A Study of the Rare Charmless Hadronic B Decay $B^+ \to a_0^+ \pi^0$ using the BABAR detector

William Panduro Vazquez

Imperial College London.

A thesis submitted for the degree of
Doctor of Philosophy
of The University of London
and the Diploma of Imperial College.

February, 2007

Para Pepe y Esther.

Abstract

An upper limit has been extracted, at the 90% C.L, for the rare charmless hadronic B meson decay $B^{\pm} \to a_0(980)^{\pm}\pi^0$, where $a_0(980)^{\pm} \to \eta\pi^{\pm}$ and $\eta \to \gamma\gamma$ or $\eta \to \pi^+\pi^-\pi^0$. The analysis was based on a sample of approximately $340.7\,\mathrm{fb}^{-1}$ of data taken at the $\Upsilon(4S)$ resonance with the *BABAR* detector at the PEP-II e^+e^- collider at SLAC from May 1999 to August 2006. The sample contains (379.9 ± 4.2) million $B\overline{B}$ pairs.

The limit, extracted using an unbinned multivariate extended maximum likelihood fit to the data, is

$$\mathcal{B}(B^{\pm} \to a_0(980)^{\pm}\pi^0) < 2.0 \times 10^{-6} \text{ (90\% C.L.)}.$$

The result does not exclude any of the current theories for the structure of scalar mesons.

Acknowledgements

When I consider the events of the past three years I count myself very lucky. I have had an opportunity to work in a fascinating field alongside some of its best and brightest. There can surely also be no better place to be sent to do your research than the San Francisco Bay Area. From the location itself to the people I had the pleasure of spending my time there with, the whole experience will be hard to beat. For all of these things and more I owe a lot of people my respect and admiration.

First, I owe enormous thanks to my supervisor, Paul Dauncey, for his immense knowledge, his encouragement and for being a true example of how science should be done. I also owe a debt of gratitude to Paul, along with Julia Sedgbeer and Peter Dornan, for the opportunity to study in the Imperial High Energy Physics group. In the same vein PPARC deserve my gratitude for their generous funding over the three years.

I must also thank Rob Flack, for his good cheer, for always being ready to go in 'mob handed' when needed and for enduring, with good grace, many depressing early mornings at the hands of the England rugby team.

To the rest of the IC *BABAR* group: Jamie, Mike, Gary, Geoff, Jordan, Ulrik, Debbie and Mark, thanks for being a top bunch to work with. Special mentions should be given here to Wahid, who put up with being my boss and sharing an office with me at SLAC with characteristic wit, and to Dan Bowerman, for all of his advice and encouragement in the early days.

A big shout-out next to the rest of the SLAC crew: Jim, Ian, Dave P, Carlos, Ed, Ali, Mitch, Matt, Manny, Nick, Alan, Tim, Chukwudi, Adam, Tom, Pablo, Henning, Eric and the many others. Thanks for making sure California never ceased to be a great time. A few people deserve special praise for their part in those 18 months. Kate Schofield, for her friendship, the countless gigs, the introduction to hiking up unfeasibly large rocks and, of course, the cups of tea! Dan Walker, my other office-mate, for patiently listening to my work-related ranting and for being such an all round nice bloke! Rebecca Daly, for introducing me to, among other things, three course fondue and gold panning. Nice. Dave Hopkins and Steve Molloy, for helping me to keep the Bay Area coffee industry going strong. Finally, I must thank the Peninsula Bar and Grill. No more words are required there!

Back home at IC I have been lucky to know a great bunch of people: Catherine, Richard H, Chris, Steve, Stuart and John, thanks for making it such a fun year to be part of! Anne-Marie, Matt Noy, Freya, Richard P, Ella, Seb, Maiko and all, thanks a million.

Away from academia, thanks to the rest of my friends for all the laughs: Matt, Mike, Claire, Martin, Joe, Mo, Alison, Jason, James, Chris and the whole posse, it's been great.

Above all I would like to thank Helen for her love and support from near and far, without which none of this would have been possible. This thesis is dedicated to you.

Declaration

The physics analysis presented in this thesis from chapter three onward was developed primarily by the author with the help and advice of Robert Flack, Paul Dauncey and the *BABAR* Charmless Quasi Two-Body Analysis Working Group. Throughout the document the author refers the reader to the relevant literature whenever work described was carried out by another party.

The description of the *BABAR* detector in chapter 2 represents the author's acquired knowledge over their period working in the operations environment. With the exception of the study presented in section 2.7.6, which was developed by the author with help from Daniel Bowerman and Paul Dauncey, the studies and plots presented in the rest of the chapter are the products of the relevant sub-detector working groups.

Contents

List of Figures

List of Tables

Chapter 1

Theoretical Background

1.1 Preamble

The analysis to be presented in this thesis is intended to shed light on a curiosity of our current physical laws. The so-called 'Standard Model' (SM) of particle physics, one of the most rigorously tested theories in history, predicts a great number of phenomena which have been observed throughout the years in laboratories the world over. However, there are a significant number of phenomena which it permits yet which have never been seen. The search for the signatures of these is as significant, if not more so, than the search for phenomena which actually violate the Standard Model.

Perhaps the most well known area of the Standard Model which has yet to be verified experimentally is the Higgs mechanism [1], through which particles attain the property we interpret as mass. A great deal of effort is currently under way to establish the true physical nature of this mechanism, or indeed whether it exists at all. Failure in this attempt will cast significant doubt over current formulations of the Standard Model.

While not so Earth shattering, the phenomenon under consideration in this thesis is a similarly curious anomaly in the otherwise impressive catalogue of experimental verifications of the Standard Model. We shall search for a potential departure from the naive quark model of hadrons. This holds that the family of particles known as mesons are made up of only two valence quarks (actually one quark and one antiquark or some superposition of such pairs). According to the Standard Model, departures from this structure are perfectly permissible, yet little evidence for such states has ever been found. The a_0 scalar meson, upon which this thesis is based, is but one candidate in a list of particles which may display some kind of novel structure. Verifying the true nature of these particles will provide an important test of the predictive potential of our established physical laws.

This chapter will begin by briefly reviewing the Standard Model of particle physics with specific reference to the phenomenon of CP violation, which is what the *BABAR* experiment was built to find in the decays of the B meson. The bulk of the chapter will concentrate on the mode under analysis in this thesis, $B^+ \to a_0^+ \pi^0$ (the charge conjugate mode is implicitly included throughout). This will begin with a review of the scalar meson sector of the Standard Model. The presence of scalar mesons in charmless decays of the B meson will be discussed, as well as how looking for these processes can help improve our knowledge of scalar meson structure.

1.2 The Standard Model of Particle Physics

The Standard Model describes the properties and interactions of particles governed by three of the four known forces of nature: the strong nuclear, weak nuclear and electromagnetic forces. The fourth force, gravity, has yet to be successfully incorporated into the model, but is of negligible effect on the extremely small distance scales under consideration in particle physics.

The Standard Model is based on relativistic quantum field theory in the Lagrangian formalism. Its development was based on the search for symmetries in nature. The importance of symmetry considerations in field theory arose originally from the work of Nöther [2], who theorised that with every symmetry there is an associated conservation law. Therefore, by imposing the symmetry on a field, the underlying law can be discovered.

Each force and the particles which feel it are represented by a component in the Lagrangian, as shown in equation 1.1. The strong force is regulated by Quantum Chromodynamics (QCD) [3], based on the $SU(3)$ symmetry group. The electromagnetic and weak forces are regulated by the Glashow-Weinberg-Salam (GWS) model [4] [5] [6], an attempt to unify the forces based on the combined $SU(2)_L \times U(1)_Y$ symmetry groups. The subscript L in the equation represents the fact that the weak force couples to left handed fermions and Y denotes weak hypercharge. The strong force couples to quark *colour*, denoted by the subscript C. In equation 1.1, the Lagrangian for the Higgs field [7] is assumed to be incorporated into that for the GWS model. The component representing the known fermions, $\mathcal{L}_{fermions}$ is composed of a sum of independent Dirac field Lagrangians, as will be described in the next section.

$$\mathcal{L}_{SM} = \mathcal{L}_{fermions} + \mathcal{L}_{SU(3)_C} + \mathcal{L}_{SU(2)_L \times U(1)_Y}. \tag{1.1}$$

1.2.1 Building the SM using Gauge Symmetry

The symmetry upon which the construction of the SM is based is that which occurs in so-called 'gauge' transformations, which we shall discuss using the example $U(1)$ symmetry. For a field ϕ there are two such transformations. Firstly there is a 'global' transformation $\phi \rightarrow e^{i\alpha}\phi$, applied simultaneously to all points in space-time for a continuous real variable α. The second type of transformation is the 'local' case, where α is dependent on the position in space-time, x_μ. Thus the transformation is described by $\phi \rightarrow e^{i\alpha(x_\mu)}\phi$.

In the case of the SM it is the symmetry requirement that the Lagrangian be invariant under such local gauge transformations which yields significant results. Consider the example of Quantum Electrodynamics (QED), which is based on a free spinor (spin-$\frac{1}{2}$) field ψ of mass m. The dynamics of this field are encoded in the Lagrangian derived from the Dirac field equation,

$$\mathcal{L} = i\overline{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\overline{\psi}\psi. \tag{1.2}$$

This Lagrangian is trivially invariant under a global gauge transformation. However, under a local transformation it becomes

$$\mathcal{L} = i\overline{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\overline{\psi}\psi - \overline{\psi}\gamma^{\mu}\psi\partial_{\mu}\alpha, \tag{1.3}$$

where the extra term in $\partial_{\mu}\alpha$ violates invariance. In order to restore the symmetry the derivatives in the Lagrangian are replaced with a so-called 'covariant' derivative, D_{μ}, chosen such that invariance is satisfied.

$$D_{\mu} = \partial_{\mu} - iqA_{\mu}, \tag{1.4}$$

where A_{μ} is a new vector 'gauge' field which is required to transform such that the offending $\partial_{\mu}\alpha$ term is cancelled:

$$A_{\mu} \rightarrow A_{\mu} + \frac{1}{q}\partial_{\mu}\alpha. \tag{1.5}$$

The introduction of the covariant derivative introduces a term into the Lagrangian which describes the interaction between the spinor and A_{μ} fields, the latter of which we will interpret as the photon field. The modified Lagrangian is therefore

$$\mathcal{L} = \overline{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi + q\overline{\psi}\gamma^{\mu}A_{\mu}\psi. \tag{1.6}$$

Finally, in order to interpret A_{μ} as the photon field an extra term must be introduced to account for the freely propagating photon. We can achieve this by considering the Lagrangian for a free vector (spin-1) field,

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}m_A^2 A^\nu A_\nu, \tag{1.7}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and m_A is the mass of the gauge boson. The second term in this Lagrangian is not invariant under a local gauge transformation, thus we require the boson to be massless for this to work. This is physically consistent since the gauge boson for the field we are considering, the photon, is indeed massless. From this we can finally construct the complete Lagrangian for QED,

$$\mathcal{L}_{QED} = \overline{\psi}(i\gamma^\mu\partial_\mu - m)\psi + q\overline{\psi}\gamma^\mu A_\mu\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{1.8}$$

Thus it can be seen that by invoking local gauge invariance the critical interaction term behind QED can be derived. Following this procedure for the relevant symmetry groups is fundamental to the construction of the Standard Model Lagrangian.

1.2.2 GWS Theory and the Higgs Mechanism

As mentioned previously, the GWS model describes the interactions mediated by the electromagnetic and weak nuclear forces. The model is constructed once again from Lagrangians describing fermionic fields (quarks and leptons), but this time invoking the symmetries of both the $U(1)_Y$ and $SU(2)_L$ groups. $U(1)_Y$ corresponds to both right and left handed quark and lepton flavour states of weak hypercharge. These exist for the three generations in the SM. This is supplemented with $SU(2)_L$ to describe the associated left handed quark and lepton doublets of weak isospin (see Table 1.1). Note that the neutrino is only included in the left handed doublet. There are no right handed neutrinos in the SM.

These symmetry considerations result in two covariant derivatives, in which are encoded the four gauge bosons of electroweak theory. For the doublets we have,

$$D^\mu_{dbl} = \partial^\mu - i\frac{g}{2}\sigma \cdot \mathbf{W}^\mu - \frac{g'}{2}YB^\mu \tag{1.9}$$

Table *1.1:* *The three fermion generations of the Standard Model and their associated electroweak quantum numbers. Note that a ′ superscript denotes a weak eigenstate rather than one of mass. I^W represents weak isospin (which are the eigenvalues of the SU(2) generators $\boldsymbol{\sigma}$). Y represents weak hypercharge and Q the electric charge.*

Fermion Generation			Quantum Numbers			
1	2	3	I^W	I_3^W	Y	Q
$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$	$+1/2$	$+1/2$ $-1/2$	-1	0 -1
e_R	μ_R	τ_R	0	0	-2	-1
$\begin{pmatrix} u \\ d' \end{pmatrix}_L$	$\begin{pmatrix} c \\ s' \end{pmatrix}_L$	$\begin{pmatrix} t \\ b' \end{pmatrix}_L$	$+1/2$	$+1/2$ $-1/2$	$\mp 1/3$	$+2/3$ $-1/3$
u_R	c_R	t_R	0	0	$+4/3$	$+2/3$
d'_R	s'_R	b'_R	0	0	$-2/3$	$-1/3$

and for the singlets,

$$D^\mu_{sgl} = \partial^\mu - \frac{g'}{2} Y B^\mu. \tag{1.10}$$

The SU(2)$_L$ group provides the three W^μ fields, two charged and one neutral. A further neutral field, B^μ, is derived from the U(1) group. It is these fields from which it will eventually be possible to derive the charged W bosons, the Z^0 boson and the photon. The fields couple to fermions with strengths g and g' and according to quantum numbers derived from weak hypercharge Y and the eigenvalues I^W of the SU(2) generators $\boldsymbol{\sigma}$, which correspond to weak isospin. The electric charge can then be expressed as $Q = I_3^W + Y/2$.

So far so good; however, there is a problem. In order to maintain gauge symmetry all four of the gauge bosons derived above must be massless, in the same way as was required when looking at the Lagrangian for QED earlier. Now, the photon is indeed massless, but the $W^\pm$ and Z^0 bosons are not [8]. In order to circumvent this problem the SM uses a form of spontaneous symmetry breaking (SSB) via what is known as the Higgs mechanism. SSB occurs when a system moves from a symmetric but unstable vacuum state into one of a number of degenerate states which are stable but, in doing so, breaks the symmetry. This is achieved in the Higgs mechanism by introducing a complex scalar field doublet,

$$\phi(x) = \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix}. \tag{1.11}$$

This doublet is assigned weak hypercharge $Y = 1$, thus requiring that the component fields have charges $+1$ and 0 as denoted by their superscripts.

The complex field is then required to obey a potential $V(\phi)$ with a shape most closely analogous to the bottom of a wine bottle.

$$V(\phi) = \mu^2\phi^\dagger(x)\phi(x) + \lambda\left[\phi^\dagger(x)\phi(x)\right]^2, \tag{1.12}$$

where the desired shape is achieved by requiring that μ^2 be negative and λ positive. The Lagrangian component for this field is built around the same symmetry considerations used for the left handed quark flavour doublets, coupling via the terms set out in the associated covariant derivative,

$$\mathcal{L}_{Higgs} = \left|(\partial^\mu - i\frac{g}{2}\sigma\cdot\mathbf{W}^\mu - \frac{g'}{2}YB^\mu)\phi\right|^2 - V(\phi). \tag{1.13}$$

The 'wine bottle' shape has an unstable minimum at $\phi = 0$ and real valued degenerate minima also exist at $\phi = \nu/\sqrt{2}$, for all values of the phase of ϕ, with $\nu = \sqrt{-\mu^2/\lambda}$. For our purposes we arbitrarily require these minima to correspond to the vacuum expectation values for the two components of the field as follows:

$$\langle 0|\phi^+(x)|0\rangle = 0,\ \langle 0|\phi^0(x)|0\rangle = \nu/\sqrt{2}. \tag{1.14}$$

Now, due to the perturbative nature of quantum field theory, any field may be considered as small fluctuations about the vacuum state. We enact this principle here by describing our complex scalar field in terms of four field perturbations about the vacuum.

$$\begin{aligned}\phi^{+}(x) &= \phi_1^{+}(x) + i\phi_2^{+}(x),\\ \phi^{0}(x) &= \nu/\sqrt{2} + \phi_1^{0}(x) + i\phi_2^{0}(x).\end{aligned}$$

When one propagates this parameterisation of the field through the Higgs Lagrangian the components $\phi_1^{+}(x)$, $\phi_2^{+}(x)$ and $\phi_2^{0}(x)$ are found to be represented by Klein-Gordon type kinetic terms with no associated mass. These massless particles are known as *Goldstone* Bosons. The component $\phi_1^{0}(x)$ performs differently, yielding a Klein-Gordon term corresponding to a massive scalar field and a massive boson ($m_H = \sqrt{2\lambda\nu^2}$). These are the Higgs boson and associated field. Terms also appear which finally allow us to describe the four electroweak gauge bosons. The charged W^{μ} fields can be used in a linear superposition to describe the charged W bosons. The neutral W_3^{μ} component mixes with the neutral B^{μ} field to describe the Z^0 boson and the photon. The three weak bosons now have associated mass terms corresponding to $M_W = \frac{g\nu}{2}$ and $M_Z = \frac{\nu}{2}\sqrt{g^2 + g'^2}$. The photon remains massless as required.

Finally, by adopting a suitable gauge the massless Goldstone bosons can be transformed away, imbuing each of the three massive electroweak gauge bosons with a longitudinal polarisation state. At the end of this process our scalar field can be represented as

$$\phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \nu + h(x) \end{pmatrix}, \tag{1.15}$$

where we now represent ϕ_1^0 as the Higgs scalar $h(x)$.

1.2.3 The Higgs Model as the Progenitor of Quark Mass and Weak Mixing

The Higgs mechanism can be used to impart mass to the quarks and leptons in the Standard Model. Once again, the original theory denies these particles mass in order to preserve local gauge invariance. The Higgs field is employed to this end through the Lagrangian term known as the Yukawa coupling (shown here for quarks only),

$$\mathcal{L}_{Yukawa} = \sum_{i,j=1}^{3} (\Gamma^u_{ij}\overline{\chi_i}_L \phi u_{jR} + \Gamma^d_{ij}\overline{\chi_i}_L \phi d_{jR} + h.c.), \tag{1.16}$$

where the Γ matrices are arbitrary couplings, χ_L refer to the left handed quark flavour doublets and u_R and d_R refer to the up-type and down-type right handed quark flavour singlets respectively. The Hermitian conjugate, denoted by '*h.c.*', is implicitly included. By substituting in our scalar field, ϕ, we recover terms of the form

$$\Gamma^u_{ij}\left(\frac{\nu}{\sqrt{2}}\overline{u_i}_L d_{jR}\right) = M^u_{ij}\overline{u_i}_L d_{jR}, \tag{1.17}$$

where M^u_{ij} is the fermionic mass matrix. A similar procedure will yield M^d_{ij}. These can be diagonalised to yield real observables via the introduction of the unitary matrices V^u_L, V^u_R, V^d_L and V^d_R. Thus we have

$$V^u_L M^u V^u_R = M^u_{diag};\ V^d_L M^d V^d_R = M^d_{diag}, \tag{1.18}$$

with the mass eigenstates of the quarks given by

$$u^m_{L(R)} = V^m_{L(R)} u_{L(R)};\ d^m_{L(R)} = V^m_{L(R)} d_{L(R)}. \tag{1.19}$$

Now that the quarks have mass via symmetry breaking, it is possible to re-express the flavour component of the SM Lagrangian in terms of the mass eigenstates. For neutral currents this has no effect but for charged currents $\mathcal{L}_{cc}$ we have (for the example of the positively charged case)

$$\mathcal{L}_{cc+} = -\frac{g}{\sqrt{2}}\overline{u_i^m}_L V^u_{ikL}\gamma^\mu W^+_\mu V^{d\dagger}_{kjL} d^m_{jL}, \tag{1.20}$$

where the matrix product $V^u_{ikL}V^{d\dagger}_{kjL} = V_{CKM}$, the Cabibbo-Kobayashi-Maskawa quark mixing matrix [9] [10]. Thus we see that up-type quarks couple to a linear superposition of down-type quarks (or vice-versa)

$$\mathcal{L}_{CC+} = \begin{pmatrix} \overline{u} & \overline{c} & \overline{t} \end{pmatrix} \frac{-g}{\sqrt{2}}\gamma^\mu W^+_\mu V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{1.21}$$

where the CKM matrix V_{CKM} can be written as

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{1.22}$$

Since the CKM matrix is a 3×3 unitary matrix it has 18 real values with 9 constraints, it can therefore be expressed using 9 parameters. However, the 5 relative phases of the 6 quark wavefunctions are unobservable, so there are only 4 physically meaningful parameters. With 3 dimensions, one of these must be a phase. The CKM matrix can be expressed in different ways but a commonly used parameterisation is that developed by Wolfenstein [11]

$$V_{CKM} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{1.23}$$

In this formulation the real parameters A, ρ and η are of order unity and $\lambda\ \sim 0.22$. It is the complex component represented by η being non-zero which makes the phenomenon of CP violation possible in the weak decays of some mesons [12]. This will be discussed in more detail in the next section. The aim of the *BABAR* experiment is to constrain the elements of this matrix and thus estimate the amount of CP violation present in the Standard Model. The experiment focuses on finding evidence of CP violation in the decays and mixing of charged and neutral B mesons. Through the work of *BABAR* and other experiments the elements of the CKM matrix are being measured with increasing precision. The current world average values of the magnitudes of the CKM matrix elements are shown below to their 90% confidence level bounds [8].

$$V_{CKM} = \begin{pmatrix} 0.9739 \text{ to } 0.9751 & 0.221 \text{ to } 0.227 & 0.0029 \text{ to } 0.0045 \\ 0.221 \text{ to } 0.227 & 0.9730 \text{ to } 0.9744 & 0.039 \text{ to } 0.044 \\ 0.0048 \text{ to } 0.014 & 0.037 \text{ to } 0.043 & 0.9990 \text{ to } 0.9992 \end{pmatrix}. \tag{1.24}$$

1.3 CP Violation

There exist three discrete symmetries which can be associated with any Lorentz invariant local field theory in the Lagrangian formalism. The charge conjugation operator 'C' converts a particle into its antiparticle. The parity operator 'P' is a space-time operation corresponding to $P : \vec{x} \to -\vec{x}$ and $P : t \to t$. Finally the time reversal operator 'T' corresponds to $t \to -t$. The combined operation CPT can be proven to be a symmetry of all such fields [13]. However, violations of the individual C and P operations have been observed [14]. Violation of the combined CP operation, first observed in the neutral kaon system in 1964 by Cronin and Fitch [15], is one of the processes postulated to explain the observed matter-antimatter asymmetry in the universe [16]. The other contributing processes are C violation, baryon number violation and a departure from thermal equilibrium in the early universe.

CP violation can arise in three different ways [12]. 'Direct' CP violation occurs when the amplitude for a given decay and for its CP conjugate are not equal. 'Indirect' CP violation occurs in the mixing of flavour eigenstates in some neutral particles. The third kind of CP violation occurs in the quantum mechanical interference of the first two processes. Direct CP violation is theoretically possible in the mode under study in this analysis, although the statistical power available renders any significant sensitivity to it extremely unlikely.

1.4 Measuring CP Violation B Meson Decays

The potential for experimental investigation of CP violation in the B meson system arises from the unitarity of the CKM matrix. By requiring unitarity it is possible to derive nine relations, six of which can be represented as triangles in the complex plane. The relation of relevance to B mesons is given by

$$V_{ud}V_{ub}^* = V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \tag{1.25}$$

This relation can be expressed as a triangle by defining the following three angles:

$$\alpha = \left[-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right],\ \beta = \left[-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right],\ \gamma = \left[-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right]. \tag{1.26}$$

Due to the sizes of the CKM matrix elements involved the angles of the triangle are expected to be close to unity and hence the B meson system is expected to exhibit large amounts of CP violation. The triangle is shown pictorially in Figure 1.1.

The three angles can be measured independently and thus we can 'over-constrain' the unitarity properties of the triangle. The primary physics goal of the *BABAR* experiment is to measure angle β. This can be done by measuring the quantity sin2β through analysis of the decays of neutral B mesons to final states involving charmonium (a $c\bar{c}$ bound state) and a neutral kaon. These modes provide a clean

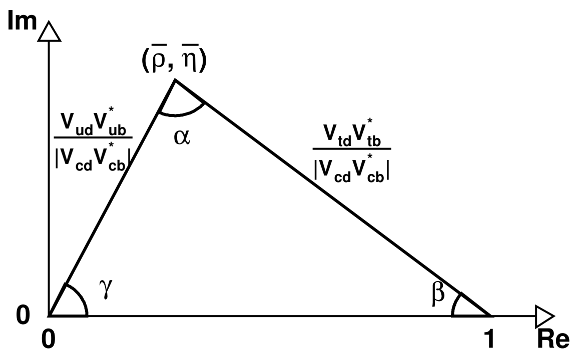

Figure *1.1:* *The unitarity triangle for the B meson system.*

experimental signature with low theoretical uncertainties due to the small contribution from higher order 'penguin' processes, which will be discussed in more detail later. These modes provide an example of the third kind of CP violation as described above, i.e. that occurring in the interference between mixing and decay processes. For these decays B^0 and $\overline{B}^0$ mesons can mix and yet decay to the same final state. Since the decay amplitudes provide no phase contribution the relative phase of $e^{-2i\beta}$, caused by the presence of CKM element V_{td} in the mixing process, is experimentally measurable. In practice the quantity measured is

$$a(t) \simeq f \sin 2\beta \sin(\Delta m \Delta t), \tag{1.27}$$

where the analysis seeks to quantify this as a function of the the lifetime difference between the B^0 and $\overline{B}^0$ mesons Δt. In this equation Δm, the frequency of the oscillation, corresponds to the mass difference between the heavy and light mass eigenstates of the B meson. The parameter f is the CP eigenstate of the final state which is $+1$ for $J/\psi K_L^0$ case and -1 for the $J/\psi K_S^0$ case. At *BABAR* this quantity is measured using a 'time-dependent' analysis, facilitated by the fact that the $B^0\overline{B}^0$ pair produced by the decay of the $\Upsilon(4S)$ is bound by angular momentum conservation to be odd under interchange. Hence neither the B^0 nor the $\overline{B}^0$ can oscillate until the other meson in the pair has decayed. *BABAR* takes advantage of this by identifying it a B^0 or $\overline{B}^0$ from the charge of certain decay products. By doing this it is then possible to 'tag' the flavour of the second meson as being the opposite to whatever the first is identified as being. The highest rate for this procedure comes from the identification of high momentum charged leptons, although the cascade decay $b \to c \to s$ is also used. In this second case the charge of final state kaons is used to identify the original meson. Finally, precise charged track information is used to reconstruct the vertex (origin) of the tracks and hence the position of the B meson.

After the decay of the first B meson, the second in the pair may then oscillate before decaying into the aforementioned $J/\psi K^0_{S/L}$ final state, which is the same irrespective of the meson flavour. This decay tree is fully reconstructed and precise vertex information also extracted. With vertex information for both B mesons it is then possible to measure the distance between the positions at which the two decayed in the z axis, labelled Δz. This is made easier by the Lorentz boost applied to the system in the laborarory frame. From the value of Δz it is then possible to calculate the difference Δt in the lifetimes of the two mesons, $\Delta t \simeq \Delta z/\beta\gamma c$.

By comparing the value of Δt obtained in the case where the second meson is tagged as a B^0 alongside the value in the $\overline{B}^0$ case it is possible to calculate and plot the difference in the relative numbers of both mesons for a given Δt value. This results in a sinusoidal distribution corresponding to $a(t)$ above, a fit to which yields the values of Δm and $\sin 2\beta$. Analyses such as the one described here have been performed by both *BABAR* and Belle collaborations and the current world average value for $\sin 2\beta$ is 0.687 ± 0.032 [8].

1.5 Scalar Mesons and Analysis Motivation

This thesis deals with a B decay involving the $a_0(980)$ meson, hereafter referred to as the a_0 unless otherwise stated. The a_0 forms part of a family of particles called 'scalar' mesons. Such particles have zero total angular momentum, $J = L + S$, a vector combination of spin ($S = 1$) and orbital components ($L = 1$). They also have 'even' parity, given by $P = -(-1)^L = 1$, therefore $P|S\rangle = +|S\rangle$ for a scalar meson state S. This is as opposed to a 'pseudoscalar' meson, such as a pion, which has zero angular momentum but 'odd' parity, i.e. $P|PS\rangle = -|PS\rangle$ for a pseudoscalar state PS. The state also has even charge-conjugation properties, given by $C = -(-1)^{L+S} = 1$, therefore in terms of J^{PC} the a_0 is commonly represented by 0^{++} [7].

Scalar mesons are interesting as there is evidence [17] that they could demonstrate structure outside the traditional two-quark model for mesons. However, scalar mesons are currently poorly understood with few experiments yielding significant insight about their structure. Scalar mesons within the two-quark model are theorised to form an isospin nonet as shown in Figure 1.2, this is due to the nine possible $q\overline{q}$ combinations possible with the three quark flavours, $q = u, d, s$ available in the low mass regime. The nature, or in some cases mere existence, of some of these particles, such as the κ or σ, is somewhat controversial [8]. Other members of the nonet, such as the f_0, have received more experimental coverage in recent times [18].

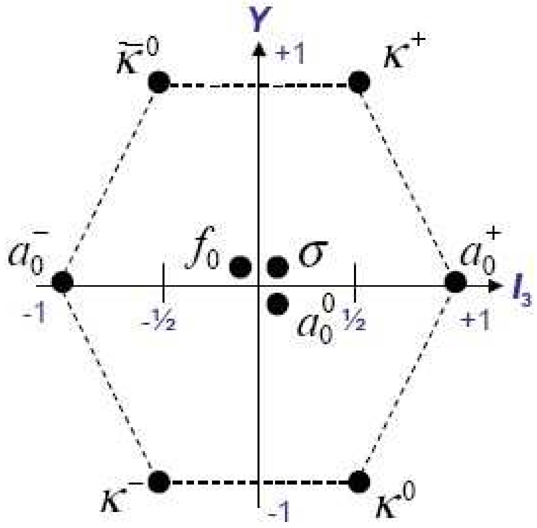

Figure *1.2:* *The proposed two-quark model scalar meson nonet.*

There are a number of other potential structure models for scalar mesons. The predominant non-standard structure theory for the a_0 is that it may contain a large four-quark component. There are a number of models for such a structure, as discussed in [18]. However, the analysis in this thesis is motivated by predictions based on the model eventually adopted in that paper, where the original $u\overline{d}$ pair is joined by an $s\overline{s}$ pair within the bound state [19]. In the two-quark case we have:

$$a_0^+ = u\overline{d},\ a_0^- = d\overline{u},\ a_0^0 = \frac{1}{\sqrt{2}}(u\overline{u} - d\overline{d}), \tag{1.28}$$

whereas for the four-quark case this becomes:

$$a_0^+ = us\overline{d}\overline{s},\ a_0^- = ds\overline{u}\overline{s},\ a_0^0 = \frac{1}{\sqrt{2}}(us\overline{u}\overline{s} - ds\overline{d}\overline{s}). \tag{1.29}$$

Other theories for a_0 structure [8] suggest a possible component of $K\overline{K}$ molecular admixture, or that there could be a substantial 'glueball' component to the state - i.e. one with no valence quarks. As yet there has been little experimental evidence to allow any of the available structure theories to be rigorously scrutinised.

Recent theoretical developments [18] suggest that the mode $B^{\pm} \rightarrow a_0^{\pm}\pi^0$ may be a good candidate for investigation. This is because the predicted (albeit model dependent) branching fractions for the two- and four-quark models differ by an order of magnitude (as will be discussed later). A measurement of this branching fraction would therefore provide an effective experimental test of this theory. Unfortunately the statistics available at *BABAR* are as yet insufficient to probe down to the level of such a rare decay. As such this analysis provides verification of the theory only in the sense that it excludes branching fractions significantly above those which it predicts.

1.6 Production of the a_0^+ Scalar Through Decays of the B Meson

When considering the decays of the B meson one must consider contributions not only from the standard electroweak tree diagrams but also from gluonic and electroweak penguin processes. The term 'penguin' is used to describe a number of second order loop processes facilitating flavour changing neutral currents. The potential diagrams for B meson decays are shown for the example of $B \to K\pi$ in Figure 1.3.

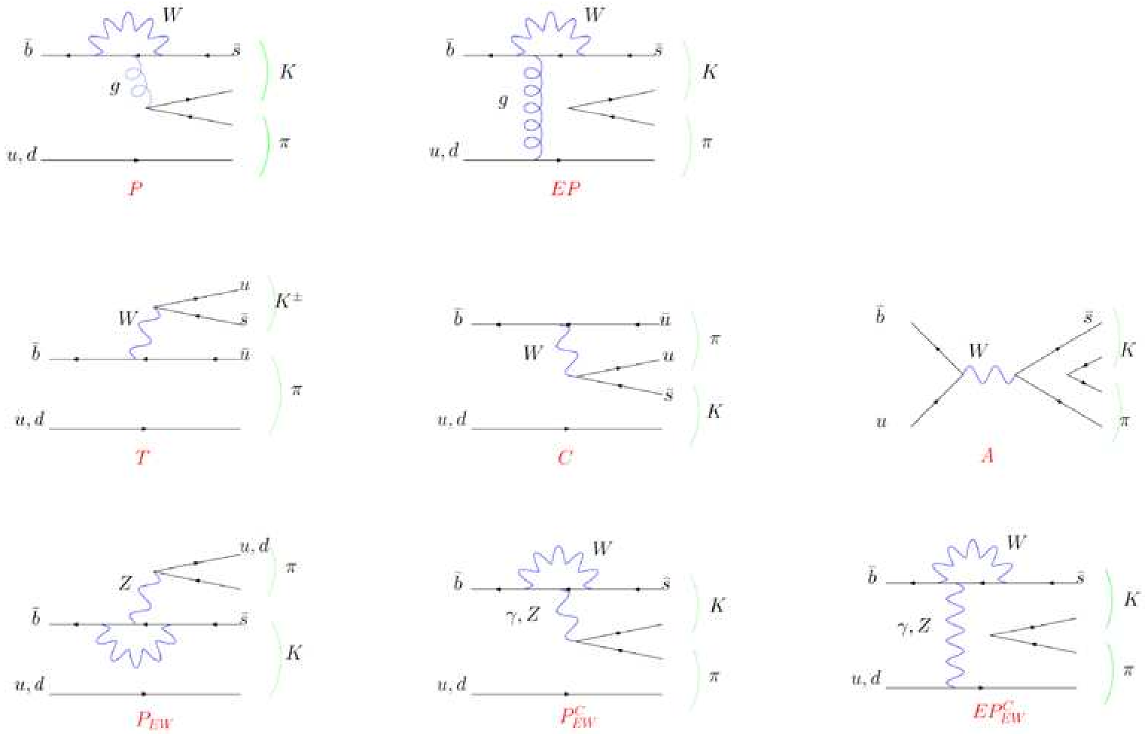

Figure *1.3:* *The gluonic and electroweak processes contributing to charmless B decays (examples are for the mode $B \to K\pi$) [20]. Diagram label T denotes the colour-allowed tree while C indicates the colour-suppressed version. The electroweak annihilation diagram is labelled A while the P is the gluonic penguin and EP is the gluonic penguin exchange. The electroweak penguins are denoted by P_{EW} for the colour-allowed case and P_{EW}^C for the colour-suppressed case with the exchange penguin labelled EP_{EW}^C.*

The final state required for the decay mode under study in this thesis requires a transition from a b quark in the initial state to either a u quark via the colour-allowed/suppressed tree diagrams (T and C in the diagram respectively) or to a d quark in the final state through any of the penguin processes. The concept of colour-suppression arises from the fact that a quark can have one of three possible colour charges. However, any bound state must be colour neutral. In the colour-allowed case this is not a problem since the a_0^+ production vertex is not internal to the original $q\overline{q}$ pair. However, in the colour-suppressed case the quarks produced from the charged boson must have opposite colour charge to those already present in order to form the required colour neutral bound state. Due to this the rate for the process is suppressed. The final diagram which potentially contributes significantly to the decay amplitude is the electroweak annihilation diagram, denoted by A above.

The penguin processes for this mode are heavily CKM suppressed since the requirement of a $b \to d$ transition means they are dominated by element V_{td} in the outgoing loop vertex. Of the penguin processes the QCD diagram P is expected to dominate over the electroweak and exchange processes.

In addition to the penguin diagrams we have the potential electroweak tree and annihilation diagrams. The annihilation process is suppressed due to helicity in the same way that the muonic channel is favoured over the electron channel in the decay of a charged pion. The weak force couples to left handed particles and right handed antiparticles. This decay would produce such a pair. However, the B meson is a spin zero particle, thus requiring the decay products of the W^+ involved in the decay to have $\pm\frac{1}{2}$ spin components along the axis of their momentum. In the rest frame of the B their momentum vectors are anti-parallel and thus they must have the same helicity. This is in conflict with the requirements of the weak force and the diagram is therefore heavily suppressed.

Due to the large suppression applied to the other contributing diagrams one would naively expect the dominant contributions to come from the colour-suppressed and colour-allowed tree diagrams. At the two-quark level the diagrams for these decays are shown explicitly in Figures 1.4a and 1.4b.

These tree processes are an example of what is known as a 'charmless' B decay. This is because they involve a $b \to u$ quark transition, as opposed to $b \to c$. This means that they are suppressed in the CKM matrix because they involve element V_{ub} rather than V_{cb}. The decays are therefore rather rare. The situation is further complicated by the fact that the supposedly dominant colour-allowed diagram is actually doubly suppressed by G-parity and vector current conservation concerns [21]. This being the case, the colour-suppressed diagram dominates and the decay is expected to be

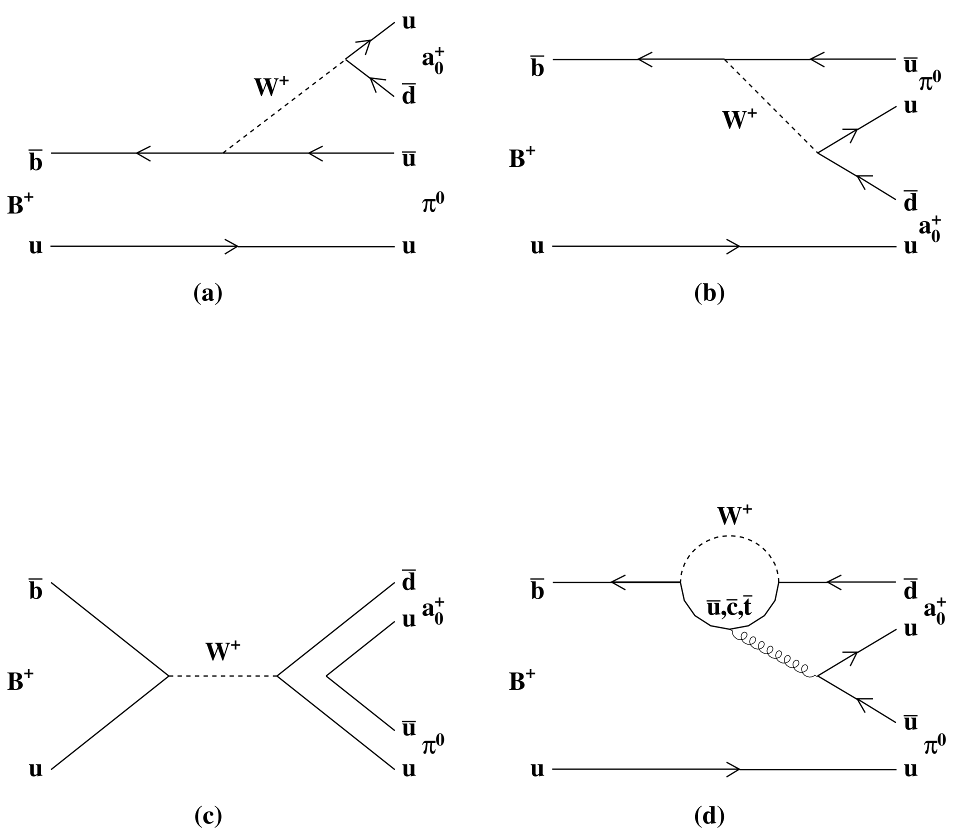

Figure *1.4:* *The Feynman diagrams contributing to the process $B^+ \to a_0^+\pi^0$. (a) is the external (colour-allowed) tree, (b) the internal (colour-suppressed) tree, (c) the annihilation process and (d) the gluonic penguin process.*

extremely rare, even for a charmless decay. Given the level of suppression, contributions from the annihilation (Figure 1.4c) and dominant gluonic penguin process (Figure 1.4d) may not be negligible [18]. In order to obtain a prediction for the branching fraction, with all of these elements in mind, the method of QCD 'factorisation' is used.

