K^+) = (0.51 \pm 0.22) \times 10^{-6} \tag{6.9}$$

7

Results: *CP* Asymmetries

In the previous chapter, the maximum likelihood was set up and used to obtain the branching fraction of the inclusive and quasi-two body modes after the model was properly validated by toy experiments. To measure the asymmetry, charge information needs to be included in the fit model, which needs to be tested again on toy MC experiments.

7.1 Toy tests for A_{CP} of inclusive mode

In these toy tests, data is generated with different values for the signal asymmetry. In the previous chapter, the purpose of running the pure and embedded toys was to

Table 7.1: *List of values for the A_{CP} used in toy MC generation and their status in the fit.*

	Generated A_{CP}	Fixed or Floating
Signal (pure toys)	-40%, 0%, 40%	Floating
Signal (embedded toys)	-30%, -15%, 0%, 15%, 30%	Floating
$q\bar{q}$	0%	Floating
All $B\bar{B}$ categories	0%	Fixed

measure the accuracy with which the model determined the inclusive yields. In the same way, these toy experiments are used to determine how accurately the model can determine the signal asymmetry. Each toy generates 500 experiments using the number of signal events obtained in the inclusive branching fraction results (1220 events). The number of events in continuum and $B\bar{B}$ backgrounds are the same as in Table 5.9. The parameters of the PDFs used are the same as those used in Section 6.1 with the exception that PDFs are split into positive and negative charged B mesons. Table 7.1 lists the values of signal and background asymmetries generated in the toy studies and whether or not these are allowed to float in the fit.

7.1.1 Pure toy studies

For each experiment, signal and background datasets are again sampled using the PDFs presented in Section 5.2. The number of events generated is varied around a mean value according to the Poisson distribution. Figure 7.1 shows the pull distributions for the signal CP asymmetries and the pull error. The distributions are all fitted with a Gaussian. The biases found in the asymmetry pull are mostly negligible.
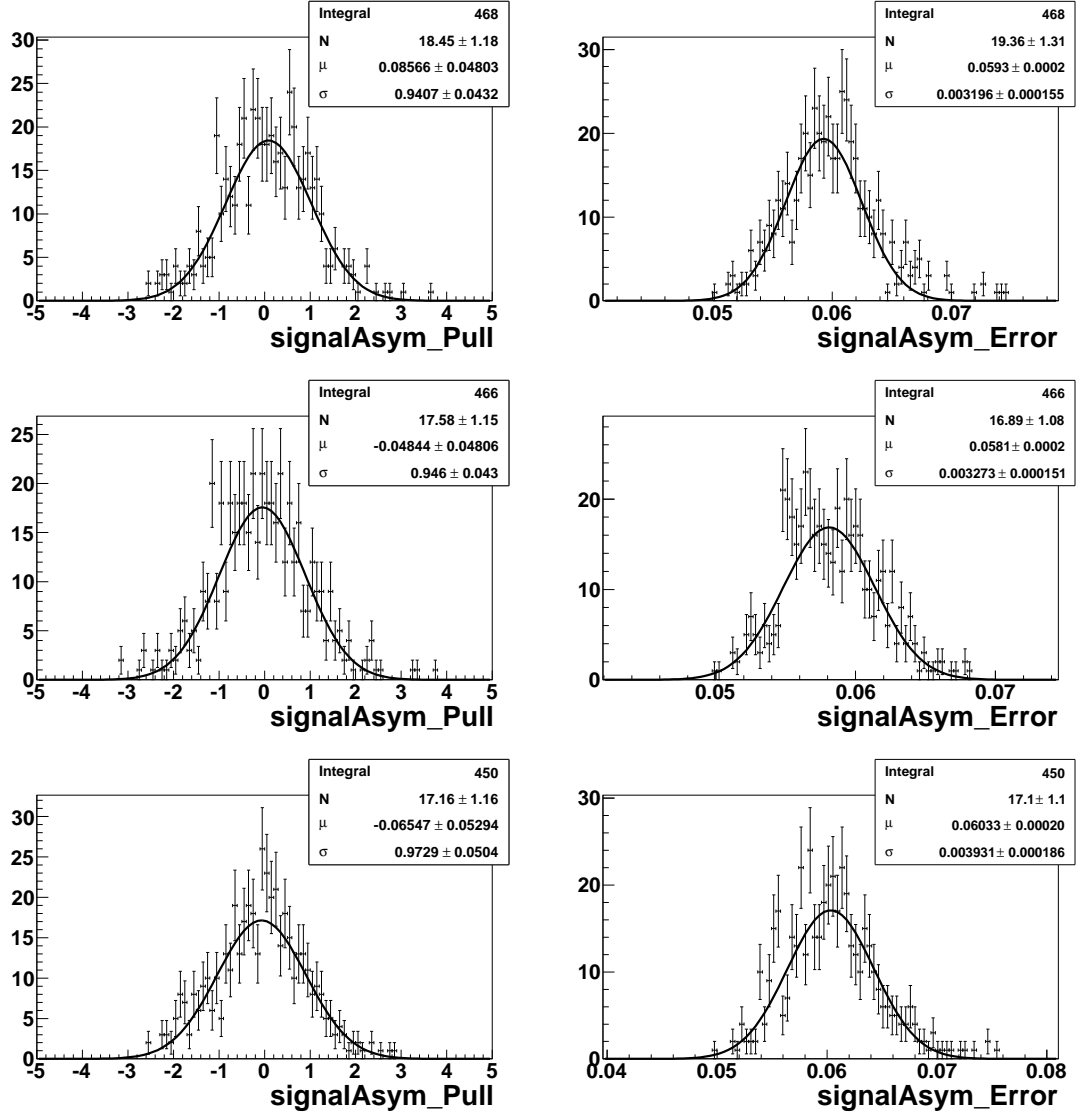
Figure 7.1: *Pull distributions for the signal asymmetry with error distributions on the right obtained from generated samples with asymmetries −40% (top row), 0% (middle row) and 40% (bottom row).*

7.1.2 Embedded toy study

For this embedded toy study, the signal events are sampled from the nonresonant signal MC with different signal asymmetries of ±30, ±15 and 0. Both the asymmetry and associated residuals are fitted with a gaussian. The signal asymmetry is statistically compatible with there being no bias as shown in Table 7.2.

Table 7.2: *List of signal asymmetries and biases on the asymmetry obtained from fits.*

Generated asymmetry	Signal asymmetry	Asymmetry bias
−30%	$(-29.9 \pm 0.3)\%$	0.1%
−15%	$(-15.0 \pm 0.2)\%$	0.0%
0%	$(-0.6 \pm 0.2)\%$	0.6%
15%	$(14.1 \pm 0.2)\%$	0.9%
30%	$(28.8 \pm 0.3)\%$	1.2%

7.2 Inclusive CP asymmetry of $B^+ \to K^+\pi^0\pi^0$

The A_{CP} fit model described above is applied to the onpeak data sample after selections. The value of f_{SCF} is fixed to the 9.7% value obtained during the unblinding of the branching fraction result in Section 6.3. The CP asymmetries of the $B\bar{B}$ categories are fixed to 0 whereas signal and continuum asymmetries are allowed to float. The fit returns the projection distributions shown in Figure 7.2 and a resultant signal asymmetry, with statistical uncertainty, of:

$$A_{CP}\left(B^+ \to K^+\pi^0\pi^0\right) = (-6 \pm 6)\,\% \tag{7.1}$$

The result is consistent with no CP violation.

7.3 CP asymmetry in the resonant decays

The same method used to extract the quasi-two-body branching fractions is applied to the decay of B^- and B^+ separately to obtain the CP asymmetries of $B^+ \to K^*(892)^+\pi^0$, $B^+ \to f_0(980)K^+$ and $B^+ \to \chi_{c0}K^+$. The statistical uncertainties in the yields are also determined using embedded toy MC pseudo-experiments. Distributions for separate charges are shown in Figure 7.3 and results for yields in
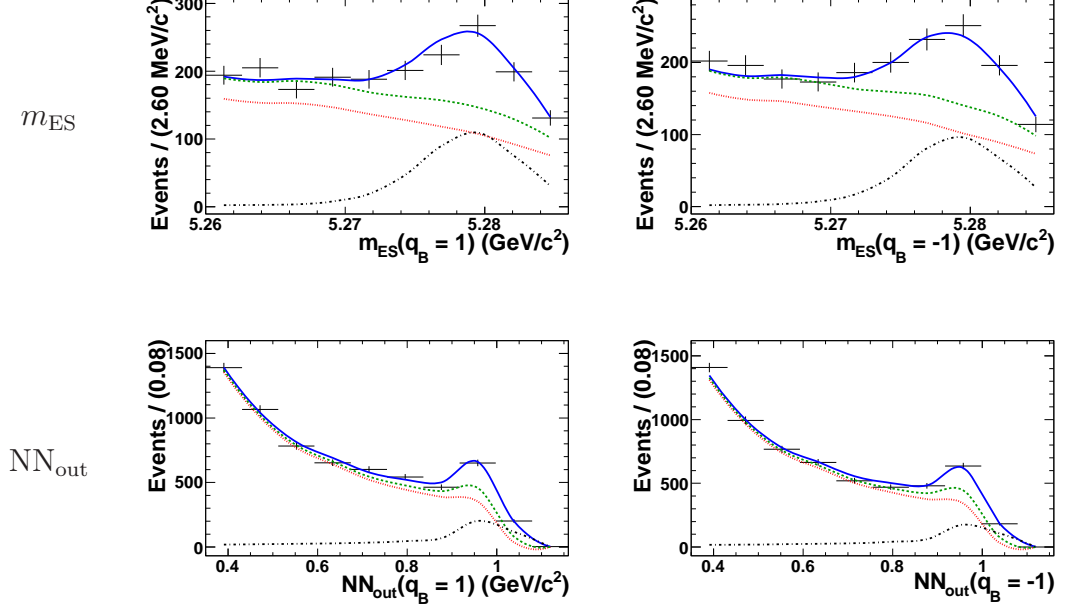
Figure 7.2: *Projection plots on m_{ES} and NN_{out} from the fit to data for the positive charged decay, $q_B = 1$ (left plot) and negative charged decay, $q_B = -1$ (right plot). Points with error bars show the data, the solid (blue) line represent the total fit result, the dashed (green) curves show the total background contribution and the dotted (red) curve is the $q\bar{q}$ component of the background. The dash-dotted curve represents the signal contribution.*

Table 7.3. This gives the following CP asymmetries with statistical uncertainties:

$$A_{CP}\left(B^+ \to K^*(892)^+\pi^0\right) \;=\; (-4 \pm 26)\,\% \tag{7.2}$$

$$A_{CP}\left(B^+ \to f_0(980)K^+\right) \;=\; (17 \pm 18)\,\% \tag{7.3}$$

$$A_{CP}\left(B^+ \to \chi_{c0}K^+\right) \;=\; (-89 \pm 37)\,\% \tag{7.4}$$
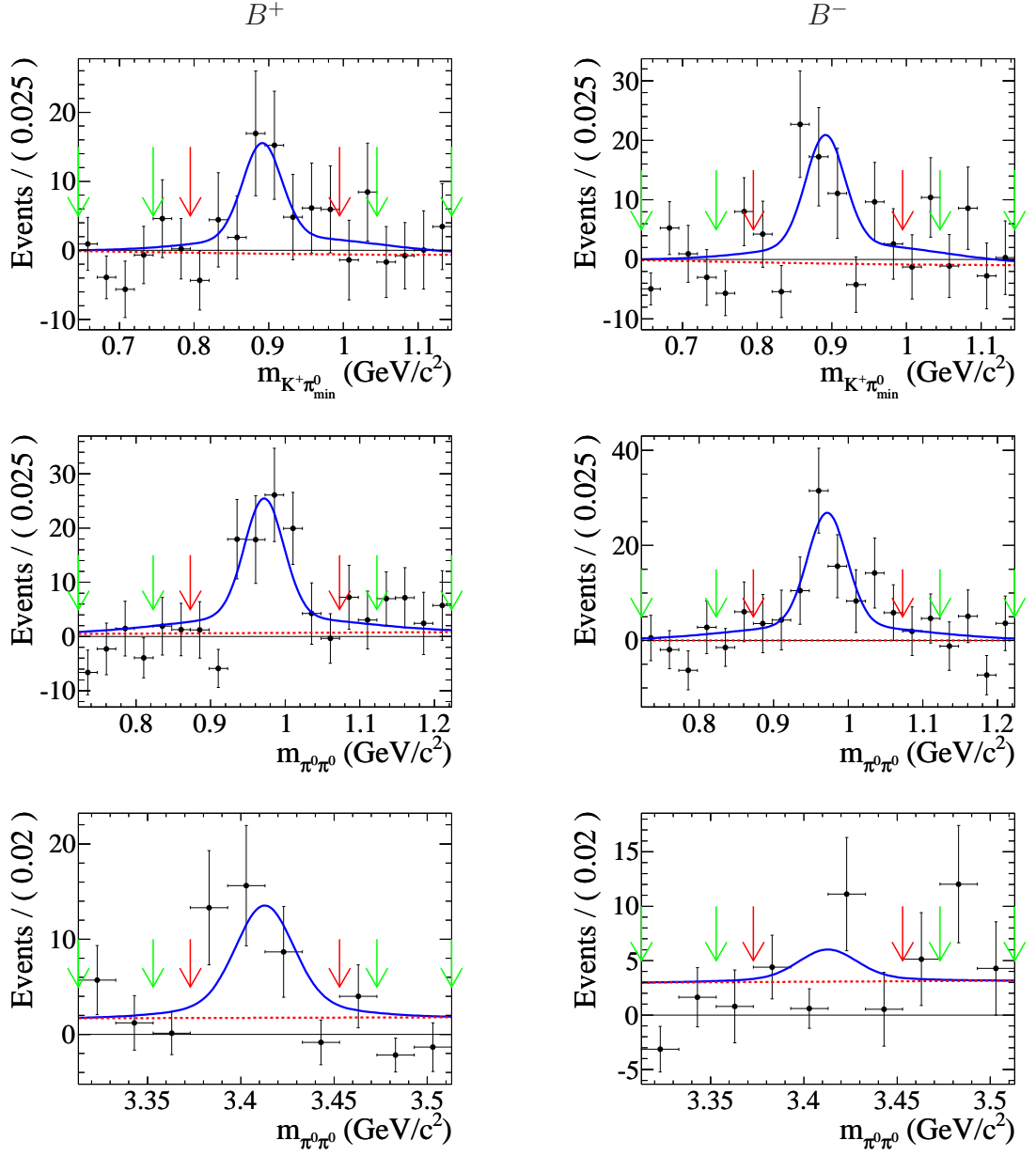
Figure 7.3: *Result of fit in the mass region of $K^*(892)^+$ (top), $f_0(980)$ (middle) and χ_{c0} (bottom) for B^+ decay (left) and B^- decay (right). The black points show the $_s\mathcal{P}$lot data, the blue continuous line is the overall fit and the red dashed line the NR contribution. Red (green) arrows indicate the signal (sidebands) region.*

Table 7.3: B^+ and B^- yields for each resonance obtained using the background subtraction method.

Resonant decay	B^+ yield	B^- yield
$B^+ \to K^*(892)^+\pi^0$	59	55
$B^+ \to f_0(980)K^+$	70	98
$B^+ \to \chi_{c0}K^+$	33	2

8

Systematic Errors Evaluation

So far, errors quoted in the results have only included statistical errors obtained from the fit covariance matrix or from pseudo-MC experiments. The errors associated to the analysis method used still need to be accounted for. These are referred to as systematic uncertainties. The two main contributions to the systematic errors in these measurements relate to the fit model and detector efficiencies.

8.1   Systematic uncertainties associated to the model

Several sources in the fit model are believed to contribute to the systematic uncertainty. These include all assumptions made in the likelihood fit model, such as fixed

parameters. These uncertainties are essential to quatify the total significance of the signal result obtained in Section 6.3.3. All the following systematic uncertainties were estimated by comparing signal yields obtained in various different fits to data, where some parameters were varied according to their errors, to the nominal signal yield obtained in the original fit.

8.1.1 Uncertainties in signal PDF shapes

The signal PDF parameters were fixed in the fit and therefore contribute to part of the systematic uncertainty. This systematic is divided into four components. Since the signal component in the fit was split to account for the SCF events, two uncertainties have to be calculated for the TM and SCF PDFs separately. Then each signal component adds a contribution to the systematic arising from the m_{ES} and $\mathrm{NN}_{\mathrm{out}}$ PDFs.

8.1.1.1 TM PDF systematics

To estimate a systematic uncertainty to associate to the m_{ES} model, the parameters of the m_{ES} PDF were varied on the basis of the errors on the shifts and scale factors determined from the $B^+ \rightarrow \overline{D}^0 \rho^+ \rightarrow (K^+ \pi^- \pi^0)(\pi^+ \pi^0)$ control sample (see Section 5.2.1.1) and listed in Table 5.4. The variations take into account the correlations between parameters also obtained from the fit to the m_{ES} distribution of the MC control sample and shown in Figure 8.1. The variation in signal yield distribution was observed to a have a width of 10 events which divided by the number of events obtained in the fit gives a systematic uncertainty of 0.8%. Since the signal yield is much larger than was originally expected, a fit was also performed where the mean and widths of the TM m_{ES} PDF were allowed to float. This cross-check gives a signal yield of 1211 ± 106, which is within the above systematic uncertainty.
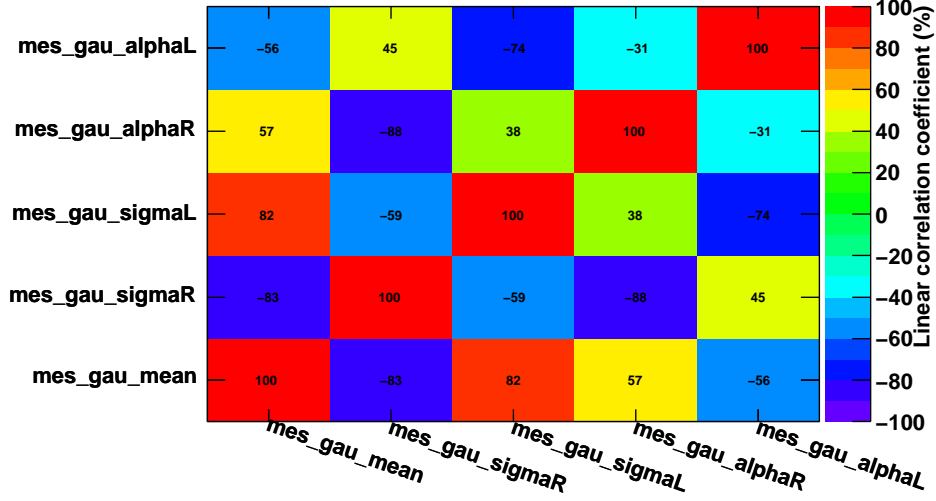
Figure 8.1: *Linear correlations coefficients obtained from the fit to the m_{ES} distribution of the $B^+ \to \overline{D}^0 \rho^+ \to (K^+\pi^-\pi^0)\,(\pi^+\pi^0)$ MC control sample.*

To estimate the systematic contribution from the $\mathrm{NN_{out}}$ PDF, the contents of the bins of the PDF were fluctuated in accordance to data/MC differences seen in the control sample and shown in Figure 8.2. The variation for each bin is sampled from a Gaussian centred on unity with width equal to the quadrature sum of the difference between the bin central value and unity and the bin uncertainty. The distribution of the fitted signal yield is shifted by approximately 30 events and has a width of 40 events. As mentioned in Section 5.2.1.1, there appears to some cross-feed between the continuum and B backgrounds in the fit to the control sample hence Figure 8.2 might not give a fully reliable correction. As such the systematic uncertainty is obtained by combining the shift and width in quadrature and gives a systematic contribution of 50 events. Dividing by the nominal yield we obtain an uncertainty of 4.1%.
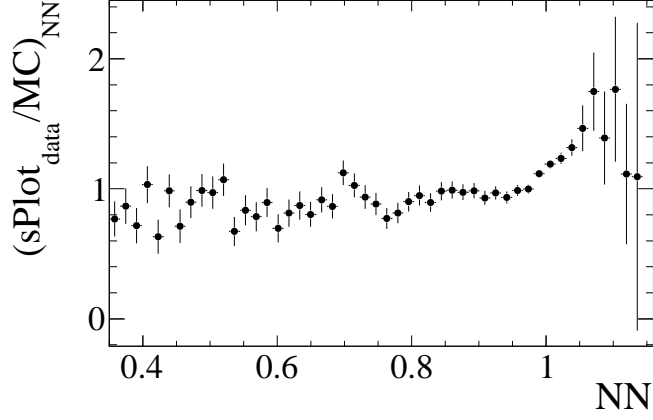
Figure 8.2: *Ratio of the* $\mathrm{NN_{out}}$ $_s\mathcal{P}lot$ *distribution and the MC based PDF for the control sample* $B^+ \to \overline{D}^0 \rho^+ \to (K^+\pi^-\pi^0)\,(\pi^+\pi^0)$.

8.1.1.2 SCF PDF systematics

Figure 8.3 shows the variations in the shape of the m_{ES} and $\mathrm{NN_{out}}$ SCF distribution for nonresonant and resonant decay modes as seen in MC. To obtain a systematic uncertainty for the m_{ES} SCF PDF, the parameters of the PDFs were extracted from each resonant MC mode and fitted to the data sample in turn. The results of the signal yield in those fits were then compared to the nominal value which used parameters extracted from nonresonant MC. The largest difference was used for the systematic and gives an uncertainty of 21 events which divided by the nominal yield contributes an extra 1.7%.

Similarly the $\mathrm{NN_{out}}$ systematic was evaluated by fitting the data using the $\mathrm{NN_{out}}$ distributions from each resonant MC modes and comparing the signal yield obtained in these fits to the nominal yield. The largest difference was used and gives an uncertainty of 9 events corresponding to a contribution of 0.7%.
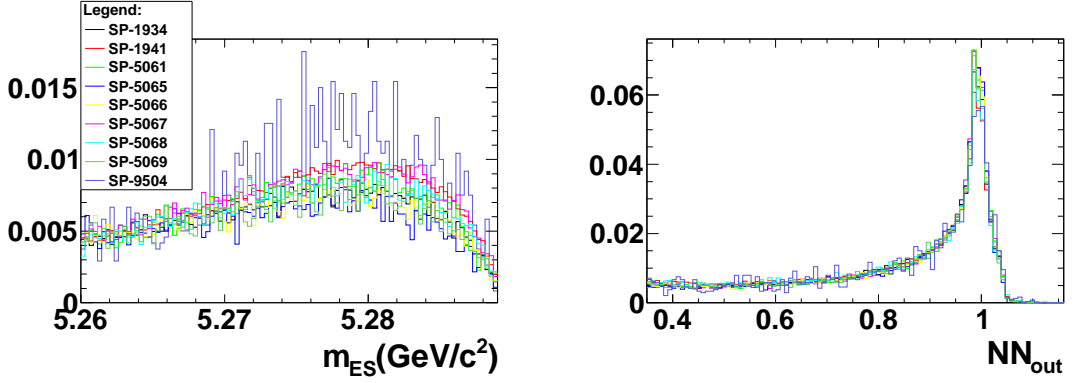
Figure 8.3: *Distributions of SCF events in m_{ES} and NN_{out} for nonresonant and resonant decay modes. Refer to Table 3.1 to match the MC number to a specific decay.*

8.1.2   Uncertainties in $B\overline{B}$ background PDFs

The NN_{out} systematic was assigned two contributions. The first uncertainty was estimated by fluctuating the bins of all the histograms used to describe the $B\overline{B}$ background PDFs. The width of the distribution of the fit results in this case is found to be about 9.5 events corresponding to an uncertainty of 0.8%. This method however is limited by the MC statistics and therefore does not give the full uncertainty contribution. To account for possible data/MC differences in the NN_{out} PDFs, the shape of the histogram is varied in the same way as for the TM signal (see Section 8.1.1.1). Because in fact there is a correlation between the shapes of the NN_{out} PDFs from TM signal and $B\overline{B}$ backgrounds, the TM signal and $B\overline{B}$ backgrounds are varied together and the systematic combined. The NN_{out} systematic increases from the 50 events found when only varying the TM signal to 60 events corresponding to a total uncertainty of 4.9%.