1.6.1 Signal Mode Predictions from QCD Factorisation

Using QCD factorisation [22] an incalculable matrix element can be expressed as the product of two effective matrix elements as per equation 1.30. This can be essentially considered as splitting the process in two with the initial state particle decaying to one of its products and with the other generated from the vacuum.

$$\langle u_1 u_2|\mathcal{O}|B\rangle = \langle u_1|\mathcal{F}_1|B\rangle\langle u_2|\mathcal{F}_2|0\rangle, \tag{1.30}$$

where u_1 and u_2 correspond to the final state hadrons and $\mathcal{F}_{1,2}$ are computed from known quantities such as form factors or branching ratios. For example, in the case of Figure 1.4(a), u_1 would be the π^0 meson and u_2 would be the a_0^+ meson. Non-perturbative effects are assumed to be negligible due to the large mass of the b quark and small mass of the a_0^+.

Since there are a number of potential contributions to the decay amplitude it is actually computed as a combination of matrix element products with the corresponding CKM matrix elements, masses and effective Wilson coefficients [23] included in the calculation. The matrix products can themselves be expressed in terms of masses, form factors (F) and decay constants (f) [7]. In the two-quark model the contributing matrix products $\mathbf{X^a_{bc}}$ for this mode are as follows,

$$\begin{aligned}
\mathbf{X}^{a_0^+}_{B^+\pi^0} &= \langle a_0^+|(\overline{d}u)_L|0\rangle\langle\pi^0|(\overline{u}b)_L|B^+\rangle \sim f_a(m_B^2 - m_\pi^2)F_0^{B^+\pi^0}(m_a^2), && (1.31)\\
\mathbf{X}^{\pi^0_u}_{B^+a_0^+} &= \langle\pi^0|(\overline{u}u)_L|0\rangle\langle a_0^+|(\overline{d}b)_L|B^+\rangle \sim \frac{f_\pi}{\sqrt{2}}(m_B^2 - m_a^2)F_0^{B^+\pi^-}(m_\pi^2), && (1.32)\\
\mathbf{X}^{B^+}_{a_0^+\pi^0} &= \langle a_0^+\pi^0|(\overline{d}u)_L|0\rangle\langle 0|(\overline{u}b)_L|B^+\rangle \sim f_B(m_a^2 - m_\pi^2)F_0^{a_0^+\pi^0}(m_B^2), && (1.33)\\
\tilde{\mathbf{X}}^{a_0^+}_{B^+\pi^0} &= \langle a_0^+|\overline{d}u|0\rangle\langle\pi^0|\overline{u}b|B^+\rangle \sim m_a\tilde{f}_{a_0^+}\frac{m_B^2 - m_\pi^2}{m_b - m_d}F_0^{B^+\pi^0}(m_a^2), && (1.34)
\end{aligned}$$

where equation 1.31 corresponds to the colour-allowed tree, equation 1.32 to the colour-suppressed tree, equation 1.33 to the electroweak annihilation diagram and equation 1.34 to the gluonic penguin diagram, as shown in Figure 1.4. The masses of the participating particles are given by: m_B for the B meson, m_π for the pion, m_a for the a_0^+ meson, m_b for the b quark and m_d for the d quark.

As can be seen, the amplitude is highly dependent on the various form factors. According to perturbative QCD [24] the form factor should be proportional to $1/q^{2(n-1)}$ for the number of valence quarks n in the relevant hadron and momentum transfer q. This is particularly significant for the annihilation term, which is dependent on the form factor $F_0^{a_0^+\pi^0}$. In the four quark model this form factor is heavily suppressed and the annihilation contribution effectively disappears. Hence the four-quark branching fraction is smaller than the two-quark alternative. The predicted branching fractions for the two modes are [18]:

$$6.4 \times 10^{-8} \leq \mathcal{B}(B^+ \to a_0^+\pi^0) \leq 2.4 \times 10^{-7} \tag{1.35}$$

for the two-quark model and

$$2 \times 10^{-9} \leq \mathcal{B}(B^+ \to a_0^+\pi^0) \leq 10^{-8} \tag{1.36}$$

for the four-quark case.

1.6.2 Scalar Mesons in $\eta\pi^\pm\pi^0$ Charmless 3-body Decays

In order to detect decays involving a_0 mesons there are a number of final state options available [8]. By far the dominant decay mode of the a_0 itself is $a_0 \to \eta\pi$, which occurs $\simeq 85\%$ of the time. The rest is largely made up of $a_0 \to K\overline{K}$. Using the dominant mode it is possible to consider the a_0 as a resonance within the final state of the three-body decay $B^+ \to \eta\pi^+\pi^0$. Two a_0 resonances exist in this final state, an a_0^0 from the combination of $\eta\pi^0$ and the a_0^+ from the combination of $\eta\pi^+$. As mentioned before this analysis is centred around the latter of the two.

A useful way to describe three-body decays is using a Dalitz phase space plot [25] (DP). This exploits the fact that the partial width for a three-body decay with transition amplitude $\mathcal{A}$ can be shown to be given by

$$d\Gamma = \frac{|\mathcal{A}|^2}{256\pi^2 M^3} dm_{12}^2 dm_{13}^2, \tag{1.37}$$

where the subscript indices 1,2,3 indicate the decay daughter in question, M is the mass of the parent and $m_{ab}^2 = (p_a + p_b)^2$ for the daughter combinations. The combinations form the axes of the plot, shown in figure 1.5. The critical point here is that there is no dependence on phase space in this formulation and if there were no dependence on the amplitude $\mathcal{A}$, it would give a uniform distribution within kinematic limits. Therefore any structure in the plot comes directly from the matrix element.

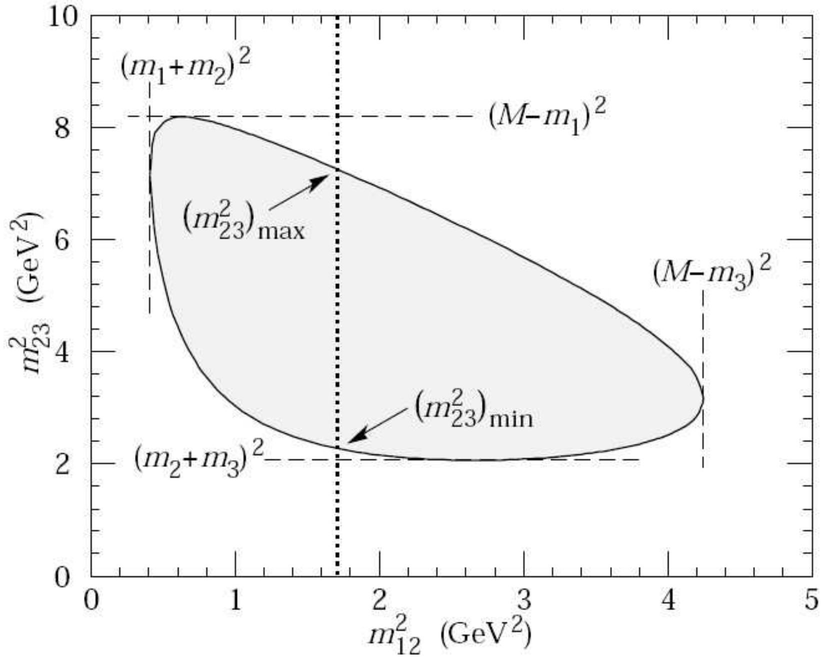

Figure *1.5:* *Pictorial representation of a Dalitz plot in which the shaded area represents the kinematically allowed region. This figure is taken from [8].*

If the amplitude is not dependent on either m_{ab}^2 the plot is uniformly populated. However, should a resonant amplitude occur this is normally visible as a highly populated linear region across the plot centred on the mass of the resonance. Hence for a_0^+ this would exist at $\simeq 0.96\,\mathrm{GeV}/c^2$ in the $m^2_{\eta\pi^+}$ axis, while a_0^0 would appear at the same point on the $m^2_{\eta\pi^0}$ axis. Other resonances also exist in the $\eta\pi^\pm\pi^0$ DP. These consist of the $\rho(770)^+$ resonance in $\pi^+\pi^0$ and the higher mass resonances of each lower mass state.

Resonances are most commonly described by a relativistic Breit-Wigner amplitude, for which the probability density takes the form

$$|\mathcal{A}(s)|^2 = \frac{m_0^2\Gamma^2}{(s-m_0^2)^2 + (m_0\Gamma)^2}, \tag{1.38}$$

where s is the square of the centre of mass energy, m_0 is the resonance pole and the characteristic width is given by Γ. For the a_0 resonance this is modified slightly to account for the contribution from $a_0 \to K\overline{K}$, resulting in the Flatté amplitude [26].

The optimal analysis method for three-body decays is to perform a full Dalitz plot fit to extract amplitudes and phases for all final state resonances, also accounting for non-resonant contributions, i.e. $B \to \eta\pi^+\pi^0$ with no intermediate resonant amplitudes in the decay. However, low statistics and large backgrounds, as well as other complications affecting resolutions and efficiencies across the DP, mean that such an analysis is not feasible in this case.

As such, we adopt what is known as a 'quasi two-body' approach to the analysis. This involves only selecting events from a narrow band around the resonance peak and considering any other overlapping resonances as backgrounds. Such overlapping resonances could potentially interfere quantum mechanically with the signal resonance. However, since their amplitude will be shown to be small relative to that for signal, the effect is considered to be negligible in this case. This thesis will document the analysis using the quasi two-body approach including fits to regions in the DP associated with potentially overlapping backgrounds to estimate their signal region contribution.

Chapter 2

The *BABAR* Detector and the PEP-II *B* Factory

2.1 Introduction and Physics Goals

The *BABAR* experiment [27] [28], located at the Stanford Linear Accelerator Center (SLAC), is one of two so-called 'B Factories' currently in operation, the other being the BELLE experiment [29] at KEK in Japan. The primary aim of these machines is to study the decays of B mesons in order to better understand and constrain the parameters of the CKM matrix. B mesons are produced for *BABAR* using the PEP-II storage rings [30] [31], which are filled from the main linear accelerator (linac). A 9.0 GeV electron beam and a 3.1 GeV positron beam are collided to give a centre of mass energy of 10.58 GeV, the mass of the $\Upsilon(4S)$ resonance, from which $B\overline{B}$ pairs are produced. Critical to the design is the asymmetric energy of the collisions, giving the resulting $\Upsilon(4S)$ a Lorentz boost in the laboratory frame of $\beta\gamma = 0.56$. This facilitates the time dependent analyses required to observe the CP violation present in the mixing of neutral B mesons. These analyses are used to constrain the values of CKM parameter sin2β, which is the primary physics goal of *BABAR*.

The high luminosity available at PEP-II (designed to run at 3×10^{33} cm^{-2}s^{-1}, although higher values have actually been achieved) also makes it possible to perform precision measurements of the decays of bottom and charm mesons as well as in the τ lepton sector. *BABAR* is also sensitive to new states created from combinations of bottom and charm quarks, as well as to a large number of rare B meson decays potentially involving new physics beyond the Standard Model. For a more detailed description of the physics available at *BABAR* the reader is encouraged to look at the *BABAR* physics book [12].

2.2 The PEP-II Storage Rings

2.2.1 Machine Overview

The PEP-II storage rings are located at the eastern end of the Stanford Linear Accelerator, as shown in Figure 2.1. Electron bunches are created using an electron gun and stored in a damping ring before being accelerated to 9 GeV along approximately one third of the linac. These are then injected into the High Energy Ring (HER). Positrons are created by colliding higher energy electrons (around 30 GeV) into a stationary tungsten-rhenium target. The positrons are then recirculated into a separate damping ring before being accelerated along the linac to 3.1 GeV and injected into the Low Energy Ring (LER).

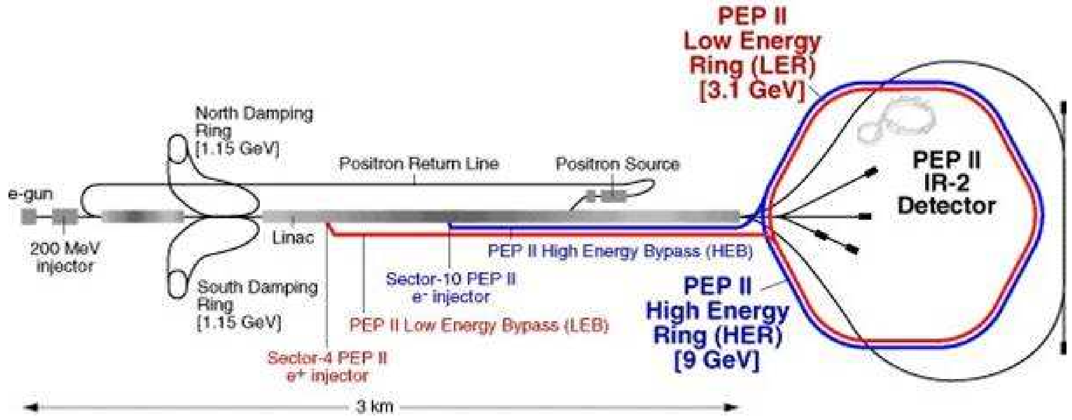

Figure 2.1: *PEP-II and the Stanford Linear Accelerator.*

2.2.2 Interaction Region (IR)

The PEP-II interaction region (IR) is shown diagrammatically in Figure 2.2. The LER and HER beams are collided head-on within the *BABAR* detector using the dipole magnets designated as B1 in the diagram. The presence of these magnets within the detector volume has an effect on detector acceptance and background conditions. Before collision the LER beam is focused using quadrupole magnet QF2 while the HER is focused using quadrupoles QD4 and QD5. Final focusing for both beams is provided by quadrupole QD1, also partially housed within the detector volume.

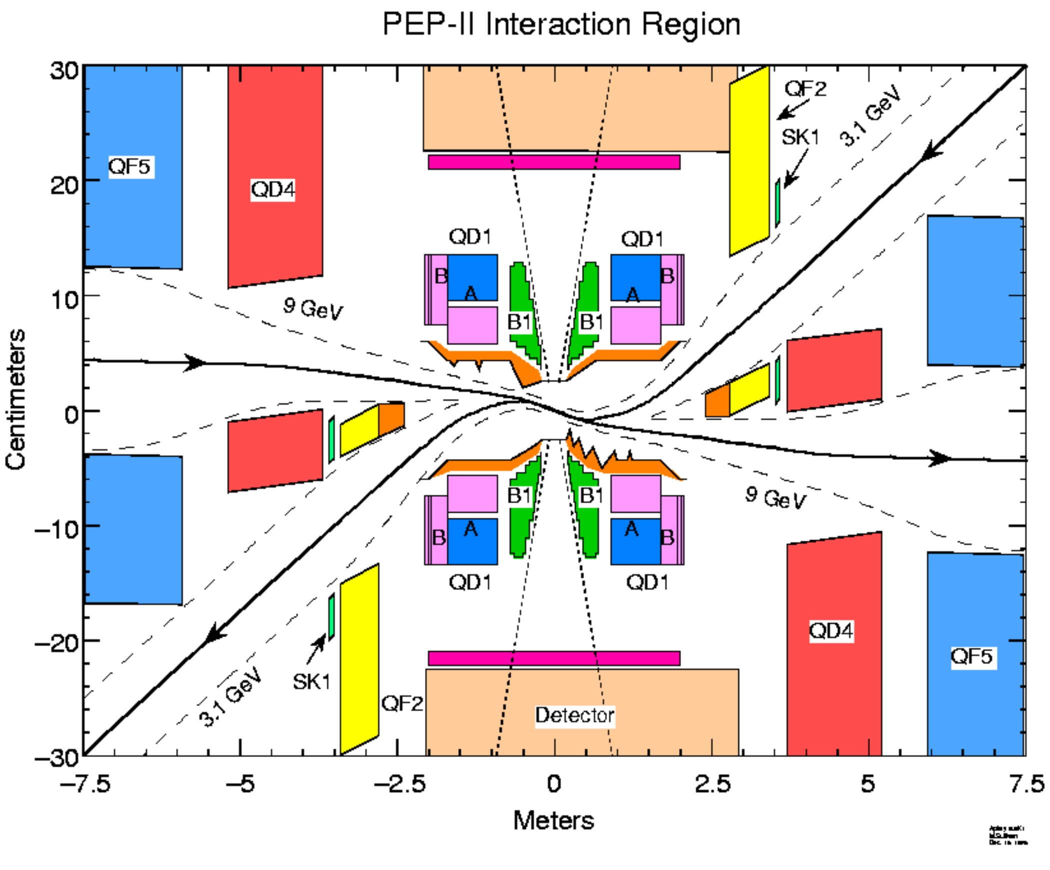

Figure 2.2: *A diagram of the PEP-II interaction region near BABAR.*

2.2.3 Performance and Luminosity Projections

Electrons and positrons are currently injected into PEP-II using 'trickle' injection [32], whereby bunches are continuously topped off. This replaced the previous, less efficient, 'fill and stop' method in 2003. The old method allowed bunches to decay over time, after which data taking would have to stopped for 15 minutes or so to allow refilling. With trickle injection there are no interruptions to data taking, with the only negative effect being a manageable increase in beam backgrounds. Trickle injection has led to substantial increases in recorded luminosity and contributed significantly to PEP-II dramatically outperforming its design specifications to reach a peak luminosity of 1.2×10^{34} cm^{-2}s^{-1} in mid 2006. The performance gain can be seen clearly in Figure 2.3.

In all, PEP-II has surpassed expectations by every measure [33]. The aim is to continue this performance improvement and reach a peak luminosity of 2×10^{34} cm^{-2}s^{-1} within the lifetime of the experiment. *BABAR* is scheduled to continue data taking until the end of 2008, at which point it is hoped that the total integrated luminosity collected will reach almost 1 ab^{-1}. Such a large dataset will facilitate extremely precise measurements of CKM matrix parameters as well as a sensitivity to rare B decays and new physics never before experimentally accessible.

2.3 Detector Layout and Overview

The *BABAR* detector is composed of five separate nested subsystems arranged radially outwards around the beam-pipe. Working outwards from the centre the Silicon Vertex Tracker (SVT), Drift Chamber (DCH), Detector of Internally Reflected Čerenkov light (DIRC) and Electromagnetic Calorimeter (EMC) are subject to a 1.5 T uniform axial magnetic field to enable momentum measurement in the tracking detectors. The outermost subsystem, the Instrumented Flux Return (IFR) provides a return path for this field. In order to maximise detector acceptance in the presence

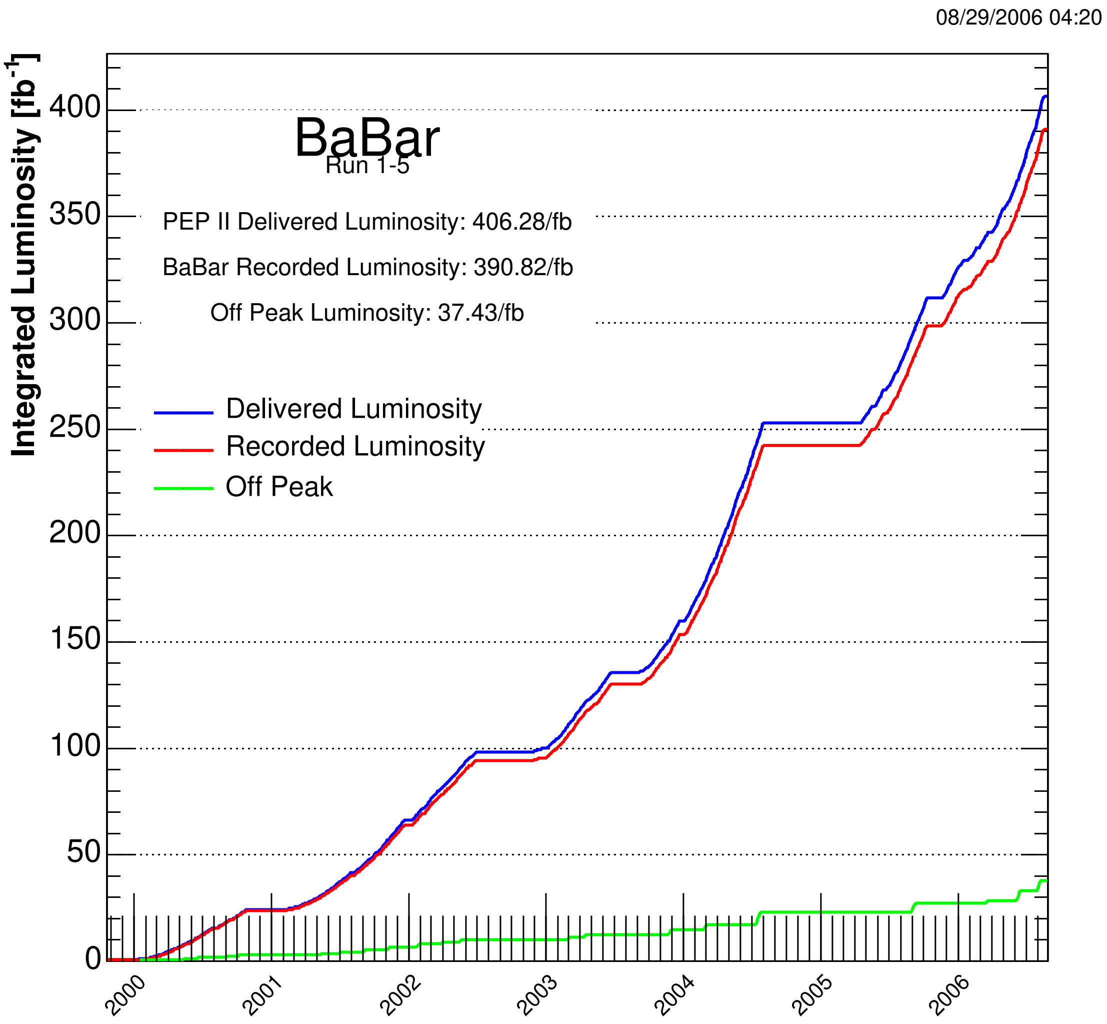

Figure 2.3: *BABAR integrated luminosity up to the end of the 5th run period in August 2006.*

of the boost applied to the $\Upsilon(4S)$ system the detector is asymmetric and offset from the Interaction Point (IP) by 0.37m in the forward direction.

Figures 2.4 and 2.5 show the detector in longitudinal and cross section. *BABAR* uses a right-handed Cartesian coordinate system with its origin at the nominal IP. The x axis points horizontally outward relative to the centre of the PEP-II ring. The y axis points vertically upward and the z axis points along the beam-pipe in the direction of the High Energy Ring (electrons). The spherical coordinates in the polar (θ) and azimuthal (ϕ) directions are defined relative to the Cartesian basis.

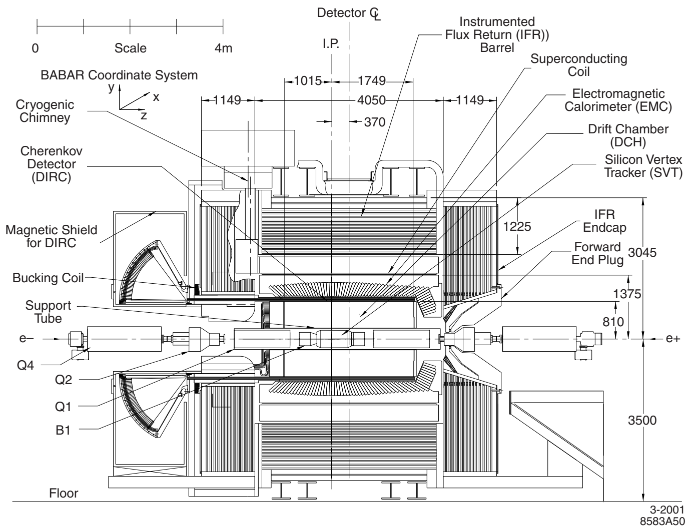

Figure 2.4: *BABAR longitudinal cross section.*

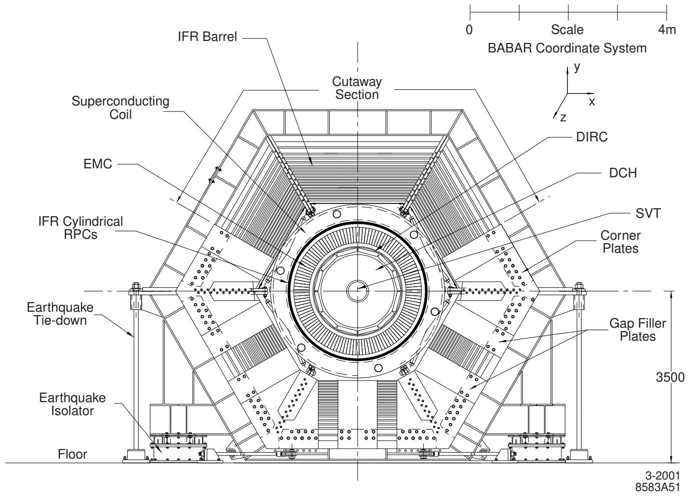

Figure 2.5: *BABAR transverse cross section.*

2.4 Silicon Vertex Tracker (SVT)

2.4.1 Physics Goals

The SVT is designed to provide precise z vertex position measurements for the decays of neutral B mesons. This information is vital in performing the time-dependent CP asymmetry analyses which are the core physics goal for *BABAR*. Due to the Lorentz boost given to the $\Upsilon(4S)$ system, the mean B vertex separation is expected to be $\sim$250 μm. In order to perform a time-dependent analysis a vertex separation resolution of less than 50% of this value is desirable. In order to achieve this a single vertex resolution of better than 80 μm is required. This is the SVT's primary design goal.

In order to reconstruct secondary decay vertices in important physics channels, such as the decays of D mesons and τ leptons, a requirement is also placed on the resolution in the $x-y$ plane. An example would be the decay $B^0 \rightarrow D^+D^-$, where the mean D vertex separation is $\sim$275 μm. Once again a vertex separation resolution of 50% of this is desirable. In order to achieve this the $x-y$ plane resolution is required to be better than 100 μm.

A number of other performance requirements are placed on the SVT in order to complement the work of other sub-systems. It is the primary source of tracking for charged particles with transverse momentum $p_T < 120\,\mathrm{MeV}/c$, which are not reliably measured by the DCH. It is important that these tracks are well measured in order to pick up the decay products of important physics processes such as $D^* \rightarrow D\pi$, which produces a slow pion. The SVT is required to deliver a tracking efficiency of better than 70% for such particles. SVT information is also important in maximising Čerenkov angle resolution in the DIRC for high p_T tracks. The system contributes further to particle identification (PID) by measuring the energy loss per unit pathlength (dE/dx) for tracks with momenta $< 700\,\mathrm{MeV}/c$.

2.4.2 Design and Layout

The design of the SVT is constrained by a number of spatial and environmental factors. Firstly the layout of the *BABAR* interaction region limits the possible acceptance of the system. Secondly, the amount of material through which tracks pass is also limited in order to reduce multiple scattering and bremsstrahlung. Finally the detector needs to be radiation hard, up to an estimated total dose of 2 MRad over the lifetime of the experiment, with an instantaneous maximum of 1 Rad/ms.

The SVT was therefore designed to be composed of five layers of double-sided silicon strip sensors, divided into modules. Layers 1-3 contain six modules each, layer 4 contains 16 modules and layer 5 contains 18. The position of the layers is shown in Figure 2.6. The modules in the inner three layers are rotated away from a symmetrical orientation in ϕ so that they overlap. In order to maintain full angular coverage while minimising material, the outer modules in layers 4 and 5 are arched towards the beam-pipe in the polar direction. Due to this the modules in these layers cannot be rotated and uniform coverage is achieved by splitting the layers into two sub-layers at slightly different radii. Figure 2.7 shows a longitudinal slice of the SVT demonstrating the arching of modules in layers 4 and 5.

The SVT design provides 0.96m^2 of active radiation-hard silicon (340 individual detectors) with total solid angle coverage of 90% in the CM frame (20.1° to 150.2° in θ). In order to further minimise the amount of material in the acceptance region all of the readout electronics are housed outside the active volume, with the layers themselves being supported by Kevlar ribs and encased in a carbon fibre tube.

The readout system for the SVT consists of 150,000 channels and is based on using the inner side of each silicon detector to measure z and the outer layer to measure ϕ. Signals coming from charge deposits in the silicon are amplified and shaped and then compared to a threshold established from beam related backgrounds during running. The time-over-threshold (TOT) for each track is then read out for further processing. The TOT is approximately logarithmically equivalent to the charge deposited but has a much better signal to noise ratio ($\sim$15), which allows a much larger dynamic range to be covered. Groups of adjacent strips with consistent timing information are then combined to form clusters, with those separated by one strip potentially being merged if appropriate. The clusters are then passed to pattern recognition software which interprets the information, forming hits expressed in terms of the physical position of the silicon sensors.

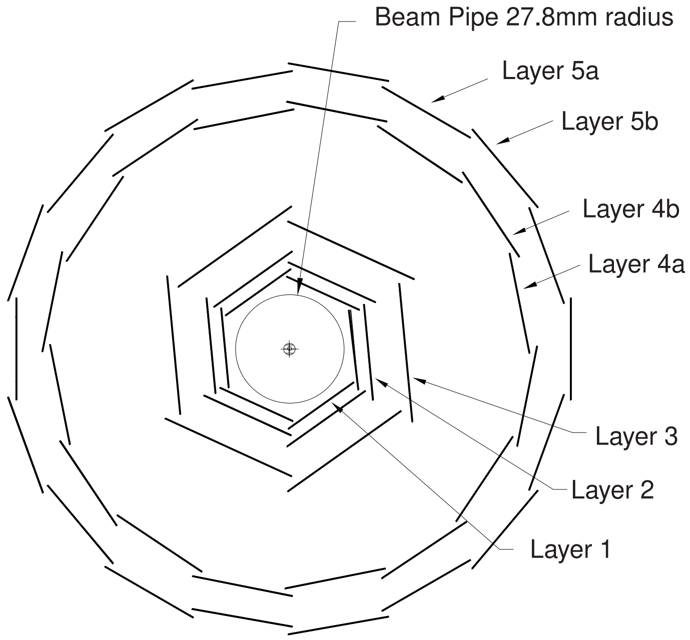

Figure 2.6: *SVT transverse cross sectional view.*

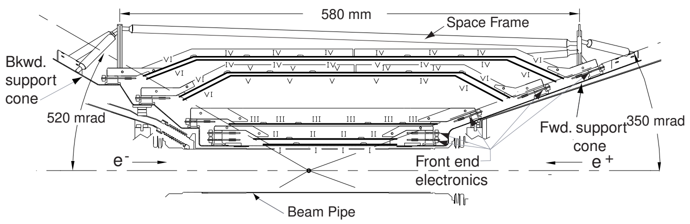

Figure 2.7: *SVT longitudinal cross sectional view.*

2.4.3 Performance

The SVT resolutions in the z and ϕ directions for each layer are shown in Figure 2.8. A resolution of better than 40 μm is achieved in the first three layers at all angles. This results in a B decay vertex resolution better than 70 μm. The SVT tracking efficiency (measured using di-muon events) is 97%. The dE/dx resolution measured for minimum ionising particles (MIPS) is 14%. This makes it possible to achieve a 2 standard deviation (2σ) separation for kaons and pions up to a momentum of 500 MeV/c and kaons and protons up to 1 GeV/c.

In total fewer than 5% of the total number of readout sections have either developed defects or been damaged during installation. However, recent studies have shown that the design assumption that the SVT would receive a total radiation dose of no more than 2 MRad over the lifetime of the experiment was somewhat optimistic. A value of 5-10 MRad is expected to be more accurate, given increasing beam luminosities over the lifetime of the detector. Despite this it is expected that any degradation in performance due to backgrounds will be manageable and therefore not make repair imperative. Replacement of damaged sections is not envisaged within the remaining lifetime of the detector since the work would take up to six months to complete.

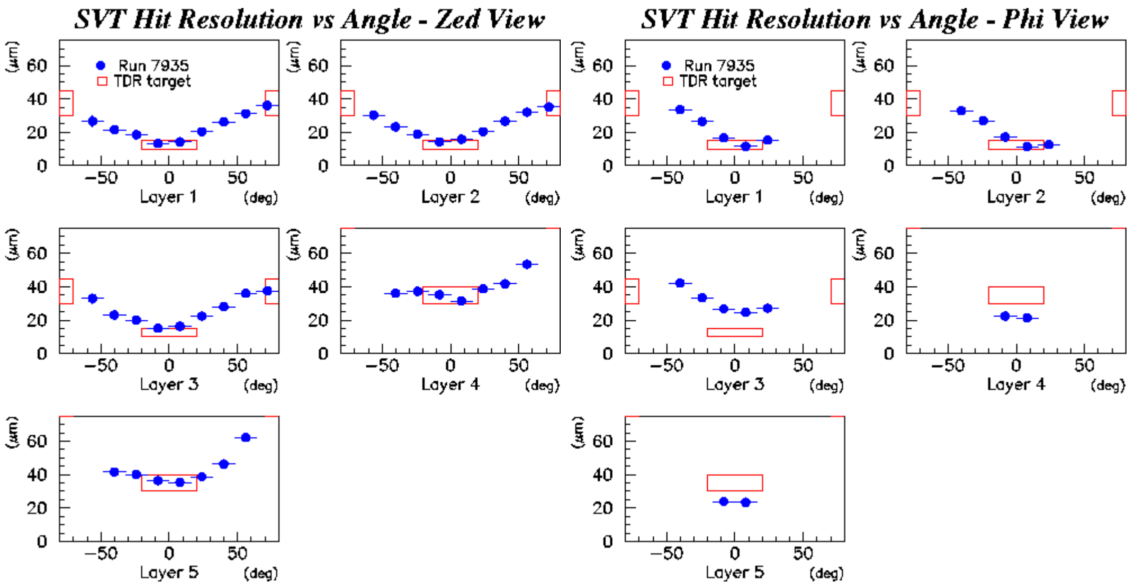

Figure 2.8: *SVT resolutions versus incidence angle for z (left) and ϕ (right).*

2.5 Drift Chamber (DCH)

2.5.1 Physics Goals

The DCH is the principal source of charged particle tracking at *BABAR*. Its primary role is to provide precision measurements of particle momenta and track angles for particles with momenta greater than 120 MeV/c and with p_T above 100 MeV/c. It is the only source of tracking for some decays involved in the *BABAR* 'Golden Modes' for measuring time dependent CP asymmetries, such $K_s^0 \to \pi^+\pi^-$, a product of the decay $B^0 \to J/\psi K_s^0$. Here, due to its average lifetime, the neutral kaon usually passes through the SVT undetected and decays inside the DCH volume. In order to reconstruct this properly a longitudinal position resolution of better than 1 mm is required. DCH information is also key in extrapolating the trajectories of such tracks into the DIRC, EMC and IFR.

Further requirements are placed on DCH performance by the need to exclusively reconstruct B and D meson decays, such as a spatial resolution $\sigma(\mathrm{R}\theta)$ better than 140 μm (averaged over all cells) and a transverse momentum resolution for particles with $p_T < 1\,\mathrm{GeV}/c$ of $\sigma_{p_T}/p_T \simeq 0.3\%$.

The DCH also plays an important role in PID in the cases where the DIRC is ineffective, such as for tracks with momentum $< 700\,\mathrm{MeV}/c$, or which are outside the DIRC acceptance which is narrower in θ than that of the DCH. PID in the DCH is based on dE/dx measurement, on which an error of better than 7% is required.

Finally, the DCH sends information directly into the Level 1 trigger at a required interval of every 269 ns with a maximum time jitter of 0.5 μs.

All of these goals must be achieved within the standard design parameters requiring maximal solid angle coverage and minimal extra material in the way of incoming particles. The DCH must operate within a 1.5 T axial field and withstand the beam related background conditions with minimal performance degradation over the lifetime of the experiment.

2.5.2 Design and Layout

The DCH is a 2.8m long cylinder placed asymmetrically about the IP in order to account for the boost in the forward direction. The chamber has an inner radius of 23.6 cm (inside which the SVT is contained) and an outer radius of 80.9 cm. These details are presented in schematic form in Figure 2.9. In order to minimise multiple scattering, while maximising spatial and dE/dx resolution, a gas mixture of helium and isobutane (ratio 4:1) was chosen. The isobutane absorbs photons thus reducing secondary ionisation while the helium reduces multiple scattering. The mixture also provides a short drift time, while its low mass means the radiation length (X_0) is up to 5 times longer than in traditional argon-based systems. Thus the total thickness of the DCH (including inner and outer walls) is $0.0108X_0$. In order to prolong the lifetime of the chamber a small amount of water vapour (0.3%) is included in the mixture.

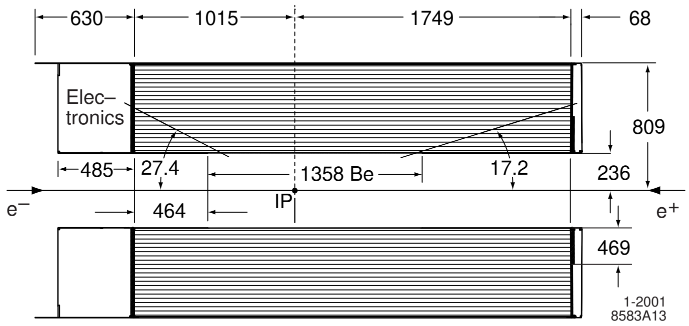

Figure 2.9: *DCH longitudinal view.*

The chamber is composed of a collection of 7104 small hexagonal drift cells (approximately 1.2×1.8 cm^2). These are arranged into 40 circular layers spreading radially outward from the inner radius. Each set of four layers is defined as a 'superlayer'. Some of the superlayers are tilted slightly with respect to the z axis in order to allow 3D position measurements. The arrangement of non-tilted 'axial' superlayers (A) and positively or negatively tilted 'stereo' superlayers (U and V respectively) follows an AUVAUVAUVA pattern, as shown in Figure 2.10.

Each DCH cell consists of a central high voltage (1960 V) sense wire surrounded by six ground wires. Therefore a field with almost circular symmetry is created over the cell. Incoming particles ionise the gas mixture causing charge avalanches within the cells which descend on the sense wire. The expected gain at the design voltage is 5×10^4. Signals are digitised at the leading edge with a resolution of 1 ns to

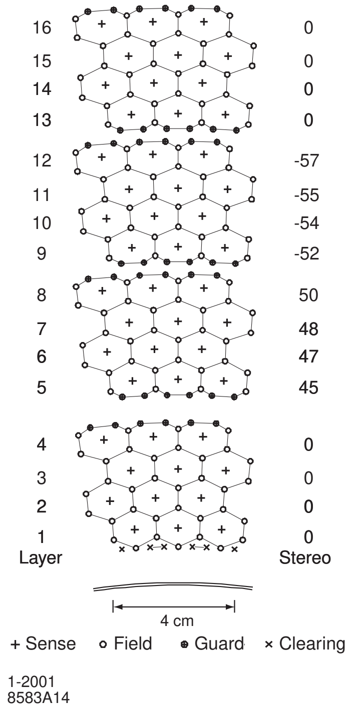

Figure ***2.10:*** *DCH superlayer arrangement for first four superlayers.*

determine the drift time. The dE/dx is determined from the time integrated charge deposited on each wire (this being proportional to the total energy loss).

Track information from the DCH is based on 5 parameters: d_0 and z_0, the distances of closest approach to the origin in the radial and z directions; ϕ_0, the azimuthal angle; ω, the curvature calculated as $1/p_T$; and tanλ, where λ is the dip angle relative to the transverse plane. All of these quantities are evaluated at the point of closest approach (POCA) to the z axis.

2.5.3 Performance

The momentum resolution of the DCH is determined from cosmic rays and is shown on the left hand side in Figure 2.11. Using a linear fit to this distribution the resolution can be parameterised as shown in equation 2.1, where p_T is in units of GeV/c,

$$\frac{\sigma_{p_T}}{p_T} = (0.13 \pm 0.01)p_T + (0.45 \pm 0.03)\%. \tag{2.1}$$

This is in good agreement with Monte Carlo simulations and close to the design resolution.

The position resolution (Rθ) for the DCH is studied using 'Bhabha' (e^+e^- scattering) and di-muon event samples and is shown as a function of the distance from a sense wire on the right hand side of Figure 2.11. The DCH performs to design specifications in this measure.

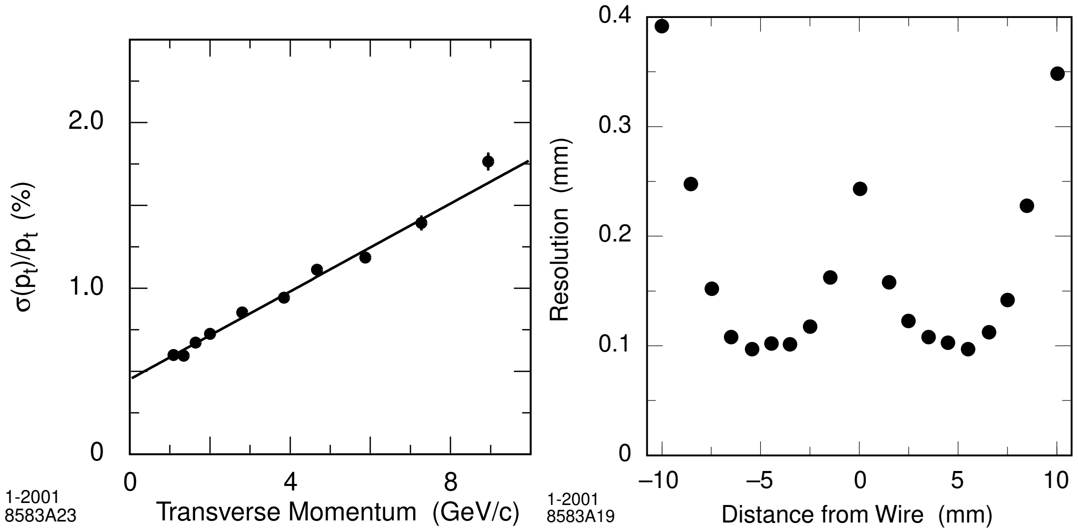

Figure 2.11: *DCH momentum resolution (left) and position resolution (right).*

The DCH's performance for dE/dx is shown in Figure 2.12 in comparison with the Bethe-Bloch predictions for given particle masses. As can be seen the data values are in excellent agreement with theory.

In summary, the DCH is performing within design requirements. Only a small number of wires are known to malfunction as a result of an accident during commissioning. Any degradation in performance due to beam related backgrounds is not expected to be significant over the lifetime of the experiment.

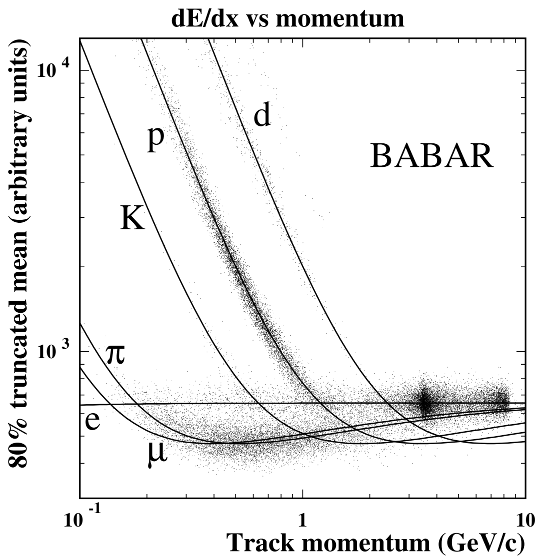

Figure ***2.12:*** *Measured dE/dx in the DCH compared with the Bethe-Bloch predictions derived from control samples for the different particle species.*

2.6 Detector of Internally Reflected Čerenkov Light (DIRC)

2.6.1 Physics Goals

The DIRC is the main source of particle identification for high momentum charged particles at *BABAR*. One of its main functions is to facilitate flavour tagging of the B meson through the cascade decay $b \rightarrow c \rightarrow s$. This produces kaons with momenta up to 2 GeV/c (although most go no higher than 1 GeV/c). This tagging is used in time dependent CP asymmetry analyses as it allows additional tagging of the second B meson in the event. Furthermore, good $K\pi$ separation is key to the study of rare hadronic B and D decays. It is from the study of rare B decays such as $B \rightarrow K\pi$ and $B \rightarrow \pi\pi$ that constraints can be placed on the value of CKM matrix angles γ and α, respectively. Tracks from these events typically have momenta up to 4.2 GeV/c and a $K\pi$ separation of at least 4σ is required in order to perform the analysis to the desired precision.

With these goals in mind, the DIRC was designed to provide particle ID in the momentum range 0.7 - 4.2 GeV/c. Below 0.7 GeV/c the DCH is responsible for this task but cannot adequately separate kaons and pions at higher momentum. Finally, the DIRC is required to complement the IFR in identifying muons of $p_T < 750$ MeV/c, a range in which the IFR is less efficient.

A number of physical constraints are also placed on the DIRC. It is required to be of small and uniform size in terms of radiation length (X_0), so as not to impair EMC resolution. Furthermore it must be physically small, so as to minimise the size of the EMC, which is the most costly part of the detector. Finally, due to the high luminosity running conditions the DIRC must provide fast readout and be tolerant of high backgrounds.

2.6.2 Design and Layout

The concept of Čerenkov radiation is well known. When a charged particle travels through a given medium faster than light, it emits Čerenkov light at an angle θ_c given by equation 2.2,

$$\beta = \frac{1}{n\cos(\theta_c)}, \tag{2.2}$$

where β is the velocity of the particle and n is the refractive index of the medium.

The DIRC is a novel design for a ring imaging Čerenkov detector based on the principle of total internal reflection. This allows all of the readout components to be stored outside the active volume minimising the amount of material impeding the progress of particles through the detector. The left hand side of Figure 2.13 shows a conceptual overview of the design while the right hand side shows a schematic of the final design.

The detection surface consists of a cylindrical arrangement of bars of synthetically fused silica, 'Spectrosil' [34], chosen because of its large refractive index of 1.473,

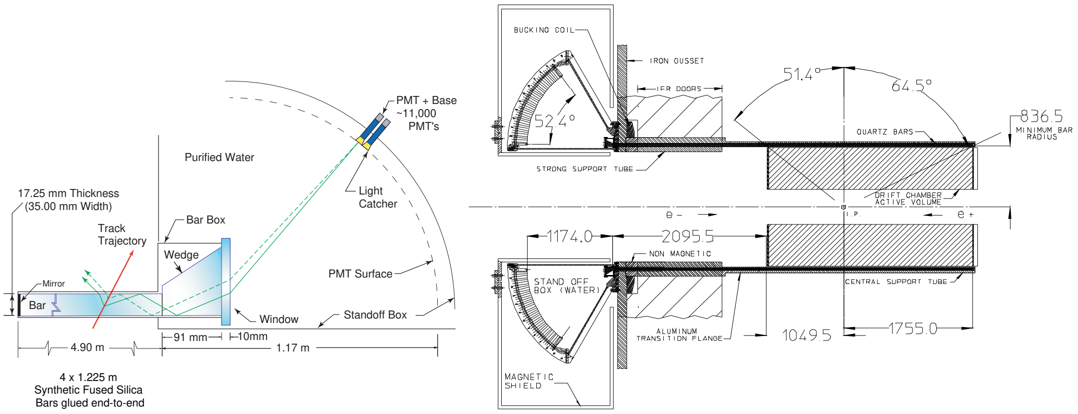

Figure *2.13:* *DIRC principle (left) and schematic (right).*

resulting in a critical angle of 47.2°. Spectrosil also provides excellent optical surface quality combined with long attenuation and chromatic dispersion lengths. The DIRC cylinder is composed of 12 groups of bars in a barrel formation. Each group is contained in a hermetically sealed container and consists of 12 bars, giving a total of 144 bars. Each bar is actually an assembly of four smaller 1.225m bars placed end to end and glued together, giving a total length of 4.9m. The entire assembly is only 8 cm thick, or 0.017 X_0 for tracks with normal incidence.

Čerenkov light is totally internally reflected along the silica bars into a large 6000 litre container of pure de-ionised water, known as the stand-off box (SOB). Water was chosen as its chromaticity and refractive index of 1.43 are close to that of the silica thus minimising refraction at the interface. The outer surface of the SOB is instrumented with 10,752 photomultiplier tubes (PMTs) of 29 mm diameter. The light pattern projected onto these PMTs is a conic section with an opening angle of θ_c, accounting for refraction. In total the DIRC has an angular acceptance of 94% in the azimuth and 83% in the cosine of the CM frame polar angle.

The DIRC measures the space-time coordinates of incoming signals and attempts to associate them with a track. SVT and DCH information are used to give track position and angles. As shown in Figure 2.14, a timing cut is applied to reduce contamination from background photons. An unbinned maximum likelihood fit is performed incorporating all space and time information to test the different particle mass hypotheses for each track. Should enough photons be associated with a given track a fit to θ_c and the number of observed signal photons is also performed.

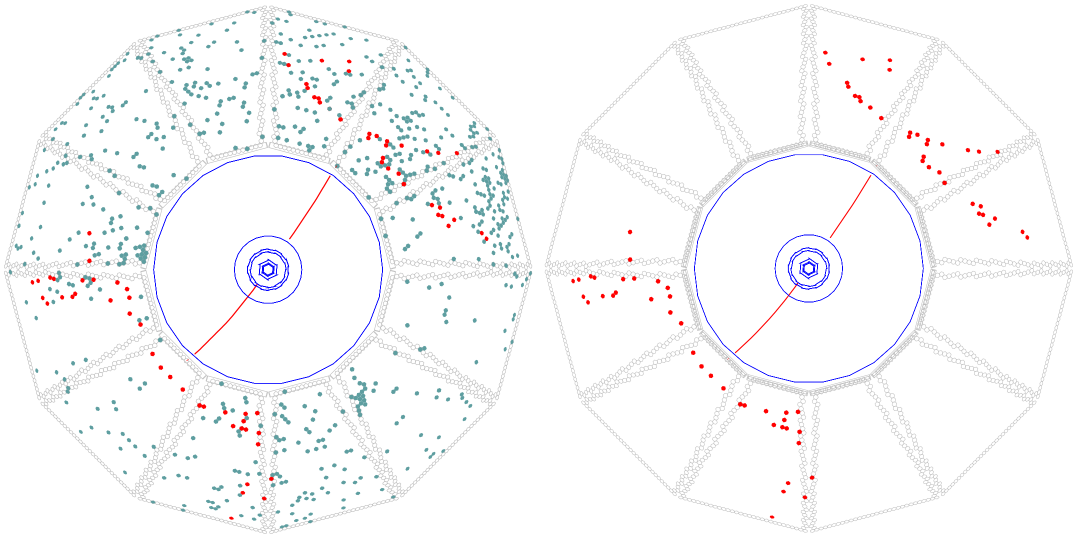

Figure 2.14: *DIRC Event with (on the left) all hits shown and (on the right) only those within 8 ns of the expected photon arrival time.*

2.6.3 Performance

In order to test the resolutions and efficiencies of the DIRC, di-muon events, such as that shown in Figure 2.14, are typically used. From the design of the detector a single photon resolution of 10 mrad is expected and a value of 10.2 mrad is measured in data, which is in good agreement. The time resolution is measured as 1.7 ns, which is close to the intrinsic time spread of the PMTs (1.5 ns).

The DIRC's measured $K\pi$ separation performance is shown on the left hand side of Figure 2.15. It is found to be in excess of 4σ for tracks with momenta below $3\,\mathrm{GeV}/c$ and is in good agreement with Monte Carlo based predictions. Finally, the efficiencies for detecting kaons and rejecting pions are shown on the right hand side of Figure 2.15. As can be seen these are in excellent agreement with performance requirements across the whole momentum range.

In summary, the DIRC is performing to design expectations. Approximately 99.7%

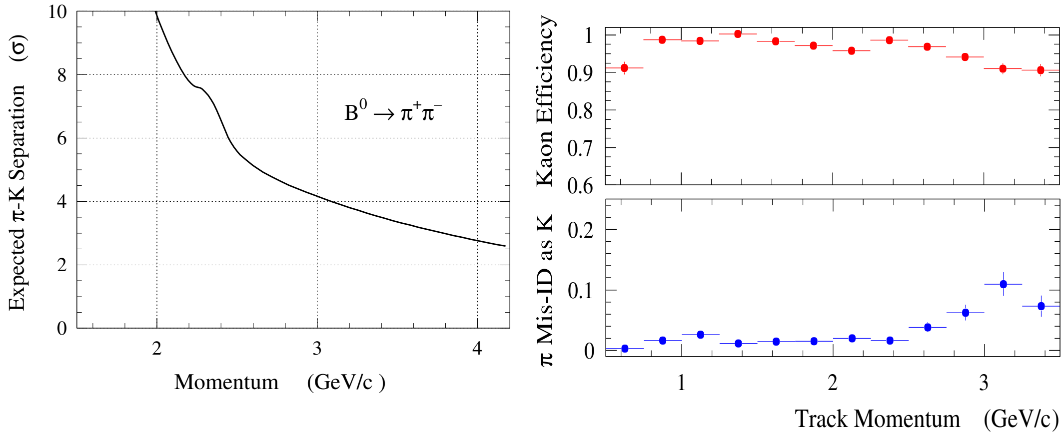

Figure 2.15: *DIRC $K\pi$ separation versus momentum (left) and kaon efficiency and pion misidentification rates versus momentum (right).*

of PMTs and readout electronics are currently fully operational and no major upgrades or refits are envisaged over the remaining lifetime of the experiment.

2.7 Electromagnetic Calorimeter (EMC)

2.7.1 Physics Goals

The EMC is designed to measure the energy of electromagnetic (EM) showers over a large energy range. In B decays 50% of photons from π^0 and η decays (which are involved in a large number of important *BABAR* physics channels, including the ones described in this thesis) have energies less than 200 MeV. Therefore sensitivity to the lower photon energy scale is essential. At the higher end of the scale it is desirable to cover the range produced from the QED processes $e^+e^- \rightarrow e^+e^-(\gamma)$ and $e^+e^- \rightarrow \gamma\gamma$, which are used in calibration and luminosity measurements. In order to fulfil these criteria the EMC is designed to span the energy range from 20 MeV to 9 GeV. The lower bound is set from beam related background estimates and a consideration of the amount of material present between the EMC and the interaction point.

Due to the rare nature of the decay processes *BABAR* seeks to study and the inherent difficulty involved in the reconstruction of π^0 and η mesons, the EMC is required to deliver high detection efficiency and excellent energy and angular resolution over the entire energy range. An example of the kind of rare high multiplicity B decay observed at *BABAR* is $B^0 \to \pi^0\pi^0$, which has a measured branching fraction of $(1.31 \pm 0.21) \times 10^{-6}$ [35] and requires an energy resolution for π^0 detection of 1-2% below a photon energy of 2 GeV. Above this energy the angular resolution dominates and is required to be of order a few milliradians.

The EMC also has an important role in electron PID, based on the energy/momentum ratio for charged tracks, which is useful in the study of semileptonic B decays as well as τ decays and the reconstruction of vector mesons such as J/ψ. Electron ID also plays an important role in the flavour tagging of neutral B decays. Energy deposits in the EMC are also used to complement IFR information in the detection of muons and K_L^0 mesons.

2.7.2 Design and Layout

The EMC is designed as a hermetic, total absorption calorimeter, providing 90% solid angle coverage in the CM frame and complete azimuthal angle coverage. The detection medium consists of 6,580 thallium doped caesium iodide (CsI(Tl)) scintillating crystals. These were chosen because of their high light yield, which is vital for the required energy resolution to be achieved. Losses through leakage into adjacent crystals are minimised by a reflective coating on the outer surfaces of each crystal. CsI(Tl) also has a short radiation length (X_0) and Molière radius (R_M), the latter being a measure of the transverse spread of an EM shower in a given medium. This is important in achieving the small angular resolution needed. A more comprehensive list of the properties of CsI(Tl) is laid out in Table 2.1. The front face of each crystal is $\sim 5 \times 5$ cm^2 in area. This is comparable with R_M which means that showers will take place across several crystals, which are then grouped together by the feature extraction algorithm. This will be discussed in more detail later.

The EMC crystals are laid out in two segments, as shown in Figure 2.16. A cylindrical barrel, consisting of 48 rings of crystals in θ and 120 in ϕ, accounts for the

Table 2.1: *Properties of CsI(Tl).*

Crystal Property	Value
Radiation Length (cm)	1.85
Molière radius (cm)	3.8
Density (g/cm^3)	4.53
Light Yield (γ/ MeV)	50,000
Peak Emission λ_{max} (nm)	565
Refractive Index (λ_{max})	1.80
Signal Decay Time (ns)	680(64%), 3340(36%)
Radiation Hardness (rad)	10^3 - 10^4

bulk of the detector area, a total of 5,760 crystals. The lengths of the crystals vary from 29.6 cm (16 X_0) in the backward direction to 32.4 cm (17.5 X_0) in the forward direction. This increase in length is to reduce shower leakage caused by the Lorentz boost applied to the Υ(4S) system in the forward direction. The remaining 820 crystals are housed in a forward endcap consisting of eight rings in θ which taper inward towards the beam-pipe with the closest three to the barrel containing 120 crystals in ϕ, the second three containing 100 and the innermost two containing 80. These inner crystals are protected from beam related backgrounds by lead blocks placed in a ring radially inward between the beam-pipe and final endcap ring. The endcap crystals are all 32.4 cm in length with the exception of those in the innermost ring, which are up to 1 X_0 shorter due to space limitations.

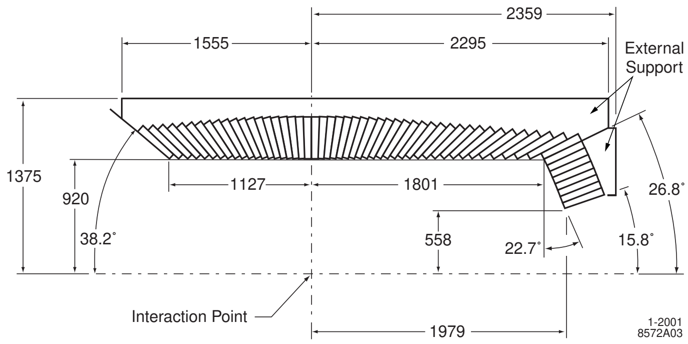

Figure 2.16: *EMC longitudinal view.*

The energy of scintillation light produced by EM showers is measured by two 1 cm^2 silicon photodiodes [36] (quantum efficiency 85%) mounted using optical epoxy at the back of each crystal. The system was designed with two diodes to provide redundancy and allow for noise reduction through averaging. The crystals are shielded from electronics noise by an aluminium coating, connected electrically to the diode assembly, forming a Faraday cage. A schematic of the crystal assembly is shown in Figure 2.17.

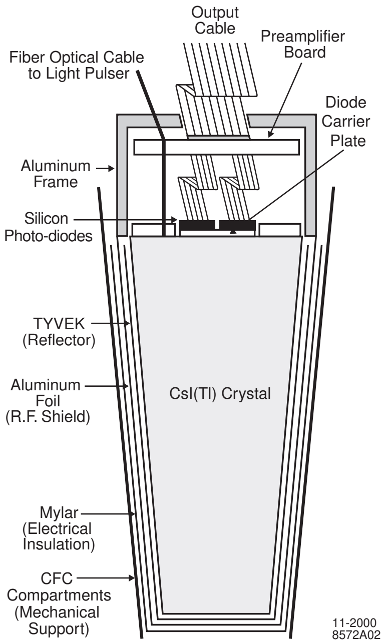

Figure *2.17:* *EMC crystal assembly layout.*

2.7.3 EMC Readout System

Signals from the two silicon photodiodes are processed first through dual gain range, low noise preamplifiers before being passed to a custom auto-range encoding (CARE) [37] ASIC (Application-Specific Integrated Circuit). This two-step process splits the signals into different amplification ranges depending on energy, as shown in Figure 2.18. The final amplification ranges correspond to: 0-50 MeV ($\times$256), 50-400 MeV ($\times$32), 0.4-3.2 GeV ($\times$4) and 3.2-13.0 GeV ($\times$1). The output for the selected range is then passed to an analogue-to-digital converter (ADC) after which the digitised signal is sent for further processing and feature extraction. Information is also sent to the Level 1 trigger.

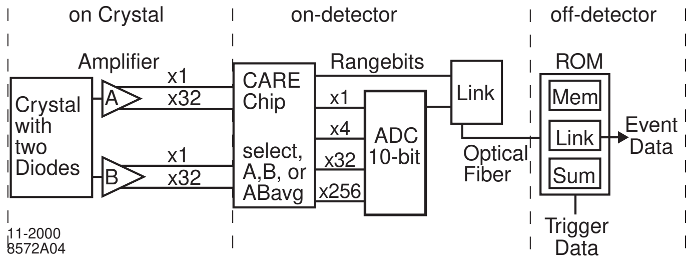

Figure *2.18:* *Signal processing paths in the EMC front end electronics (FEE).*

2.7.4 EMC Reconstruction

Electromagnetic showers in the EMC typically occur across several crystals. The feature extraction software looks for groups of adjacent crystals with a total energy of greater than 20 MeV. Should these contain at least one crystal with an energy deposit of greater than 10 MeV, and also satisfy other constraints placed on lower energy neighbours, the group is labelled as a cluster. A pattern recognition algorithm is then run to search for any local maxima, known as 'bumps', which may be present within the cluster. This can occur when two particles shower close together forming one large cluster. Corrections are then applied to account for energy loss through leakage or particles showering before the calorimeter. Finally an algorithm is run to ascertain if any EMC bumps can be associated with tracks in the event. This is done by extrapolating the track impact point on to the calorimeter inner surface and checking if it overlaps with the relevant bump centroid.

2.7.5 Performance

In order to measure the performance of the EMC a number of criteria can be tested. First and foremost are the energy and angular resolutions, which can be described empirically by equations 2.3 and 2.4 respectively:

$$\frac{\sigma_E}{E} = \frac{a}{E(\,\mathrm{GeV})^{1/4}} \oplus b, \tag{2.3}$$

$$\sigma_\theta = \sigma_\phi = \frac{c}{\sqrt{E(\,\mathrm{GeV})}} \oplus d, \tag{2.4}$$

where E and σ_E represent the photon energy and its r.m.s. error, measured in GeV, and σ_θ and σ_ϕ are the polar and azimuthal angular errors. The constant parameters a, b, c, and d are determined through fits to data, which constitute a test of detector performance. The design expectations were $a = 1\%$ and $b = 1.2\%$ for the energy resolution, and $c = 2$ mrad and $d = 3$ mrad for the angular resolution.

In practice the energy resolution below 2 GeV is measured using 6.13 MeV photons from a radioactive source within the detector and yields $\sigma_E/E = 5.0 \pm 0.8\%$. Above 2 GeV the resolution is derived from Bhabha scattering processes giving $\sigma_E/E = 1.90 \pm 0.07\%$. The results of energy resolution studies, mainly from π^0 decays, using photons of similar energies, are shown in the left hand plot in Figure 2.19. By fitting this curve below 2 GeV values for a and b can be extracted. The current measurements give $a = (2.32 \pm 0.30)\%$ and $b = (1.85 \pm 0.12)\%$. These are clearly not as good as the design expectations, but have been shown to agree with Monte Carlo (MC) studies with detailed modelling of the impact of electronics noise and beam related backgrounds.

The angular resolution, which depends on transverse crystal size and distance from the IP, is measured from decays of π^0 and η mesons to two photons of approximately the same energy. The result varies from 12 mrad at low energies to 3 mrad for high energies. The measured values for the resolution for given photon energies are shown in the right hand plot of Figure 2.19. Once again a fit to the data gives values for the constant parameters in the resolution function. The current best fit gives $c = (3.87 \pm 0.07)$ mrad and $d = (0.00 \pm 0.04)$ mrad.

The mass distributions for π^0 and η mesons reconstructed from $B\overline{B}$ pairs are shown in Figure 2.20. The mass resolutions in both cases are found to be within analysis requirements.

Overall the EMC is performing within design specifications, with only two permanently dead channels out of 6580 (due to damage during installation).

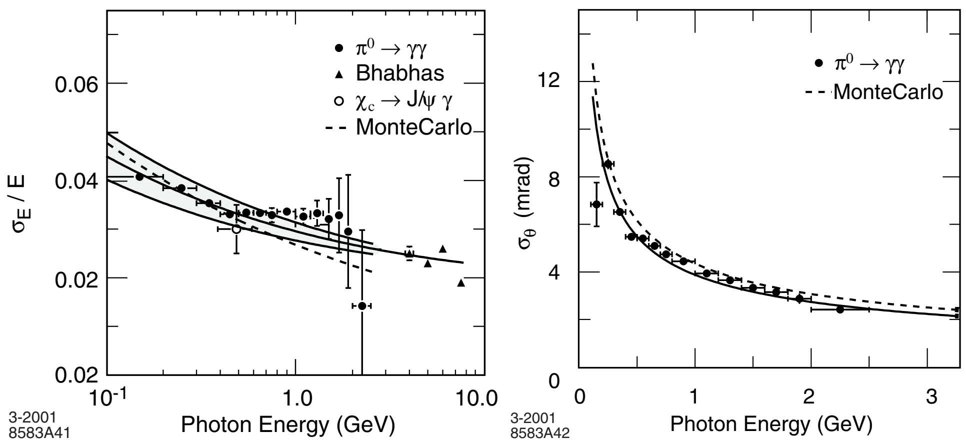

Figure 2.19: *EMC energy resolution versus photon energy for a number of interesting physics processes (left). Angular resolution (theta) versus photon energy for pion decays and Monte Carlo simulation (right). The solid curves indicate the fits used to extract the resolution measurements. For the energy resolution the curves above and below the central fit denote the* $\pm 1\sigma$ *bound for the result.*

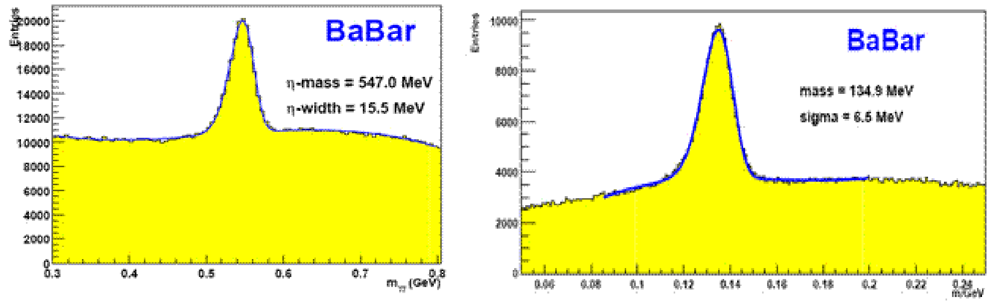

Figure 2.20: η *mass spectrum for* η *mesons with energy* $> 1\,\mathrm{GeV}$ *(left) and the* π^0 *mass spectrum for* π^0 *mesons with energy* $> 300\,\mathrm{MeV}$ *formed from photons of energy* $> 30\,\mathrm{MeV}$ *(right).*

2.7.6 Non-linearities in EMC Electronics Response

As part of the author's service work on the *BABAR* detector a study was undertaken to add the effects of historically known distortions in the EMC electronics response to the Monte Carlo simulation of the system. From this it was possible to ascertain whether the distortions had any effect on some of the physics quantities needed by *BABAR* analyses. This section includes a description of the study with results and conclusions.

Background to the investigation

Since the initial running of *BABAR* problems have been discovered with the EMC electronics response, resulting in non-linear distortions to the crystal energy spectra measured in data taking as shown on the left in Figure 2.21. These distortions were found to coincide with the amplification range changes for different signal energies described previously. The bulk causes of these distortions were isolated and corrected in hardware and software by 2001 [38], but a smaller ∼2% effect remained as shown on the right in Figure 2.21. It has been observed (although never explained) that the distortions appear worse in the central barrel region of the EMC.

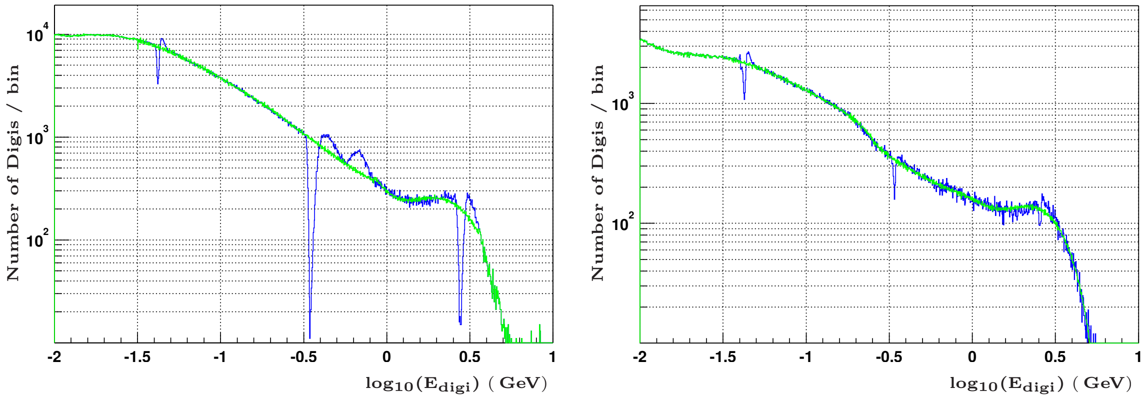

Figure 2.21: *Crystal energy (E_{digi}) spectra before hardware corrections in 2001 (left) and after corrections (right). The distorted data are shown by the blue line with the green line demonstrating the idealised extrapolation used to correct the effect in processing.*

The physical cause of the problem is unknown but it manifests itself in a discrepancy in the performance of the CARE chip in data taking and calibration states. For calibration purposes each amplification range is tested independently with known simulated signal energies (fixed mode operation), whereas in data taking the CARE chip is required to decide which range is to be used to process a signal (auto-ranging mode). A new calibration was developed to alleviate this problem and successfully deployed in April 2002 [39].

Motivation for the Study

Although now largely removed, the effects of the non-linearities on uncorrected data may have a significant impact on reconstructed physics quantities. The EMC quantity expected to be most affected by the distortions is the reconstructed π^0 width. Any problem here would have a serious impact on a number of important *BABAR* physics channels. As such, this study was centred on looking for shifts in this quantity. Previous studies [40], based on correcting the data and looking for shifts in the π^0 width, have found there to be no such shifts.

In order to be fully satisfied that the problem was small it was decided to also test the effect of the distortions on the Monte Carlo simulation for the EMC, which does not model any non-linear behaviour as standard. The purpose of this study was therefore to alter the Monte Carlo to model the distortions at a data-realistic level. From this it would then be possible to ascertain whether the distortions were having an effect on the π^0 width.

It was decided to base the study solely around the distortion caused by the lowest energy range change (at 50 MeV), since the effects in the other regions were small and difficult to separate from background fluctuations, as well as containing potential interference from other known electronics problems. In case of any π^0 energy dependence, Monte Carlo samples were generated for single π^0 decays in 100 MeV intervals from 100-600 MeV and for 1 GeV. Generated π^0 mesons were reconstructed with very loose selections applied so as to maximise the fitting sample. The photons combined form a π^0 were required to have a laboratory frame energy greater than 0.03 GeV with lateral moment less than 0.8. The π^0 candidates were then required to have mass satisfying the range $0.09 < m_{\gamma\gamma} < 0.165\,\mathrm{GeV}/c^2$. The π^0 candidates do not have any corrections applied to account for deficiencies in the Monte Carlo simulation of the detector, nor are calibrations applied to correct for detector conditions, normally based on radioactive source tests and high energy 'Bhabha' events.

Method for modifying Monte Carlo to include distortions

The Monte Carlo was modified by applying a correction to the generated crystal energies, based on a sample of data containing the distortions. The data used were collected between 1999 and 2002, the period during which the effects of the distortions were expected to be most pronounced. The data are fully calibrated to account for detector conditions. For the purposes of the study, the EMC was split up into 8 ring segments, starting at the endcap, in order to model any variation across the detector in polar angle while minimising processing time.

The correction itself was extracted using an inverse transform method [41]. This is a method of mapping between two separate continuous distributions. Take the distorted crystal energy to be parameterised by the variable E_d, with the distribution modelled by the function $f(E_d) = dN/dE_d$. Through an analogous representation take the true, undistorted, energy to be parameterised by the variable E_t and the distribution modelled by $g(E_t) = dN/dE_t$. The undistorted and distorted energy values can be related [41] using the cumulative probabilities for each distribution, given by equations 2.5 and 2.6.

$$F(E_d) = \int_{E_{min}}^{E_d} f(E_d^{'})dE_d^{'} \tag{2.5}$$

$$G(E_t) = \int_{E_{min}}^{E_t} g(E_t^{'})dE_t^{'}. \tag{2.6}$$

This integrated distribution $F(E_d)$ is shown for a sample of real data in Figure 2.22. The depopulated region from -1.38 to -1.28 in the x axis is caused by crystals shifted upward in energy by the distortion. The distortion disappears once the integration has passed over these shifted crystals. Now, the idealised distribution $G(E_t)$ is not known *a priori*, but can be approximated by extrapolating over the distorted area in $F(E_d)$ from the end points, which are assumed to be true to the ideal spectrum. This extrapolation is shown in Figure 2.22 as the blue line.

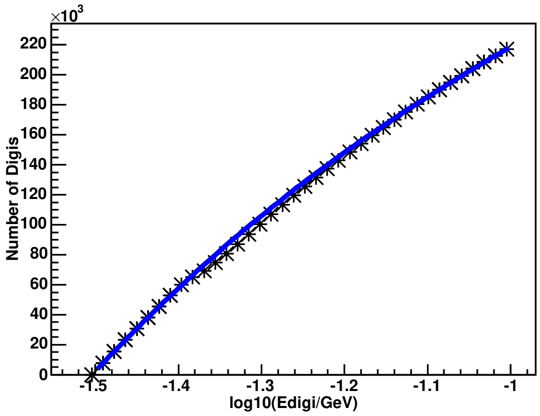

Figure 2.22: *Integrated crystal energy spectrum from data (data points) with idealised extrapolation (solid blue curve).*

Since the purpose of this study is to apply the real data distortion to the MC, we would like to be able to calculate the shifted energy for a given ideal value. This can be achieved mathematically by requiring $G(E_t) = F(E_d)$, represented by the y-axis values of both functions being equivalent. From this it can be seen that the distorted energy E_d at any point can therefore be expressed as $E_d = F^{-1}(G(E_t))$. Thus we have a way of extracting the distorted value from the true spectrum. In order to express this shift as a continuous function in E_t we use the fact that the value E_d can be described in terms of some unknown function of E_t, i.e. $E_d = C(E_t) + E_t$, where $C(E_t)$ is the correction function.

Thus by solving for E_d we can parameterise the correction solely in terms of the ideal data, as shown in equation 2.7.

$$C(E_t) = F^{-1}(G(E_t)) - E_t. \tag{2.7}$$

Therefore, by scanning along the ideal integrated curve for a given sample we can numerically construct the correction function in order to reproduce any distortion which may be present in the energy spectrum.

Using this fact a simple linear interpolation was run across the bins of the ideal spectrum to compare points on the curves and extract the shifts in energy. The shift at any point can then be plotted against the ideal E_t value to give an additive correction for crystals of a given energy. The resulting continuous functions in each θ range were then modelled using simple polynomial functions which could be included in the Monte Carlo model without much processing overhead. An example of such a correction function for rings 25-32, in the centre of the EMC, is shown in Figure 2.23.

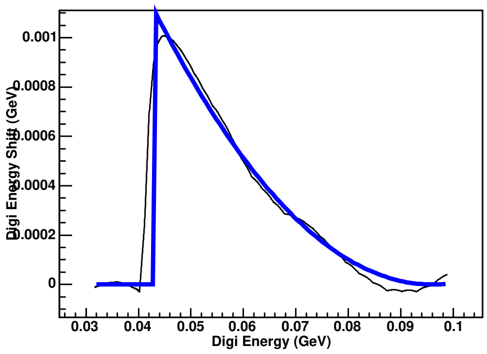

Figure 2.23: *A sample correction function extracted from data and the polynomial function used to parameterise it (solid blue curve).*

A comparison of the generated distortion for 500 MeV π^0 mesons with the real data is presented in Figure 2.24. As can be seen the size of the low energy distortion is reasonably well modelled.