For the m_{ES} PDFs, the histograms were convoluted with a Gaussian reflecting the data/MC differences found in the control sample and refit. The systematic is taken to be the difference between the results obtained during this new fit and the nominal value and is found to be 20 events, corresponding to an uncertainty of 1.6%.

8.1.3   Correcting for fixed parameters and biases

The likelihood model contains a substantial number of parameters that were fixed. In Section 5.4, the iterative procedure was tested on MC and shown to be able to determine the SCF fraction to within $\pm 3.5\%$. The fit was repeated by varying f_{SCF} by this amount and the difference from the nominal fit result established. This returned a difference of 31 events corresponding to a systematic error contribution of 2.5%. The other fixed parameters to consider are the $B\overline{B}$ background yields. To estimate this contribution, the yield of each $B\overline{B}$ category was varied in turn using the uncertainties on the expected yields given in Table 5.9. All the differences with the nominal yield, given in Table 8.1, were calculated and added in quadrature to give the uncertainty in terms of the number of events. The systematic uncertainty assigned to the fixed $B\overline{B}$ yields is 1.4%.

Table 8.1: *Results of the fits made by varying the $B\overline{B}$ background yields within their uncertainties. This table lists the differences between signal yields obtained in these fits and the nominal value.*

$B\overline{B}$ category	$\left\| n_{\mathrm{sig}}^{\sigma_{B\bar{B}_i}} - n_{\mathrm{sig}}^{\mathrm{nom}} \right\|$
$B\overline{B}$ 1	3 events
$B\overline{B}$ 2	10 events
$B\overline{B}$ 3	6 events
$B\overline{B}$ 4	13 events
Combination	18 events

Finally, another systematic that needs to be considered is due to the biases that were observed in the embedded toy study using a MC cocktail where the signal MC content matched the SCF fraction obtained in data (see Section 6.3.3). Taking into account that systematic effects in determining the fit bias are approximately 50% of the bias, a systematic error is assigned corresponding to half of the bias added in quadrature with the uncertainty with which the bias is known. This amounts to an

uncertainty of 22 events corresponding to a systematic of 1.8%.

8.1.4 Summary of fit model systematics

The subtotal for the systematic uncertainties due to the fit model and listed in Table 8.2 is 6.5%. This systematic represents the uncertainty on the yield and is used in Section 9.1 to calculate the significance of the inclusive branching fraction result.

Table 8.2: *Summary of systematic uncertainties associated to the fit model for the inclusive branching fraction measurement of* $B^+ \to K^+ \pi^0 \pi^0$.

Source	Uncertainty (%)
TM signal m_{ES} PDF	0.8
TM signal and $B\bar{B}$ background $\mathrm{NN_{out}}$ PDF	4.9
SCF signal m_{ES} PDF	1.7
SCF signal $\mathrm{NN_{out}}$ PDF	0.7
SCF fraction	2.5
$B\bar{B}$ background PDFs (MC statistics)	0.8
$B\bar{B}$ background m_{ES} PDFs	1.6
$B\bar{B}$ background yields	1.4
Fit bias	1.8
Subtotal from fit model	6.5

8.2 Systematic uncertainties due to efficiency

There are several systematic uncertainties associated with the selection efficiency and particle identification. Some of these uncertainties are standard *BABAR* estimates or have been already calculated in other analyses such as:

⬦ 0.4% tracking efficiency associated to the single kaon track: The recommended strategy is to combine in quadrature the correction factor found in Ref. [80] with the track uncertainty [81].

⬦ 1.0% particle identification for the kaon selector used (see Section 4.1.2): This uncertainty has been estimated in a previous analysis done on three charged kaons [82].

⬦ 0.6% the number of $B\overline{B}$ pairs in the onpeak sample: This is calculated using the B counting technique, described in Section 3.1.3 [54].

8.2.1 The π^0 efficiency systematic

The π^0 efficiency used to be by far the largest standard systematic uncertainty (6% for 2 π^0 mesons in final state). This uncertainty was based on a π^0 efficiency using τ decays and Run 1 to 3 data [83]. As part of this thesis, a similar π^0 efficiency study was performed using the full dataset to establish a more up-to-date systematic uncertainty. The outcome of the study deemed necessary to apply a momentum dependent correction. For further details on this π^0 efficiency study see Appendix B. By approximating the systematic effects involved in the determination of the π^0 efficiency corrections to be 50% of the corrections, the systematic uncertainty is calculated in the same way as the systematic uncertainty for the bias, *i.e.* it is taken as half the correction, calculated in Section 6.3.3, added in quadrature to the error with which it is known. This gives a total systematic uncertainty for the π^0 efficiency of 2.4%.

8.2.2 Efficiencies due to selection criterias

Additional uncertainties have to be assigned to the nonstandard selections and corrections. Two selections have been applied to the discriminating variables, the first was a tight cut on the ΔE distribution between $-0.15 < \Delta E < 0.05$ GeV and the

second, $NN_{out} > 0.35$. The uncertainties of these two selections are obtained by comparing the efficiencies of these selections between data and MC in the control sample. The efficiency correction for the ΔE selection is found by the ratio of signal yields obtained by fitting data and MC with a tight ΔE cut and a looser signal region of $-0.3 < \Delta E < 0.3$ GeV. The efficiency correction is then added in quadrature to the error in the ratio and the result is taken as the systematic. For NN_{out}, the systematic uncertainty is an estimate of the possible error in the 90% efficiency of the selection requirement (see Table 4.2). These two systematic uncertainties due to selection efficiencies amount to:

- ⋄ ΔE cut: 4.0%;

- ⋄ NN_{out} cut: 3.0%.

A further uncertainty in this category was mentioned previously in Section 6.3.3 and is due to the efficiency of the K_S^0 veto. This was found to contribute approximately 2% to the total systematic.

8.2.3 Summary of efficiency systematics

Combining in quadrature the systematic uncertainties due to efficiencies in Table 8.3 with the subtotal of 6.5% obtained from the fit model systematics in Table 8.2 gives a total systematic uncertainty for the inclusive branching fraction measurement of 8.9% of the final result.

8.3 Systematic uncertainties for the asymmetry

The majority of the sources of systematic uncertainties found in branching fraction measurements cancel in the calculation of the CP asymmetry. The sources of systematics that remain are effects that could potentially induce a bias between B^+ and B^- decays.

Table 8.3: *Summary of systematic uncertainties due to PID and selection efficiencies in the inclusive branching fraction measurement of* $B^+ \to K^+\pi^0\pi^0$.

Source	Uncertainty (%)
Tracking efficiency	0.4
Particle identification	1.0
Neutral pion efficiency	2.4
ΔE cut efficiency	4.0
$\mathrm{NN_{out}}$ cut efficiency	3.0
K_S^0 veto	2.0
$N_{B\bar{B}}$	0.6
Total	8.9

8.3.1 Detector and selection induced asymmetry

The cross-section for the interaction of kaons with protons and neutrons can differ with charge. At low momenta this effect can introduce a bias to the observed charge asymmetry due to interactions within the detector. This effect was estimated in previous studies, see Refs. [3, 84]. These analyses found that the error due to imperfect understanding of the DIRC and DCH particle identification performance was too small compared to the current precision at which CP asymmetries are measured. The systematic uncertainty assigned for this bias based on these studies is 0.5%.

In addition to detector induced asymmetries, selection requirements can also accidentally introduce a difference in the efficiency between charged decays. To estimate an uncertainty, the CP asymmetry in the control sample is measured where the asymmetry is know to be negligible. The A_{CP} fit to the control sample returned projection plots for B^+ and B^- shown in Figure 8.4 and a signal asymmetry of $(3 \pm 1)\%$. The systematic assigned for selection asymmetries is therefore 3%.
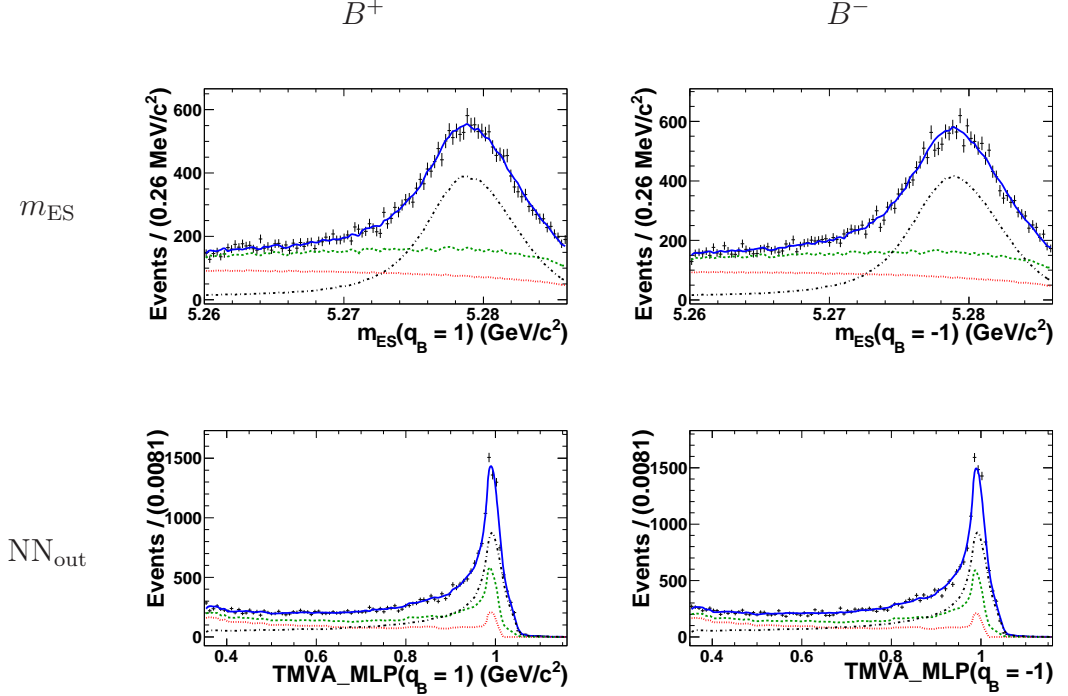
Figure 8.4: *Projection plots on* $m_{\rm ES}$ *and* $\rm NN_{out}$ *from the fit to the control sample for the positive charged decay (left plot) and negative charged decay (right plot). Points with error bars show the data, the solid (blue) line represent the total fit result, the dashed (green) curves show the total background contribution and the dotted (red) curve is the* $q\overline{q}$ *component of the background. The dash-dotted curve represents the signal contribution.*

8.3.2   Background asymmetries

The $B\overline{B}$ backgrounds are constrained in the fit to have zero asymmetry. However Table 5.10 shows that asymmetries in some modes making up the $B\overline{B}$ backgrounds exist and are already measured. This can cause a bias in the fit results. The fit was therefore repeated after introducing non-zero asymmetries in each background category where the variation is taken as an approximate weighted average of the known A_{CP} values in each category. Table 8.4 shows the results of the fits when varying the asymmetries of each $B\overline{B}$ categories. For each category the biggest variation is taken and then added in quadrature to obtain a systematic variation of

1.78%.

Table 8.4: *Variation in the fitted signal A_{CP} with varying $B\bar{B}$ background asymmetries.*

$B\bar{B}$ Category	$B\bar{B}$ background asymmetry (%)	Signal asymmetry (%)	Variation (%)
$B\bar{B}$ 1	10	−6.32	**0.34**
	−10	−5.70	0.31
$B\bar{B}$ 2	5	−6.12	**0.11**
	−5	−5.91	0.10
$B\bar{B}$ 3	5	−7.71	**1.70**
	−5	−4.32	1.69
$B\bar{B}$ 4	5	−6.39	**0.38**
	−5	−5.63	0.38

8.3.3   Fit biases

Finally the fit itself can introduce a bias in the asymmetry. This is dealt with in the same way as for the bias in the inclusive branching fraction. The biases obtained in the embedded toy experiments for the A_{CP} fit and listed in Table 7.2 show that the fit does not introduce a bias larger than 1.2% on the asymmetry. This value is therefore assigned for this uncertainty.