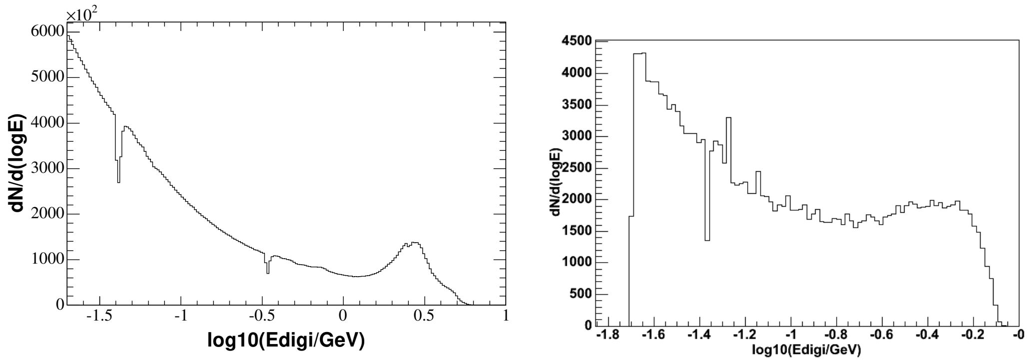

Figure 2.24: *A comparison of a distorted crystal energy spectrum in data (left) and Monte Carlo (right). Note the different x-axis scales since the data spectrum includes contributions from higher energy π^0 mesons.*

Assessing shifts in the π^0 mass

In order to compare the π^0 masses for distorted and undistorted samples a fit was performed to the separate mass distributions using the combination of a third order polynomial function and a Novosibirsk function [42] (a Gaussian modified to include a tail). A sample fit, taken on real experimental data with very loose selections applied, is shown in Figure 2.25; as can be seen, a distinct peak with an associated width can be extracted.

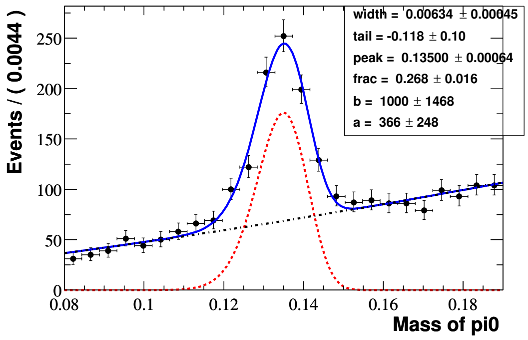

Figure 2.25: *A sample fit to the π^0 mass in data (GeV/c^2).*

Using this method it was possible to compare the fitted mass and width of the sample of undistorted Monte Carlo π^0 mesons with a sample with the distortion applied. Three separate subsets of π^0 mesons from both undistorted (standard) and distorted (non-linear) datasets were compared: those extracted from the entire data sample, those extracted only from crystals with energy falling outside the predefined non-linear region and those with at least one crystal falling into the non-linear region.

Figures 2.26 and 2.27 show a comparison of the fitted π^0 mass peak and width values for the different samples, subdivided by π^0 energy. As can be seen, any shifts in the mass and width are small. The measured mass does increase slightly in the presence of non-linearities. The effect is largest for lower energy π^0 mesons, which is to be expected since higher energy π^0 mesons would result in higher energy crystals in the spectrum and thus a lower fraction in the area affected by the distortion being modelled. In any case, the worst shift is < 1 MeV/c^2, which is not seen as a problem for physics analysis. Indeed any systematic error incurred from the accuracy of the mass fit model would probably be larger than this value. A brief investigation of the potential systematic error, involving replacing the Novosibirsk function with other Gaussian-type peaking functions suggests this lies in the region of $\pm 1\,\mathrm{MeV}/c^2$.

The π^0 width is largely unaffected by the non-linearities, apart from a slight downward shift for lower energy π^0 mesons of $< 1\,\mathrm{MeV}$. Once again, any systematic error in the model would render this negligible. Studies have shown that this error lies in the region of $\pm 1.5\,\mathrm{MeV}/c^2$.

Analysis of both mass peak and width would appear to show a decreasing trend as the π^0 meson energy increases. This is believed to be an effect of the fit model whereby at lower energies the larger fraction of misreconstruction results in a larger background component which artificially increases the width of the peaking, as well as pushing the measured peak value upward. Since the study is only interested in the comparison of distorted and undistorted samples this trend was not seen as affecting the result since the background levels are comparable for both distortion scenarios at a given energy.

Conclusions

The aim of this study was to accurately simulate non-linear effects in the EMC electronics response. The distortions were modelled by extracting a map of the effect from a sample of experimental data. Fits to the resulting generated samples suggest little significant effect on the measured properties of π^0 mesons. As such, the non-linearities have now been demonstrated, through analyses based independently on data and MC, to have no significant effect on physics analysis at *BABAR*.

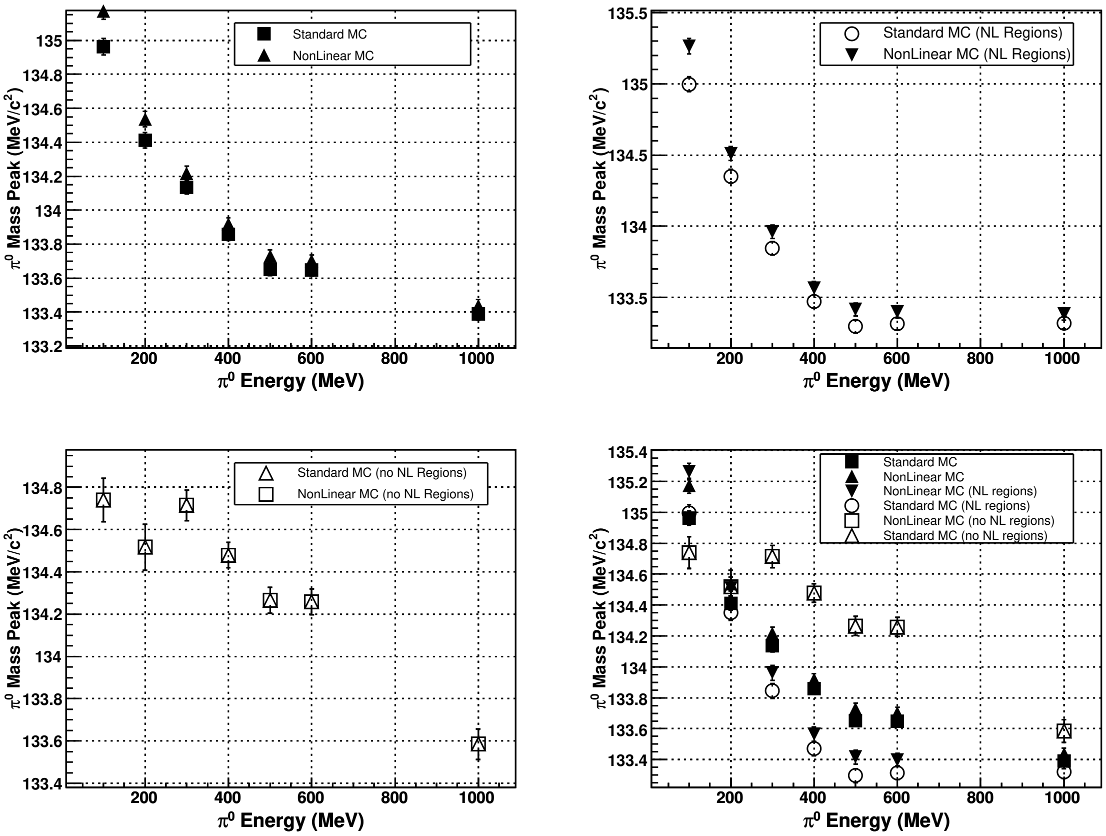

Figure 2.26: *Comparison of the fitted π^0 mass peak (MeV/c^2) vs. generated π^0 energy (MeV) for the different regions of the crystal energy spectrum with and without distortion. The top left plot shows π^0 mesons taken from the entire crystal range. The top right shows π^0 mesons including at least one crystal in the distortion region and the bottom left plot shows π^0 mesons including no crystals from the distortion region. The bottom right plot is an overlay of the other three to aid comparison.*

2.8 Instrumented Flux Return (IFR)

2.8.1 Physics Goals

The IFR provides muon and neutral hadron detection for *BABAR* as well as providing a flux return for the 1.5 T solenoidal magnet. Good muon detection is important since a number of important CP modes decay through $J/\psi \rightarrow \mu^+\mu^-$. Furthermore, muons are used to tag the flavour of B mesons in semileptonic decays, as well as facilitating similar studies of the decays of D mesons and τ leptons. In order to

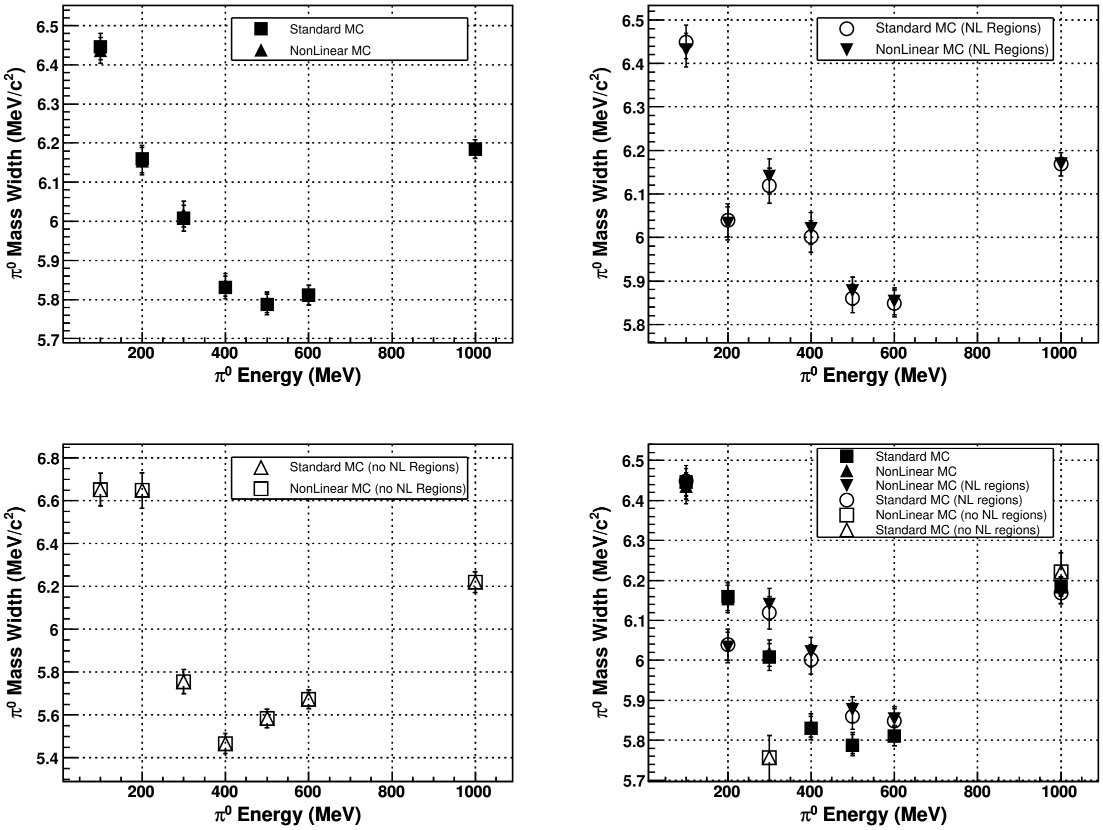

Figure *2.27:* *Comparison of the fitted π^0 mass width (MeV/c^2) vs. generated π^0 energy (MeV) for the different regions of the crystal energy spectrum with and without distortion. The top left plot shows π^0 mesons taken from the entire crystal range. The top right shows π^0 mesons including at least one crystal in the distortion region and the bottom left plot shows π^0 mesons including no crystals from the distortion region. The bottom right plot is an overlay of the other three to aid comparison.*

achieve this the IFR must provide high purity, high efficiency muon detection with high background rejection down to a momentum of 1 GeV/c.

Another important role for the IFR is the detection of neutral hadrons with high efficiency, most significantly the K_L^0. This is important in the reconstruction of the decay $B^0 \to J/\psi K_L^0$, which is used to measure the CKM parameter sin2β. In order to detect K_L^0 mesons the IFR must have good angular resolution. Finally, the IFR contributes to neutrino reconstruction by improving missing energy measurements.

Physically the IFR is required to have the standard large solid angle coverage, while also housing the support structure for the detector and performing as a flux return.

2.8.2 Overall Design and Layout

The IFR consists of a hexagonal barrel with a forward and backward endcap. The sections are composed of steel flux return layers interleaved with active detector surfaces. The steel sections were optimised for muon filtering and hadron absorption using Monte Carlo simulation. As such the layers begin with a 2 cm thickness at the inner radius and expand up to 10 cm thickness at the outer radius. The detector covers a solid angle extending to 300 mrad in the forward direction and 400 mrad in the backward. The mechanical structure of the IFR is shown in Figure 2.28.

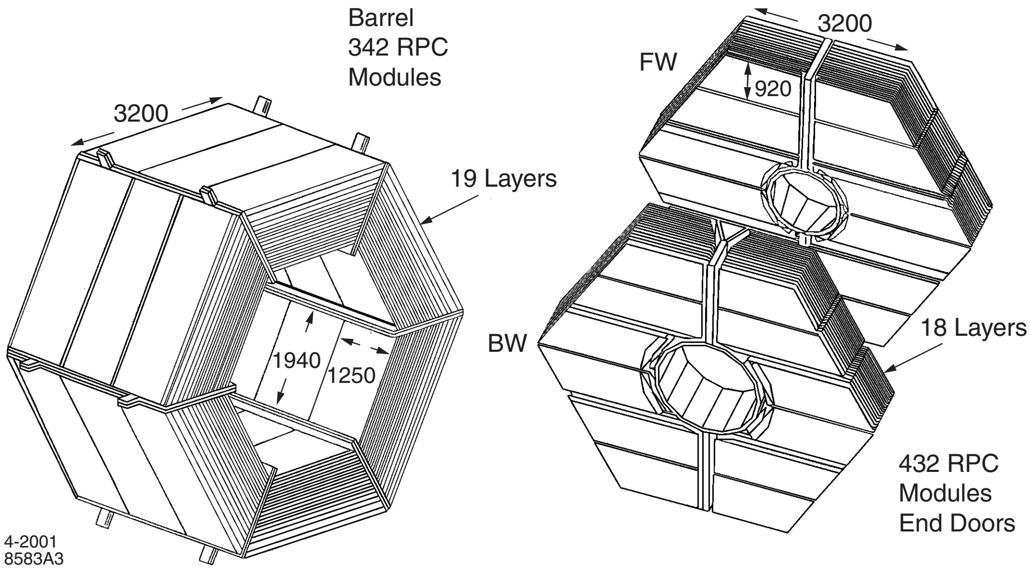

Figure 2.28: *IFR overview.*

The original active detector component in the IFR was made up of Resistive Plate Chambers (RPCs) [43]. However, due to a number of factors, including a flaw in the construction process, the performance of the RPCs began to degrade alarmingly (and unarrestably) by the end of the first year of *BABAR* running. As such they are in the process of being replaced with Limited Streamer Tubes (LSTs). The upper and lower barrel sextants were replaced in the 2004 shutdown and the remaining four were replaced in late 2006. An overview will be given of both detector technologies and their performance to date.

2.8.3 Resistive Plate Chambers (RPCs): Design

RPCs were originally chosen due to their low cost and large signal coupled with fast time response (1-2 ns). The RPC chambers consist of two 2 mm bakelite strips separated by a 2 mm gap filled with a non-flammable gas mixture of argon (56.7%), freon134a (38.8%) and isobutane (4.5%). The outer surfaces of the bakelite are coated with graphite, one layer connected to an 8 kV potential and another to ground. Particles passing through the chamber ionise the gas creating a conductive path across which a discharge takes place. This is picked up through capacitative coupling by aluminium strips located outside the graphite layers and behind an insulator. The two layers are arranged orthogonally to extract 3D position information. The design of an RPC is presented in Figure 2.29.

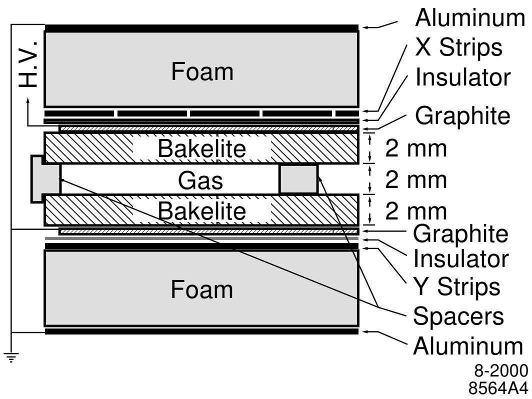

Figure 2.29: *RPC cross section.*

IFR information is read out by combining signals from individual strips to form clusters, which themselves are grouped to form tracks. These are then matched with tracks projected from the SVT and DCH. Muon candidates are required to satisfy the conditions for a minimum ionising particle in the EMC. Neutral hadrons, such as the K_L^0, are identified from clusters which do not match SVT or DCH tracks. Hit information is also passed to the Level 1 trigger every 269 ns.

In total the original IFR consisted of 806 RPCs, with a total active area of 2,000m^2. There were 19 RPC layers in the barrel and 18 in the endcaps. In addition there were two cylindrical RPCs installed between the EMC and the magnet cryostat in order to link tracks and EMC clusters to IFR clusters.

2.8.4 Resistive Plate Chambers (RPCs): Performance

Muon ID efficiency was measured during initial running by studying muons from reconstructed $\mu\mu ee$ and $\mu\mu\gamma$ final states and comparing them with pion samples taken from $K^0_S \rightarrow \pi^+\pi^-$ and three prong τ decays. The results of the studies are presented in Figure 2.30 and show an efficiency of $\sim$ 90% over the momentum range $1.5 < p < 3.0\,\mathrm{GeV}/c$ with a pion fake rate of 6-8%.

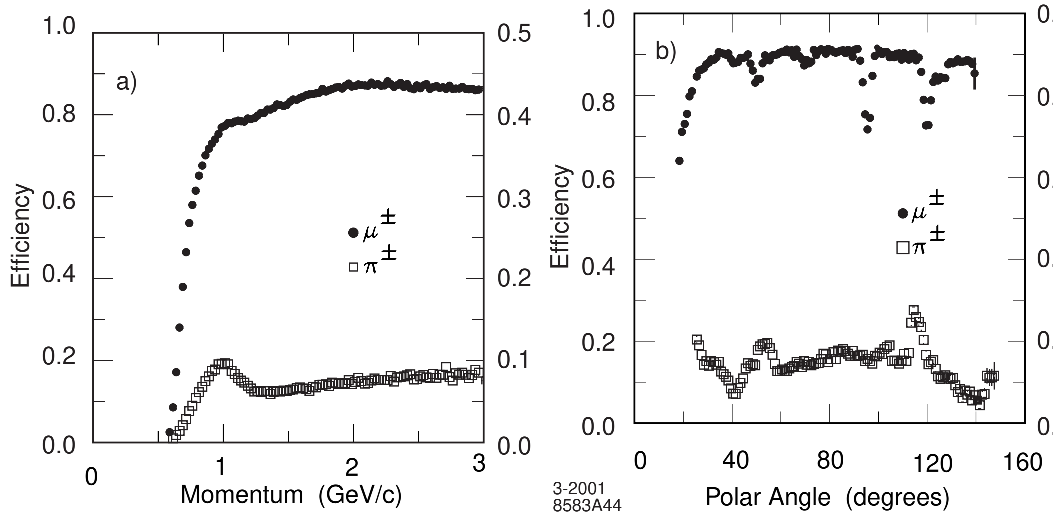

Figure 2.30: *IFR muon detection efficiency (left hand scale) and pion mis-identification rate (right hand scale) versus momentum (left) and polar angle (right).*

The angular resolution for the K^0_L was found to be in the region of 60 mrad (better if coupled with EMC information). This was tested using the decay $\phi \rightarrow K^0_S K^0_L$.

Towards the end of the first year of *BABAR* running a large number of RPCs showed rapid decreases in efficiency. This was traced to a number of environmental factors, as well as an error in construction, which adversely affect the resistivity of the bakelite. As a result large sections of RPC have become effectively useless (muon ID efficiency $< 10\%$), hence the decision to move to LST technology.

2.8.5 Limited Streamer Tubes (LSTs): Design

The limited streamer tube is a simple, robust and well known detector technology. The design is based on groups of gas filled tubes or 'cells' with a central anode wire, held at ground, bounded by high voltage connectors. Particles passing through the tube ionise the gas creating an electrical path to ground resulting in a discharge (streamer). These are read out via capacitors on the HV connection.

At *BABAR* the LSTs are composed of 7 or 8 cells of 17 x 15 mm area and 380 mm length. The gas is a mixture of CO_2, argon and isobutane at a ratio of (89:3:8), which is non-flammable and reduces secondary ionisation thanks to the isobutane content. A cross sectional sketch of a *BABAR* LST is shown in Figure 2.31.

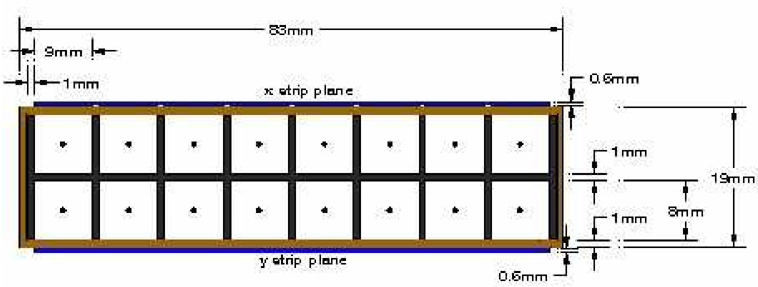

Figure *2.31:* *LST cross sectional view.*

LST information is read out in the same format as for the RPCs using largely the same criteria and electronics.

2.8.6 Limited Streamer Tubes (LSTs): Performance

The LSTs installed in 2004 have thus far outperformed the RPCs in their first year of operation (2000), with higher pion rejection for a given muon efficiency. This information is presented in Figure 2.32.

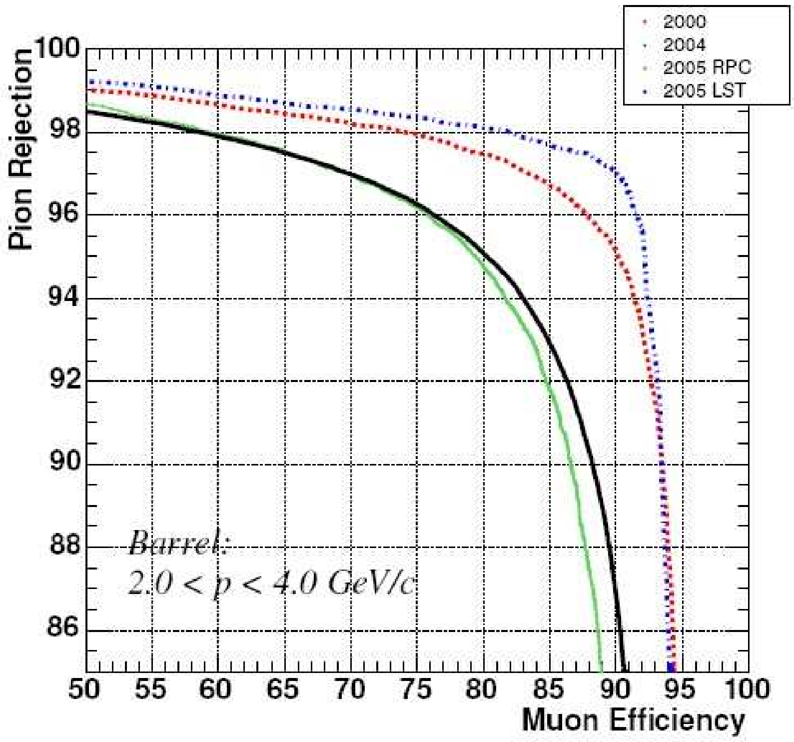

Figure 2.32: *LST muon detection efficiency (%) compared to RPC efficiency (%) in the 2005 run period and RPC efficiency in previous years. LST performance in 2005 is described by the dot-dashed curve, RPC performance in the same period by the lighter solid curve. RPC performance in 2000 is described by the dotted curve and in 2004 by the darker solid curve.*

In total $\simeq$99.7% of channels are operating normally, with the impact of any dead channels expected to be negligible due to the high granularity of the system. The analysis presented in this thesis contains approximately 200 fb^{-1} of data taken using LST technology in the upper and lower sextants of the IFR.

2.9 Event Triggers and Data Acquisition System (DAQ)

The *BABAR* trigger system is designed to reject background events while selecting potentially interesting physics events with a high, stable and well understood efficiency. The system consists of a hardware based component (Level 1) and a software based component (Level 3).

The physics performance goals for the trigger are extremely demanding. For $B\overline{B}$ events an efficiency of greater than 99% is desirable. This is in order to minimise any bias in the sample. Such an efficiency is achieved using multiple redundant trigger sources. For continuum and $\tau^+\tau^-$ events an efficiency of 95% is required and for Bhabha events, which are used for luminosity studies, the efficiency must be known to an accuracy of $\pm 0.5\%$.

The trigger is required to work in the presence of high beam background rates (up to 10 times the rate projected at design luminosity) with minimal performance deterioration over time. It must also be robust enough to function in the presence of dead or noisy channels in any detector subsystem. Finally the trigger is required to contribute no more than 1% of the total detector dead-time at any point during normal operation.

2.9.1 Level 1 Trigger

The Level 1 trigger (L1T) is a hardware based system which operates using input from the drift chamber (DCT), electromagnetic calorimeter (EMT) and instrumented flux return (IFT), the latter being used solely for cosmic ray detection and calibration. A flow diagram of the interface between the L1T and front end electronics (FEE) for each subsystem is shown in Figure 2.33. Each of these systems generates trigger *primitives*, summaries of data in terms of position and either energy or momentum, which are passed to the global trigger (GLT). The GLT produces triggers within a latency window of 11-12 μs after a bunch crossing.

The DCT forms primitives by grouping together hits in adjacent DCH cells using a Track Segment Finder (TSF). These are then passed to a Binary Link Tracker (BLT), which groups the segments into tracks. Momentum requirements are then placed on the tracks depending on their penetration into the DCH.

EMT primitives are based on combinations of EMC 'towers', which are pre-defined groups of EMC crystals. Simple feature extraction is run to extract the peak energy for a given group, which is then compared to thresholds for certain physics processes of interest.

Finally, IFT primitives are based simply on single or back-to-back IFR clusters.

When the GLT receives trigger primitives, it combines them to search for any one of 24 pre-defined physics signatures, known as trigger lines. Should one of these signatures be found, the information is passed to the Fast Control and Timing System (FCTS), which decides whether to accept, reject or prescale the trigger. Prescaling means only accepting every n^{th} event of a given type.

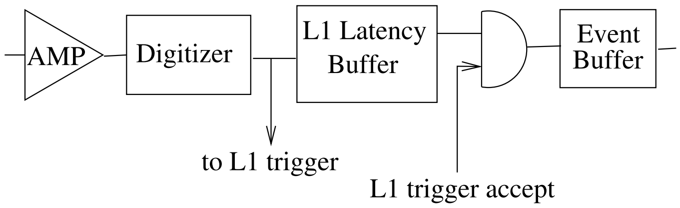

Figure 2.33: *Path to the BABAR L1 trigger for DCH, EMC and IFR.*

Trigger lines are prioritised so that, for example, high multiplicity multihadronic events are always triggered whereas Bhabha events, which are used for luminosity measurements, are prescaled and therefore read out at an arbitrarily reduced rate. This is required to reduce the data bandwidth passed to Level 3. Prescaling can be safely applied in this case since there is, by definition, no danger of biasing the sample. Should the FCTS decide to accept an event, a Level 1 accept (L1A) is produced which orders all subsystems to read out the contents of their latency buffers for further selection and processing.

Overall the L1T is performing well with a measured $B\overline{B}$ efficiency for the DCT and EMT combined of 99.9%. Individually both subsystem triggers operate at 99% efficiency, thus ensuring the required redundancy.

2.9.2 Level 3 Trigger and OPR

The Level 3 trigger is designed to reduce the Level 1 rate (nominally 1kHz, although 4kHz has been achieved without incurring unacceptable deadtime) to around 100Hz for further processing. This is run on a set of designated online event processing (OEP) farms. At Level 3 the system has access to the full event information and can thus run more sophisticated algorithms to select interesting events, resulting in simple pass/fail flags. Patterns in algorithm results are looked for and combined into Level 3 output lines, in a way analogous to Level 1. Once again, given event types can be prescaled to reduce the number passed on to the next processing phase. All in all, the Level 3 trigger is meeting its design criteria with a measured $B\overline{B}$ efficiency of 99%.

Events passing Level 3 are stored in an intermediate event store before being picked up by the Offline Prompt Reconstruction (OPR) system for the final stage of processing before the data are made available to analysts. An overview of the whole DAQ process is shown in Figure 2.34.

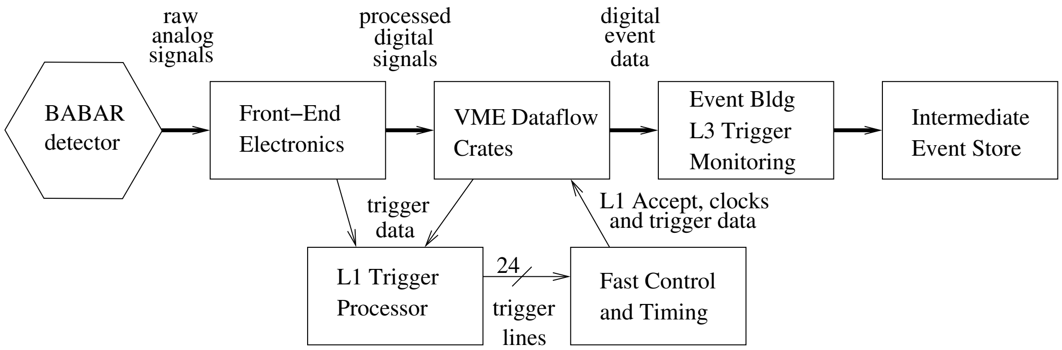

Figure 2.34: *Global flow diagram of the BABAR data acquisition system.*

Chapter 3

Analysis Fundamentals

3.1 Analysis Strategy

The goal of the analysis described in this thesis is to measure, or set a limit on, the branching fraction for the $B^{\pm} \rightarrow a_0^{\pm}\pi^0$ decay mode (charge conjugate assumed throughout) with the maximum statistical power. To this end a Maximum Likelihood Fit technique is employed using variables based on particle kinematics and event topology designed to separate background processes from the desired signal decay mode. In order to avoid biasing the result the analysis is prepared 'blind', where the experimental data are not used to tune event selection criteria or to model the distributions for the given fit variables. Instead, large samples of simulated events generated using Monte Carlo (MC) techniques are used. These are based upon a model of the *BABAR* detector built using the Geant4 package [44].

This chapter will present the datasets used for the analysis as well as discussing the methods used to reconstruct events from detector information in data and MC. Finally the variables used to parameterise physics events are presented and discussed.

3.2 Analysis Datasets

3.2.1 Experimental Datasets

This analysis makes use of the full dataset taken during the first five *BABAR* run periods, from startup in 1999 to August 2006. The total data sample comes to 340.67 fb^{-1} taken at the $\Upsilon(4S)$ resonance and 36.18 fb^{-1} 'off-resonance' taken 40 MeV below $\Upsilon(4S)$. The on-resonance sample is estimated to contain (379.9$\pm$4.2) million $B\overline{B}$ pairs. This estimate is produced using the method known as 'B-counting' [45].

3.2.2 Simulated Data

The simulated datasets used for the analysis are designed to at least match the real experimental dataset in size. Simulated events are used to model both signal and background processes. The sample sizes for signal and the bulk of the background MC are presented in Table 3.1.

Backgrounds from continuum light quark production ($e^+e^- \rightarrow q\overline{q}, q = u, d, s, c$) are modelled on QCD fragmentation [46], giving large centrally produced samples. Similar samples exist for B decay backgrounds, which are generated using the latest measured branching fractions. However, these samples are only used in the analysis to model the bulk charm decay ($b \rightarrow c$) component, which is the dominant B background source. This type of background is not expected to peak in most of the chosen analysis variables.

Contributions from charmless B decays may peak in the kinematics-based analysis variables and are as such modelled separately using specific MC samples for each mode, with all charmless decays therefore removed from the bulk sample used for the charm model. (Note that in Tables 3.1 and 3.2 the numbers quoted include both charm and charmless contributions.)

The current MC simulation techniques employed at *BABAR* have been shown in many situations to not model the experimental data perfectly. In order to improve

Table 3.1: *Bulk Monte Carlo datasets used in the analysis.*

Signal Mode	Number of Events
$B^{\pm} \rightarrow a_0^{\pm}\pi^0$ ($\eta \rightarrow \gamma\gamma$)	531k
$B^{\pm} \rightarrow a_0^{\pm}\pi^0$ ($\eta \rightarrow \pi^+\pi^-\pi^0$)	531k
Background Mode	
$q\overline{q}$ ($q = u, d, s$)	696M
$c\overline{c}$	581M
B^+B^-	556M
$B^0\overline{B}^0$	552M

the situation for neutral particles in the EMC, a smearing is applied during reconstruction to MC particles. This helps improve modelling of the data and reduce systematic errors from data/MC disagreement.

3.3 Event Reconstruction

3.3.1 Overview

Candidate events are reconstructed in *BABAR* from the core event store processed by OPR, as discussed in Chapter 2. Composite objects are formed from lists of more fundamental particle candidates found in the event, such as photons and charged tracks. At each stage candidates are required to satisfy various quality control criteria. Before more detailed reconstruction, samples undergo a pre-selection procedure known as a 'skim' in order to reduce processing overheads by rejecting events which are background-like at the cost of only a small fraction of signal.

For this analysis two reconstruction paths are required, the difference occurring when considering decays of the η meson, to which the a_0 is thought to decay $\sim$85% of the time [8]. Explicitly, we have $B^{\pm} \rightarrow a_0^{\pm}\pi^0$, where the π^0 decays to a pair of photons and $a_0^{\pm} \rightarrow \eta\pi^{\pm}$. Thus both decay channels under study require at least one charged track and two photons to be reconstructed. The η meson then decays either to a pair of photons or to a $\pi^+\pi^-\pi^0$ final state.

3.3.2 Charged Track Selection Criteria

The charged tracks used in the analysis are taken from the `GoodTracksLoose` list. This requires tracks to have momentum $\leq 10\,\mathrm{GeV}/c$ with transverse momentum $\geq 0.1\,\mathrm{GeV}/c$. Tracks are also required to be composed from ≥ 12 DCH hits and with a distance of closest approach to the interaction region $< 1.5\,\mathrm{cm}$ in the transverse xy direction and $< 10\,\mathrm{cm}$ along the 'z' axis of the detector.

3.3.3 Photon Selection Criteria

The photons used in the analysis are mostly taken from the `GoodPhotonsLoose` list. This requires photons to originate from single bump clusters in the EMC not associated with a track (so called 'CalorNeutral' objects). The photons are then required to have a raw energy greater than 30 MeV. The lateral moment (LAT) [12] of the photons is required to be less than 0.8.

Photons from η meson decays can also originate from 'CalorClusterNeutral' objects, where the EMC cluster contains more than one bump. These are accessed separately from the `GoodPhotonsLoose` list.

In order to reduce mis-reconstructions, the algorithm does not allow two or more composites in an event to contain the same photon.

3.3.4 Low Level Composite Candidate Reconstruction

π^0 meson candidates

In order to form part of a π^0 candidate further requirements are placed on photons beyond the core list specified above. Their raw energy is required to be lower than 10 GeV. The photons are combined through four-vector addition with the mass of the combined pair required to be between 0.115 and $0.150\,\mathrm{GeV}/c^2$. The π^0 candidate must also have a laboratory frame energy above 0.2 GeV.

Should a photon pair satisfy these conditions the resulting candidate is refitted to constrain its mass to the PDG value for the π^0 mass [8]. This is acceptable since the measured width of the particle is dominated by detector resolution and so refitting will provide an improvement in the resolution of reconstructed π^0 resonances while not biasing the sample. The unconstrained mass of the π^0 is stored for further analysis but it is the refitted four-momentum which is passed to the kinematic calculations for higher resonances. Candidates satisfying all of the requirements are stored in the `pi0DefaultMass` list.

η meson candidates

As described above, η meson candidates are reconstructed via two different routes. In the first case, $\eta \to \gamma\gamma$, `GoodPhotonsLoose` photons are further required to have a laboratory frame energy of between 0.05 and 10.0 GeV, the lower limit being to reduce combinatorial background. Pairs of photons are combined using four-vector addition, with the resulting candidate being required to have a laboratory frame momentum between 0.2 and 10.0 GeV/c and mass between 0.47 and 0.62 GeV/c^2.

In the $\eta \to \pi^+\pi^-\pi^0$ case, charged tracks are taken from the `GoodTracksLoose` list described above and the π^0 from the `pi0AllLoose` list. This is similar to `pi0DefaultMass` except the mass constraint from the two-photon pair is relaxed to between 0.10 and 0.16 GeV/c^2. Photons in this case can also come from multi-bump CalorClusterNeutral objects. The tracks and π^0 forming the η meson candidate are then combined using four-vector addition with the constraint that daughters originate from the primary vertex. The resulting candidate is required to be within the mass range 0.515 to 0.575 GeV/c^2.

In both cases, once an η meson candidate is identified, it is refitted to have the PDG mass of the η, using the same principles applied in the π^0 case.

3.3.5 Skim Selection Criteria

With the lowest level composites reconstructed, the skim used for this analysis is applied, known as `BToCXX`. This requires events to contain at least one charged track and two other particle candidates. Individual tag-bits are set for specific decay final states such as $\eta\pi^+\pi^0$. A requirement is also made that the energy of the final state not exceed 20.0 GeV. The reduced data sample contains lists of objects from events satisfying the skim criteria. These are then made available for higher level reconstruction. The efficiencies of the skim for data and MC are shown in Table 3.2.

Table *3.2:* *Skim efficiencies for data and Monte Carlo, defined for MC as the number of events passing the skim selection as a fraction of the total generated. For data it is the number of events passing the skim selection as a fraction of those passing OPR.*

Data	Skim Efficiency (%)
On-Resonance	7.4
Off-Resonance	6.5
Signal MC	
$B^{\pm} \to a_0^{\pm}\pi^0$ ($\eta \to \gamma\gamma$)	58.7
$B^{\pm} \to a_0^{\pm}\pi^0$ ($\eta \to \pi^+\pi^-\pi^0$)	68.2
Background MC	
$q\overline{q}$ ($q = u, d, s$)	29.3
$c\overline{c}$	30.8
B^+B^-	11.8
$B^0\overline{B}^0$	9.0

3.3.6 High Level Composite Candidate Reconstruction

$a_0^{\pm}$ Candidate Reconstruction

Reconstruction of $a_0^{\pm}$ candidates proceeds identically for both η decay modes. Tracks from the `GoodTracksLoose` list are combined with the relevant η candidate and a vertex-constrained fit is applied.

$\boldsymbol{B}$ Candidate Reconstruction

The final step of the reconstruction is to create a candidate B meson for further analysis. This is achieved by combining $a_0^{\pm}$ candidates with π^0 candidates taken from the `pi0DefaultMass` list. A vertex-constrained fit is also applied to the resulting candidates.

Due to the nature of the reconstruction process, multiple candidate B mesons can exist for each event. The reason behind this and treatment of the ensuing candidates is discussed in the next chapter.

3.4 Analysis Variables

3.4.1 Strategy for Background Discrimination

When studying the decays of B mesons the particles and their decay products can be characterised in a number of different ways. The key goal of any such variable is that it should provide a significantly different distribution in signal and background, thus making it possible to differentiate between the two. This analysis makes use of variables based on particle kinematics, the shape of events and particle identification (PID) assessments of charged tracks.

3.4.2 Kinematic Variables

The principal kinematic quantities used for the analysis are the two most widely used for B physics analysis in *BABAR* [47]. They are the beam energy-substituted B mass, m_{ES}, and the beam and B candidate energy difference, ΔE. Both make use of the fact that we know the centre of mass (CM) energy of the $\Upsilon(4\text{S})$ system to high precision. Since we are sitting on the resonance we know there is too little energy to create a $B\overline{B}$ pair plus a further particle. Therefore we expect the energy of B candidates to be half the beam energy, which also fixes the momentum magnitude to be $p_B = \sqrt{E_B^2 - m_B^2}$, where E_B and m_B are the energy and mass of the B candidate. Explicitly, m_{ES} and ΔE are defined as in Equations 3.1 and 3.2 below. Calculations for m_{ES} are done in the laboratory frame whereas ΔE is in the centre of mass frame.

$$m_{\mathrm{ES}} = \sqrt{\frac{(\frac{1}{2}s + \vec{p}_0.\vec{p}_B)^2}{E_0^2} - |\vec{p}_B|^2}, \tag{3.1}$$

$$\Delta E = E_B^* - \frac{1}{2}\sqrt{s}, \tag{3.2}$$

where $\vec{p}_0$ and E_0 are the momentum and energy of the initial (e^+, e^-) state and the * indicates this quantity is expressed in the centre-of-mass system. s is the square of the CM energy. Events which are not true B decays, or mis-reconstructed B decays other than our desired signal mode, may have m_{ES} or ΔE values outside the expected range and can therefore be classified as background.

For m_{ES}, true signal events should form an approximately Gaussian distribution centred around the B mass, $5.28\,\mathrm{GeV}/c^2$. For ΔE the calculation is performed assuming a pion hypothesis for charged tracks and in this case the Gaussian should be centred around zero for a true B. For different particle types in the decay the Gaussian will contain components with shifts to lower mean values.

In the case of continuum background events the distribution is spread over the entire range for both variables. For m_{ES} this can be parameterised with an ARGUS threshold function [48] and for ΔE with a basic low order polynomial. Signal and *uds* continuum MC distributions for both of these variables are presented in Figure 3.1.

Decays of B mesons other than the desired signal mode will respond differently depending on the final state of the decay. If the final state is different from that which is sought, a greater degree of mis-reconstruction has occurred and the distribution is more likely to be non-Gaussian. If the final state is the same as that which is sought the decays are indistinguishable from the desired signal mode and are likely to peak at $5.28\,\mathrm{GeV}/c^2$ and zero for m_{ES} and ΔE, respectively. It is in these cases

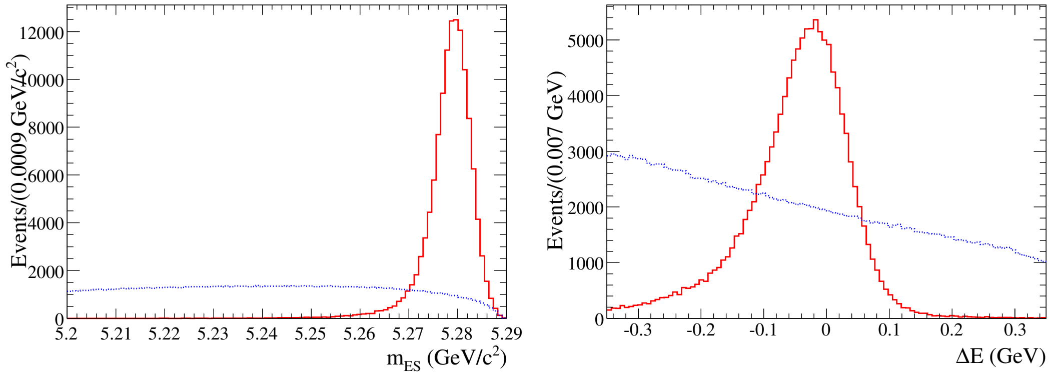

Figure 3.1: *Comparison of signal and uds MC distributions for $m_{\rm ES}$ (left) and ΔE (right). In both cases the signal is denoted by the red solid line and the uds by the blue dotted line. The distributions are arbitrarily normalised to aid comparison and are not comparable across the two variables. The MC used to make these plots has all selections described in Section 3.3 applied with the exception of those on the mass of the η, which is instead required to satisfy $0.5 < m_\eta < 0.6\,{\rm GeV}/c^2$.*

that these variables have the least discriminating power. It should be noted that, as can be seen on the right hand side of Figure 3.1, the peak for the ΔE variable appears shifted downward by a small amount (up to 10 MeV) from zero. This effect is seen in both η channels, although only to a smaller extent in $\eta \to \pi^+\pi^-\pi^0$. At this point there explanation available for this shift beyond it being related to the EMC calibration - a cause which could be consistent with the effect being worse in the mode EMC dependent $\eta \to \gamma\gamma$ channel. The final result will account for any uncertainty in the shape of the distribution as a systematic error.

The other kinematic variables used in the analysis are the resonance masses of the $a_0^\pm$ and η mesons. Since the η meson mass is fixed to the PDG value when reconstructing the $a_0^\pm$, its true mass distribution is effectively uncorrelated with that of the $a_0^\pm$ and thus the two can be used together without the risk of bias.

The resonance masses are useful in reducing contributions from B background decays above and beyond what is provided by $m_{\rm ES}$ and ΔE. In cases where there is no η decay there should be no well defined peak in the distribution. The same is true for the $a_0^\pm$ resonance mass. This is particularly useful in the case of a background with a similar final state since a large fraction will not contain $a_0^\pm$ decays. Distributions for each of these variables for signal and *uds* continuum MC are presented in Figures 3.2 and 3.3.

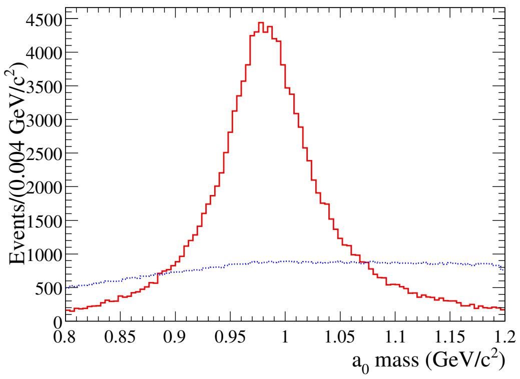

Figure 3.2: *Comparison of signal and uds MC distributions for the a_0 resonance mass. The signal is denoted by the red solid line and the uds as the blue dotted line. The distributions shown are from $\eta \to \gamma\gamma$ samples, but are not dependent on the η resolution due to the mass constrained refitting discussed in the text. The distributions are arbitrarily normalised to aid comparison. The MC used to make this plot has all selections described in Section 3.3 applied with the exception of those on the mass of the η, which is instead required to satisfy $0.5 < m_\eta < 0.6\,\mathrm{GeV}/c^2$.*

A complication to the signal model is the shape of the a_0 mass distribution, which is distorted due to the $K\overline{K}$ decay channel. The MC has been generated assuming a relativistic Breit-Wigner shape. Statistical limitations of the available dataset mean we would not be sensitive to any model improvement at this stage. The technically correct model would be a Flatté lineshape [26], which is a modified Breit-Wigner accounting for the $K\overline{K}$ partial width in the complex term.

3.4.3 Event Shape Variables

The next set of variables used in the analysis exploit the fact that signal and background events can have differing spatial distributions. For example, considering the centre of mass frame, continuum events tend to have a jet like structure due to the low mass of the decay products. B meson decays, on the other hand, tend to be isotropically distributed about the interaction point. This fact is exploited in the analysis using a quantity known as thrust [49], defined in Equation 3.3.

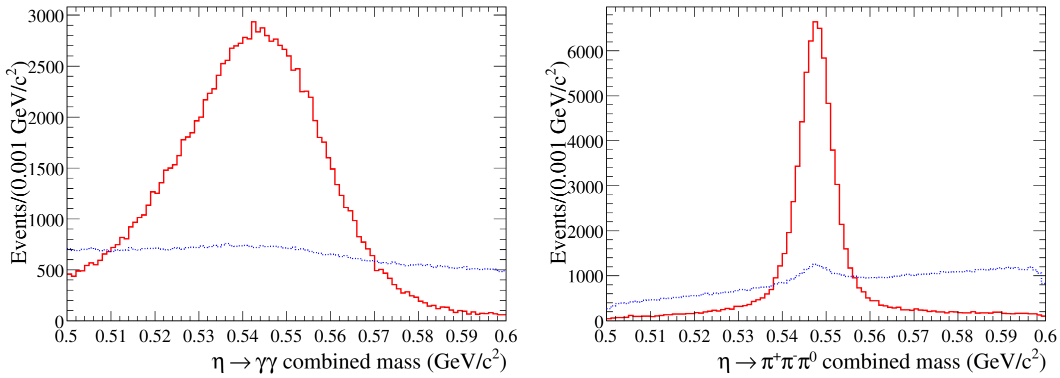

Figure 3.3: *Comparison of signal and uds MC distributions for the η mass in the $\gamma\gamma$ case (left) and η mass in the $\pi^+\pi^-\pi^0$ case (right). In both cases the signal is denoted by the red solid line and the uds as the blue dotted line. The η mass peak is narrower in the $\pi^+\pi^-\pi^0$ case due to the presence of charged tracks in the final state, which can be reconstructed with greater accuracy than photons. The distributions are arbitrarily normalised to aid comparison and are not comparable across the two channels. The MC used to make these plots has all selections described in Section 3.3 applied with the exception of those on the mass of the η, which is instead required to satisfy 0.5 $< m_\eta <$0.6* GeV/c^2.

$$T = \frac{\sum_i |\vec{P_i} \cdot \hat{n}|}{\sum_i |\vec{P_i}|}, \tag{3.3}$$

where $\vec{P_i}$ is the three momentum of the particle i and $\hat{n}$ is the unit vector for which T is a maximum. The sums are over all of the charged particles and photons in the event.

Thrust is essentially a measure of the fraction of the momentum of the products of an event or decay projected on to a given axis. By taking the cosine of the angle between the thrust axis of a B candidate with that of the rest of the event (θ_{TB}), it is possible to distinguish between a background-like event, where the two axes will be highly correlated and thus lead to a distribution peaking at ± 1, and signal-like events, where the value of the cosine should be distributed evenly over the range. All calculations are performed in the centre of mass frame. By placing a selection requirement on this quantity it is possible to remove a large fraction of mainly continuum events from the sample while removing a much smaller fraction of signal.

In practice this selection is applied to the absolute value of this variable, hereafter referred to as $|cos(\theta_{TB})|$. The first exploitation of its large discriminating power is performed when first reconstructing B candidates, where a selection is applied requiring $|cos(\theta_{TB})| < 0.9$ to reduce file size. It is later used to aid event selection when preparing samples for fitting. Signal and *uds* continuum MC distributions for this variable are presented in Figure 3.4.

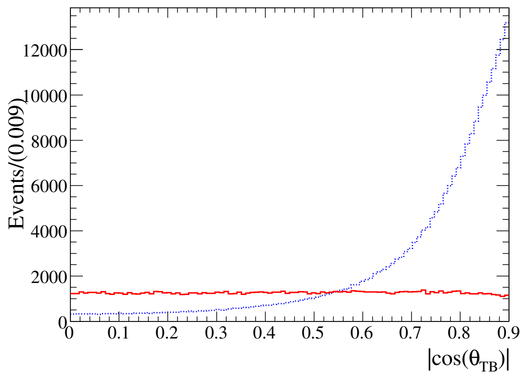

Figure 3.4: *Comparison of signal and uds MC distributions for* $|cos(\theta_{TB})|$. *The signal is denoted by the red solid line and the uds by the blue dotted line. The distributions are arbitrarily normalised to aid comparison. The MC used to make this plot has all selections described in Section 3.3 applied with the exception of those on the mass of the* η*, which is instead required to satisfy* $0.5 < m_\eta < 0.6\,\mathrm{GeV}/c^2$.

The second event shape variable used in the analysis is a Fisher discriminant [50]. This is a generic name for any linear function of the properties of a given sample whose coefficients are selected to maximise the separation between given species [51]. The implementation of the discriminant used for the analysis is based upon four different event shape quantities: the cosine of the angle between the direction of the B candidate momentum and the beam axis; $cos(\theta_{TB})$, as defined above; and the L_0 and L_2 Legendre polynomial projections [52] of the energy flow of the event with respect to the B candidate thrust axis.

For the physics processes relevant to this analysis the Fisher is found to display Gaussian distributions in differing parts of the variable range. Given the coefficient values used at *BABAR* continuum background will tend to a more positive value with signal tending to the negative side. Events containing B background decays will occupy the space in-between. In data the separation is not so great as to be able to distinguish between the Gaussians and a combined Gaussian-like distribution is formed, with background events tending to the positive side and signal to the negative. In order to reduce sample size events are required to have a Fisher discriminant value within the range -3 to 1, which eliminates mostly continuum events on the positive side. On the negative side of the distribution only severely mis-reconstructed outliers are removed. Distributions for signal, generic B^+B^- background and *uds* continuum MC are presented in Figure 3.5.

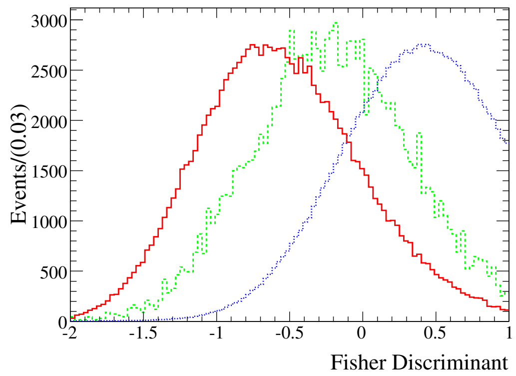

Figure ***3.5:*** *Comparison of signal, B^+B^- and uds MC distributions for the Fisher discriminant. The signal is denoted by the red solid line, the B^+B^- by the green dashed line and the uds by the blue dotted line. The distributions are arbitrarily normalised to aid comparison. The MC used to make these plots has all selections described in Section 3.3 applied with the exception of those on the mass of the η, which is instead required to satisfy $0.5 < m_\eta < 0.6\,\mathrm{GeV}/c^2$.*

It should be noted that, by definition, the Fisher discriminant and $|cos(\theta_{TB})|$ will be highly correlated. This is taken into account when constructing the fit model and discussed further in the next chapter.

3.4.4 Particle Identification (PID)

The final source of information used to characterise events is PID information, which applies to charged tracks. In *BABAR* this information is provided by the DIRC and DCH detector subsystems. Maximum likelihood-based techniques [53] are used to convert this information into a probability of a track satisfying a given PID hypothesis, i.e. whether it is a kaon, pion, electron or proton. It is then possible to select tracks which are most likely to satisfy the desired particle hypothesis. In the case of this analysis, we seek to minimise contamination of our pion track sample from kaons. As such, a veto selection is placed on the likelihood of the track being a kaon. The severity of this selection, which can be varied using selectors in different likelihood ranges, is discussed further in the next chapter.

Chapter 4

Maximum Likelihood Fit

4.1 Overview

This chapter will present the core analysis method used to extract a yield for $B^{\pm} \rightarrow a_0^{\pm}\pi^0$ from the experimental data. An unbinned multivariate extended maximum likelihood fit is employed making use of Monte Carlo simulated data samples to model the distributions for signal and background modes with respect to the chosen fit variables. Contributions from charmless B decays are identified initially from a selection-based exploratory analysis using MC samples. These selections are then re-optimised with the full fit model in order to fully exploit the statistical power of the maximum likelihood method. Fits are also run in other regions of the $\eta\pi^+\pi^0$ final state Dalitz plane in order to estimate charmless background contributions for a number of poorly measured modes.

4.2 The Principle of Maximum Likelihood

In order to derive information from any sample of data, procedures called *estimators* are used. Estimators, for example the arithmetic mean, are designed to parameterise a particular property of the sample, or more accurately its parent distribution. A likelihood [54] is a statistical term used to describe the probability of a given value of an estimator result producing a desired statistical distribution. For example a sample may have a distribution in a certain variable which can be parameterised as a 1^{st} order polynomial function. An estimator would exist which would yield a value for the x coefficient in the function. The polynomial can be considered as a probability density function (PDF) for the sample with respect to that variable, with the value of the estimator result defining the slope of the function. Varying this result will change the slope and thus alter the probability that a given value of the variable be represented in the distribution. This information can be computed into a likelihood function,

$$\begin{aligned} \mathcal{L}(x_1, x_2, \ldots, x_N; a) &= \mathcal{P}(x_1; a)\mathcal{P}(x_2; a) \cdots \mathcal{P}(x_N; a) \\ &= \prod \mathcal{P}(x_i; a), \end{aligned} \tag{4.1}$$

for PDFs '$\mathcal{P}$' defined for values of the observable x and estimator result a. If the PDF is based on a poor estimate of a the cumulative probability will be lower than that for a sample with a more accurate value. So for example if a sloped distribution is modelled as flat the probability assigned to the more populated end will be too small and thus the cumulative value will be lower than it would be if the data were modelled better, with higher probabilities for the more populous region. The value of a yielding the 'maximum likelihood' will be that which best models the distribution.

4.2.1 Implementation and Applications

In practice the PDF for a given event is modelled as the weighted sum of the PDFs for all the possible event hypotheses, i.e. signal or any of the various backgrounds. Therefore the overall likelihood for N events is defined as

$$\mathcal{L} = \prod_{i=1}^{N} \left[\sum_{j=1}^{m} n_j \mathcal{P}_j(\mathbf{x}_i) \right] , \tag{4.2}$$

where n_j are the yield weights assigned for each hypothesis present in each event.

This function can be modified to better model the case where the number of events of any given species is actually unknown (such as is common in particle physics). This is done by requiring that the total PDF integrates not to 1, but to the total number of events for all hypotheses, which shall be referred to as $\sum n_j$ here. This is implemented by multiplying the likelihood by a term accounting for the Poisson probability of obtaining N events from a mean value of $\sum n_j$,

$$\mathcal{L} = \frac{e^{-(\sum n_j)}}{N!} \prod_{i=1}^{N} \left[\sum_{j=1}^{m} n_j \mathcal{P}_j(\mathbf{x}_i) \right] , \tag{4.3}$$

thus defining the so-called *extended* maximum likelihood [54]. By floating any n_j we can therefore obtain a yield estimate from the fit itself. Variations in n_j will alter the likelihood due to the exponential term. However, the associated change in the normalisation of $\mathcal{P}_j$ will counteract this effect. The fitter balances the two competing factors and thus the maximum $\mathcal{L}$ will yield the most probable number of events, $\sum n_j$, for a given dataset.

It is simpler to compute the natural logarithm of the likelihood (since the numbers involved would be minute for large distributions) and thus fitting programs actually operate by minimising $-\ln \mathcal{L}$, which is equivalent to maximising $\mathcal{L}$ itself. This also means that the calculation is a sum rather than a product, which is computationally more convenient.

All likelihood fitting in the analysis was conducted using the RooFit toolkit for data modelling [55], which is based on the Minuit [56] function minimisation core.

4.2.2 Ensemble 'Toy' Monte Carlo Generation

In order to test a particular fit model a method is used based on ensemble, or 'toy', Monte Carlo studies. These are based on constructing a simulated dataset from

MC which matches the yield profile expected for all signal and background modes in the real data. This can be done either through generating events from the PDF model, the 'pure' toy, or by selecting and using events from the MC based on the full physics and detector simulation, the 'embedded' toy.

In a pure study 'experiment', a given PDF parameter is tested by generating events from a known PDF model with a desired value of that particular parameter required. Events are generated randomly throughout the allowed parameter space of the particular fit variables in use for that model, for example m_{ES} and ΔE. The final dataset is generated to match the PDF distributions in each variable of the initial model. Once complete, this dataset is re-fitted with the original PDF model with the value of the parameter under test allowed to float. If the model is internally consistent, i.e. all shapes are sufficiently distinct so as not to cause bias, the refitted value of the parameter should be close to, if not the same as, the original value. This similarity is described using a quantity known as the 'pull', which is a measure of the difference between the initial and re-fitted values of the parameter in terms of the statistical error on the re-fitted value. The pull P is defined as:

$$P = \frac{\text{Fitted Value} - \text{Generated Value}}{\text{Fit Error}} \ . \tag{4.4}$$

The toy experiment is repeated a number of times, ensuring that the random distributions generated each time are all distinct, to produce a distribution of pulls. If the model is unbiased and the error calculated properly the pull distribution should be Gaussian, centred on zero with a width of one. Any divergence from this indicates a problem with the fit model.

Embedded toy studies are similar to pure studies except that, instead of generating events from the PDFs, they use fully simulated MC to make up the event sample for the part of the model under test. This provides an extra test of the model as the PDF hypothesis may not correctly model correlations between variables present in data and MC. Any biases resulting from such a study, indicated by an abnormal pull distribution, suggest the presence of a correlation which has either been poorly modelled, or not modelled at all.

Toy studies will be used throughout the analysis to validate a number of aspects of the model. These will be presented in the relevant sections of the text. Where an event yield is generated it will be smeared according to Poisson statistics in keeping with the extended maximum likelihood method. The validation section will present a review of the most comprehensive, full model toy studies.

4.3 Configuration of Variables

The maximum likelihood fit implemented for the analysis was based on four variables. For some modes no correlations exist between these variables; in these cases the PDFs for each decay hypothesis were constructed as products of the PDFs for each of the four variables. Where a significant correlation is expected to occur between two variables the relevant distributions are modelled two-dimensionally with the resulting PDF multiplied with those for the two uncorrelated variables.

The variables used for the fit were m_{ES}, ΔE, the a_0 resonance mass m_{a_0} and the Fisher discriminant $\mathcal{F}$. These variables were expected to give the maximum discriminating power to the fit while minimising inter-variable correlations, which can impair fit performance. The remaining variables; $|\cos(\theta_{TB})|$, the η resonance mass and the PID selectors, were thus used in event selection to reduce the size of the fitting sample. The reasons for these choices are discussed below. The area of the data allowed by the combined selection criteria will be referred to as the 'signal box' throughout the text.

4.3.1 $|\cos(\theta_{TB})|$ vs. $\mathcal{F}$

Due to the large correlation between $|\cos(\theta_{TB})|$ and $\mathcal{F}$, it was decided that both could not be used in the fit. $\mathcal{F}$ was chosen due to its superior discriminating power for B background modes as well as being easier to parameterise effectively than $|\cos(\theta_{TB})|$.

4.3.2 a_0 and η Resonance Masses

It was decided to leave the η resonance mass out of the fit due to two main factors. Firstly, a large number of background decays contain η mesons and so its discriminating power as a fit variable will not be that large. This is in contrast to the a_0 resonance mass where only a few background modes were expected to give a peak. Secondly, the stability of a fit tends to decrease with increasing numbers of variables. Any increase in fit power from including the η mass was not considered worth the resulting loss in stability. As with $|\cos(\theta_{TB})|$, the η mass was therefore used to select events so as to reduce the sample size.

4.3.3 The Role of PID

Particle Identification information was used by placing a selection on a particular severity of a likelihood-based kaon veto in order to reduce the data sample. The particular severity was chosen from an optimisation process for all selection variables which will be described shortly. As far as the veto is concerned, the tighter the selection, the more likely that the tracks removed from the sample will be kaons, and not mis-identified pions. However the efficiency of the selection for kaons will be lower than for a looser case, therefore a lower fraction of the overall number of kaons would be removed.

4.3.4 Correlations Between Kinematic Variables

Small correlations between m_{ES} and ΔE were expected in most decay types since they are calculated from the same fundamental kinematic properties of an event. However, any effect is small compared to the correlation between $|\cos(\theta_{TB})|$ and

$\mathcal{F}$. As such this did not affect the decision to use both variables in the fit. Any correlations were expected to be most evident in background charmless B decays and a solution was implemented using non-parametric PDFs to model the distributions. These will be explained in more detail when describing the charmless B background model. For the signal mode m_{ES} and ΔE were taken as uncorrelated.

4.4 Signal Model

In the following section the procedure by which the signal model for the fit is constructed will be discussed. This model is based on a novel approach to identify mis-reconstructed signal events using the discriminating power of the fit.

4.4.1 Multiple Reconstructed Candidates Per Event

In any given event there may be multiple reconstruction paths available to the software. For example, two separate combinations of photons could provide separate η candidates both of which satisfy the conditions prescribed before. Instead of attempting to decide which η candidate to use the software splits the reconstruction hypothesis into two, one based on the first η and the other on the second. This will result in two separate $a_0^{\pm}$ candidates and eventually two separate B meson candidates for a given event. This process can occur a number of times in the reconstruction of an event leading to multiple B candidates.

Since there is only one real physics event associated with multiple B candidates, a way must be found to avoid multiple counting. This can be done by selecting one B candidate from each event, either randomly or based on various candidate quality factors. It is often the case for signal events that one of the final B candidates is the true decay and the others are mis-reconstructions, although it is possible that all of the B candidates in an event are mis-reconstructions. Clearly, for non-signal events, all candidates are mis-reconstructions. For signal, techniques exist which attempt to distinguish between the true B candidate and the mis-reconstructed ones, which will be referred to hereafter as self-crossfeed (SxF) candidates for true signal events. While useful these methods are never 100% efficient.

Another way of dealing with multiple candidates is to model all of the B candidates in the final sample and find a way of extracting the true event number based on either estimates of the multiplicity or the shape difference between true signal and SxF distributions for given variables. This analysis uses the latter technique. Discussion and validation of this method will now be presented, as well as a review of the potential ways to select a 'best' B candidate from an event.

4.4.2 Methods of Separating Signal and SxF

As mentioned previously, SxF B candidates are defined as the result of a true signal decay being mis-reconstructed to varying degrees. An example of this would be the substitution of a photon from the other B decay for a true signal photon. Such a mis-reconstruction has the effect of distorting the distribution for reconstructed signal events in the variables used in the analysis.

Typically, each reconstructed event will contain the true reconstructed B meson along with a number of SxF candidates , although it is possible that no SxF candidate will exist. Events also exist where the true decay is not properly reconstructed at all and therefore all candidates are SxF.

When constructing a fit model it is not always necessary to model the SxF explicitly. This depends on whether it is distinct enough from the other backgrounds to bias the fit if left unaccounted for, since the fitter has to associate all events present in the sample with one of the available PDFs. In any case, it is always desirable to remove SxF contamination from the signal shape in order to make it as distinct from background as possible and therefore maximise the discriminating power of the fit.

In order to distinguish between signal and SxF, the method often used elsewhere involves what is known as 'Truth Matching' (TM). This is a procedure which is only possible in Monte Carlo events and exploits the fact that the MC generator will know the exact momenta, energies and decay trajectories of all generated particles. By comparing the signatures (e.g. DCH hits) of reconstructed particles with the generated ones, it is possible to identify their 'truth partners', i.e. those for which the properties match within a certain error bound.

There are a number of ways to implement truth matching for a particle decay tree. The one used for this analysis is to require a match to all the fundamental decay daughters for the desired mode, i.e. the lowest level photons and tracks. By requiring that these all satisfy the truth matching algorithm and that their truth partners are all part of the correct signal decay tree it is possible to identify SxF candidates with relatively good accuracy. The procedure is, however, imperfect. As such this analysis used a method which, while using truth matching as a start point, is ultimately independent of it.

4.4.3 Iterative Fit Parameterisation Method

Current truth matching methods are not perfect, there is an associated inefficiency with the fundamental procedure. The inclusion of radiative corrections for charged tracks in the MC has also been known to cause true signal events to be incorrectly labelled as SxF, thus reducing the efficiency of the truth matching process. The purity of the TM sample can be affected by errors which result in SxF candidates leaking into it. These effects are undesirable since they firstly result in the signal PDF containing what is effectively a background component, and thus being less distinct from the other PDFs, reducing the discriminating power of the fit. Secondly, true signal events can be lost into the SxF sample through inefficiency. In general, the inefficiency is a larger effect than the impurity for truth matching. Should the SxF component be included in the fit with the described effects in place, the PDF may have a strong correlation with the signal, which will cause biases in the fit results.

In order to circumvent problems with truth matching the analysis used a novel approach based on modelling the shapes of signal and SxF separately and using the fit itself to decide what is signal and what is SxF. The procedure is based on the assumption that true signal decays will adopt a distribution which can be described by a particular shape in each variable, independent of whether it is truth matched or not. SxF candidates are modelled following the assumption that they will adopt a background-like distribution, again with a shape independent of truth matching.

The PDF functions used here are listed in Table 4.1 and are identical (albeit with differing shape parameter values) for both η decay modes. For the true signal case ΔE is modelled using a Novosibirsk function [42], which is a Gaussian with a power law tail. For m_{ES} two independent Gaussians are used, referred to in the table as a 'Double Gaussian'. Here all five parameters: two means, two widths and the relative fraction are allowed to float in the fit. The Fisher discriminant is parameterised as a 'Bifurcated Gaussian', which is one with three free parameters: a mean and two asymmetric width components (dependent of which side of the mean is being considered). The a_0 resonance mass is modelled with a Breit-Wigner shape (for reasons discussed in the previous chapter). In general, the SxF is modelled with shapes similar to those expected for background distributions, i.e. with Chebychev polynomials [52] and, in the m_{ES} case, an ARGUS threshold function [48]. Chebychev polynomials are used where a smooth function is required throughout the analysis for all decay modes since they are designed to minimise the correlations between polynomial coefficients and thus increase fit stability. The SxF Fisher is modelled with a Double Gaussian.

The shapes for each PDF were modelled based on signal MC. The generated samples for both η decay modes were each comprised of 531,000 events. In order for the procedure to work, truth matching is used to separate the signal and SxF distributions to the best possible accuracy before the fit procedure is applied. The assumption is made (which the TM must be good enough to achieve) that the truth matched signal component has high purity, with little SxF contamination. This is fair since errors in truth matching are observed to be more likely to reject true signal events than promote false ones. Therefore the TM sample is used to extract initial parameters for the ideal PDFs. The functional form for the ideal PDF will be referred to as $S(p_i)$ where p_i are the parameters used in the fit to the four variables.

Table 4.1: *The PDFs used to model each variable used to fit signal and SxF in the MC. The model is identical for both η subdecay modes.*

Variable	Signal Function	SxF function
ΔE	Novosibirsk	3^{rd} order Chebychev polynomial
m_{ES}	Double Gaussian	ARGUS
m_{a_0}	Breit-Wigner	2^{nd} order Chebychev polynomial
$\mathcal{F}$	Bifurcated Gaussian	Double Gaussian

As the truth matching algorithm is not 100% pure, the truth matched sample derived from it will have a very small component of SxF with a different functional form from the true signal component. The key to this procedure is that this SxF contribution is not large enough to prevent an accurate parameterisation of the signal peak being acquired. As shown in Figures 4.1 and 4.2, the SxF contamination is small and thus does not cause a problem.

The rest of the sample of signal events, made up of those which do not satisfy the truth matching algorithm, is a mixture consisting mainly of SxF with some true signal events incorrectly interpreted as SxF by the algorithm. This sample is used to parameterise the SxF shape using the function $(1-f_{\overline{TM}})S(p_i)+f_{\overline{TM}}X(q_j)$, where $X(q_j)$ corresponds to the SxF PDF with shape parameters q_j. The fit fraction $f_{\overline{TM}}$ corresponds to the fraction of SxF in the non-TM sample. In this case, the shape parameters for the signal component, p_i, are held fixed at the values obtained from the fit to the pure TM sample. The SxF shape parameters however, are allowed to float along with $f_{\overline{TM}}$. This has the effect of allowing the SxF shape to be accurately modelled without the distortion caused by the true signal events which have leaked into the sample. This can be thought of as 'background subtracting' the true signal distortion from the desired SxF distribution.

Once this is complete the function $(1-f_{TM})S(p_i)+f_{TM}X(q_j)$ is used to fit the original TM sample, which includes the previously discussed small SxF distortion. This time the SxF shape parameters q_j are fixed and the signal shape parameters p_i and fit fraction f_{TM} are floated. Here, f_{TM} refers to the fraction of SxF in the TM sample. Thus it is possible to improve the signal shape by removing some

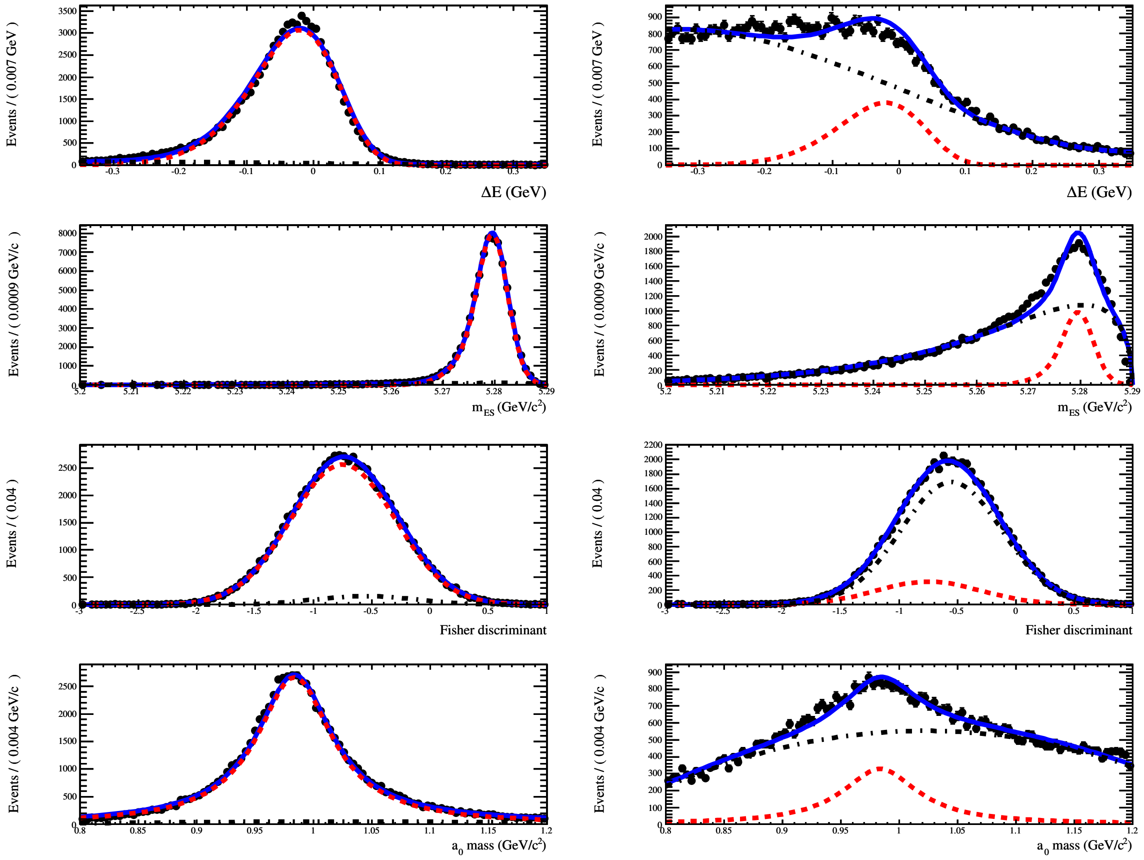

Figure 4.1: *Fits to the separate signal (left column) and SxF (right column) samples for the $\eta \to \gamma\gamma$ subdecay mode. The total PDF is the blue solid line, the pure signal PDF is the red dashed line and the SxF PDF is the black dot-dashed line.*

of the distortion caused by the small number of incorrectly truth matched SxF events. This results in a small change to $S(p_i)$. The process is then iterated, fitting truth matched and non-truth matched samples alternately as described, until the fit fractions and shape parameters stabilise. In practice this takes five iterations to complete satisfactorily.

The procedure is initially run separately on each variable (for stability purposes) to get good estimates for p_i and q_j. The PDFs are then combined into the full 4D model and the iteration is re-run, once again with five steps. The shape parameters are mostly stable after the 1D phase but the errors are reduced at 4D. As a cross check, the stable model is fitted to the entire MC sample, containing all signal and SxF components, $(1-f)S(p_i) + fX(q_j)$. Here all shape parameters are fixed and only

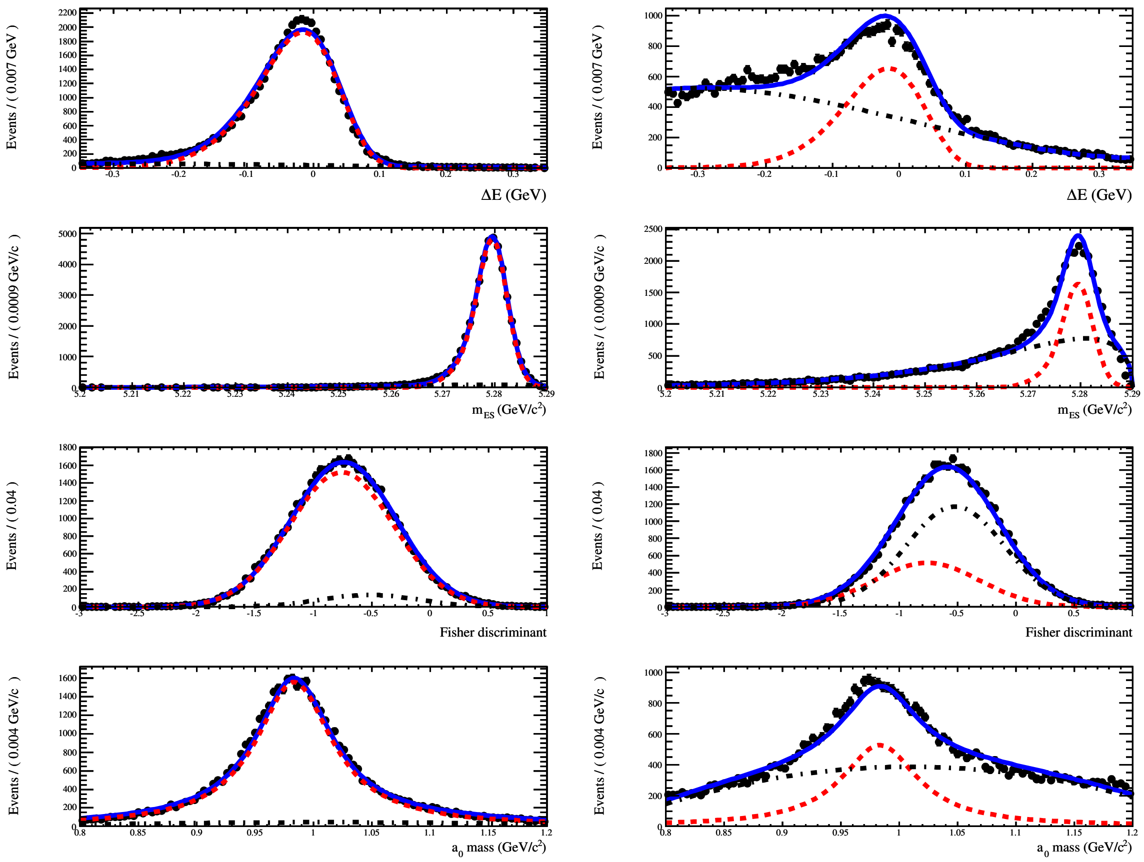

Figure 4.2: *Fits to the separate signal (left column) and SxF (right column) samples for the $\eta \to \pi^+\pi^-\pi^0$ subdecay mode. The total PDF is the blue solid line, the pure signal PDF is the red dashed line and the SxF PDF is the black dot-dashed line.*

the fit fraction f is floated. This corresponds to the overall SxF fraction in the MC sample. Thus the overall number of signal and SxF candidates can be estimated. Plots of the resulting distributions are presented in Figures 4.3 and 4.4 for both η decay modes. As will be seen the negative shift away from zero in the ΔE variable for both η channels, as discussed in Section 3.4, is still present. This is thought to be an EMC calibration issue since the effect appears reduced in the $\eta \to \pi^+\pi^-\pi^0$ channel.