8.3.4   Summary of systematic uncertainties for A_{CP}

The systematic uncertainties described above and reported in Table 8.5 are added in quadrature and give a total systematic uncertainty on the CP asymmetry measurement of 3.7%.

Table 8.5: *Summary of systematic uncertainties for asymmetry measurement.*

Source	Uncertainty
Intrinsic detector asymmetry	0.5%
Selection induced asymmetries	3.0%
Background asymmetry	1.8%
Fit Bias	1.2%
Total	3.7%

8.4 Systematic uncertainties for the resonances

Systematic uncertainties on the resonant decay branching fractions arise from all the same sources as for the inclusive branching fraction measurement. All the uncertainties listed in Table 8.2 and Table 8.3 are scaled by the ratio of the resonant branching fraction to the inclusive branching fraction. However some systematic sources have to be either added since they are specific to the method used to obtain the resonant branching fractions or re-evaluated like the fit bias and π^0 efficiency.

8.4.1 Variations from inclusive systematics

The K_S^0 veto is only included in the systematic uncertainty of the $K^*(892)^+$ resonance since the signal regions for both the other two resonances lie outside the veto range. The K_S^0 systematic is evaluated from the efficiency of the K_S^0 veto in the $B^+ \rightarrow K^*(892)^+\pi^0$ MC given in Table 4.2 and was assigned an uncertainty of 2.0%

The fit bias systematic is calculated from the biases listed in Table 6.11. The systematic uncertainty is assigned as half the bias added in quadrature with the uncertainty with which the bias is known. As a percentage of the corrected branching fraction, it is found to be 6.6% for the $K^*(892)^+\pi^0$, 2.1% for the $f_0(980)K^+$ and 6.8% for the $\chi_{c0}K^+$ decay. Same method applies for the π^0 efficiency, where the systematic

is calculated from the corrections in Section 6.4.2. Therefore the systematics are found to be 2.1% for the $K^*(892)^+\pi^0$, 3.1% for the $f_0(980)K^+$ and 2.0% for the $\chi_{c0}K^+$ decay.

8.4.2 Additional systematics for the quasi-two body decays

The uncertainty on the results due to the nonresonant subtraction from the sideband method is determined by using reduced sidebands going from $\pm 3\sigma \to \pm 4\sigma$ (inner sideband) and $\pm 4\sigma \to \pm 5\sigma$ (outer sideband) on either side. Results of these variations are shown in Table 8.6 where the biases are found to have equal magnitude and opposite sign. The systematic assigned on the basis of the shift from the nominal branching fractions are 3.0% for $K^*(892)^+\pi^0$, 11.5% for $f_0(980)K^+$ and 14.3% for $\chi_{c0}K^+$.

Table 8.6: *Variations in the nominal result using reduced sidebands.*

Resonance	Sidebands used	BF($\times 10^{-6}$)
$K^*(892)^+\pi^0$	Inner	2.37
	Outer	2.21
$f_0(980)K^+$	Inner	2.22
	Outer	2.82
$\chi_{c0}K^+$	Inner	0.45
	Outer	0.59

Data/MC differences can affect the efficiency applied for the invariant mass selection. This effect is heavily enhanced by variations in the SCF fraction where the distributions are smeared out. In the inclusive measurement, the value of the SCF fraction in data was found to be 9.7% and an uncertainty of 3.5% was assigned for the iteration procedure. Therefore the value of f_{SCF} in the MC is varied by $\pm 35\%$ of the values listed in Table 4.3. The variations in the signal cut efficiency for each mode are listed in Table 8.7. The largest variation is taken and gives a contribution

of 5.6% for $K^*(892)^+\pi^0$, 3.8% for $f_0(980)K^+$ and 0.4% for $\chi_{c0}K^+$.

Table 8.7: *Variation of the signal cut efficiency in the MC with f_{SCF}.*

Resonance	Nominal signal cut effciency (%)	Signal cut efficiency with higher f_{SCF} (%)	Signal cut efficiency with lower f_{SCF} (%)
$K^*(892)^+\pi^0$	73.9 ± 0.3	70.0 ± 0.3	77.7 ± 0.3
$f_0(980)K^+$	79.9 ± 0.6	77.0 ± 0.7	82.6 ± 0.7
$\chi_{c0}K^+$	90.4 ± 0.7	90.2 ± 0.8	90.8 ± 0.7

8.4.3 Summary of systematics for the resonances

The subtotal for the systematic uncertainties taken from the inclusive branching fraction is 8.1%. Adding in quadrature this total to the other systematic uncertainties, shown in Table 8.8, gives a total systematic of 13.2% for $K^*(892)^+\pi^0$, 15.4% for $f_0(980)K^+$ and 17.1% for $\chi_{c0}K^+$.

Table 8.8: *Summary of systematic uncertainties on the branching fraction measurement for each quasi-two-body mode.*

Source	Uncertainty (%)		
	$K^*(892)^+\pi^0$	$f_0(980)K^+$	$\chi_{c0}K^+$
Subtotal from sources identical to inclusive inclusive	8.1	8.1	8.1
Background subtraction	3.0	11.5	14.3
Fit bias	6.6	2.1	6.8
Neutral pion efficiency	2.1	3.1	2.0
Mass cut efficiency	5.6	3.8	0.4
K_S^0 veto	2.0	-	-
Total	13.2	15.4	17.1

The additional systematic uncertainties listed above cancel in the asymmetry measurements. Hence the systematic uncertainties for the resonant decay asymmetries are taken to be exactly the same as reported in Table 8.5 with a total of 3.7%.

9

Discussion and Conclusion

The previous chapter assigned various uncertainties to the method used to obtain the measurements given in Chapter 6 and Chapter 7. The conclusion contains the final results for the branching fractions, corrected where necessary to give the overall branching fraction for the B decay, and the CP asymmetry including this time any uncertainty arising from systematics. Additionally the impact of the results is discussed together with potential improvements expected using data from current and future experiments.

9.1   Final results

The total systematic uncertainty on the branching fraction is the sum of all contributions from the fit model and efficiencies added in quadrature and gives a percentage uncertainty on the branching fraction of 8.9%. Similarly the total percentage systematic uncertainty on the inclusive CP asymmetry is found to be 3%. The full result for the inclusive mode including systematics is:

$$\mathcal{B}\left(B^+ \rightarrow K^+\pi^0\pi^0\right) \;=\; (16.2 \pm 1.2 \pm 1.4) \times 10^{-6}, \tag{9.1}$$

$$A_{CP}\left(B^+ \rightarrow K^+\pi^0\pi^0\right) \;=\; (-6 \pm 6 \pm 4). \tag{9.2}$$

This is the first measurement of this decay and therefore a significance for the signal observed is calculated. The total systematic uncertainty on the fit model alone represents the uncertainty on the yield and amounts to 6.5% or 79 events. As shown in Figure 6.5, the statistical significance obtained from the change in the likelihood with and without a signal component included in the fit is 15.6σ. The negative log likelihood is approximately Gaussian therefore a conservative estimate for the significance including systematic is estimated by:

$$N_{\mathrm{SD}} \approx \frac{n_{\mathrm{sig}}}{\sqrt{\sigma_{n_{\mathrm{sig}}}^2 + \sigma_{\mathrm{sys}}^2}} \tag{9.3}$$

where N_{SD} represents the significance in standard deviations, n_{sig} is the uncorrected signal yield with its statistical error $\sigma_{n_{sig}}$ and σ_{sys} the systematic uncertainty on the yield. The total significance on the signal yield is found with the above equation to be greater than 10σ. For the quasi-two-body decays, the total systematic is found to be 13% for the $K^*(892)^+\pi^0$, 15% for the $f_0(980)K^+$ and 17% for the $\chi_{c0}K^+$ decay. This gives the following results for the product branching fractions:

$$\mathcal{B}\left(B^+ \rightarrow K^*(892)^+\pi^0\right) \times \mathcal{B}\left(K^*(892)^+ \rightarrow K^+\pi^0\right) \;=\; (2.72 \pm 0.50 \pm 0.34) \times 10^{-6}, \tag{9.4}$$

$$\mathcal{B}\left(B^+ \rightarrow f_0(980)K^+\right) \times \mathcal{B}\left(f_0(980) \rightarrow \pi^0\pi^0\right) \;=\; (2.77 \pm 0.56 \pm 0.43) \times 10^{-6}, \tag{9.5}$$

$$\mathcal{B}\left(B^+ \rightarrow \chi_{c0}K^+\right) \times \mathcal{B}\left(\chi_{c0} \rightarrow \pi^0\pi^0\right) \;=\; (0.51 \pm 0.22 \pm 0.09) \times 10^{-6}. \tag{9.6}$$

The overall branching fractions for the B decays alone are obtained by applying corrections for the branching fraction of $K^*(892)^+ \rightarrow K^+\pi^0$ and $\chi_{c0} \rightarrow \pi^0\pi^0$ as

follows [38]:

$$\mathcal{B}\left(K^*(892)^+ \to K^+\pi^0\right) = 1/3, \tag{9.7}$$

$$\mathcal{B}\left(\chi_{c0} \to \pi^0\pi^0\right) = 1/3\,(8.4 \pm 0.4), \tag{9.8}$$

where the factor of 1/3 is due to isospin. The overall results with the full result for the A_{CP} measurement is given in Table 9.1 together with the current world averages. Only the CP asymmetry of $B^+ \to f_0(980)K^+$ is quoted since the value of $\mathcal{B}\left(f_0(980) \to \pi^0\pi^0\right)$ is unknown. As mentioned in the introduction to the thesis, the ratio of this product branching fraction to the world average of $\mathcal{B}\left(B^+ \to f_0(980)K^+\right) \times \mathcal{B}\left(f_0(980) \to \pi^+\pi^-\right) = \left(9.4^{+0.9}_{-1.0}\right) \times 10^{-6}$ [33, 38] is used instead. This gives:

$$\frac{\mathcal{B}\left(B^+ \to f_0(980)K^+\right) \times \mathcal{B}\left(f_0(980) \to \pi^0\pi^0\right)}{\mathcal{B}\left(B^+ \to f_0(980)K^+\right) \times \mathcal{B}\left(f_0(980) \to \pi^+\pi^-\right)} = 0.30 \pm 0.07. \tag{9.9}$$

This result shows consistency within 3σ with the prediction of 0.5 from isospin symmetry.

9.2 Discussion of the results

This analysis successfully produced results for the unmeasured inclusive branching fraction and CP asymmetry of the charmless 3-body B decay, $B^+ \to K^+\pi^0\pi^0$. These two results are the first measurements of this mode.

For the quasi-two-body B decay to $K^*(892)^+\pi^0$ the branching fraction and A_{CP} measurements are found to be in very good agreement with the only previous measurement of this mode based on a smaller *BABAR* dataset. No major improvement is observed in the statistical error, despite the larger dataset used compared to the previous study. In the previous measurement the $K^*(892)^+\pi^0$ yield was found by applying a similar cut in the invariant mass region of the K^* but the method used to account for nonresonant decay and higher K^* resonances in that mass window leads to underestimation of the contributions since the assumption is made that the

Table 9.1: *Comparison of our results to previous measurements [33, 38]. The world averages of the branching fraction and CP asymmetry of $B^+ \to K^*(892)^+\pi^0$ come from a sole prior measurement by BABAR [3] and are superceded by our results. Note that the decay $B^+ \to \chi_{c0}K^+$ and A_{CP} of $B^+ \to f_0(980)K^+$ have been studied in the $K^+\pi^-\pi^+$ final state, giving more precise results.*

	Our result	Previous world average
	$B^+ \to K^*(892)^+\pi^0$	
$\mathcal{B}$	$(8.2 \pm 1.5 \pm 1.1) \times 10^{-6}$	$(6.9 \pm 2.3) \times 10^{-6}$
A_{CP}	$(-4 \pm 26 \pm 4)\%$	$(4 \pm 29)\%$
	$B^+ \to f_0(980)K^+$	
A_{CP}	$(17 \pm 18 \pm 4)\%$	$(-9.5^{+4.9}_{-4.2})\%$
	$B^+ \to \chi_{c0}K^+$	
$\mathcal{B}$	$(182 \pm 78 \pm 31 \pm 8) \times 10^{-6}$	$(133^{+19}_{-16}) \times 10^{-6}$
A_{CP}	$(-89 \pm 37 \pm 4)\%$	$(-11 \pm 12)\%$

$K_0^{*+}(1430)$ resonance is the only contribution to the $K\pi$ S-wave. From analyses of other $B \to K\pi\pi$ decays [4,85], it is now clear that a substantial nonresonant contribution is also present. It is likely therefore that such events could have contributed to what was considered to be the $K^*(892)^+\pi^0$ signal, which would have led to a smaller statistical error. The results for $K^*(892)^+\pi^0$ found from this analysis are still more precise then the previous BABAR measurement and therefore supersede it.

Both the CP asymmetry for the quasi-two-body decay $B^+ \to f_0(980)K^+$ and the branching fraction of $B^+ \to \chi_{c0}K^+$ are in very good agreement within error with the previous world averages. The CP asymmetry for $B^+ \to \chi_{c0}K^+$ is found to be 2σ away from the world average and therefore marginally agrees with the world average and predictions of $b \to c$ transitions that the A_{CP} in this mode should be consistent with zero. The ratio of branching fractions for the quasi-two-body decay $B^+ \to f_0(980)K^+$ for which results are shown in Eq. 9.9, also marginally agrees with

the predictions.

9.3   Improvements using future experiments

This analysis could benefit from much larger statistics than available at *BABAR* to improve the statistical error in $B^+ \to K^*(892)^+\pi^0$. The Belle experiment at KEK (Japan) did not measure this mode despite having almost double the *BABAR* statistics. However much larger dataset would be needed to make a substantial improvement to the measurements presented in this thesis. B physics experiments at hadron colliders, such as LHCb at CERN, would have severe difficulty reconstructing this final state due to the presence of two π^0 mesons [86]. Consequently, despite the very large statistics that LHCb will collect, it is highly unlikely that any improvements on these measurements will be possible.

A Super B Factory could provide the large statistics in a much cleaner environment than the LHCb experiment and therefore could expect to improve on the measurements presented in this thesis. An example of a super B Factory is the SuperB experiment, to be hosted at the Laboratori Nazionali di Frascati in Italy and currently in planning phase. The experiment will consist of an e^+e^- asymmetric collider operating at a the $\Upsilon(4S)$ resonance with a target luminosity of 10^{36} and an expected total dataset of approximately $75\,\mathrm{ab}^{-1}$ [87]. The SuperB detector itself is based on extensive reuse of the current *BABAR* detector, therefore an approximation of the expected yield of $B^+ \to K^+\pi^0\pi^0$ events can be made by assuming that the efficiencies are similar to those at *BABAR*. With the SuperB dataset, approximately $200,000$ events can be expected, almost 200 times more events than were observed at *BABAR*. The upgrade of the Belle experiment, BelleII, is also a Super B Factory, albeit a slightly less ambitious programme than SuperB with an expected final dataset of $50\,\mathrm{ab}^{-1}$. This experiment will also certainly be able to measure this mode with a much larger precision than both *BABAR* and Belle [88].

A full amplitude analysis of $B^+ \to K^+\pi^0\pi^0$ would in fact have multiple advantages. Because the three body decay is fully modelled and interferences within the Dalitz plot are quantified, all contributions to the Dalitz plot can be studied and to a higher degree of accuracy, *i.e.* can measure phases and amplitudes which gives more sensitivity to localised CP violation effects. In addition to allowing a more in-depth study of the Dalitz plot, it also allows to improve on the measurements made in this thesis by solving many of the issues encountered during this analysis, such as the Punzi effect. With the Dalitz plot parameters in the fit, it is possible to accommodate PDF parameters for ΔE which are dependent on the Dalitz plot parameters. This allows to better model the $B\bar{B}$ backgrounds and reduce the systematic uncertainty. The variation of the SCF fraction over the Dalitz plot can also be included in the fit and again reduce the systematic uncertainty. It also makes the measurement less reliant on the correct reconstruction from the ${}_s\mathcal{P}lots$ of the Dalitz plot distribution. However there are numerous difficulties associated with a full amplitude analysis of $B^+ \to K^+\pi^0\pi^0$.

Even not taking into account the various difficulties in modelling, a Dalitz plot analysis done on just a few 1000 events would not improve the measurement. Therefore with the current statistics available in the *BABAR* dataset, it would be impossible to improve the measurements in this thesis. At a Super B Factory, the level of statistics is more than sufficient but other factors have to be taken into account and studied. The large SCF fraction makes it vital to model the migration of events accurately. This is the difference between the true position and the reconstructed position of an event in the Dalitz plot. Because the resolution tends to be worse for two π^0 mesons than charged tracks even for correctly reconstructed events, migration effects are also non-negligeable in truth matched events. An illustration of these migration effects obtained from nonresonant MC of this analysis is shown in Figure A.2. Consequently, the MC needs to be extremely accurate to be able to model these effects. Even if this can be achieved, the poorer resolution and smearing effects degrade the sensitivity to the physics. A further problem is that larger

statistics implies a larger background. These backgrounds have to be eliminated as much as possible before the fitting stage, *i.e.* very good displaced vertex resolution and continuum background separation.

The conclusion is that, even with very large statistics, all these resolution, migration and background effects have to be studied prior to attempting a full amplitude analysis of $B^+ \rightarrow K^+\pi^0\pi^0$ to determine whether or not it is feasible.

A

Effect of Punzi bias in fit model

An effect observed during various stages of this analysis is a bias arising during the fit due to the very high correlations of variables across the Dalitz plot. This effect is known as a Punzi bias after the physicist Giovanni Punzi who discussed the pitfalls in Maximum Likelihood fits due to the use of variables in a PDF that depend on another set of variables, known as "conditional variables" [89]. An illustration of such a PDF is shown below:

$$P_{sig}\left(t,\sigma_t\right) = \frac{1}{2\tau}e^{-t/\tau} \otimes G\left(0,\sigma_t\right) \qquad P_{bkg}\left(t,\sigma_t\right) = \delta(t) \otimes G\left(0,\sigma_t\right) \qquad \text{(A.1)}$$

where $G\left(0,\sigma_t\right)$ represents a normalised Gaussian function with mean zero and an event-by-event dependent width σ_t. In an unbinned extended maximum likelihood fit, the optimal value of the fit parameter τ is obtained by minimising the log

likelihood of the function:

$$\mathcal{L}(t) = e^{-\left(N_{sig}+N_{bkg}\right)} \prod_i \left(N_{sig}P_{sig}\left(t^i; \sigma_t^i\right) + N_{bkg}P_{bkg}\left(t^i; \sigma_t^i\right)\right) \tag{A.2}$$

The product is over the events in the fit sample and superscript i were added to t and σ_t to explicitely state that the normalisation of the PDFs is an event-by-event quantity, which differs from the normalisation used in Section 3.3. The number of signal and background events are typically obtained as free parameters in the fit. However such conditional PDFs can lead to biases if the underlying distributions of the conditional variables are different for the different event types in the fit (signal and background). The reason for this is that the PDF expressions for signal and background in Eq. A.1 are actually conditional probability distributions and the full PDF should in fact be written as:

$$\mathcal{P}_{\text{sig}}\left(t, \sigma_t\right) = P_{\text{sig}}\left(t|\sigma_t\right) \times P_{\text{sig}}\left(\sigma_t\right) \qquad \mathcal{P}_{\text{bkg}}\left(t, \sigma_t\right) = P_{\text{bkg}}\left(t|\sigma_t\right) \times P_{\text{bkg}}\left(\sigma_t\right) \tag{A.3}$$

In this analysis, two discriminating variables, ΔE and the initial NN_{out} considered at the beginning of the analysis with 5 input variables (all the variables listed in Section 4.2 together with the absolute value of Δt over its uncertainty), were correlated with Dalitz plot coordinates. The parameters for the PDFs of these variables will therefore vary according to the position in the Dalitz plot on an event-by-event basis. The NN_{out} dependence was solved by eliminating the strongly Dalitz plot dependent variable from the neural network input. The major issue was the dependence in ΔE. It is a known fact that if the final state involves neutral pions, the ΔE distribution is dependent on the momentum of the pion, which in turn is directly correlated to Dalitz plot position. Self-cross-feed effects tend to be the cause of Dalitz plot dependencies since these events tend to be more frequent at the corner of the Dalitz plot where one particle is slow (see Figure 4.12). Therefore these events tend to have worse ΔE and m_{ES} resolutions than correctly reconstructed events (TM). The dependence of these variables in the Dalitz plot can be observed in Figure A.1.

A number of possible fit strategies were considered [90]:

152
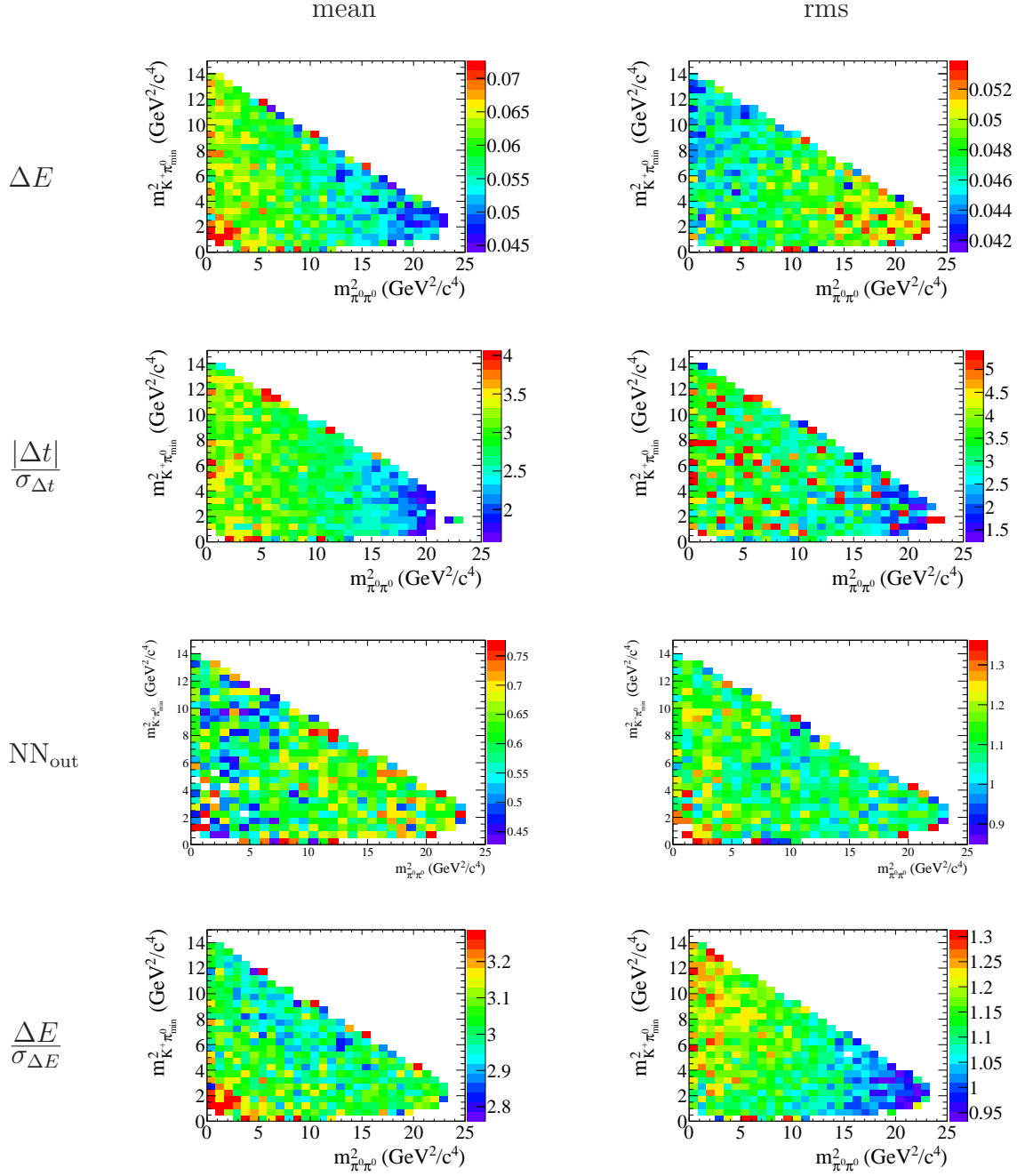
Figure A.1: *Dalitz plot dependences of the mean and width of the distributions for ΔE, $\frac{|\Delta t|}{\sigma_{\Delta t}}$, the $\mathrm{NN}_{\mathrm{out}}$, as a consequence of the dependence in $\frac{|\Delta t|}{\sigma_{\Delta t}}$, and finally $\frac{\Delta E}{\sigma_{\Delta E}}$.*