The results of the 4D iterations are presented in Table 4.2 for both η subdecay modes. A number of quantities can be extracted from the process in order to test both the validity of the model and the quality of the truth matching algorithm. Firstly, the purity of the truth matched sample can be estimated from the fraction of signal

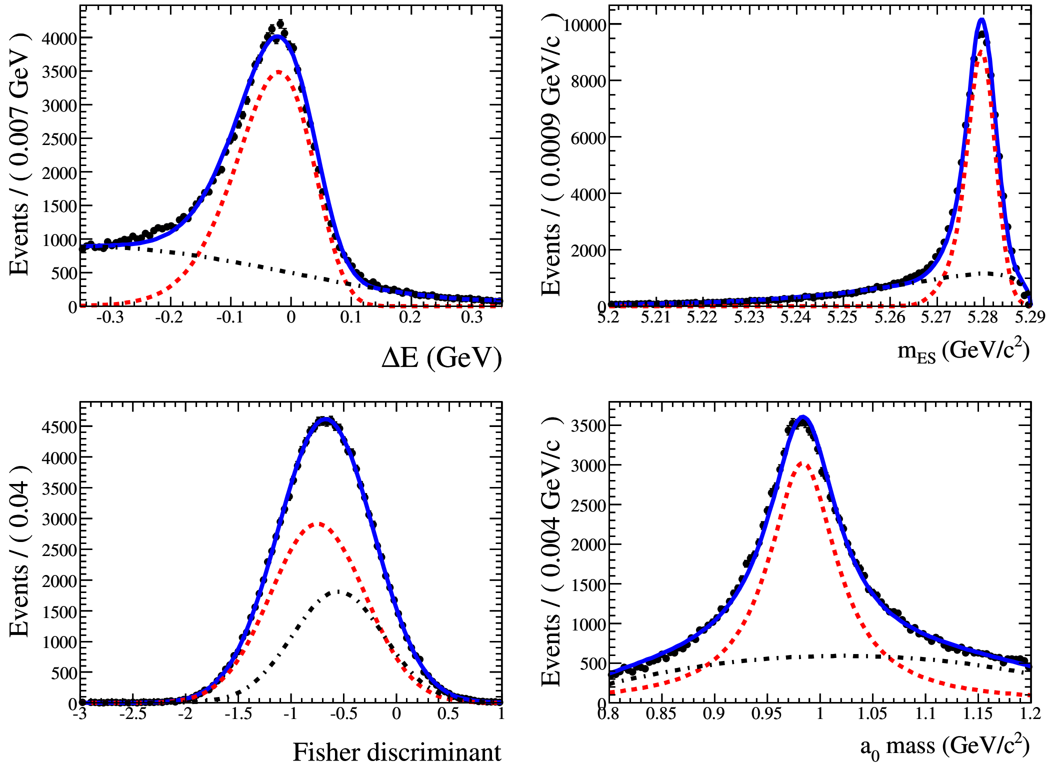

Figure 4.3: *Fits to the whole signal MC sample for the $\eta \to \gamma\gamma$ subdecay mode using the parameterisations extracted from the iterative procedure. The total PDF is the blue solid line, the pure signal PDF is the red dashed line and the SxF PDF is the black dot-dashed line.*

candidates identified in it by the iterative fit. For $\eta \to \gamma\gamma$ this quantity was found to be 94.4±0.1% and for $\eta \to \pi^+\pi^-\pi^0$ it was slightly lower at 92.2±0.2%. In addition to this the fraction of signal events in the non-TM sample can be estimated by similar means. Thus by combining these it is possible to estimate the total number of true candidates in the entire sample (TM and non-TM). For the procedure to be a fair estimate this combined number cannot be greater than the number of events. This is a key test as there cannot be more than one true candidate per event (as verified in Table 4.2). This being the case, the yield associated with the true signal PDF will automatically represent the true number of signal events even in the presence of multiple candidates. Finally the efficiency of the truth matching can be estimated by calculating the number of true signal candidates fitted in the TM sample as a fraction of those fitted in the overall sample. For $\eta \to \gamma\gamma$ this was found to be 87.9±0.8%, while for $\eta \to \pi^+\pi^-\pi^0$ it was worse, measured at

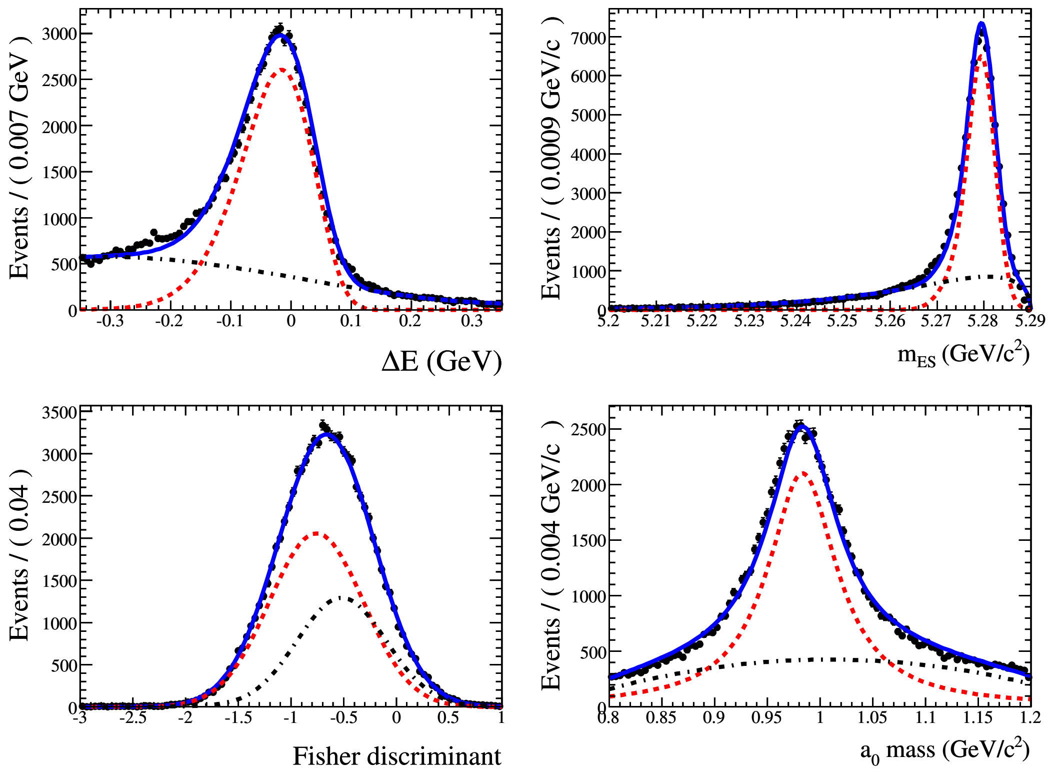

Figure 4.4: *Fits to the whole signal MC sample for the $\eta \to \pi^+\pi^-\pi^0$ subdecay mode using the parameterisations extracted from the iterative procedure. The total PDF is the blue solid line, the pure signal PDF is the red dashed line and the SxF PDF is the black dot-dashed line.*

74.1±0.3%. From the cross-check fits to the total sample the overall signal fraction for $\eta \to \gamma\gamma$ was calculated to be 62.4±0.2%. From this the efficiency for true signal MC was estimated to be 15.44±0.06%, this efficiency will be used later to calculate the branching fraction for the mode resulting from our fit results. The efficiency is defined as the number of signal events fitted in the overall signal MC sample as a fraction of the total number of events generated (531,000). It is noted that in this case the combined number of signal candidates from both fits does not agree within error with the number from the fit to the total sample. This may indicate some systematic issues but since the agreement is still good to within 1% no further investigation was carried out. For $\eta \to \pi^+\pi^-\pi^0$ the signal fraction comes out to be 62.5±0.2%, giving a true signal MC efficiency of 10.83±0.05%. The results of the iteration method are clearly dependent on the selection optimisation procedure and so the signal MC efficiencies are also presented in Table 4.12, Section 4.6.6.

Table 4.2: *The results, for both η decay modes, of the 4D iterative fit to determine signal model shape parameters and separate out signal and SxF candidates. All quantities are expressed in terms of numbers of candidates.*

Fit Quantity	$\eta \to \gamma\gamma$	$\eta \to \pi^+\pi^-\pi^0$
Number of events in sample	96545	65443
Number of B candidates in TM sample	76393	46214
Fitted signal fraction in TM sample (purity) (%)	94.4±0.1	92.2±0.2
Number of fitted signal candidates in TM sample	72115±76	42609±92
Number of B candidates in non-TM sample	55011	45776
Fitted signal fraction in non-TM sample (%)	16.2±0.2	31.5±0.3
Number of fitted signal candidates in non-TM sample	8912±110	14419±137
Combined number of signal candidates from both fits	81027±134	57029±166
Total B candidates in sample	131404	91990
Overall fitted signal fraction in total sample (%)	62.4±0.2	62.5±0.2
Fitted number of signal candidates in total sample	81996±263	57494±184
Efficiency of TM algorithm (%)	87.9±0.8	74.1±0.3

4.4.4 Strategy for SxF Component in Final Fit and the Treatment of Multiple Candidates

As a result of the iteration process we have achieved, as shown in Figures 4.1 and 4.2, independent parameterisations for true signal and SxF and the signal model is therefore maximally distinct from background.

Candidates covered by the SxF model can be treated via one of three possible options:

- Fit the combined signal/SxF model, $(1-f)S(p_i)+fX(q_j)$, with fixed parameters p_i, q_j and overall SxF fraction f as defined previously, to extract a yield for all candidates N_{tot} and define the signal yield in terms of the signal fraction based on the results from MC, i.e. $N_{sig} = (1-f)N_{tot}$
- Fix the SxF yield to a given value, allowing only the signal yield to float. Since a low yield is expected, the SxF would be fixed to zero, i.e. only using the signal shape, $S(p_i)$, in the fit. It would be assumed that any SxF events will be absorbed by other background PDFs, mainly continuum. The signal yield would therefore be given directly by the fit.

- Allow both signal and SxF yields to float independently of any common fraction, $N_{sig}S(p_i)+N_{SxF}X(q_j)$. The signal yield N_{sig} would once again be resolved directly from the fit.

The first option was originally preferred as it made full use of the information extracted from the iteration procedure while modelling any SxF events that might be present in the final sample. It was pointed out, however, that in a low background limit such a procedure may violate Poisson statistics. This could occur in such a situation since the error is estimated by scaling down the fitted error by the signal fraction from the iterative fit. Thus in a signal dominated situation this scaled error estimate could be lower than $\sqrt{N}$ for signal yield N. Although in this analysis the backgrounds were large, it was decided that the most correct procedure statistically is number three, floating both signal and SxF yields independently. However the fit was found to be too unstable for this to be viable. This was mainly due to large correlations between the SxF component and the continuum model. The decision was finally made, in the light of this, to proceed with option two, where the SxF shape is fixed in the fit. Since we expect no signal (and hence no SxF) the SxF yield was finally fixed to zero. Any SxF events which did occur would then be absorbed by the continuum model. Thus the only signal model component used in the final fit was the pure component, the yield from which would directly correspond to the number of signal events, even in the presence of multiple candidates.

The lack of signal contamination from SxF for this model was verified using a pure toy MC study whereby a known quantity of events were generated from the SxF PDF. Should these leak into the signal shape a pull bias would be evident in the result. The results of this study are presented in Table 4.3. No significant bias as a result of SxF contamination was detected. As such the decision was taken to proceed with this fitting strategy. The plots associated with this study are presented in Appendix A.1.

Table 4.3: *The results of the toy studies to assess the potential bias arising from not explicitly including an SxF component in the fit. The signal yield pull is represented by P_{sig} and the signal yield statistical error by $\sigma_{N_{sig}}$. The width of the pull distribution is given by $\Gamma_{P_{sig}}$. The fitted and expected signal yields are given by N_{sig} and G_{sig} respectively. The generated SxF yield is given by G_{sxf}. Finally the number of toy experiments run is given by N_{toy}.*

Toy type	G_{sig} (events)	G_{sxf} (events)	N_{toy}	N_{sig} (events)	P_{sig}	$\Gamma_{P_{sig}}$	$\sigma_{N_{sig}}$ (events)
$\eta \to \gamma\gamma$	0	10	489	1.15±0.85	−0.01±0.04	1.00±0.03	18.81±0.09
$\eta \to \pi^+\pi^-\pi^0$	0	10	499	−0.31±0.58	−0.13±0.05	1.09±0.03	12.01±0.09

4.5 Continuum and Charmed B Decay Background Models

In this analysis the continuum light quark background shape was modelled in two steps. In the initial step, a parameterisation was based on a combination of large fully simulated MC samples for $e^+e^- \to q\overline{q}$ $(q = u, d, s, c)$. This ensures that the shapes are based solely on continuum-like events and allows us to model events which lie kinematically underneath the signal peak in m_{ES} and ΔE. The MC is not perfect but is sufficient to provide an initial estimate for the shape parameters, providing the fit sensible starting values to ensure convergence for the next step of the process. This second step of the parameterisation was to float the majority of continuum shape parameters in the final fit to the on-resonance data. Therefore the final parameterisation was based on the on-resonance data sidebands i.e. events outside the expected signal region in the fit variables. Since all B background shapes are held fixed in the final fit any correlations between continuum and B backgrounds are therefore minimised since the continuum shape should adapt to exclude contamination from these modes. The continuum model is tested using fits to off-resonance data, which are expected to contain no B decays. The results of this are presented in the validation section.

The backgrounds from charmed B decays, i.e. those involving a $b \to c$ quark transition, are not expected to peak in the m_{ES} and ΔE or m_{a_0} variables. This is because

Table 4.4: *The PDFs used to model each variable to fit the bulk background distributions for both η subdecay modes. BW = Breit-Wigner, BG = Bifurcated Gaussian, SG = Single Gaussian and Pn = n^{th} Order Polynomial.*

Decay channel	ΔE	$m_{\rm ES}$	$a_0^{\pm}$ mass	Fisher
$\eta \to \gamma\gamma$ subdecay mode				
B^+B^-	P3	Arg	P2	BG
$B^0\overline{B}^0$	P3	Arg	P2	BG
Combined $q\overline{q}$ $(q = u, d, s, c)$	P2	Arg	BW + P2	SG + BG
$\eta \to \pi^+\pi^-\pi^0$ subdecay mode				
B^+B^-	P3	Arg	P2	BG
$B^0\overline{B}^0$	P3	Arg	P2	BG
Combined $q\overline{q}$ $(q = u, d, s, c)$	P2	Arg	BW + P4	SG + BG

it is unlikely that a significant number of events of this type will produce a combination of decay products with the correct kinematics to emulate the signal mode. They are therefore treated in much the same way as continuum. They are modelled using centrally produced MC, stripping out all charmless decay contributions. Charged and neutral B decays are modelled separately and all shape parameters are held fixed in the final fit, unlike the continuum case.

The models for both background sources are almost identical for the two η subdecay modes and are presented in Table 4.4. Kinematic variables $m_{\rm ES}$ and ΔE are modelled using an ARGUS threshold function and n^{th} order Chebychev polynomial, respectively.

The Fisher variable is modelled in the charmed B case using a Bifurcated Gaussian. For the continuum this is modelled with a more complex function in order to correctly model the low side tail often present in this variable in continuum events. A central Gaussian is used in conjunction with a Bifurcated Gaussian to model the tail. Both Gaussians are required to have the same central mean.

The a_0 mass variable is modelled for charmed B decays using an n^{th} order Chebychev polynomial, since no peaking behaviour is observed in the MC. For continuum the situation is complicated by the fact that there may be real a_0 mesons produced through processes other than the desired signal mode. In order to minimise any bias in this case a Breit-Wigner component is included, with shape parameters taken from the signal MC (pure shape component), together with the standard polynomial. The Breit-Wigner parameters are then fixed in the final fit, but the fraction of the PDF taken up by this component is allowed to float since the amount of a_0 production is not known *a priori*. Plots showing MC distributions and PDFs for continuum and charmed B backgrounds are presented in Appendix B.

In the final fit the yield for continuum events is floated. For toy MC purposes a prediction is calculated as follows:

$$N_{udsc} = (\mathcal{L}_{INT} \times \sigma_{uds} \times \epsilon_{uds}) + (\mathcal{L}_{INT} \times \sigma_{c\overline{c}} \times \epsilon_{c\overline{c}}) \ , \tag{4.5}$$

where the total contribution N_{udsc} is the sum of the separate contributions estimated from $u\overline{u}, d\overline{d}, s\overline{s}$ and $c\overline{c}$ MC. $\mathcal{L}_{INT}$ is the total integrated luminosity of the data sample, σ_X is the production cross-section for each mode and ϵ_X is the efficiency for each mode passing the selection criteria.

It should be noted that for the $\eta \to \gamma\gamma$ case the predicted yield used for toy experiments is scaled down by 20% to account for a previously observed problem modelling low multiplicity events in the MC leading to an incorrectly large sample passing the event selection criteria. This was discovered in an earlier iteration of the analysis when the background contributions predicted from the MC were found to to be significantly higher than the number of events present in the on-resonance dataset for the same selection criteria.

The expected charmed B decay yield in the final fit and for toy studies is calculated as follows:

$$N_{charm} = N_{B\overline{B}} \times 0.5 \times \epsilon_{charm} \ , \tag{4.6}$$

where N_{charm} is the yield for either charged or neutral mode, $N_{B\overline{B}}$ is the total number of $B\overline{B}$ events in the data, calculated using a process known as B-counting [45] and

ϵ_{charm} is the efficiency of each mode for passing the selection cuts. The sample is assumed to contain 50% B^+B^- and 50% $B^0\overline{B}^0$, hence the factor of 0.5 above. Breakdowns of the predicted contributions from these background components for both η decay modes are presented in Tables 4.5 and 4.6.

Table 4.5: *The efficiencies and predicted contributions from charmed B decay and continuum backgrounds for the optimised selection criteria in the $\eta \to \gamma\gamma$ case. The errors on the B candidate multiplicities are calculated as RMS/$\sqrt{N}$.*

Mode	MC sample size	Events passing Selection	Efficiency ($\times 10^{-5}$)	cross section (nb)	Signal box contribution (events)	B Candidate Multiplicity
uds	696M	79693±282	11.45± 0.04	2.09	65238†	1.184±0.002
$c\bar{c}$	581M	14400±120	2.48± 0.02	1.30	8773†	1.172±0.004
B^+B^-	556M	1425±38	0.26± 0.01	0.535*	487	1.14±0.01
$B^0\overline{B}^0$	552M	832±29	0.15± 0.01	0.535*	286	1.20±0.02

**The cross-sections for B decays are included for informational purposes only, the number of B mesons used is actually taken from B-counting. †The continuum contributions for this mode are scaled down by 20% to correct for a previously observed MC problem.*

Table 4.6: *The efficiencies and predicted contributions from charmed B decay and continuum backgrounds for the optimised selection criteria in the $\eta \to \pi^+\pi^-\pi^0$ case. The errors on the B candidate multiplicities are calculated as RMS/$\sqrt{N}$.*

Mode	MC sample size	Events passing Selection	Efficiency ($\times 10^{-5}$)	cross section (nb)	Signal box contribution (events)	B Candidate Multiplicity
uds	696M	18423±136	2.65± 0.02	2.09	18852	1.245±0.005
$c\bar{c}$	581M	8484±92	1.46± 0.02	1.30	6461	1.23±0.01
B^+B^-	556M	1201±35	0.22± 0.01	0.535*	411	1.19±0.02
$B^0\overline{B}^0$	552M	856±29	0.16± 0.01	0.535*	294	1.20±0.02

**The cross-sections for B decays are included for informational purposes only, the number of B mesons used is actually taken from B-counting.*

4.6 Charmless B Decay Background Model

Charmless decays, i.e. those which do not involve a $b \to c$ quark transition, form a potentially peaking background component in the analysis. As such they are the hardest to distinguish from the desired signal decay. In some cases, where the final state of the background mode is the same as for signal, the situation can be further complicated by quantum mechanical interference between overlapping signal and background resonances.

4.6.1 Decays to the Same Final State as Signal

Some of the expected major charmless backgrounds for the analysis can be predicted by looking at the Dalitz plane (DP) for the $\eta\pi^+\pi^0$ final state, shown in Figure 4.5.

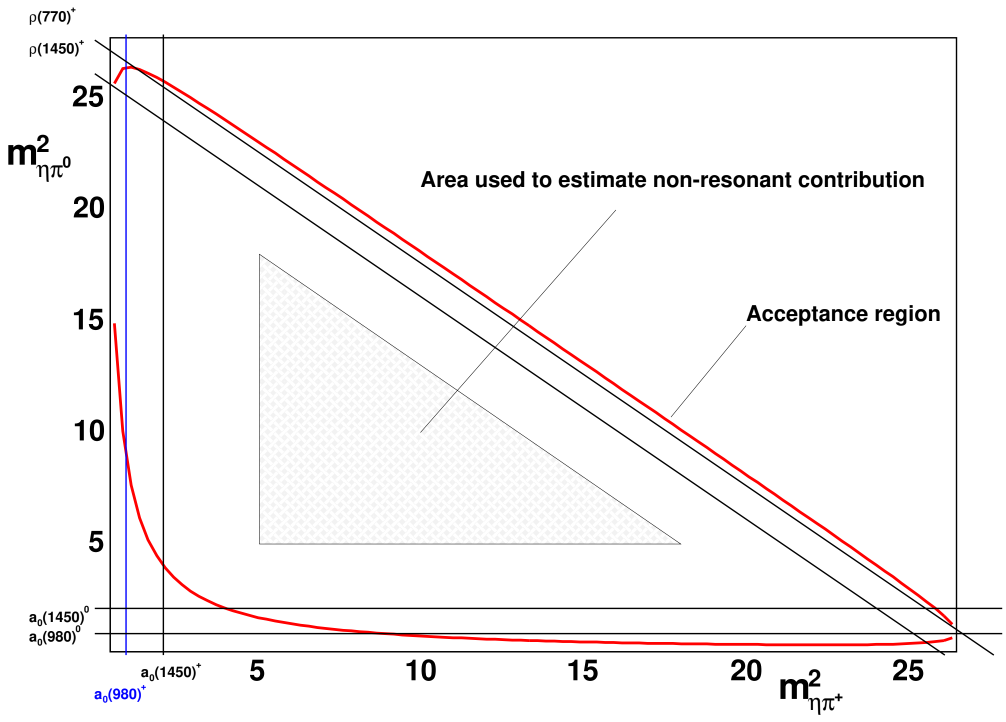

Figure 4.5: *A sketch of the $\eta\pi^+\pi^0$ Dalitz plane showing all major interfering backgrounds. The signal mode is shown in blue with the kinematic acceptance window in red.*

Our signal mode is a resonance in the $m^2_{\eta\pi^+}$ variable, shown in blue in the plot. Two other low-mass resonances exist along the $\eta\pi^0$ and $\pi^+\pi^0$ axes. The first corresponds to the neutral a_0^0 in the decay $B^+ \to a_0(980)^0\pi^+$ and the second to the charged ρ^+ meson in the decay $B^+ \to \rho(770)^+\eta$. Resonances are also thought to exist corresponding to higher mass states of each of the above; these are the $a_0(1450)^{+/0}$ and $\rho(1450)^+$ respectively. The B decay processes involved are analogous to those for the lower mass resonances in each case. There is also a contribution from the theorised non-resonant $B^+ \rightarrow \eta\pi^+\pi^0$ decay, events from which are evenly distributed over the available phase space.

When DP resonances overlap the two decay paths are indistinguishable, rendering it difficult to separate signal from background. The same goes for the non-resonant background, events from which are present underneath all resonances. These background decays will peak in the same way as the signal decay in m_{ES}, ΔE and Fisher, leaving the a_0 mass variable as the principal basis for discrimination. In this analysis these modes are tackled by modelling them in the four fit variables using exclusive MC samples and including the PDFs in the final fit at a fixed yield, the yield being based on the latest measurements or theoretical predictions. Therefore peaking events not accounted for by the background PDFs remain as candidate signal events, assuming they match the parameters required by the signal PDF.

Fortunately for these modes the DP predicts that, with the exception of the non-resonant decay, only one resonance will overlap our signal mode. The overlap regions for the $\rho(770)^+$, $a_0(980)^0$ and $a_0(1450)^0$ are excluded by the kinematic limits of phase space, shown by the red line in the figure. Thus only the decay $B^+ \rightarrow \rho(1450)^+\eta$ should provide an undistorted peaking background. As such the non-resonant and $\rho(1450)^+$ modes are individually modelled in the final fit. Due to the proximity of the $\rho(770)^+$ resonance to the kinematic boundary it is possible that mis-reconstructed events from this decay mode will contaminate the signal region near the boundary. As such this decay was also modelled explicitly in the fit. It is also theoretically possible for the $a_0(1450)^+$ resonance to interfere quantum mechanically with our signal resonance due to their overlapping tails. However, since the yields for both modes will be shown to be consistent with zero, it was decided to neglect any interference effects.

4.6.2 Decays to Other Final States

The other potential source of charmless background comes from mis-reconstructed events with a different final state from the signal mode. In order to identify these backgrounds, two methods were employed.

Firstly, the decays were identified which could constitute the same final state as signal with a simple degree of mis-reconstruction. The categories chosen were as follows: decays containing an extra particle in the final state, decays with one less particle in the final state, decays which are the same except for particle ID (i.e. contain a charged kaon instead of a charged pion) and finally those which contain a mis-reconstructed η.

The second method used to identify decays was aimed at detecting those not in the above categories. This was done by applying tight selections in the seven analysis variables (m_{ES}, ΔE, $\mathcal{F}$, m_{a_0}, m_η, $|cos(\theta_{TB})|$ and PID) to reduce the generic B MC sample down to only the handful of events which are underneath the signal peak, and therefore hardest to distinguish. The events actually generated by the MC were then identified. A number of background modes were discovered. Different backgrounds were discovered for the two η subdecay modes. Note that the selection values yielded from this optimisation were only used for background identification, the main selection optimisation is done with the fit itself for greater accuracy.

All background modes discovered from these two methods were tested by applying the optimal fit selection values to exclusive MC samples and predicting their signal box yield. If the yield was less than one event, the mode was assumed to be negligible and not included in the fit.

4.6.3 Expected Charmless Background Yields

The expected charmless background modes and their associated predicted yields are presented for both η subdecay modes in Tables 4.7 and 4.8. The yields are calculated as follows:

$$N_{charmless} = N_{B\overline{B}} \times \mathcal{B} \times \epsilon_{charmless} \ , \tag{4.7}$$

where $N_{charmless}$ is the yield for the mode in question, $N_{B\overline{B}}$ is the total number of

$B\overline{B}$ events in the data, $\mathcal{B}$ is the branching fraction for that particular mode and $\epsilon_{charmless}$ is its efficiency for passing the selection cuts.

Where a measurement or upper limit for the branching fractions of the different modes existed, they were used as the basis for the estimated yield. In the upper limit case a value of 50%±50% of the limit was taken as the central value.

Where no measurement or limit existed the branching fractions were estimated in a number of ways. The first method of obtaining an estimate is by comparing the mode in question with a similar mode with a known BF or limit, a reference mode. This method is applied to, for example, $B^0 \to a_0^{\pm}\rho^{\mp}$ and $B^{\pm} \to a_0^0\rho^{\pm}$. The reference mode should have one particle different and $B^0 \to a_0^{\pm}\pi^{\mp}$ and $B^{\pm} \to a_0^0\pi^{\pm}$, respectively, are used. For all three of these modes BFs are not available and limits have been set [35]. These limits are multiplied by the square of the ratio of the form factors [59] [60] to scale the limit of the reference mode to give a preliminary estimate of the limit for the mode in question. The calculations are:

- For $B^0 \to a_0^{\pm}\rho^{\mp}$, $B^0 \to a_0^{\pm}\pi^{\mp}$ is used as a reference mode. The current limit for the reference mode has been set at 2.8×10^{-6} [35]. Therefore the preliminary estimate of the limit on the BF is

$$(209/131)^2 \times 2.8 \times 10^{-6} = 7.1 \times 10^{-6}. \tag{4.8}$$

- For $B^{\pm} \to a_0^0\rho^{\pm}$, $B^{\pm} \to a_0^0\pi^{\pm}$ is used as a reference mode. The limit for the reference mode has been reported as 5.8×10^{-6} [35]. Therefore the calculated preliminary limit is

$$(209/131)^2 \times 5.8 \times 10^{-6} = 1.5 \times 10^{-5}. \tag{4.9}$$

Table 4.7: *The charmless B decay modes considered as potential backgrounds that survive the selection cuts and appear in the signal region for the $\eta \to \gamma\gamma$ decay mode. The size of the samples and the efficiency of the skim cuts along with efficiency of the selection cuts are presented. The values for the assumed BF are obtained in several ways as explained in the text. For all of the modes where only an upper limit is available a value of 50% of the limit ± 50% (as per text) is used to calculate the final contribution to the signal box. All of the modes which include an η have had their BF multiplied by 0.4 when calculating their final contribution to take account of the BF of $\eta \to \gamma\gamma$. The errors on the B-candidate multiplicities are calculated as RMS/$\sqrt{N}$. Branching Fraction Limits marked 'DP' have been extracted from Dalitz Plot fits which will be discussed later. * Includes $\mathcal{B}(a_0(1450)^+ \to \eta\pi^+)$.*

Decay channel	Number of events in the sample	Number of events after all cuts	Efficiency of the selection cuts (%)	BF/Limit	Number of events in signal box	B Candidate Multiplicity
Same final state						
$B^\pm \to \rho^\pm\eta$	134000	4732±69	3.53±0.05	$5.3^{+1.2}_{-1.1} \times 10^{-6}$ [35]	28±6	1.44±0.01
$B^\pm \to \eta\pi^\pm\pi^0$ (non-res)	175000	1003±32	0.57±0.02	$< 6.0 \times 10^{-6}$ (DP)	3±3	1.31±0.02
$B^\pm \to \rho^\pm(1450)\eta$	208000	9340±97	4.49±0.05	$2.65 \pm 2.65 \times 10^{-6}$	18±18	1.51±0.01
$B^\pm \to a_0^\pm(1450)\pi^0$	134000	5054±71	3.77±0.05	$< 4.3 \times 10^{-6}$ (DP)*	12±12	1.34±0.01
Same final state +1						
$B^0 \to a_0^\pm\rho^\mp$	134000	4221±65	3.15±0.05	$< 7.13 \times 10^{-6}$	17±17	1.21±0.01
$B^\pm \to a_0^0\rho^\pm$	134000	990±31	0.74±0.02	$< 14.8 \times 10^{-6}$	8±8	1.29±0.02
Same final state −1						
$B^0 \to \eta\pi^0$	143000	12342±111	8.63±0.08	$< 1.3 \times 10^{-6}$ [35]	9±9	1.50±0.01
Mis-reconstructed η						
$B^\pm \to \pi^\pm\pi^0\pi^0$(non-res)	1541000	8849±94	0.57±0.01	$< 4.6 \times 10^{-5}$	50±50	1.29±0.01
$B^\pm \to \rho^\pm\pi^0$	2882000	63808±253	2.21±0.01	$10.8^{+1.4}_{-1.5} \times 10^{-6}$ [35]	91^{+12}_{-13}	1.491±0.003
$B^\pm \to a_1^\pm\pi^0$	143000	1741±42	1.22±0.03	$1.66 \pm 0.36 \times 10^{-5}$	77±17	1.25±0.01
$B^0 \to \pi^0\pi^0$	2882000	117279±342	4.07±0.01	$1.31 \pm 0.21 \times 10^{-6}$	20±3	1.599±0.003
Inclusive$B \to X_{sd}\gamma$	1331000	229±15	0.017±0.001	$1.64^{+0.58}_{-0.45} \times 10^{-4}$ [57]	11^{+4}_{-3}	1.35±0.04
Inclusive$B \to X_{su}\gamma$	1463000	279±17	0.019±0.001	$1.64^{+0.58}_{-0.45} \times 10^{-4}$ [57]	12^{+4}_{-3}	1.41±0.05

Table 4.8: *The charmless B decay modes considered as potential backgrounds that survive the selection cuts and appear in the signal region for the $\eta \to \pi^+\pi^-\pi^0$ mode. The size of the samples and the efficiency of the skim cuts along with efficiency of the selection cuts are presented. The values for the assumed BF are obtained in several ways as explained in the text. For all of the modes where only an upper limit is available a value of 50% of the limit ± 50% (as per text) is used to calculate the final contribution to the signal box. All of the modes which include an η have had their BF multiplied by 0.23 when calculating their final contribution to take account of the BF of $\eta \to \pi^+\pi^-\pi^0$. The errors on the B-candidate multiplicities are calculated as $RMS/\sqrt{N}$. Branching Fraction Limits marked 'DP' have been extracted from Dalitz Plot fits which will be discussed later. * Includes $\mathcal{B}(a_0(1450)^+ \to \eta\pi^+)$.*

Decay channel	Number of events in the sample	Number of events after all cuts	Efficiency of the selection cuts (%)	BF/Limit	Number of events in signal box	B Candidate Multiplicity
Same final state						
$B^\pm \to \rho^\pm\eta$	134000	3026±55	2.26±0.04	$5.3^{+1.2}_{-1.1} \times 10^{-6}$ [35]	10^{+2}_{-2}	1.42±0.02
$B^\pm \to \eta\pi^\pm\pi^0$ (non-res)	175000	404±20	0.23±0.01	$< 6.0 \times 10^{-6}$ (DP)	1±1	1.51±0.01
$B^\pm \to \rho^\pm(1450)\eta$	143000	4307±66	3.01±0.05	$2.65 \pm 2.65 \times 10^{-6}$	7±7	1.35±0.04
$B^\pm \to a_0^\pm(1450)\pi^0$	143000	4102±64	2.87±0.04	$< 4.3 \times 10^{-6}$ (DP)*	5±5	1.38±0.01
Same final state +1						
$B^0 \to a_0^\pm\rho^\mp$	143000	3062±55	2.14±0.04	$< 7.13 \times 10^{-6}$	7±7	1.26±0.01
$B^\pm \to a_0^0\rho^\pm$	143000	662±26	0.46±0.02	$< 14.8 \times 10^{-6}$	3±3	1.30±0.03
Same final state −1						
$B^0 \to \eta\pi^0$	143000	8077±90	5.65±0.06	$< 1.3 \times 10^{-6}$ [35]	3±3	1.44±0.01
Mis-reconstructed η						
$B^\pm \to \pi^\pm\pi^0\pi^0$(non-res)	1541000	298±17	0.019±0.001	$< 4.6 \times 10^{-5}$	2±2	1.25±0.04
$B^\pm \to \rho^\pm\pi^0$	2882000	1615±40	0.056±0.001	$10.8^{+1.4}_{-1.5} \times 10^{-6}$ [35]	2±0	1.38±0.02
$B^\pm \to \rho^\pm\omega$	675000	578±24	0.086±0.003	$10.6^{+2.6}_{-2.3} \times 10^{-6}$ [58]	3±1	1.24±0.02
$B^\pm \to b_1^\pm\pi^0$	143000	572±24	0.40±0.02	$1.1 \pm 1.1 \times 10^{-5}$	17±17	1.30±0.03
Inclusive$B \to X_{sd}\gamma$	1331000	65±8	0.005±0.001	$1.64^{+0.58}_{-0.45} \times 10^{-4}$ [57]	3±1	1.32± 0.09
Inclusive$B \to X_{su}\gamma$	1463000	89±9	0.006±0.001	$1.64^{+0.58}_{-0.45} \times 10^{-4}$ [57]	4±1	1.42±0.10

- For $B^{\pm} \to a_1^{\pm}\pi^0$ the BF is expected to be 50% of the measured value for the related mode $B^0 \to a_1^{\pm}\pi^{\mp}$ [35] from isospin symmetry. The error is naively assumed to scale as $\sqrt{2}$. The calculated BF is,

$$0.5 \times (33.2 \pm 5.1) \times 10^{-6} = (1.7 \pm 0.4) \times 10^{-5} \tag{4.10}$$

- For $B^{\pm} \to b_1^{\pm}\pi^0$ the BF is estimated by scaling the expected value for $B^{\pm} \to a_1^{\pm}\pi^0$ using the ratio of theoretical predictions for other a_1 and b_1 modes. The resulting prediction is $(11 \pm 11) \times 10^{-6}$.

- For non-resonant $B^{\pm} \to \pi^{\pm}\pi^0\pi^0$ the BF is estimated by scaling the expected value for $B^{\pm} \to \pi^{\pm}\pi^{\mp}\pi^{\pm}$. The upper limit for this mode is 4.6×10^{-6} [8]. The two modes can be related using isospin symmetry, however, there are multiple isospin amplitudes associated with this procedure. As such, a conservative estimate of the limit is taken to be 4.6×10^{-5}.

The second method for estimating charmless background yields relates solely to the modes $B^+ \to a_0(1450)^+\pi^0$ and non-resonant $B^+ \to \eta\pi^+\pi^0$. These modes, with the same final state as signal, can be analysed using specific regions of the Dalitz plane where they are expected to dominate. Fits to these regions are expected to yield a good estimate of their contribution in the signal $a_0(980)^+$ region. The method employed for these fits is discussed in detail later. These modes are expected to give a small contribution in the signal region and as such were excluded from the fit optimisation.

4.6.4 PDFs used to Model Charmless Backgrounds

The PDFs used to model the specific charmless backgrounds involved in the analysis are listed in Tables 4.9 and 4.10 for each η subdecay mode. All were based on exclusive decay MC samples for each mode. Plots of the PDFs for each background mode are presented in Appendix B.

In most cases the PDFs for Fisher were constructed from asymmetric Gaussians and those for the a_0 mass from combinations of Chebychev polynomials and Gaussian or Breit-Wigner peaking components where appropriate.

For the $B^{\pm} \rightarrow a_0^{\pm}(1450)\pi^0$ and non-resonant $B^{\pm} \rightarrow \eta\pi^{\pm}\pi^0$ cases ΔE was modelled using combinations of Chebychev polynomials and Novosibirsk peaking components, while $m_{\rm ES}$ was modelled using combinations of the ARGUS threshold function and Gaussian peaking components. In all other cases the expected correlations between the variables led to the employment of a different method to model the distributions two-dimensionally and thus correctly account for the effect. To do this a non-parametric 'KEYS' PDF was used [61]. This system uses the MC sample to model a given distribution as the summation of n Gaussians for a desired granularity. The method can be implemented in any number of dimensions.

4.6.5 Techniques for Best B Candidate Selection

As previously discussed, SxF is the cause of multiple B candidates in any given signal event. A standard way by which SxF can be reduced is by selecting a 'best' B candidate from each event and discarding the rest. This is done by selecting on any number of candidate quality factors, the assumption being that the true candidate will be closest to the ideal quantity in each chosen variable.

While useful, this method does, by definition, lose a certain fraction of true signal events as well as potentially causing biases due to artificial peaks arising in the background. If, for example, you select on a quantity related to the a_0 mass you essentially exclude background events which are far from the peak, thus only allowing through those which display the most similar structure to the signal. This can cause a signal-like peak in the mass distribution.