1. Fitting m_{ES} and ΔE without Dalitz plot dependence in the ΔE PDF;

2. Fitting m_{ES} and ΔE including Dalitz plot dependences in the ΔE PDF, *i.e.* so that the functional form of the PDF has parameters such as mean and width that depend on the Dalitz plot variables, but without fitting the Dalitz plot variables themselves;

3. Fitting m_{ES} and $\frac{\Delta E}{\sigma_{\Delta E}}$, as a replacement for ΔE. This solution was used in the analysis $B^0 \to \pi^0 \pi^0$ to improve the resolution of the ΔE distribution [91]. Events in regions of the Dalitz plot where the ΔE resolution is poorer should in principle have larger event-by-event uncertainties on ΔE ($\sigma_{\Delta E}$) and therefore the effects of resolution should cancel in $\frac{\Delta E}{\sigma_{\Delta E}}$;

4. Fitting m_{ES}, ΔE and the Dalitz plot variables including the Dalitz plot dependence in the ΔE PDF;

5. Fitting m_{ES} only and placing a tight cut on the ΔE variable.

Embedded toy studies have clearly shown that solutions 1 and 2 are indeed sensitive to this effect and produce large biases. Some examples of these large biases are shown in the signal yield distributions from the signal embedded toy MC experiments performed for the nonresonant and resonant modes of $B^+ \to K^+ \pi^0 \pi^0$ in Figure A.3. Due to the different Dalitz plot distributions of signal and background, to have solution 2 work would require including the Dalitz plot variables in the fit itself *i.e.* solution 4 and therefore doing a full amplitude analysis. Since this is the first measurement of this mode, the signal yield expected was not well known and could have been too small for a Dalitz plot analysis. Together with the large levels of self-cross-feed, the phases and amplitudes at interference regions at the corners of the Dalitz plot where self-cross-feed events are dominant would be difficult to determine. In addition, SCF events tend to migrate dramatically across the Dalitz plot. The migration of an event is defined as the difference between its true position and reconstructed position in the Dalitz plot. The nonresonant MC was used to study
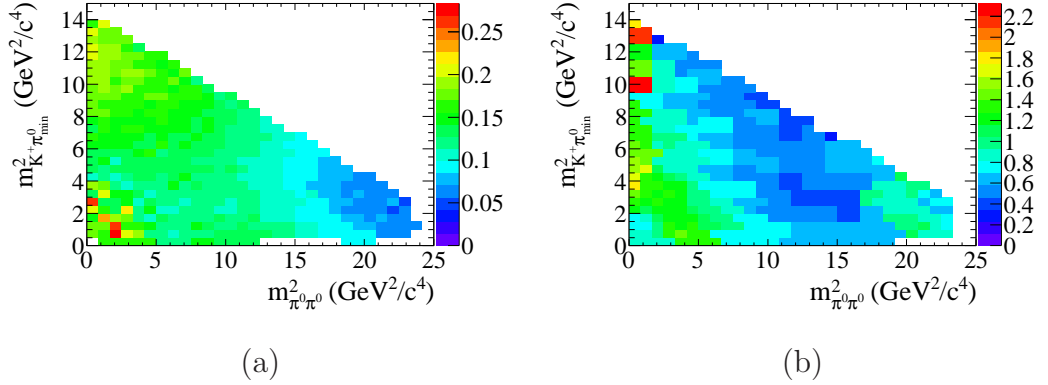
Figure A.2: *Migration of (a) TM and (b) SCF events within the Dalitz plot, plotted at their MC truth coordinates. Note the different scales on the z-axes.*

the migration of events and the difference between truth and reconstructed values is calculated. Figure A.2 shows the Dalitz plot distributions for TM and SCF events weighted by the migration distance defined as:

$$d_{\text{migration}} = \sqrt{(m_{13}^{2\,\text{truth}} - m_{13}^{2\,\text{reco}})^2 + (m_{23}^{2\,\text{truth}} - m_{23}^{2\,\text{reco}})^2}\,. \tag{A.4}$$

where m_{13} and m_{23} are the invariant masses for the two-particle combinations $K\pi^0$ and $\pi^0\pi^0$. An amplitude analysis of the $B^+ \to K^+\pi^0\pi^0$ decay would require detailed understanding of these resolution effects not available with the current statistics. This excludes solution 4. Solution 3 replaces ΔE by $\frac{\Delta E}{\sigma_{\Delta E}}$. However in this case, as shown in Figure A.1, the variation of this variable across the Dalitz plot exhibits slightly worse correlations then the variation of ΔE. This could be due to the uncertainty being itself strongly dependent on the Dalitz plot position more so then ΔE itself. In which case the ratio would also be dependent on the Dalitz plot coordinates. This leaves only solution 5 described in detail throughout this thesis.

Each of the six plots shows the fitted parameters:

Top-left plot:

Integral	500
N	19.37 ± 1.24
μ	717.5 ± 3.8
σ	74.58 ± 3.51

$B^+ \to K^+\pi^0\pi^0$ nonresonant

Top-right plot:

Integral	500
N	18.63 ± 1.13
μ	619.7 ± 3.9
σ	80.15 ± 3.39

$B^+ \to K^*(892)^+\pi^0$

Middle-left plot:

Integral	500
N	18.09 ± 1.11
μ	593.1 ± 3.9
σ	78.57 ± 3.56

$B^+ \to K^{*+}(1680)\pi^0$

Middle-right plot:

Integral	500
N	16.95 ± 1.04
μ	600 ± 4.3
σ	83.99 ± 3.85

$B^+ \to K^{*+}(1410)\pi^0$

Bottom-left plot:

Integral	500
N	17.11 ± 1.05
μ	778.8 ± 4.2
σ	80.61 ± 3.71

$B^+ \to f_2(1270)K^+$

Bottom-right plot:

Integral	500
N	19.27 ± 1.19
μ	845 ± 3.5
σ	72.28 ± 3.12

$B^+ \to f_0(980)K^+$

Figure A.3: *Signal yield distributions from embedded signal MC toy with 500 experiments. The red arrow indicated the expected yield with a loose ΔE signal region of 673 events. Biases range from 40 events to as high as 171 for resonances decaying to $\pi^0\pi^0$.*

B

Neutral pion efficiency study

By far the largest systematic uncertainty is expected to be due to the two π^0 mesons in the final state. A separate study was carried out to precisely measure the efficiency with which neutral pions are detected by *BABAR*. More on this study can be found in Ref. [92]. Since the branching fraction of τ leptons to final states containing one charged track with or without additional neutral particles are quite well known (see Table B.1), these decays provide the potential to study the efficiency to percent precision.

Decay mode	Branching fraction	Uncertainty as percentage
$\tau \to \pi\nu$	$(10.91 \pm 0.07) \times 10^{-2}$	0.6%
$\tau \to K\nu$	$(6.95 \pm 0.23) \times 10^{-3}$	3.3%
$\tau \to \mu\nu\nu$	$(17.36 \pm 0.05) \times 10^{-2}$	0.3%
$\tau \to e\nu\nu$	$(17.85 \pm 0.05) \times 10^{-2}$	0.3%
$\tau \to \mu\nu\nu\gamma$	$(3.6 \pm 0.4) \times 10^{-3}$	11%
$\tau \to e\nu\nu\gamma$	$(1.75 \pm 0.18) \times 10^{-2}$	10%
$\tau \to \pi\pi^0\nu$	$(25.52 \pm 0.10) \times 10^{-2}$	0.4%
$\tau \to K\pi^0\nu$	$(4.28 \pm 0.15) \times 10^{-3}$	3.5%
$\tau \to \pi\eta\nu$	$< 1.4 \times 10^{-4}$	n/a
$\tau \to K\eta\nu$	$(2.7 \pm 0.6) \times 10^{-4}$	22%
$\tau \to \pi\pi^0\pi^0\nu$	$(9.27 \pm 0.12) \times 10^{-2}$	1.3%
$\tau \to K\pi^0\pi^0\nu$	$(6.3 \pm 2.3) \times 10^{-4}$	37%
$\tau \to \pi\pi^0\eta\nu$	$(1.81 \pm 0.24) \times 10^{-3}$	13%
$\tau \to K\pi^0\eta\nu$	$(1.8 \pm 0.9) \times 10^{-4}$	50%

B.1 Measuring the π^0 efficiency using τ decays

The efficiency is measured by reconstructing τ decays to "$\pi^\pm\nu$" and "$\pi^\pm\pi^0\nu$", where the π^0 candidate is combined to the charged pion to form a ρ candidate. The efficiency is calculated using minimal selection criteria so as to minimise possible differences in selection efficiencies between data and MC. The other τ in the event is required to be reconstructed as a $\tau \to e\nu\nu$ decay, to enable effective selection of $e^+e^- \to \tau^+\tau^-$ events. Single ratios of number of events are then formed in both data and MC as follows:

$$R_{\mathrm{data}} = \frac{N\left(\tau \to \rho^\pm\nu\right)_{\mathrm{data}}}{N\left(\tau \to \pi^\pm\nu\right)_{\mathrm{data}}} \tag{B.1}$$

$$R_{\text{MC}} = \frac{N\left(\tau \to \rho^{\pm}\nu\right)_{\text{MC}}}{N\left(\tau \to \pi^{\pm}\nu\right)_{\text{MC}}} \tag{B.2}$$

This can in principle be used to calculate the π^0 efficiencies in data and MC respectively but corrections need to be applied for the relevant branching fractions. Since it is expected that the π^0 efficiency will have some momentum dependence, the ratios are formed in terms of the π^0 momentum as such:

$$R_{\text{data}}\left(P\left(\pi^0\right)\right) = \frac{\left(\frac{dN\left(\tau \to t\pi^0\nu\right)}{dP(\pi^0)}\right)_{\text{data}}}{N\left(\tau \to t\nu\right)_{\text{data}}} \tag{B.3}$$

$$R_{\text{MC}}\left(P\left(\pi^0\right)\right) = \frac{\left(\frac{dN\left(\tau \to t\pi^0\nu\right)}{dP(\pi^0)}\right)_{\text{MC}}}{N\left(\tau \to t\nu\right)_{\text{MC}}} \tag{B.4}$$

The data/MC efficiency is then extracted from a double ratio formed using the single ratios in Eqs. B.3–B.4. This ratio becomes:

$$\mathcal{R}\left(P\left(\pi^0\right)\right) = \frac{R_{\text{data}}\left(P\left(\pi^0\right)\right)}{R_{\text{MC}}\left(P\left(\pi^0\right)\right)} \tag{B.5}$$

The strength of this technique is that the systematic uncertainty for the efficiency measurement cancel in the double ratio to good approximation. Effects due to luminosity or MC statistics, trigger and tracking efficiencies should cancel in the single ratios and effects due to branching fractions should cancel again in the double ratio. As demonstrated in Table B.1, these branching fractions are quite well known. There are however a number of effects that need to be taken into account and minimised. These are listed below:

◇ Selection efficiencies can vary between data and MC. The majority of selection requirements are applied to both numerator and denominator in single and double ratios. Therefore to first order any data/MC differences should cancel in the double ratio but second order effects can persist and cause biases comparable to the level of precision required. To minimise these effects only minimal selection requirements were applied.

◇ Some detector effects may not be accurately modelled in the MC *e.g.* hadronic interactions. This effect was studied by comparing the distributions of the separation between clusters in data and MC.

◇ Particle identification performances can be different between data and MC. Standard PID weighting algorithms were applied to correct these discrepancies.

◇ The π^0 efficiency may vary as a function of time, either due to dependances on the instantaneous luminosity or due to localised inefficiencies *e.g.* bad crystals in the EMC. Luminosity related effects should cancel to first order due to the nature of the double ratio but second order effects can again persist. The MC is weighted in order to accurately match the data luminosity after checking that the spatial distribution of clusters in MC follows same variations as in data.