In order to ascertain whether such a selection would be beneficial in this analysis, toy MC studies were carried out for fit models constructed using a simple and robust best candidate selection hypothesis. This was done independently for both η decay

Table 4.9: *The full fit model for the $\eta \to \gamma\gamma$ subdecay mode. BW =Breit-Wigner, BG = Bifurcated Gaussian, SG = Single Gaussian, DG = Double Gaussian, NS = Novosibirsk, Arg = ARGUS, Pn = n^{th} order Polynomial and Keys = non-parametric KEYS PDF. The $B^+ \to \mathrm{a}_0^+\pi^0$ and continuum yields are floated in the fit as well as all continuum shape parameters, with the exception of the mean and width of the BW modelling the peaking component in the $\mathrm{a}_0^\pm$ mass variable.*

Decay channel	ΔE	m_{ES}	$\mathrm{a}_0^\pm$ mass	Fisher	Yield (cands)
Signal Model					
$B^+ \to a_0^+\pi^0$	NS	DG	BW	BG	–
Bulk Background Model					
B^+B^-	P3	Arg	P2	BG	557
$B^0\overline{B}^0$	P3	Arg	P2	BG	343
Combined $q\overline{q}$ $(q = u, d, s, c)$	P2	Arg	BW + P2	SG + BG	–
Charmless B Decays					
$B^\pm \to \rho^\pm\eta$	2D Keys		P2	BG	41
$B^\pm \to \rho^\pm(1450)\eta$	2D Keys		P2	BG	27
$B^\pm \to a_0^\pm(1450)\pi^0$	P2 + DG	Arg + SG + BG	P3	BG	17
$B^\pm \to \eta\pi^\pm\pi^0$ (non-res)	P2 + NS	Arg + SG	P3	BG	3
$B^0 \to \pi^0\pi^0$	2D Keys		P3	BG	32
$B^\pm \to a_1^\pm\pi^0$	2D Keys		P3	BG	96
$B^0 \to a_0^\pm\rho^\mp$	2D Keys		BW + P2	BG	21
$B^\pm \to a_0^0\rho^\pm$	2D Keys		P2	BG	11
$B^0 \to \eta\pi^0$	2D Keys		P1	BG	13
$B^\pm \to \rho^\pm\pi^0$	2D Keys		BW + P2	BG	135
$B^\pm \to \pi^\pm\pi^0\pi^0$ (non-res)	2D Keys		P3	BG	64
Inclusive $B \to X_s\gamma$ (comb)	2D Keys		P2	BG	31

Table 4.10: *The full fit model for the $\eta \rightarrow \pi^+\pi^-\pi^0$ subdecay mode. BW =Breit-Wigner, BG = Bifurcated Gaussian, SG = Single Gaussian, DG = Double Gaussian, NS = Novosibirsk, Arg = ARGUS, Pn = n^{th} order Polynomial and Keys = non-parametric KEYS PDF. The $B^+ \rightarrow a_0^+\pi^0$ and continuum yields are floated in the fit as well as all continuum shape parameters, with the exception of the mean and width of the BW modelling the peaking component in the $a_0^\pm$ mass variable.*

Decay channel	ΔE	$m_{\rm ES}$	$a_0^\pm$ mass	Fisher	Est. Yield (cands)
Signal Model					
$B^+ \rightarrow a_0^+\pi^0$	NS	DG	BW	BG	–
Bulk Background Model					
B^+B^-	P3	Arg	P2	BG	490
$B^0\overline{B}^0$	P3	Arg	P2	BG	354
Combined $q\overline{q}$ ($q = u, d, s, c$)	P2	Arg	BW + P2	SG + BG	–
Charmless B Decays					
$B^\pm \rightarrow \rho^\pm\eta$	2D Keys		P1	BG	15
$B^\pm \rightarrow \rho^\pm(1450)\eta$	2D Keys		P2	BG	11
$B^\pm \rightarrow a_0^\pm(1450)\pi^0$	P2 + DG	Arg + BG	P2	BG	8
$B^\pm \rightarrow \eta\pi^\pm\pi^0$ (non-res)	P2 + NS	Arg + SG	P3	BG	1
$B^\pm \rightarrow b_1^\pm\pi^0$	2D Keys		P3	BG	22
$B^\pm \rightarrow \rho^\pm\omega$	2D Keys		P3+SG	BG	4
$B^0 \rightarrow a_0^\pm\rho^\mp$	2D Keys		BW + P2	BG	8
$B^\pm \rightarrow a_0^0\rho^\pm$	2D Keys		P2	BG	4
$B^0 \rightarrow \eta\pi^0$	2D Keys		P1	BG	5
$B^\pm \rightarrow \rho^\pm\pi^0$	2D Keys		BW + P2	BG	3
$B^\pm \rightarrow \pi^\pm\pi^0\pi^0$ (non-res)	2D Keys		P2	BG	2
Inclusive $B \rightarrow X_s\gamma$ (comb)	2D Keys		P2	BG	9

Table 4.11: *The results of ensemble MC studies comparing a model based on a best candidate selection with a model with no selection applied. The signal efficiency is represented by ϵ_{sig} and the fitted signal yield error by $\sigma_{N_{sig}}$. Note that this error dominated by background statistics and hence is not dependent on any assumed signal branching fraction.*

	$\eta \to \gamma\gamma$		$\eta \to \pi^+\pi^-\pi^0$	
Analysis quantity	No Selection	$\chi^2_{\eta\pi^0}$ Selection	No Selection	$\chi^2_{\eta\pi^0}$ Selection
ϵ_{sig} (%)	15.44	12.45	10.83	8.24
$\sigma_{N_{sig}}$ (events)	17.74	15.44	11.70	9.93
$\epsilon_{sig}/\sigma_{N_{sig}}$ (%/event)	0.87	0.81	0.93	0.83

modes for samples derived from the optimised selection cuts. For each case the significance parameter was taken to be the signal efficiency over the fitted signal yield error, since a zero signal yield is expected. The efficiency was estimated in each case from fits to signal MC for events matching the pure signal shape, thus excluding SxF candidates.

The best candidate selection was based on calculating the combined χ^2 of the η and B daughter π^0 masses with respect to the PDG values of the two. In the case where two or more B candidates are composed of the same η and π^0, i.e. only the charged pion is different, a random selection is made. The combined χ^2 was calculated as follows:

$$\chi^2 = \left(\frac{m_\eta - m_\eta^{PDG}}{\sigma_\eta}\right)^2 + \left(\frac{m_{\pi^0} - m_{\pi^0}^{PDG}}{\sigma_{\pi^0}}\right)^2 , \qquad (4.11)$$

where m_X is the resonance mass in each case and the widths σ_X were estimated from Gaussian fits to the relevant signal MC sample in all cases. The results for the comparison studies are shown in Table 4.11.

As can be seen from the table, the reduced fit error resulting from a best candidate selection is mitigated by a corresponding reduction in signal efficiency. This occurs because the best candidate selection is rejecting a fraction of true signal events and thus reducing overall sensitivity. Based on these results it was decided to proceed without selecting a best candidate. The plots associated with this study are presented in Appendix A.2.

4.6.6 Optimisation of Event Selection

The event selection variables used in the analysis were optimised using a large number of toy MC studies based on models derived from the differing selection hypotheses. Four of the variables were not optimised and were left with a loose selection for fit statistics purposes. These were ΔE, m_{ES}, the a_0 resonance mass and the Fisher discriminant $\mathcal{F}$. The selection requirements placed on the remaining variables; $|\cos(\theta_{TB})|$, η mass and the PID Selector, were optimised using a nested loop. For each configuration of variable values, reduced MC samples were made and the signal efficiency calculated, along with predicted yields for each background mode. New values for the PDF parameters for signal and background modes were calculated in each case for the full fit model. The $B^+ \to a_0(1450)^+\pi^0$ and non-resonant $B^+ \to \eta\pi^+\pi^0$ modes were left out of the optimisation since their contribution was expected to be small and thus not worth the significant processing overhead incurred in including them.

For each selection hypothesis approximately 270 toy experiments were then run and an optimisation parameter (signal efficiency)/(mean fitted signal yield error) was calculated from the results. The efficiency is used instead of the signal yield since we are working under the assumption of zero signal and thus wish to select the point with the best sensitivity in the limit of small signal where background dominates. The mean fitted signal yield error was taken from a Gaussian fit of the distribution of errors derived from the multiple toy experiments. In analysing the results of the optimisation only those with non-significant mean pulls were considered to be valid. A mean pull was defined as significant if the fitted mean pull of a set of toy experiments was larger than three times the statistical error in the fitted value. For $\eta \to \pi^+\pi^-\pi^0$ this excluded 29 out of 625 possible selection hypotheses. For $\eta \to \gamma\gamma$ 21 out of 500 were excluded in the same way.

The values taken for these variables are listed in Table 4.12. For the PID selector increasing number refers to a looser kaon veto. The looser the veto the more efficient the removal of kaons, but the higher the probability of mistakenly rejecting a pion. Scans of the differing optimal values for the four fit variables are presented in Figures 4.6 and 4.7 for the two η decay modes. Each point corresponds to the optimal value with the given selection on the variable under consideration in place calculated over all possible values of the other three variables. This is included for illustrative purposes and the optimal value taken forward for the rest of the analysis is actually calculated over all four variables, shown in the scans as the highest point for each particular variable.

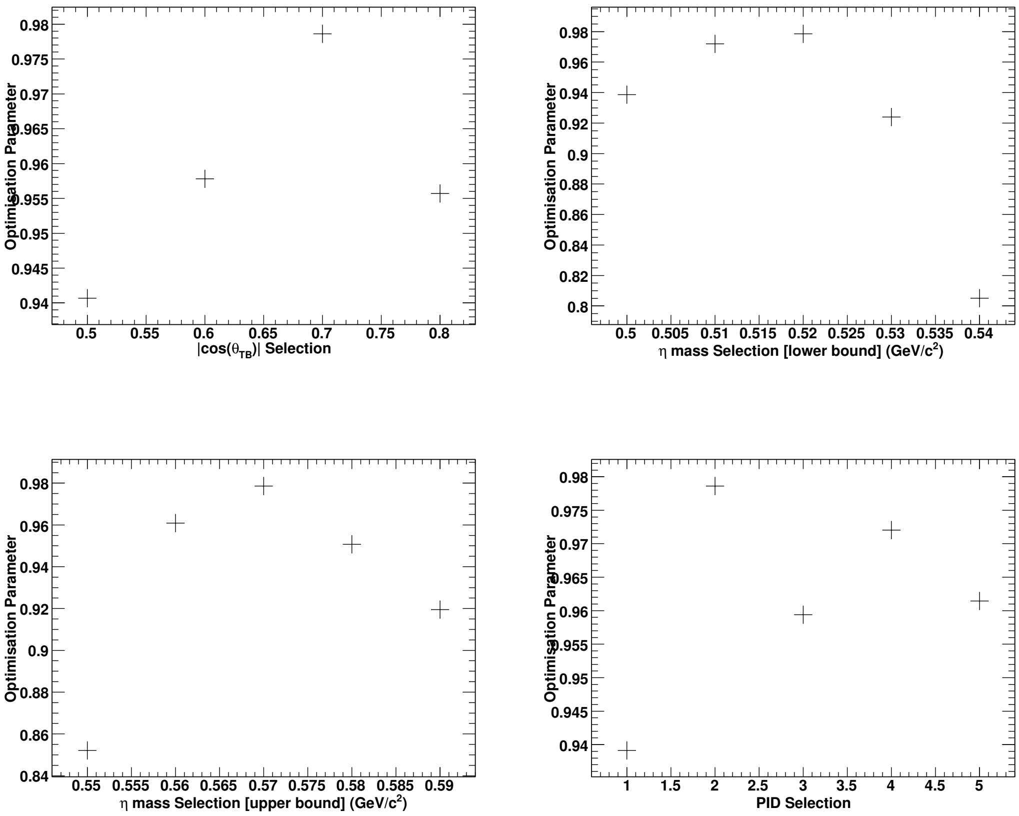

Figure 4.6: *$\eta \to \gamma\gamma$ mode optimisation scans for each relevant analysis variable: (a) $|\cos(\theta_{TB})|$, (b) η mass (lower bound), (c) η mass (upper bound) and (d) Particle ID. Each point corresponds to the optimal value with that particular selection in place calculated over all possible values of the other three variables.*

For both modes the optimisation removes the tails of the η mass distribution. Since

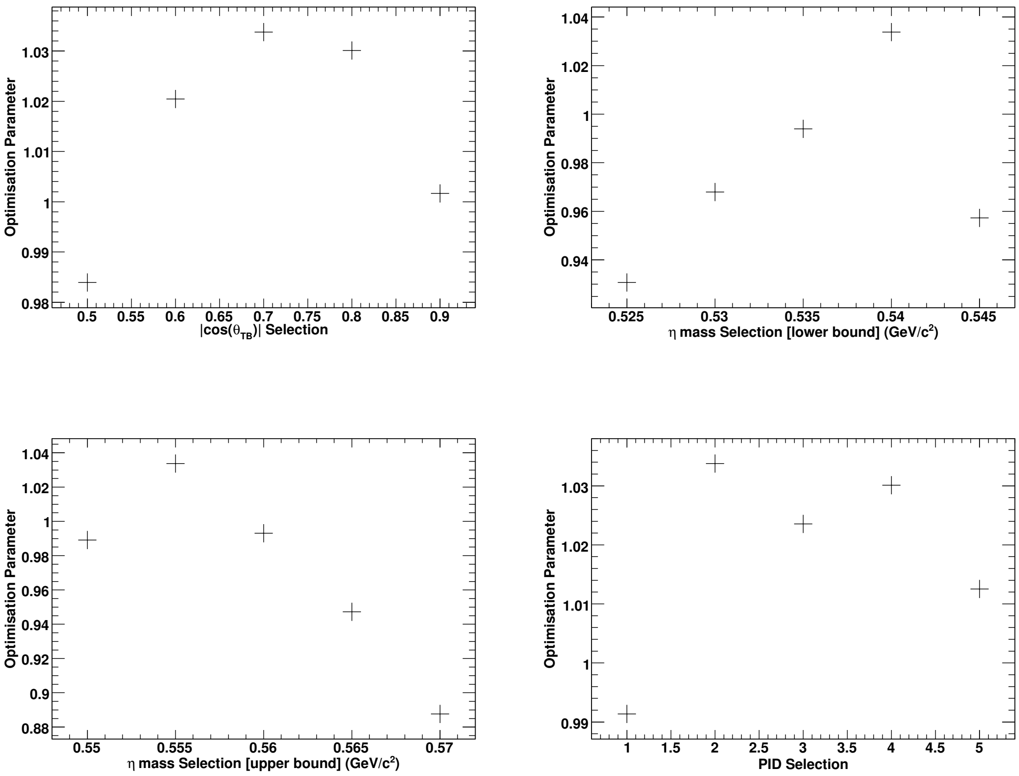

Figure 4.7: *$\eta \to \pi^+\pi^-\pi^0$ mode optimisation scans for each relevant analysis variable: (a) $|\cos(\theta_{TB})|$, (b) η mass (lower bound), (c) η mass (upper bound) and (d) Particle ID. Each point corresponds to the optimal value with that particular selection in place calculated over all possible values of the other three variables.*

the $\eta \to \pi^+\pi^-\pi^0$ resolution is significantly better then the $\eta \to \gamma\gamma$ case the selection is predictably tighter. Both optimisations select the tightest possible kaon veto which is to be expected since this removes the most obvious kaons while not affecting the efficiency too severely. It is noted that a dip in the optimal value occurs in both optimisations for the intermediate kaon veto. A satisfactory explanation for this does not exist although it is not thought to be due to a technical problem with the procedure. Finally, the optimisations reach the same selection for the $|\cos(\theta_{TB})|$ variable with only events satisfying $|cos(\theta_{TB})| < 0.7$ passing. This will remove the vast majority of continuum events while reducing the signal efficiency by a relatively small 30%.

As will be obvious from the plots for the $\eta \to \gamma\gamma$ optimisation, the $|\cos(\theta_{TB})| < 0.9$

Table 4.12: *Final selection and fit variable values after optimisation.*

Analysis Variable	Selection Value	
Fixed	$\eta \to \gamma\gamma$	$\eta \to \pi^+\pi^-\pi^0$
m_{ES} (GeV/c^2)	5.2 – 5.3	5.2 – 5.3
ΔE (GeV)	$< \lvert 0.35 \rvert$	$< \lvert 0.35 \rvert$
$\mathcal{F}$	−3 – 1	−3 – 1
a_0 mass (GeV/c^2)	0.8 – 1.2	0.8 – 1.2
Optimised	$\eta \to \gamma\gamma$	$\eta \to \pi^+\pi^-\pi^0$
$\lvert \cos(\theta_{TB}) \rvert$	< 0.7	< 0.7
η mass (GeV/c^2)	0.52 – 0.57	0.540 – 0.555
PID Selector	1 (veto `trkisKLHVT`)	1 (veto `trkisKLHVT`)
MC Efficiency (%)	15.44±0.06	10.83±0.05

selection hypothesis was omitted from the results. This is because a high proportion of the toys were failing and those which passed reported unexpectedly high efficiency/error scores. This led to suspicions that the results were therefore biased, with the most stable fits being the ones which have the lowest errors. The problem is most likely to be down to it becoming increasingly difficult for the fitter to handle the number of floated continuum shape parameters with such a large sample. The fact that the $\eta \to \pi^+\pi^-\pi^0$ optimisation toys are successful for all cuts adds credence to this since on average 2.5× fewer continuum events were generated for this mode.

4.7 Background Estimation from Dalitz Plane Fits in Data

4.7.1 Preamble

As previously discussed, there exist a number of background modes for which no reliable limit or branching fraction measurement exists. In many cases this problem is circumvented by estimating the contribution using measurements for other analogous modes. In some cases, however, we have the opportunity to obtain a better estimate for the mode from the fit itself. This is achieved by applying the fit based on the same selection criteria as for our signal region but moving the a_0

mass window to a desired alternative region in the $\eta\pi^+\pi^0$ Dalitz plane. In this analysis such a method is used to estimate background contamination from two modes: $B^\pm \to a_0(1450)^\pm\pi^0$ and the non-resonant decay $B^\pm \to \eta\pi^\pm\pi^0$. By using these fits to estimate upper limits for the branching fractions of the modes in question it was possible to predict how many events of each type contribute a background to the analysis of the $a_0(980)$ signal region.

It should be noted that any mode with the same final state as signal can, in principle, quantum mechanically interfere with the signal resonance. However, at the small yields considered here this is expected to be a negligible effect. As such, no steps are taken to estimate it. In all cases the signal mode is assumed to provide negligible contamination to any of the background fits.

4.7.2 $a_0(1450)^+$ Resonance Region Model

This mode, the higher mass a_0 resonance in $\eta\pi^+$, is modelled by shifting the a_0 mass fit window to $1.2 - 2.0\,\mathrm{GeV}/c^2$. This excludes the majority of the $a_0^+(980)$ resonance while also minimising any contribution from the centre of the Dalitz plane, which is mainly populated by non-resonant events. PDFs are constructed for all background modes contributing one or more candidates within this region. There is, in principle, a background contribution from non-resonant $B^\pm \to \eta\pi^\pm\pi^0$, which is the other background mode under study in this section. Any potential contamination is estimated using an iterative fit procedure which will be discussed shortly. The PDF models employed for this fit are presented in Tables 4.13 and 4.14 for $\eta \to \gamma\gamma$ and $\eta \to \pi^+\pi^-\pi^0$ decay modes respectively. The fit model is applied to the full on-resonance dataset processed in an identical way to that used for the main $a_0(980)$ fit.

4.7.3 Non-Resonant Region: $\eta\pi^+\pi^0$ Model

The central region of the DP, away from resonances, is used to estimate the background contribution from the non-resonant decay of the B meson to the signal mode

Table 4.13: *The $B^{\pm} \rightarrow a_0^{\pm}(1450)\pi^0$ fit model for the $\eta \rightarrow \gamma\gamma$ subdecay mode. BW =Breit-Wigner, BG = Bifurcated Gaussian, SG = Single Gaussian, DG = Double Gaussian, NS = Novosibirsk, Arg = ARGUS, Pn = n^{th} order Polynomial and Keys = non-parametric KEYS PDF. The $B^{\pm} \rightarrow a_0^{\pm}(1450)\pi^0$ and continuum yields are floated in the fit as well as all continuum shape parameters.*

Decay channel	ΔE	m_{ES}	$\mathrm{a}_0^{\pm}$ mass	Fisher	Est. Yield (cands)
Signal Model					
$B^{\pm} \rightarrow a_0^{\pm}(1450)\pi^0$	P3 + DG	Arg + SG + BG	P3 + BW	BG	–
Bulk Background Model					
$B^{+}B^{-}$	P3	Arg	P3	BG	5836
$B^0\overline{B}^0$	P3	Arg + SG	P2 + SG	BG	2767
Combined $q\overline{q}$ $(q = u, d, s, c)$	P2	Arg	P3	BG	–
Charmless B Decays					
$B^{\pm} \rightarrow \rho^{\pm}\eta$	2D Keys		P3	BG	132
$B^{\pm} \rightarrow \rho^{\pm}(1450)\eta$	2D Keys		P3	BG	71
$B^{\pm} \rightarrow \eta\pi^{\pm}\pi^0$ (non-res)	P2 + NS	Arg + SG	P2	BG	14
$B^{\pm} \rightarrow \pi^0\pi^0$	2D Keys		P3	BG	56
$B^{\pm} \rightarrow a_1^{\pm}\pi^0$	2D Keys		P3	BG	191
$B^0 \rightarrow a_0^{\pm}\rho^{\mp}$	2D Keys		BW	BG	15
$B^{\pm} \rightarrow a_0^0\rho^{\pm}$	2D Keys		P2	BG	36
$B^0 \rightarrow \eta\pi^0$	2D Keys		P3	BG	24
$B^{\pm} \rightarrow \rho^{\pm}\pi^0$	2D Keys		BW	BG	316
$B^{\pm} \rightarrow \pi^{\pm}\pi^0\pi^0$ (non-res)	2D Keys		P3	BG	361
Inclusive $B \rightarrow X_s\gamma$ (comb)	2D Keys		P2	BG	108

Table 4.14: *The $B^{\pm} \to a_0^{\pm}(1450)\pi^0$ fit model for the $\eta \to \pi^+\pi^-\pi^0$ subdecay mode. BW =Breit-Wigner, BG = Bifurcated Gaussian, SG = Single Gaussian, DG = Double Gaussian, NS = Novosibirsk, Arg = ARGUS, Pn = n^{th} order Polynomial and Keys = non-parametric KEYS PDF. The $B^{\pm} \to a_0^{\pm}(1450)\pi^0$ and continuum yields are floated in the fit as well as all continuum shape parameters.*

Decay channel	ΔE	$m_{\rm ES}$	$a_0^{\pm}$ mass	Fisher	Est. Yield (cands)
Signal Model					
$B^{\pm} \to a_0^{\pm}(1450)\pi^0$	P2 + NS	Arg + SG + BG	P2 + BW	BG	–
Bulk Background Model					
B^+B^-	P3	Arg	P3	BG	3192
$B^0\overline{B}^0$	P3	Arg + SG	P2	BG	1813
Combined $q\overline{q}$ $(q = u, d, s, c)$	P2	Arg	P3	BG	–
Charmless B Decays					
$B^{\pm} \to \rho^{\pm}\eta$	2D Keys		P3	BG	51
$B^{\pm} \to \rho^{\pm}(1450)\eta$	2D Keys		P3	BG	27
$B^{\pm} \to \eta\pi^{\pm}\pi^0$ (non-res)	P2 + NS	Arg + SG	P2	BG	3
$B^{\pm} \to b_1^{\pm}\pi^0$	2D Keys		P3	BG	12
$B^{\pm} \to \rho^{\pm}\omega$	2D Keys		P3	BG	2
$B^0 \to a_0^{\pm}\rho^{\mp}$	2D Keys		BW	BG	6
$B^{\pm} \to a_0^0\rho^{\pm}$	2D Keys		P2	BG	12
$B^0 \to \eta\pi^0$	2D Keys		P3	BG	8
$B^{\pm} \to \rho^{\pm}\pi^0$	2D Keys		BW	BG	18
$B^{\pm} \to \pi^{\pm}\pi^0\pi^0$ (non-res)	2D Keys		P3	BG	23
Inclusive $B \to X_s\gamma$ (comb)	2D Keys		P2	BG	18

Table 4.15: *The $B^{\pm} \to \eta\pi^{\pm}\pi^{0}$ fit model for the $\eta \to \gamma\gamma$ subdecay mode. BW =Breit-Wigner, BG = Bifurcated Gaussian, SG = Single Gaussian, DG = Double Gaussian, NS = Novosibirsk, Arg = ARGUS, Pn = n^{th} order Polynomial and Keys = non-parametric KEYS PDF. The $B^{\pm} \to \eta\pi^{\pm}\pi^{0}$ and continuum yields are floated in the fit as well as all continuum shape parameters.*

Decay channel	ΔE	$m_{\rm ES}$	$a_0^{\pm}$ mass	Fisher	Est. Yield (cands)
Signal Model					
$B^{\pm} \to \eta\pi^{\pm}\pi^{0}$ (non-res)	P3 + NS	Arg + DG	P4	BG	–
Bulk Background Model					
$B^{+}B^{-}$	P3	Arg	P3	BG	938
$B^{0}\overline{B}^{0}$	P3	Arg	P3	BG	397
Combined $q\overline{q}$ $(q = u, d, s, c)$	P3	Arg	SG + P3	BG	–
Charmless B Decays					
$B^{\pm} \to \rho^{\pm}(1450)\eta$	P1 + SG	Arg + SG	P3	BG	1
$B^{\pm} \to a_1^{\pm}\pi^{0}$	2D Keys		P2	BG	2
$B^{\pm} \to \rho^{\pm}\pi^{0}$	2D Keys		BW	BG	1
$B^{\pm} \to \pi^{\pm}\pi^{0}\pi^{0}$ (non-res)	2D Keys		P2	BG	207
Inclusive $B \to X_s\gamma$ (comb)	2D Keys		P3	BG	13

final state, $\eta\pi^{+}\pi^{0}$.

The fit to this mode is prepared in a way analogous to that for the $a_0(1450)^{+}$ resonance, the exception being that in this case, a selection is not solely made on the reconstructed a_0^{+} mass, but also on the invariant mass combinations $M_{\eta\pi^0}$ and $M_{\pi^+\pi^0}$, which form the other 'axes' of the DP. The selection required that the a_0^{+} mass be greater than $2.25\,\mathrm{GeV}/c^2$, with $M_{\pi^+\pi^0}$ and $M_{\eta\pi^0}$ both required to be greater than $\sqrt{5}(\,\mathrm{GeV}/c^2)$. Thus the resonant regions of the plot are stripped away leaving only the central region. Due to this selection, the vast majority of charmless B decays do not contribute significant backgrounds and are therefore excluded from the fit. The models used for the $\eta \to \gamma\gamma$ and $\eta \to \pi^{+}\pi^{-}\pi^{0}$ subdecay modes are presented in Tables 4.15 and 4.16 respectively.

The yield for the non-resonant contribution is extracted using an iterative fit method, to be discussed shortly, from the full on-resonance dataset. This fit also accounts for potential contamination from the tail of the $a_0(1450)^{+}$ resonance.

Table 4.16: *The $B^{\pm} \rightarrow \eta\pi^{\pm}\pi^{0}$ fit model for the $\eta \rightarrow \pi^{+}\pi^{-}\pi^{0}$ subdecay mode. BW =Breit-Wigner, BG = Bifurcated Gaussian, SG = Single Gaussian, DG = Double Gaussian, NS = Novosibirsk, Arg = ARGUS, Pn = n^{th} order Polynomial and Keys = non-parametric KEYS PDF. The $B^{\pm} \rightarrow \eta\pi^{\pm}\pi^{0}$ and continuum yields are floated in the fit as well as all continuum shape parameters.*

Decay channel	ΔE	m_{ES}	$\mathrm{a}_0^{\pm}$ mass	Fisher	Est. Yield (cands)
Signal Model					
$B^{\pm} \rightarrow \eta\pi^{\pm}\pi^{0}$ (non-res)	P2 + NS	Arg + DG	P2	BG	–
Bulk Background Model					
$B^{+}B^{-}$	P3	Arg	P3	BG	913
$B^{0}\overline{B}^{0}$	P3	Arg	P3	BG	411
Combined $q\overline{q}$ $(q = u, d, s, c)$	P2	Arg	P4	BG	–
Charmless B Decays					
$B^{\pm} \rightarrow \pi^{\pm}\pi^{0}\pi^{0}$ (non-res)	SG + P3	Arg + SG	1D Keys	BG	5

4.7.4 Iteration of Background DP Fit and Final Background Estimates

The yields for the two background modes extracted from fits to the DP are, in principle, backgrounds to each other. Therefore, in order to obtain a fair estimate for each mode, an iterative fit approach is used. The procedure begins with the non-resonant fit: The $a_0^{+}(1450)$ contribution is fixed to zero and the fit is run to extract a yield for $\eta\pi^{+}\pi^{0}$. Fits are run for both η subdecay modes and a 'likelihood scan' is carried out in each case. This is done by fixing the $\eta\pi^{+}\pi^{0}$ yields to a range of values about the minimum. Thus a negative log likelihood curve can be plotted. The likelihood curves for both η decay modes are then combined and integrated numerically in order to extract an estimate of the upper limit. Note that in all cases the integration begins at zero so even though we, in this case, have a negative yield we do not bias the resulting upper limit to a lower value than is statistically permissible. The strongly negative yield in the $\eta \rightarrow \gamma\gamma$ case, shown in Table 4.17, is thought to be due to insufficient sensitivity in the fit model, which is not entirely unexpected since the selection criteria were not optimised for this mode. The combined upper limit was found to be $< 6.0 \times 10^{-6}$ at 90% C.L.

Once an upper limit for $\eta\pi^+\pi^0$ was established, it was used to estimate the contribution for this mode in the $a_0^+(1450)$ resonance region. The limit was treated in the same way as all others in the analysis, with the contribution calculated using 50% of the limit as a central value. For the $\eta \to \gamma\gamma$ case a contribution of 14 candidates was predicted with only 3 candidates in the $\eta \to \pi^+\pi^-\pi^0$ case. With the background estimates in place the $a_0^+(1450)\pi^0$ fit was run and likelihood scans for both η subdecay modes performed as for the $\eta\pi^+\pi^0$ fit. From this a combined 90% C.L. upper limit for $\mathcal{B}(a_0(1450)^+\pi^0) \times \mathcal{B}(a_0(1450)^+ \to \eta\pi^+)$ was set at $< 4.3 \times 10^{-6}$.

Using this upper limit for $a_0^+(1450)\pi^0$ it was now possible to look once again at the $\eta\pi^+\pi^0$ to see whether there is any background contribution in the non-resonant region from the tail of the $a_0^+(1450)$ resonance. In the case of a non-zero contribution the $\eta\pi^+\pi^0$ would be re-run with the new background information and the entire procedure described above repeated iteratively until the yields stabilised across the two modes. However, the background contribution from $a_0^+(1450)\pi^0$ in the non-resonant region is predicted to be consistent with zero using the upper limit determined above. As such no further iteration was required and the two limits could be taken for use in the main fit to the $a_0^+(980)$ resonance region. The results of the iterative procedure are presented in Tables 4.17 and 4.18. It should be noted that the limit extracted for $B^\pm \to a_0(1450)^\pm\pi^0$ is taken as $< 4.3\ \times 10^{-6}$ for the purposes of the analysis. This is because we are only interested in the case where $a_0(1450)^+ \to \eta\pi^+$, which we estimate occurs $\sim$45% of the time [8], although the exact value is not relevant here as only this decay mode is used. Plots of the fits to data and likelihood scans for both modes are presented in Appendix C.

4.8 Fit Validation

4.8.1 ‘Toy’ MC Ensemble Studies Results

As previously described, fits were run on ‘toy’ MC datasets derived from the full fit model in order to detect any biases which may exist. Two types of toy study were

Table 4.17: *Results of fits to the non-resonant $B^{\pm} \to \eta\pi^{\pm}\pi^{0}$ DP Region. The yield in candidates is scaled down by the multiplicity in order to calculate the limit.*

Required quantity/result	$\eta \to \gamma\gamma$	$\eta \to \pi^+\pi^-\pi^0$
Candidates to fit	30689	10667
Background from $a_0(1450)^{\pm}\pi^0$	0	0
Signal Yield (cands)	-188 ± 36	-21 ± 22
Signal Multiplicity	1.19	1.26
Continuum Yield (cands)	29319.8±179.6	9326.4±105.8
Accepted eff. and BFs		
Uncorr ϵ (%)	1.74	0.78
Neutrals Eff Corr	0.968	0.938
Corr ϵ (%)	1.68	0.73
$\mathcal{B}(\eta \to X)$ (%)	39.43	22.60
Combined Mode Results		
Upper Limit 90% C.L. ($\times 10^{-6}$)	$<$ 6.04 (statistical error only)	

Table 4.18: *Results of fits to the $B^{\pm} \to a_0(1450)^{\pm}\pi^0$ DP Region. The yield in candidates is scaled down by the multiplicity in order to calculate the limit.*

Required quantity/result	$\eta \to \gamma\gamma$	$\eta \to \pi^+\pi^-\pi^0$
Candidates to fit	166457	84612
Background from $\eta\pi^{\pm}\pi^0$	14	3
Signal Yield (cands)	32±95	-1 ± 75
Signal Multiplicity	1.58	1.64
Continuum Yield (cands)	156619±419	79367 ± 300
Accepted eff. and BFs		
Uncorr ϵ (%)	16.19	11.60
Neutrals Eff Corr	0.968	0.938
Corr ϵ (%)	15.67	10.88
$\mathcal{B}(\eta \to X)$ (%)	39.43	22.60
Combined Mode Results		
Upper Limit 90% C.L. ($\times 10^{-6}$)	$<$ 4.34 (statistical error only)	

made: firstly, the 'pure' toy, where all events are generated from the fit model, and secondly, the 'embedded' toy, where events used to describe charmed and charmless B background contributions were selected from a fully simulated MC sample. This second test was used to uncover any biases present from poorly modelled correlations in the fit variables, which are known to exist in B decays. In both studies we assume zero signal and therefore generate/embed no signal events. The results of the toy

Table 4.19: *The results of the toy studies to assess the stability/bias of the fit. The signal yield pull is represented by P_{sig} and the signal yield statistical error by $\sigma_{N_{sig}}$. The width of the pull distribution is given by $\Gamma_{P_{sig}}$ The fitted and expected signal yields are given by N_{sig} and G_{sig} respectively. Finally the number of toy experiments run is given by N_{toy}.*

Toy type (Full fit)	G_{sig} (events)	N_{toy}	N_{sig}	P_{sig}	$\Gamma_{P_{sig}}$	$\sigma_{N_{sig}}$ (events)
$\eta \rightarrow \gamma\gamma$						
Pure toy	0	487	2.17±0.79	0.05±0.04	0.94±0.03	18.92±0.09
$B\overline{B}$ embedded	0	242	4.50±1.30	0.16±0.07	1.07±0.05	19.06±0.14
$\eta \rightarrow \pi^+\pi^-\pi^0$						
Pure toy	0	499	−0.43±0.57	−0.14±0.05	1.09±0.04	11.95±0.08
$B\overline{B}$ embedded	0	247	−2.64±0.71	−0.32±0.06	0.98±0.04	11.71±0.12

studies are presented in Table 4.19. A total of ∼500 experiments were run in the pure case and ∼250 experiments in the embedded case so as to minimise any oversampling of the MC. The pull means were deemed to be acceptable in all cases. The only significant value, for the $\eta \rightarrow \pi^+\pi^-\pi^0$ embedded study, is not expected to generate a large bias. Plots of the generated distributions are presented in Appendix A.3.

4.8.2 Fits to Off-Resonance Data

In order to test the distinctness of the signal and continuum background models, fits were run on a sample of off-resonance data. The model used contained only the signal and continuum PDFs from the main fit model. Since no $B\overline{B}$ production should occur at this centre-of-mass energy the expected signal yield should be zero and no B backgrounds were included in the fit. The effective luminosity ratio of on/off-resonance data for the samples used in the analysis comes to $340.67/36.18 = 9.42$. In this calculation the correction due to the small variation in cross-section with centre of mass energy is neglected By scaling the resulting continuum yield from the fit by this quantity it is possible to predict the amount of continuum background that will be yielded by the on-resonance data. This will be used later as a cross check of the final fitted continuum yield. The results of this study, for both η decay modes, are presented in Table 4.20. Data distributions with overlaid PDFs are

Table 4.20: *The results of fits to the off-resonance data sample with only the signal and continuum background models.*

Fit Quantity	$\eta \to \gamma\gamma$	$\eta \to \pi^+\pi^-\pi^0$
Signal Yield (events)	-11.77 ± 4.08	-6.30 ± 5.74
Continuum Yield (cands)	8906 ± 95	3302 ± 58
Pred. *udsc* Yield at $\Upsilon(4S)$ (cands)	83856 ± 890	31094 ± 544

presented in Figures 4.8 and 4.9. It is interesting to compare the results of the fits with the continuum yield values predicted from the MC. For $\eta \to \gamma\gamma$ the prediction is 87535 candidates and for $\eta \to \pi^+\pi^-\pi^0$ it is 31451 candidates. The prediction for $\eta \to \pi^+\pi^-\pi^0$ is consistent with the data value, whereas the prediction for $\eta \to \gamma\gamma$ is slightly high. However this prediction had already been scaled down by 20% to correct a previously observed MC generation problem, as described in Section 4.5. The true discrepancy is therefore large, but is not an issue for the analysis since the yield is floated in the final fit, thus these results are only significant as a quality check on the MC. To conclude, the presence of the MC problem for the $\eta \to \gamma\gamma$ case is confirmed and, as expected, the problem is not seen in the $\eta \to \pi^+\pi^-\pi^0$ case.

4.8.3 Blind Fits

A further tool available when performing a maximum likelihood fit is the so-called 'blind' fit. This is where the true value of a given parameter is known internally to the fitting program but reported to the user after being shifted by a replicable offset (the value of which is unknown to the user). Blind fits make it possible to verify the stability of a fitting model on the real data without prematurely knowing the true fit result. Another key aspect to using blind fits is that the reported yield responds to changes in the model in quantitatively the same way as the true value. Thus it is possible to test aspects of the model and estimate systematic error contributions from fit-related sources.