◇ Background levels from other $e^+e^- \rightarrow \tau^+\tau^-$ events can be different in either MC or data. A check was done at generator level information for these background types. Selection requirements should however be highly effective at removing most type of backgrounds from other type of events.

◇ Machine-related background can be an issue as they result in additional clusters in the EMC that might not be well modelled in the MC. However the production of MC events includes overlaying of randomly triggered background frames so that this effect should be accounted for to first order. This effect was further studied by raising the minimum cluster energy threshold.

B.2   Data sample and selections

This study uses Runs 1–5 onpeak data and $\tau^+\tau^-$ generic MC samples. The total luminosity for data was about $346\,\mathrm{fb}^{-1}$ and the MC contained 326×10^6 generated events. No standard π^0 list selection is applied but all neutral clusters to be considered as part of the event must have energy greater than $100\,\mathrm{MeV}$ and be in the detector acceptance defined as $0.28 < \theta < 2.4$, where θ is the polar angle of the track in the laboratory frame. The following selection requirements are made on all events:

◇ events are required to have two oppositely charged tracks with one track passing the electron likelihood selector while the other is required to fail the same selectors;

◇ the event is required to have passed either the DCH or the EMC L3 trigger (see Section 2.8 on the two types of trigger);

◇ each track is required to pass through the barrel region of the detector, defined as $0.42 < \theta < 2.4$;

◇ the ratio of visible (E_{vis}) to total energy (E_{tot}) is required to be $0.4 < E_{vis}/E_{tot} < 0.7$;

◇ the missing energy vector needs to satisfy $-0.76 < \cos\theta_{miss} < 0.92$ where θ_{miss} is the cosine of the polar angle of the missing energy;

◇ the difference between the two tracks for both 3D and azymuthal angles needs to be less than $174°$;

◇ the event shape variable $R2$ is required to satisfy $0.6 < R2 < 0.99$. $R2$ is defined as the ratio of the second and zeroth Fox-Wolfram moment which equations consist of:

$$H_i = \sum_{i,j} \frac{|p_i||p_j|}{E_{vis}^2} \times P_l \left(cos\left(\phi_{ij}\right)\right) \tag{B.6}$$

where indices i and j run over the hadrons produced in the event, phi_{ij} is the angle between them, E_{vis} the visual energy in the event and P_l the corresponding Lagrange polynomial. Isotropic events have $R2 \to 0$, whereas jet-like events like in $\tau^+\tau^-$ decays have $R2 \to 1$ [93];

◇ the ratio of the transverse component of the sum of the momenta of the tracks to the total missing energy in the centre of mass frame is:

$$\frac{(p_1^* + p_2^*)}{\sqrt{2} - |p_1^*| - |p_2^*|} > 0.2 \tag{B.7}$$

◇ the momentum of each track is required to be greater than $0.6\,\mathrm{GeV}/c$ in the laboratory frame and less than $4.5\,\mathrm{GeV}/c$ in the centre of mass frame;

◇ the signal-side track is also required to satisfy $E/p < 0.8$.

Additional requirements pertaining to the selection of the neutral clusters in 1 and 2 bumps events (see Section 2.6 for a full definition of bumps in the EMC) are applied to specifically choose π^0 candidates to be in the same hemisphere as the charged pion. Neutral candidates are require to have at most two neutral clusters. Events with one neutral cluster are assumed to be merged π^0 candidates. Two neutral clusters are assumed to be two photons from a π^0 decay. Finally π^0 and ρ candidates are further selected by:

◇ the momentum of the π^0 candidate in the laboratory frame must be greater than $100\,\mathrm{MeV}/c$;

◇ the π^0 candidate mass must be less than $250\,\mathrm{MeV}/c^2$;

◇ the ρ mass must satisfy $0.55 < m_\rho < 1.0\,\mathrm{GeV}/c^2$;

◇ the modulus of the cosine of the ρ decay angle must be less than 0.9;

◇ the distance between the intersection of the charged pion track with the EMC and the nearest neutral cluster must be greater $20\,\mathrm{cm}$.

B.3 Single/Double ratio results and efficiency extraction

The momentum double ratio is given by dividing the single ratio distribution of the π^0 momentum in one and two bump events by the ratio of the pion momentum in the zero bump sample from data and MC. Before calculating the double ratio, the single ratios of these quantities was studied to look for discrepancies between data and MC.
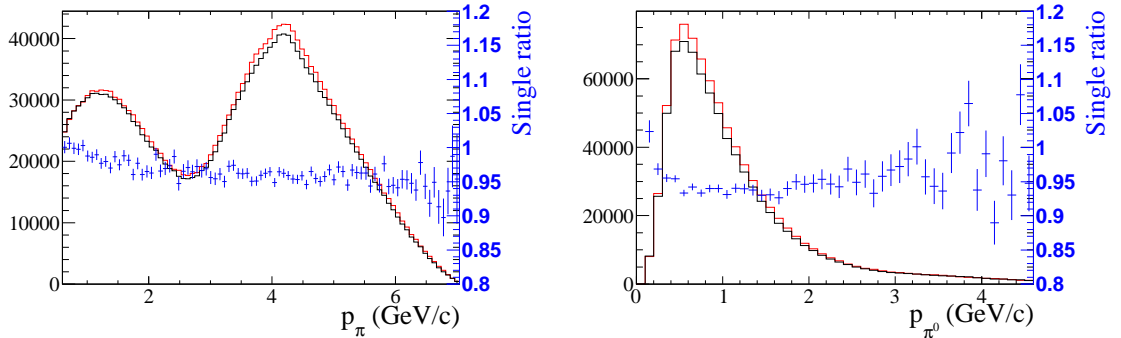
Figure B.1: *Single ratio momentum distributions of (left) charged pions in the zero bump sample and (right) neutral pions in the one and two bumps samples. The red line is the distribution obtained from MC and black line data. The blue data points with blue axis to the right show the result of the single ratio (data/MC).*

These are shown in Figure B.1. Overall there is fairly good agreement between data and MC. A small discrepancy is observed in the charged pion momentum between $3.5 - 5.5$ GeV/c and in the π^0 distribution at low momenta. These discrepancies were studied with variations on the selection criteria and documented in Ref. [92]. The resultant double ratio distribution is fitted with a first order polynomial up to $4\,\text{GeV}/c$. Figure B.2 shows the double ratio distribution adding to the selection the `PiOAllLoose` list requirements. At higher momenta the double ratio becomes unstable due to lack of statistics so the fit is applied up to a mass of $4\,\text{GeV}/c^2$.

B.4   Systematic variations

The minimum cluster energy and ρ mass selections are varied to determine a systematic variation of the double ratio. To achieve this, double ratios are constructed with the following changes in requirements:

◇ the minimum energy for a neutral cluster is lowered to $50\,\text{MeV}$,
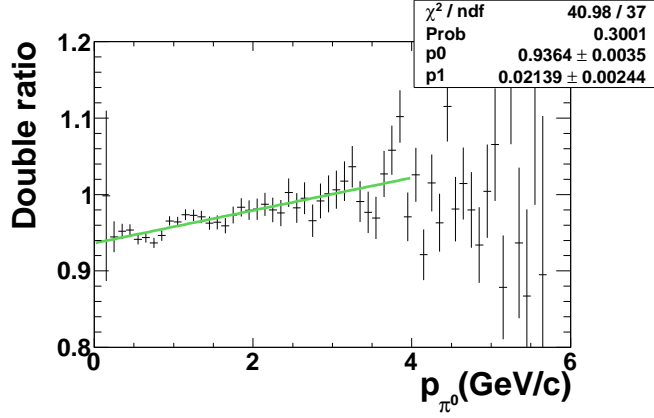
Figure B.2: *Double ratio as a function of π^0 momentum using additional requirements from the* `PiOAllLoose` *list. The black data points show the ratio of the π^0 momentum of 1 and 2 bump events to pion momentum for events with no bumps. The green line indicates the first order polynomial fit to the double ratio.*

◇ looser and tighter selections on the ρ mass are applied:

– $0.50\,\text{GeV}/c^2 < m_\rho < 1.05\,\text{GeV}/c^2$,

– $0.60\,\text{GeV}/c^2 < m_\rho < 0.95\,\text{GeV}/c^2$.

The difference between each of these double ratio formulations and the nominal double ratio is calculated bin-by-bin. In addition to the sources mentioned above, a 2% systematic uncertainty is assigned for the electron and PID requirement (veto) on the electron (pion) track. All of these uncertainties are added in quadrature to give a total uncertainty for each bin. This is then added in quadrature to the statistical uncertainty of the bin in the double ratio and the new distribution, shown in Figure B.3, is fitted again. The fitted function is:

$$f(p_{\pi^0}) = (0.936 \pm 0.003 \pm 0.007) + (0.022 \pm 0.002 \pm 0.003)p_{\pi^0} \qquad \text{(B.8)}$$

where the first error on each parameter is statistical and the second is systematic. A momentum dependence of the efficiency correction is clearly observed. Therefore a

164

π^0 momentum dependent correction and systematic need to be obtained using this double ratio distribution.
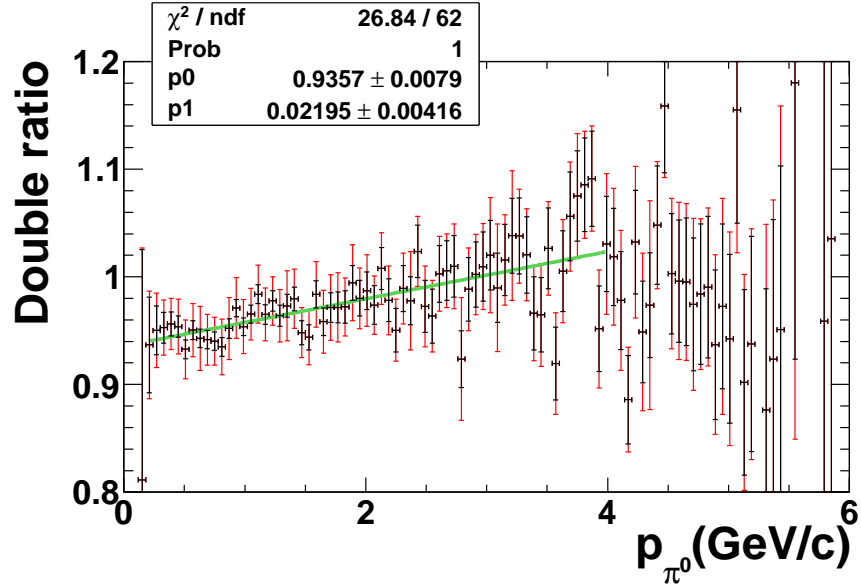

Figure B.3: *Double ratio as a function of π^0 momentum using additional requirements from the* `PiOAllLoose` *list. The black data points show the nominal double ratio distribution with statistical errors and the red extended error bars the additional uncertainty from the systematics. The green line is the linear fit to the double ratio that includes systematics.*

Bibliography

[1] Ciuchini, M., Franco, E., Martinelli, G., Pierini, M., Silvestrini, L., *Phys. Lett.*, B674, 197-203 (2009), `0811.0341`.

[2] Gronau, M., Pirjol, D., Zupan, J., *Phys. Rev.*, D81, 094011 (2010), `1001.0702`.

[3] Aubert, B., et al. (*BABAR*), *Phys. Rev.*, D71, 111101 (2005), `hep-ex/0504009`.

[4] Aubert, B., et al. (*BABAR*), *Phys. Rev.*, D78, 012004 (2008), `0803.4451`.

[5] Garmash, A., et al., *Phys. Rev. Lett.*, 96, 251803 (2006), `hep-ex/0512066`.

[6] Alford, M.G., Jaffe, R.L., *Nucl. Phys.*, B578, 367-382 (2000), `hep-lat/0001023`.

[7] Maiani, L., Piccinini, F., Polosa, A.D., Riquer, V., *Phys. Rev. Lett.*, 93, 212002 (2004), `hep-ph/0407017`.

[8] Weinstein, J.D., Isgur, N., *Phys. Rev.*, D41, 2236 (1990).

[9] Deandrea, A., Gatto, R., Nardulli, G., Polosa, A.D., Tornqvist, N.A., *Phys. Lett.*, B502, 79-86 (2001), `hep-ph/0012120`.

[10] Ecklund, K.M., et al. (CLEO), *Phys. Rev.*, D80, 052009 (2009), `0907.3201`.

[11] Christenson, J.H., Cronin, J.W., Fitch, V.L., Turlay, R., *Phys. Rev. Lett.*, 13, 138–140 (1964).

[12] Sakharov, A.D., *Pisma Zh. Eksp. Teor. Fiz.*, 5, 32-35 (1967).

[13] Bigi, I., Sanda, A., *CP Violation*, 2nd ed. (Cambridge University Press, 2009), pp. 143–158.

[14] Sozzi, M.S., *Discrete Symmetries and CP Violation*, 1st ed. (Oxford University Press, 2008), pp. 383–415.

[15] Höcker, A., Ligeti, Z., *Annual Review of Nuclear and Particle Science*, 56, 501-567 (2006).

[16] Ceccucci, A., Ligeti, Z., Sakai, Y., *LBNL*, 447 (2008).

[17] Cabibbo, N., *Phys. Rev. Lett.*, 10, 531-533 (1963).

[18] Kobayashi, M., Maskawa, T., *Prog. Theor. Phys.*, 49, 652-657 (1973).

[19] Charles, J., et al. (CKMfitter Group), *Eur. Phys. J.*, C41, 1-131 (2005), `hep-ph/0406184`, URL `http://ckmfitter.in2p3.fr`.

[20] Sozzi, M.S., *Discrete Symmetries and CP Violation*, 1st ed. (Oxford University Press, 2008), pp. 1–260.

[21] Kirkby, D., Nir, Y. (Particle Data Group), *PDG review*, 12 (2007).

[22] Angelopoulos, A., et al. (CPLEAR), *Phys. Lett.*, B444, 43-51 (1998).

[23] Lande, K., Booth, E.T., Impeduglia, J., Lederman, L.M., Chinowsky, W., *Phys. Rev.*, 103, 1901–1904 (1956).

[24] Albrecht, H., et al. (ARGUS), *Phys. Lett.*, B192, 245 (1987).

[25] Aubert, B., et al. (*BABAR*), *Phys. Rev. Lett.*, 98, 211802 (2007), `hep-ex/0703020`.

[26] Starič, M., et al., *Phys. Rev. Lett.*, 98, 211803 (2007).

[27] Aubert, B., et al. (BaBar), *Phys. Rev. Lett.*, 93, 131801 (2004), `hep-ex/0407057`.

[28] Lin, S.W., et al. (Belle), *Nature*, 452, 332-335 (2008).

[29] Aubert, B., et al. (*BABAR*), *Phys. Rev.*, D76, 091102 (2007), `0707.2798`.

[30] Gronau, M., *Phys. Lett.*, B627, 82-88 (2005), `hep-ph/0508047`.

[31] Gronau, M., Hernandez, O.F., London, D., Rosner, J.L., *Phys. Rev.*, D50, 4529-4543 (1994), `hep-ph/9404283`.

[32] Chiang, C.W., London, D., *Mod. Phys. Lett.*, A24, 1983-1994 (2009), `0904.2235`.

[33] Barberio, E., et al. (Heavy Flavor Averaging Group) (2008), `0808.1297`.

[34] Bauer, C.W., Pirjol, D., *Phys. Lett.*, B604, 183-191 (2004), `hep-ph/0408161`.

[35] Chang, Q., Li, X.Q., Yang, Y.D., *JHEP*, 09, 038 (2008), `0807.4295`.

[36] Peskin, M.E., *Nature*, 452, 293-294 (2008).

[37] Wolfenstein, L., *Phys. Rev. Lett.*, 51, 1945 (1983).

[38] Amsler, C., et al. (Particle Data Group), *Phys. Lett.*, B667 (2008).

[39] Dalitz, R.H., *Phil. Mag.*, 44, 1068-1080 (1953).

[40] Latham, T.E., *Amplitude analysis of the charmless decays of charged B mesons to the final states $K^{\pm}\pi^{\mp}\pi^{\pm}$ using the BABAR detector*, Ph.D. thesis, SLAC-R-834.

[41] Gershon, T., Latham, T., Mohanty, G., "Dalitz Plot Analysis of $B^+ \to K^+\pi^-\pi^+$ decays", *BABAR* Analysis Document #1512.

[42] Kozanecki, W., *Nucl. Instrum. Meth.*, A446, 59-64 (2000).

[43] URL   http://www.slac.stanford.edu/BFROOT/www/Detector/Images/
peprings.gif.

[44] Vance, E., Harris, D., *Symmetry*, 03.03 (2006), URL http://www.
symmetrymagazine.org/cms/?pid=1000310.

[45] Turner, J.L., et al., "Trickle-charge: a new operational mode for PEP-II", *in*
"9th European Particle Accelerator Conference", (2004), p. 881.

[46] Seeman, J., et al., "Results and plans of the PEP-II B-factory", *in* "9th Euro-
pean Particle Accelerator Conference", (2004), p. 875.

[47] URL http://www.slac.stanford.edu/BFROOT/detector/online/archive/
bbr-nfs01/www/babarrc/LumInt-Run_1-7.eps.

[48] Aubert, B., et al. (*BABAR*), *Nucl. Instrum. Meth.*, A479, 1-116 (2002), hep-ex/
0105044.

[49] Schwiening, J., "Performance of the *BABAR*-DIRC", *in* "5th Workshop on RICH
Detectors (RICH04)", (2005), p. 875.

[50] Khan, A., Schofield, K., Wright, D., "Validation of geant4 electromagnetic and
hadronic processes in the babar detector", *in* "Nuclear Science Symposium
Conference Record, 2005 IEEE", , vol. 2 (2005), pp. 844 –846.

[51] Menges, W., *IEEE Nucl. Sci. Symp. Conf. Rec.*, 5, 1470-1474 (2006), physics/
0609039.

[52] Burchat, P., et al. (Blind Analysis Task Force and BaBar Publication Board),
"Draft guidelines for blind analyses in *BABAR*", *BABAR* Analysis Document #91.

[53] Harrison, ed., P.F., Quinn, ed., H.R. (*BABAR*), *The BABAR physics book: Physics
at an asymmetric B factory*, pp. 111–125, papers from Workshop on Physics at
an Asymmetric B Factory (BaBar Collaboration Meeting), Rome, Italy, 11-14
Nov 1996, Princeton, NJ, 17-20 Mar 1997, Orsay, France, 16-19 Jun 1997 and
Pasadena, CA, 22-24 Sep 1997.

[54] Hearty, C., "Measurement of the number of $\Upsilon(4S)$ mesons produced in run1 (B counting)", *BABAR* Analysis Document #134.

[55] Ford, W., "Choice of kinematic variables in B meson reconstruction", *BABAR* Analysis Document #53.

[56] Höcker, A., Speckmayer, P., et al., CERN-OPEN-2007-007, `physics/0703039v5`.

[57] Barlow, R., *Nucl. Instrum. Meth.*, A297, 496 - 506 (1990).

[58] James, F., Roos, M., *Comput. Phys. Commun.*, 10, 343-367 (1975).

[59] "Laura++", URL `http://www.slac.stanford.edu/BFROOT/www/Physics/Analysis/AWG/chrmls_hadronic/ThreeBody/Software/Laura++/index.html`.

[60] "The roofit toolkit for data modeling", URL `http://roofit.sourceforge.net/docs/index.html`.

[61] "Root", URL `http://root.cern.ch/drupal/content/documentation`.

[62] Pivk, M. (2006), `physics/0602023`.

[63] Pivk, M., Le Diberder, F.R., *Nucl. Instrum. Meth.*, A555, 356-369 (2005), `physics/0402083`.

[64] Allen, M., Mitchell, N., Roodman, A., Banerjee, S., "A measurement of π^0 efficiency using $\tau \to \rho\nu$ and $\tau \to \pi\nu$", *BABAR* Analysis Document #870.

[65] Dujmic, D., "Merged π^0 reconstruction: from cluster shape to mass", *BABAR* Analysis Document #514.

[66] "Btacandidate lists", URL `http://www.slac.stanford.edu/BFROOT/www/doc/workbook/eventinfo/BtaCandInfo/CandLists.html`.

[67] "π^0 lists", URL `http://www.slac.stanford.edu/BFROOT/www/Physics/Analysis/AWG/Neutrals/doc/pi0lists-release14.html`.

[68] Ongmongkolkul, P., Mishra, K., Telnov, A., Gaz, A., "Particle identification using error correction output code multiclass classifier", *BABAR* Analysis Document #2199.

[69] URL `http://www.slac.stanford.edu/BFROOT/www/Physics/Tools/Pid/UltimatePID/ultimate_PID_recipe.html#kaonBDT`.

[70] Punzi, G. (2003), `physics/0308063`.

[71] Pivk, M., et al.

[72] "Btathrust", URL `http://www.slac.stanford.edu/BFROOT/dist/releases/newest/BetaCoreTools/BtaThrust.hh`.

[73] Beringer, J., "B Tagger: A multivariate tagging algorithm with categories based on the physics of the B_{tag} decay", *BABAR* Analysis Document #317.

[74] Gershon, T., Latham, T., "Inclusive branching fraction measurements of SM-suppressed modes $B^+ \rightarrow K^- \pi^+ \pi^+$ and $B^+ \rightarrow K^+ K^+ \pi^-$", *BABAR* Analysis Document #1980.

[75] Hulsbergen, W., Jawahery, H., Kovalskyi, D., Pierini, M., "Measurement of the branching ratio and the time-dependent CP asymmetries for $B^0 \rightarrow K_S^0 \pi^0$ decays on run 1–4 data", *BABAR* Analysis Document #904.

[76] Verkerke, W., Kirkby, D., "RooFit Users Manual", (2006).

[77] Albrecht, H., et al. (ARGUS), *Phys. Lett.*, B241, 278-282 (1990).

[78] Allen, M., Krishnamurthy, M., Roodman, A., Spanier, S., "Measurements of *CP* (t) Asymmetries in $B^0 \rightarrow K_S^0 \pi^0 \pi^0$ Decays", *BABAR* Analysis Document #1182.

[79] Ilic, J., Mohanty, G., Gershon, T., "Search for the decay $B^0 \rightarrow K_S^0 K^\pm \pi^\pm$", *BABAR* Analysis Document #2154.

[80] "Tracking efficiency study in r24", URL `http://www.slac.stanford.edu/BFROOT/www/Physics/TrackEfficTaskForce/TauEff/R24/TauEff.html`.

[81] "Tracking efficiency recipes for run1-7 in r24", URL `http://www.slac.stanford.edu/BFROOT/www/Physics/TrackEfficTaskForce/TrackingTaskForce-2009-R24.html`.

[82] Back, J., "Studies of $B^+ \to h^+h^-h^+$ decays ($h = \pi/K$) on the full Summer 2002 dataset.", *BABAR* Analysis Document #505.

[83] Allen, M., Naisbit, M., Roodman, A., Banerjee, S., "A measurement of π^0 Efficiency Using $\tau \to \rho\nu$ and $\tau \to \pi\nu$", *BABAR* Analysis Document #870.

[84] Aubert, B., et al. (*BABAR*), *Phys. Rev. Lett.*, 99, 021603 (2007), `hep-ex/0703016`.

[85] Aubert, B., et al. (*BABAR*), *Phys. Rev.*, D80, 112001 (2009), `0905.3615`.

[86] Alves, A.A., et al. (LHCb), *JINST*, 3, S08005 (2008).

[87] Bevan, A.J. (2010), `1009.5963`.

[88] Abe, T., et al. (Belle II) (2010), `1011.0352`.

[89] Punzi, G. (2004), `physics/0401045`.

[90] Gershon, T., Latham, T., Puccio, E., "A discussion of the "Punzi effect" in Dalitz plot analyses", *BABAR* Analysis Document #2261.

[91] Aubert, B., et al. (BABAR), *Phys. Rev. Lett.*, 91, 241801 (2003), `hep-ex/0308012`.

[92] Gershon, T., Latham, T., Puccio, E., "Study of Neutral Pion Efficiency Using τ Lepton Decays in the *BABAR* Detector", *BABAR* Analysis Document #1897.

[93] Fox, G.C., Wolfram, S., *Nucl. Phys.*, B149, 413 (1979).