A blind fit was run on the full on-resonance dataset using the full model for both η subdecay modes. Since only the central value of the signal yield is required to be

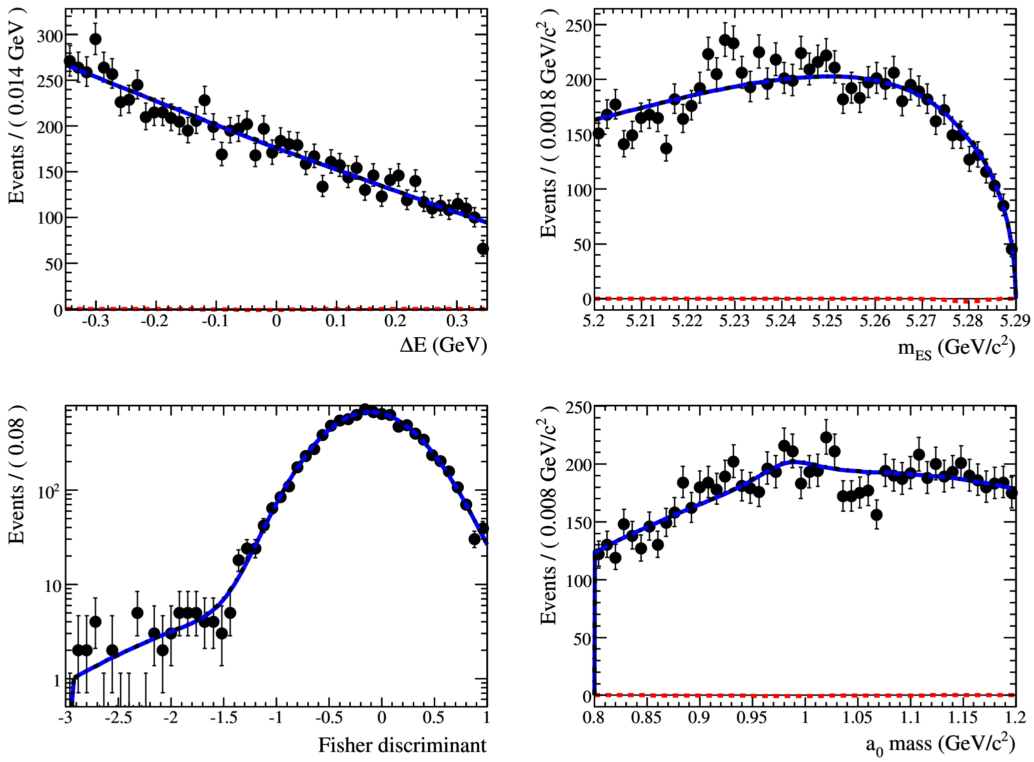

Figure 4.8: *Fits to the off-resonance data for the $\eta \to \gamma\gamma$ subdecay mode. The total PDF is the blue solid line, the signal PDF is the red dashed line and the continuum PDF is the black dot-dashed line. The Fisher variable is plotted on a logarithmic scale to render the negative side tail visible.*

blind one can make use of other quantities, such as the signal yield error and the continuum yield, in their unblind forms. The results for the blind fits to the data are presented in Table 4.21. As can be seen, the predicted continuum yield from the fit to off-resonance data is consistent with that which is measured on-resonance.

Test floating charmed B background PDFs

A study was carried out using a blind fit in order to ascertain whether it would be desirable to float the total charmed B yield in the final fit. As a first step the combined charmed B yield was floated with the charged to neutral B ratio fixed at 50%. This was then compared to two other cases. In the first instance, the charged and neutral charmed B yields are fixed but the contributions are scaled up by 10%

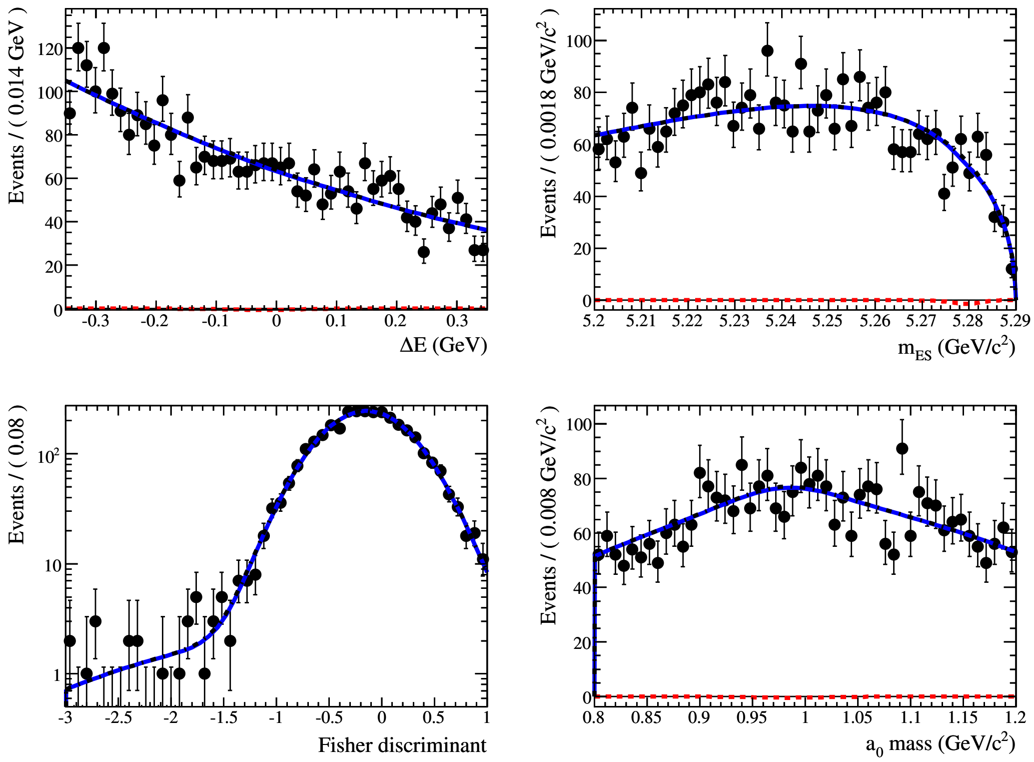

Figure 4.9: *Fits to the off-resonance data for the $\eta \to \pi^+\pi^-\pi^0$ subdecay mode. The total PDF is the blue solid line, the signal PDF is the red dashed line and the continuum PDF is the black dot-dashed line. The Fisher variable is plotted on a logarithmic scale to render the negative side tail visible.*

from the original MC predictions. In the second instance the two yields are both scaled down by 10% from the MC predictions. These ±10% shifts are larger than the known uncertainty on the B^+B^- branching fraction but include the uncertainty in the efficiency as well as the rate of the $b \to c$ decay modes. The results of this study are presented in Table 4.22.

As can be seen, floating the charmed B yield in the blind fit gives no improvement over holding it fixed. As such the main fit strategy will be unchanged and the ±10% shifted yield values will be taken as systematics.

Table 4.21: *The results of blind fits to the on-resonance data sample with the full fit model.*

Fit Quantity	$\eta \to \gamma\gamma$	$\eta \to \pi^+\pi^-\pi^0$
Signal Yield Error (events)	$\pm$18.36	$\pm$13.22
Continuum Yield (cands)	85649$\pm$296	30380$\pm$177
Pred. *udsc* Yield from Off-res fit (cands)	83856$\pm$890	31094$\pm$544

Table 4.22: *The results of fixing and floating the charmed B yield in a blind fit to on-resonance data. The shift in the blind signal yield is given by ΔN_{sig}, the signal yield statistical error by $\sigma_{N_{sig}}$, the charmed B yield is given by $N_{b\to c}$ and the continuum yield by N_{udsc}.*

Fit type	ΔN_{sig} (events)	$\sigma_{N_{sig}}$ (events)	$N_{b\to c}$	N_{udsc}
$\eta \to \gamma\gamma$				
Float charmed B yield	-4.56	18.27	-121.2 ± 327.0	86663.3$\pm$440.5
Fix charmed B yield -10%	-2.22	18.35	810	85737.5$\pm$295.5
Fix charmed B yield $+10\%$	-1.78	18.37	990	85560.3$\pm$295.5
$\eta \to \pi^+\pi^-\pi^0$				
Float charmed B yield	-0.36	13.26	693.5$\pm$202.7	30528$\pm$267.0
Fix charmed B yield -10%	-0.20	13.22	760	30462.4$\pm$177.3
Fix charmed B yield $+10\%$	$+0.19$	13.26	929	30297.3$\pm$177.3

Chapter 5

Results

5.1 Overview

This chapter will present the final results of the maximum likelihood fit analysis for the mode $B^+ \to a_0^+ \pi^0$.

The branching fraction was calculated using the following formula

$$\mathcal{B} = \frac{Y - Y_B}{\epsilon N_{B\overline{B}} \Pi_i \mathcal{B}_i} \tag{5.1}$$

where Y is the signal event yield from the fit, Y_B is the fit bias, ϵ is the efficiency for the B decaying via the studied mode, $N_{B\overline{B}}$ is the number of produced $B\overline{B}$ mesons and $\Pi_i \mathcal{B}_i$ is the product of the daughter branching fractions. The decay rates for $\Upsilon(4S) \to B^+B^-$ and $\Upsilon(4S) \to B^0\overline{B}^0$ are assumed to be equal.

The fit bias was estimated using a toy MC study embedding signal and B background MC according to the yields extracted/fixed from the main fit. The continuum contribution was generated according to the fitted yield and using the shape parameters taken from on-resonance data. Plots for the relevant toy study quantities are presented in Appendix A.4.

The results for the two η decay channels are combined by a multiplying together their associated likelihood functions. The are produced by inverting the negative log likelihood curves which are computed for both fits by calculating the minimum negative log likelihood for a given range of fixed signal yields. The 90% confidence level (C.L.) upper limit is then calculated by integrating this combined function for increasing yield, starting at zero. The upper limit is then defined as the yield which bounds 90% of the area under the combined likelihood curve when one starts from zero.

Before considering the final fit results, the sources of systematic uncertainty in the analysis are discussed and estimates for the resulting error in each η decay mode are presented. Systematic errors which are uncorrelated between the two η channels are accounted for by convolving the likelihood curves for each channel with a Gaussian function representing the error. Once the likelihood curves have been combined, as per the procedure described above, the systematics are included. These errors are are correlated between the η channels and are accounted for by convolving the combined likelihood curve with a representative Gaussian. The final limit will be quoted with and without the systematic contributions in order to quantify their effect on the result.

5.2 Estimating Systematic Uncertainty

The systematic error contributions which are about to be discussed are presented in Table 5.1. In this section the source and treatment of each error is discussed and the largest contributions to the total identified.

5.2.1 Uncertainties in the Fitted Yield

The uncertainties which are considered in this section are those which arise from the fit and directly affect the signal yield in an additive sense. The combined error from each of these sources will therefore be treated as an additive systematic error on the final result.

PDF Parameter Uncertainty

An uncertainty exists due to the statistical errors in the fitted parameters of the PDFs which make up the fit model and which are held fixed in the final analysis. This is fundamentally due to the limited MC available upon which to base the parameterisations for signal and background.

In order to estimate the error derived from this uncertainty the fit was re-run testing any fixed parameters in the signal, charmed B and continuum shapes. Any contribution due to parameter uncertainty in the charm and charmless models was expected to be small and therefore neglected. The test was conducted by varying each PDF parameter individually within its statistical error bounds. The only exception to this was for the signal a_0 mass Breit-Wigner width. In the MC the a_0 is assumed to have a nominal width of $80\,\mathrm{MeV}/c^2$ but, due to the poorly known lineshape for the resonance, it was decided that limits of 50-100 MeV/c^2 would be required to adequately model the uncertainty. This is by far the dominant systematic error of this type, contributing roughly $^{+4}_{-3}$ events for $\eta \to \gamma\gamma$ and $^{+0.5}_{-2}$ events for $\eta \to \pi^+\pi^-\pi^0$.

The changes in the fitted signal yield for each parameter test were then added in quadrature to form a total systematic error estimate.

Charmless *B* Decay Uncertainty

The branching fractions used to estimate the contributions from charmless B decay processes each have an associated error. These form a source of systematic uncertainty in the final result. To estimate the size of the resultant error the fit was re-run varying each charmless mode normalisation in turn according to the error limits of its branching fraction. In the cases where only an upper limit exists the central value (calculated using 50% of the limit) was varied by ±100% of its central value, i.e. to zero and the upper limit. The largest systematic shifts in the $\eta \to \gamma\gamma$ channel were produced by the $\rho(1450)^+\eta$ and $\pi^+\pi^0\pi^0$ modes, both contributing of order ±2 events. In the $\eta \to \pi^+\pi^-\pi^0$ channel the largest contribution comes from $\rho(1450)^+\eta$

with ± 1 event. The second largest contribution here comes from $b_1^+\pi^0$, giving ± 0.5 events.

Once again, the changes in the fitted signal yield for each test were combined in quadrature to provide a total systematic error estimate.

Charmed *B* Decay Yield Uncertainty

The uncertainty in the yields for backgrounds from charmed B decays was dominated by reconstruction efficiency and B decay branching fractions rather than the production cross-section. As such a conservative estimate of $\pm 10\%$ was taken and the fit re-run with the yields varied individually by this fraction. The changes in the fitted signal yields were combined in quadrature with the overall additive systematic error estimate.

Fit Bias Uncertainties

A systematic error must be assigned to account for the uncertainty in estimating the bias in the fit result. The bias was calculated by running an embedded toy MC study where all floating shape parameters are generated according to their values from the final fit. The systematic error was then taken to be the error on the measured bias. The resulting error was then combined in quadrature with the overall additive systematic error estimate. The bias was measured to be 3.5 ± 1.3 for $\eta \to \gamma\gamma$ and -1.1 ± 0.8 for $\eta \to \pi^+\pi^-\pi^0$.

5.2.2 Uncertainties in the Efficiency

There are a number of sources of uncertainty which affect the signal efficiency and are applied as a multiplicative correction to the final result. The errors from each source are combined in quadrature to give a final multiplicative systematic error estimate.

Tracking and Neutrals Efficiency

There exists an uncertainty in the tracking efficiency of charged particles. This was estimated to account for an error of $\pm 0.45\%$ per track based on the results of dedicated studies by the *BABAR* tracking group [62].

An uncertainty also exists in the reconstruction efficiency of neutral particles [63]. The *BABAR* neutrals group estimates this to be $\pm 3\%$ each for every π^0 and photon composite η in the final state, i.e. for $\eta \to \pi^+\pi^-\pi^0$ the error comes not from the η but from its π^0 daughter.

For both of these sources of uncertainty the resulting errors are correlated and hence added directly, rather than in quadrature. Thus the combined errors in the $\eta \to \pi^+\pi^-\pi^0$ case are 1.4% for tracks and 6% for neutral particles.

Data/MC Agreement in the $|\cos(\theta_{TB})|$ and η mass variables

Due to the imperfect agreement between data and MC samples, the selections placed on $|\cos(\theta_{TB})|$ and the η mass variables require the assignment of a systematic error. Figure 5.1 shows continuum MC compared with on-resonance data for the two variables. As can be seen the discrepancy is most marked at the higher end of the $|\cos(\theta_{TB})|$ distribution with our optimised selection at 0.7 excluding most of the problem region. Control sample studies [64] have shown that a tighter selection in this variable incurs a larger error. As such a conservative estimate of $\pm 3\%$ is taken for this analysis. Since any discrepancy in the η mass variable was deemed negligible a systematic error was not assigned to account for it.

Statistical Error

Finally, the error due to limited MC statistics in the efficiency had to be accounted for. This was calculated from the error in the fitted number of true signal events from the iterative signal/SxF reduction procedure.

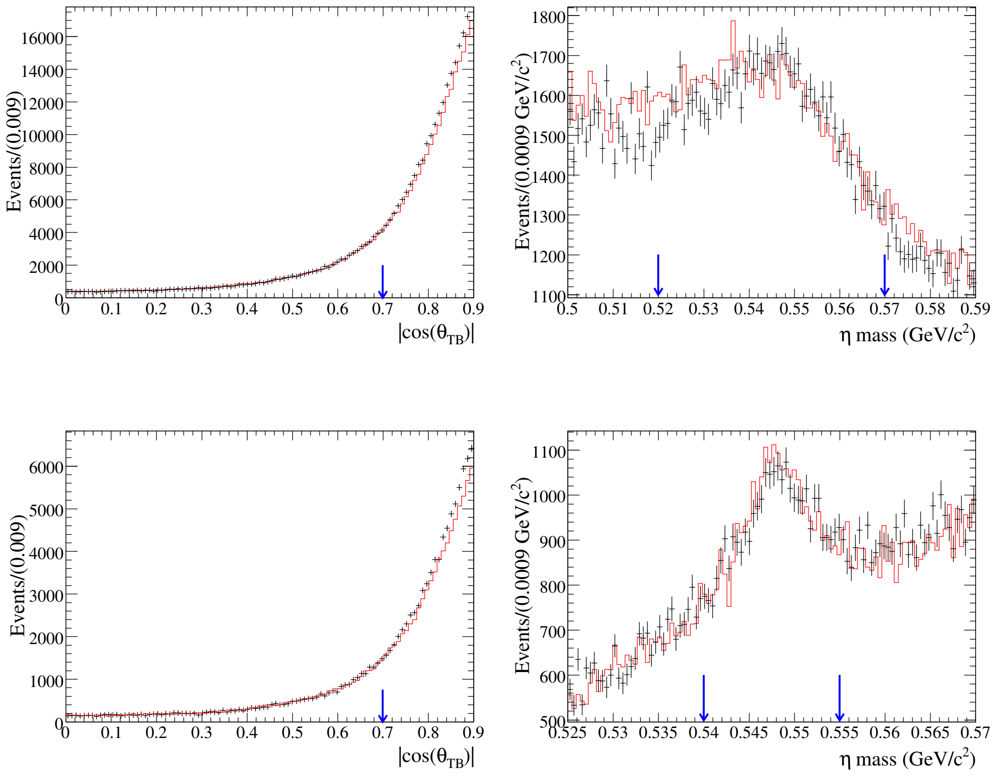

Figure 5.1: *Data MC comparisons where continuum MC (red line) is overlaid with on-resonance data (black points). The top row of plots are for the $\eta \to \gamma\gamma$ case and the bottom row for $\eta \to \pi^+\pi^-\pi^0$. The left hand column shows $|\cos(\theta_{TB})|$ while the right hand column shows the η mass. The locations of selections applied for the analysis are indicated by solid blue arrows.*

5.2.3 Uncertainties in the Branching Fraction Calculation

These systematic contributions, affecting the conversion of the fit yield to a BF using Equation 5.1, are also multiplicative errors.

Total Number of $B\overline{B}$ Events

The total number of $B\overline{B}$ events in the dataset was estimated to be $(379.9 \pm 4.2) \times 10^6$ [45]. The error was taken as a systematic uncertainty.

Uncertainties in the Daughter Decay BFs

The errors in $\mathcal{B}(a_0^+ \to \eta\pi^+)$, $\mathcal{B}(\eta \to \gamma\gamma)$ and $\mathcal{B}(\eta \to \pi^+\pi^-\pi^0)$ were taken from the Particle Data Group [8]. The error in $\mathcal{B}(a_0^+ \to \eta\pi^+)$ was calculated by taking the ratio of the partial widths of the two dominant a_0^+ decay modes; $a_0^+ \to \eta\pi^+$ and $a_0^+ \to K\overline{K}$. The measured ratio is 0.183 ± 0.024 which, assuming all other decay modes are negligible, gives $\mathcal{B}(a_0^+ \to \eta\pi^+) = 0.85 \pm 0.02$. The value taken for $\mathcal{B}(\eta \to \gamma\gamma)$ was 0.3943 ± 0.0026 and for $\mathcal{B}(\eta \to \pi^+\pi^-\pi^0)$ we take 0.226 ± 0.004.

5.2.4 Summary of Systematic Errors

The results of all of the studies to estimate systematic errors are presented in Table 5.1.

Table *5.1:* *Estimated systematic errors in the final fit result. Correlated and Uncorrelated error sources are denoted by [C] and [U] respectively.*

Source of Uncertainty	$\eta \to \gamma\gamma$	$\eta \to \pi^+\pi^-\pi^0$
Additive (Events)		
Fit Parameters [U]	$^{+4.3}_{-3.4}$	$^{+0.6}_{-2.1}$
Charmless Yields [U]	$^{+3.3}_{-3.2}$	$^{+1.3}_{-1.3}$
Charm Yields [U]	$^{+0.0}_{-2.8}$	$^{+0.2}_{-0.2}$
Fit Bias [U]	± 1.3	± 0.8
Total Additive (Events)	$^{+5.6}_{-5.6}$	$^{+1.7}_{-2.6}$
Multiplicative (%)		
Neutral efficiency [C]	± 6.0	± 6.0
Tracking efficiency [C]	± 0.5	± 1.4
$\lvert\cos(\theta_{TB})\rvert$ Selection [C]	± 3.0	± 3.0
MC Statistics [U]	± 0.4	± 0.4
Number of $B\overline{B}$ Events [C]	± 1.1	± 1.1
Daughter a_0 Decay BF [C]	± 2.0	± 2.0
Daughter η Decay BF [U]	± 0.7	± 1.8
Total Multiplicative (%)	± 7.2	± 7.5

5.3 Fit Results

The complete analysis results are presented in Table 5.2. We fit small positive signal yields in both modes, both of which are consistent with zero within their statistical errors. Likelihood scans for separate and combined η decay modes are presented in Figure 5.2. From these we are able, using the procedure outlined in Section 5.1, to set an upper limit on the branching fraction of $< 2.0 \times 10^{-6}$ when we include both statistical and systematic error. As such the limit is clearly statistics dominated with the $\eta \to \gamma\gamma$ mode providing the most significant contribution to the limit. The significance of the result is calculated as the square root of the difference in the value of -2ln$\mathcal{L}$ between the negative log likelihood curve minimum and the value at zero yield.

Table 5.2: *The results of the fit to the full data set and other values required for calculating the branching fraction. All B background yields were held fixed to the values described in Sections 4.5 and 4.6. The upper limit is shown first with only the statistical error and then with the total error.*

Required quantity/result	$\eta \to \gamma\gamma$	$\eta \to \pi^+\pi^-\pi^0$
Candidates to fit	87022	31326
Signal Yield (events)	4.3 ± 18.4	14.2±13.3
Continuum Yield (candidates)	85649±296	30381±177
ML Fit bias (events)	3.5±1.3	−1.1±0.8
Accepted eff. and BFs		
Uncorr ϵ (%)	15.44	10.83
Neutrals Eff Corr	0.968	0.938
Corr ϵ (%)	14.95	10.15
$\mathcal{B}(\eta \to X)$ (%)	39.43	22.60
$\mathcal{B}(a_0^+ \to \eta\pi^+)$ (%)	84.5	84.5
Branching Fraction ($\times 10^{-6}$)	$0.1^{+1.0}_{-0.9}(stat)^{+0.3}_{-0.3}(syst)$	$2.1^{+1.9}_{-1.8}(stat)^{+0.3}_{-0.4}(syst)$
Combined Mode Results		
Branching Fraction ($\times 10^{-6}$)	$0.6^{+0.9}_{-0.8}(stat)^{+0.2}_{-0.2}(syst)$	
Significance	0.7σ (stat + syst)	
Upper Limit 90% C.L. ($\times 10^{-6}$)	< 1.95 (statistical error only)	
Upper Limit 90% C.L. ($\times 10^{-6}$)	< 2.01 (total error)	

Projection plots for each of the four fit variables are presented in Figures 5.3 and 5.4. The dataset for these plots has been background-reduced by requiring that the

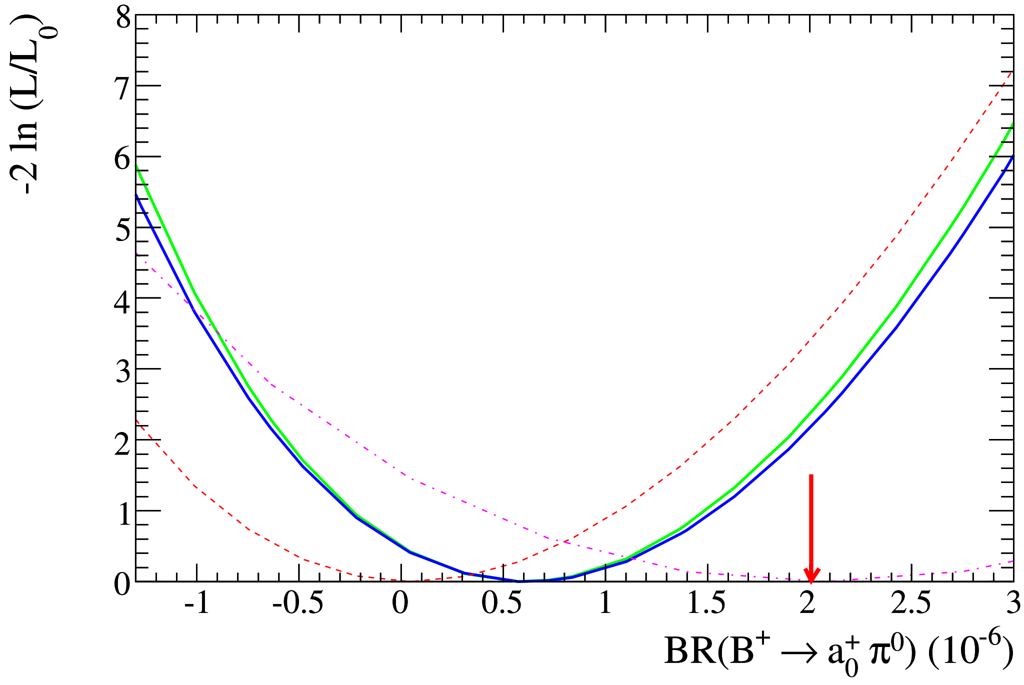

Figure 5.2: *Likelihood scan for a range of branching fractions for both η decay modes. The red dashed curve represents the $\eta \to \gamma\gamma$ contribution with the $\eta \to \pi^+\pi^-\pi^0$ represented by the pink dot-dashed curve. The green solid curve represents the combined likelihood for the two modes and the blue solid curve represents the combined likelihood with systematic errors convolved in. The upper limit on the branching fraction is set using the blue curve. The position of the limit is indicated by the red solid arrow.*

likelihood ratio $\mathcal{L}_{sig}/[\mathcal{L}_{sig} + \Sigma\mathcal{L}_{bkg}]$ for any event be > 0.6. As can be seen there is no significant signal peak for either mode.

The curves are made by generating toy MC from the final fit PDFs with all parameters at their values corresponding to the real data. The likelihood ratio selection is then applied to this generated dataset and it is overlaid on the background-reduced experimental data. The reason this is done in this manner is that creating true projections has been found to be extremely difficult in practice due to the large processing demands required for this fit model. The source of this inefficiency is thought to be the KEYS PDFs used to model the charmless backgrounds.

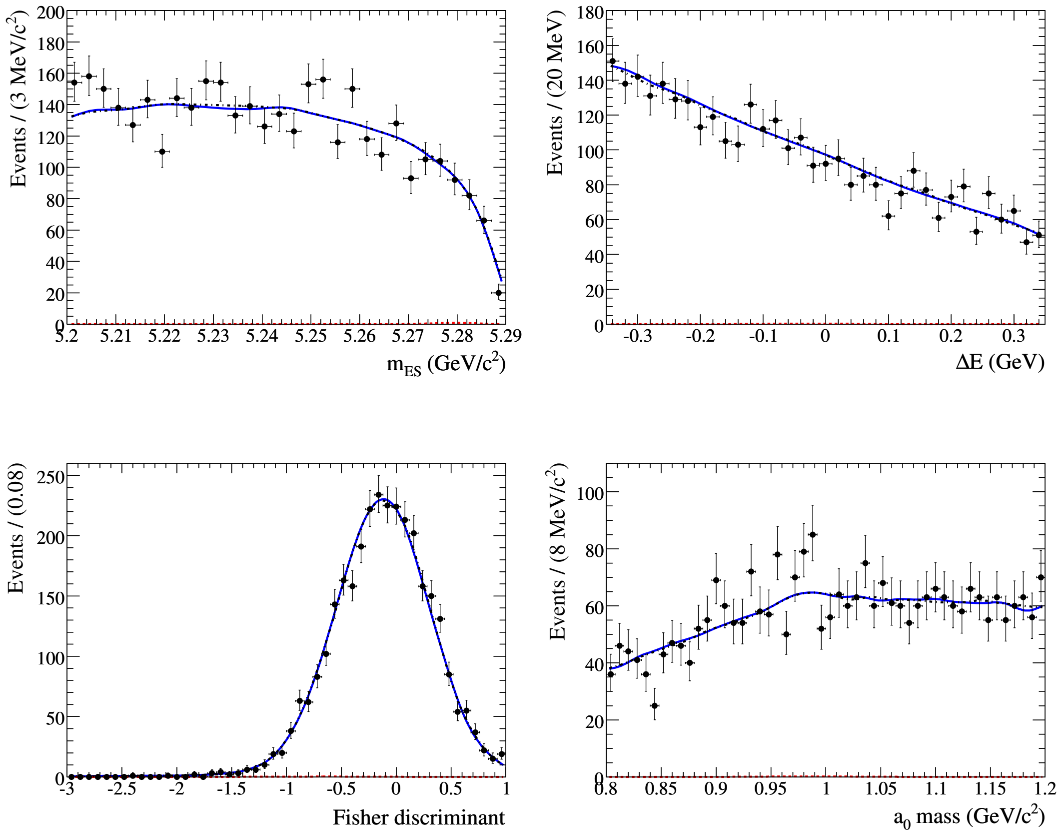

Figure 5.3: *Likelihood ratio enhanced projection plots for the $\eta \to \gamma\gamma$ case. These plots have been made by requiring that $\mathcal{L}_{sig}/[\mathcal{L}_{sig} + \Sigma\mathcal{L}_{bkg}]$ for any event be > 0.6. (top left) shows the $m_{\rm ES}$ variable, (top right) ΔE, (bottom left) Fisher and (bottom right) a_0 mass. The experimental data are represented by the black points while the overall PDF is represented by the blue curve. The combined background component is represented by the black dash-dotted curve and the signal component by the red dashed curve. The average efficiency of the likelihood ratio selection is 4% for background and 88% for signal.*

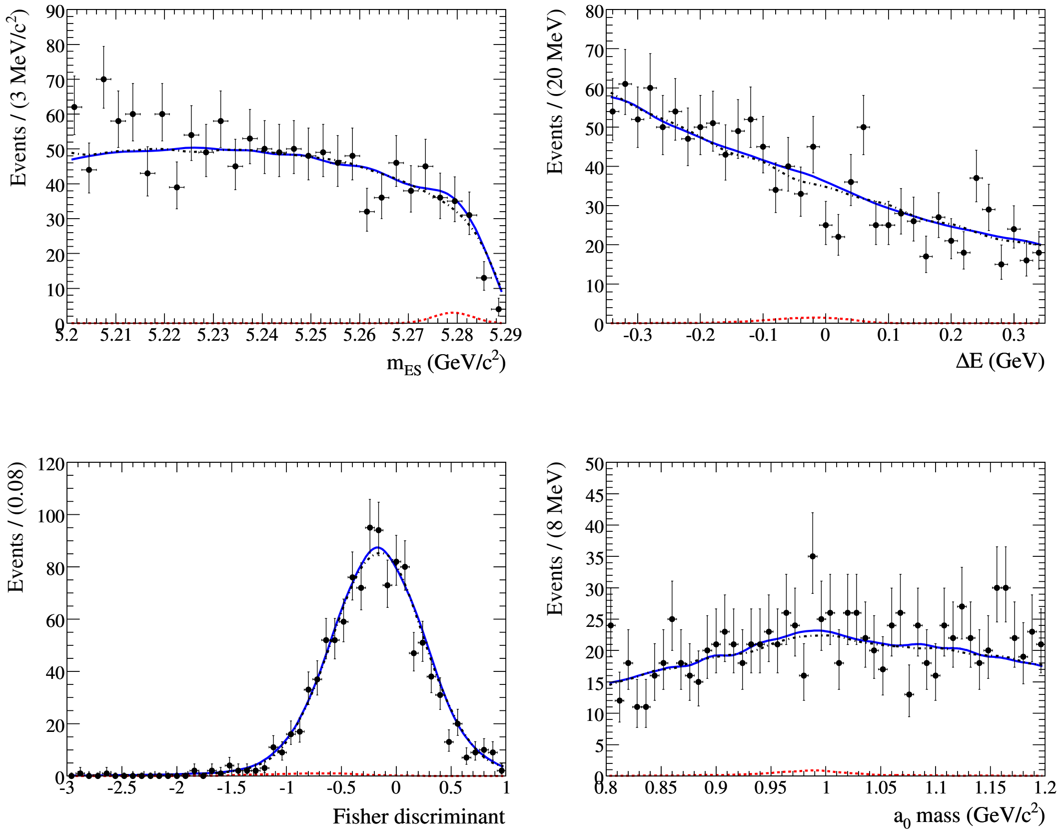

Figure 5.4: *Likelihood ratio enhanced projection plots for the $\eta \to \pi^+\pi^-\pi^0$ case. These plots have been made by requiring that $\mathcal{L}_{sig}/[\mathcal{L}_{sig} + \Sigma\mathcal{L}_{bkg}]$ for any event be > 0.6. (top left) shows the $m_{\rm ES}$ variable, (top right) ΔE, (bottom left) Fisher and (bottom right) a_0 mass. The experimental data are represented by the black points while the overall PDF is represented by the blue curve. The combined background component is represented by the black dash-dotted curve and the signal component by the red dashed curve. The average efficiency of the likelihood ratio selection is 4% for background and 79% for signal.*

Chapter 6

Conclusions

An upper limit has been extracted, at the 90% C.L, for the decay $B^{\pm} \to a_0(980)^{\pm}\pi^0$, where $a_0(980)^{\pm} \to \eta\pi^{\pm}$ and $\eta \to \gamma\gamma$ or $\eta \to \pi^+\pi^-\pi^0$. The result was obtained using an unbinned multivariate extended maximum likelihood fit. The dominant background source came from continuum light quark production although a non-negligible contribution from charmed and charmless B meson decays was also present. The number of continuum events was floated in the final fit with the B background contributions estimated from Monte Carlo simulation and held fixed. The branching fraction extracted for the combined η modes has a significance of 0.7σ and was found to be

$$\mathcal{B}(B^{\pm} \to a_0(980)^{\pm}\pi^0) = 0.6^{+0.9}_{-0.8}(stat)^{+0.2}_{-0.2}(syst).$$

The statistical component is by far the largest contribution to the error. The dominant systematic errors come from background modelling and the uncertainty in the a_0 lineshape. The upper limit for this mode is set as

$$\mathcal{B}(B^{\pm} \to a_0(980)^{\pm}\pi^0) < 2.0 \times 10^{-6} \text{ (90\% C.L.)}.$$

In light of these results, one can look back to the theoretical predictions made in Chapter 1. Here different structure models for the scalar a_0 were shown to provide significantly different branching fraction predictions. Specifically this analysis sought to test two different structure models for which predictions could be made using QCD factorisation as described by Delepine *et al.* [18]. They predicted a branching fraction of

$$6.4 \times 10^{-8} \leq \mathcal{B}(B^+ \rightarrow a_0^+ \pi^0) \leq 2.4 \times 10^{-7}$$

for the case where this a_0 is a simple quark-antiquark state and

$$2 \times 10^{-9} \leq \mathcal{B}(B^+ \rightarrow a_0^+ \pi^0) \leq 10^{-8}$$

for the more complicated case where the a_0 is in fact a four quark state where the original quarks are joined by an $s\bar{s}$ pair. It is clear from the results that the analysis is not sensitive enough, given the currently available statistics, to show which (if either) of these models is correct. However the result is consistent with current theories since no evidence of a signal at a significantly higher branching fraction has been found.

The short term prospects for measurements of this mode depend entirely on the amount of data which the two B Factories will record during the remainder of their operational lives. Current predictions state that the combined data sample for both *BABAR* and BELLE will approach 2 ab^{-1} by late 2008/early 2009 (at which point *BABAR* is expected to cease running). If achieved this data sample will be over 4$\times$ that available for the analysis which has been presented in this thesis. Since the upper limit on a given mode can be expected to decrease approximately linearly with increasing luminosity, the combined world average for both B Factories would lie in the region of $\sim 5 \times 10^{-7}$ by the end of 2008. Given the current theoretical predictions analysis techniques would therefore have to improve such that up to a five-fold increase in sensitivity were achieved if we are to be able to begin to include interesting parameter space, which seems highly unlikely.

Longer term, progress with this mode will be dependent on two future experiments: the LHCb detector [65], currently under construction at CERN, and the proposed "Super B Factory" [66].

At LHCb the interaction environment will provide an extremely challenging barrier to the effective reconstruction of this mode, which relies on high efficiency for detection of neutral particles. That said, the number of $b\bar{b}$ pairs expected to be produced per year at design luminosity is $\sim 10^{12}$. Of these 40% will hadronise to form charged B mesons. With such a large production rate it is not impossible that some improvement to the limit for this mode could be made over the lifetime of the LHCb experiment. Needless to say this mode is not an analysis priority for LHCb and no studies have yet been undertaken on the feasibility of such work.

At a Super B Factory, with its clean environment and extremely high luminosity (dataset estimates range from $\sim 10-30$ ab^{-1}), it should be possible to make an observation for this mode within the projected lifetime of the machine. Simply scaling our limit would give a value of order 7×10^{-8} at 10 ab^{-1} and 2×10^{-8} at 30 ab^{-1} assuming a detector of similar sensitivity to *BABAR* or BELLE and that no evidence for a signal is found. These values are within the predicted two-quark range. However this experiment is still at the design and approval phase with the most optimistic start date for data taking at $\sim$2012. Taking all of these factors into consideration the conclusion must be that the best (although remote) chance of seeing interesting physics results for this mode within the next few years remains at the current B Factories.

References

[1] P. W. Higgs, "Broken symmetries, massless particles and gauge fields", *Phys. Lett.* **12** (1964) 132–133.

[2] E. Noether, "Invariant Variation Problems (English Translation)", `physics/0503066`.

[3] H. Fritzsch and M. Gell-Mann, "Current algebra: Quarks and what else?", `hep-ph/0208010`.

[4] S. L. Glashow, "Partial Symmetries of Weak Interactions", *Nucl. Phys.* **22** (1961) 579–588.

[5] S. Weinberg, "A Model of Leptons", *Phys. Rev. Lett.* **19** (1967) 1264–1266.

[6] A. Salam and J. C. Ward, "Gauge theory of elementary interactions", *Phys. Rev.* **B136** (1964) 763–768.

[7] F. Halzen and A. D. Martin, "Quarks and Leptons: An Introductory Course in Modern Particle Physics",. New York, Usa: Wiley (1984) 396p.

[8] W. M. Yao *et al.*, "Review of particle physics", *J. Phys.* **G33** (2006) 1–1232.

[9] N. Cabibbo, "Unitary Symmetry and Leptonic Decays", *Phys. Rev. Lett.* **10** (1963) 531–532.

[10] M. Kobayashi and T. Maskawa, "CP Violation in the Renormalizable Theory of Weak Interactions", *Prog. Theor. Phys.* **49** (1973) 652–657.

[11] L. Wolfenstein, "Parameterization of the Kobayashi-Maskawa matrix", *Phys. Rev. Lett.* **51** (1983) 1945.

[12] *BABAR* Collaboration, *ed.* P. F. Harrison and H. R. Quinn, "The *BABAR* physics book: Physics at an asymmetric B factory." SLAC-R-0457.

[13] S. Weinberg, "The Quantum Theory of Fields." Cambridge University Press, 1995.

[14] T. D. Lee and C. N. Yang, "Parity Nonconservation and a Two-Component Theory of the Neutrino", *Phys. Rev.* **105** (1957) 1671–1675.

[15] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, "Evidence for the $2\,\pi$ Decay of the K_2^0 Meson", *Phys. Rev. Lett.* **13** (1964) 138–140.

[16] A. D. Sakharov, "Variation of CP Invariance, C Asymmetry, And Baryon Asymmetry Of The Universe", *JETP. Lett.* **5** (1967) 24–27.

[17] V. Baru *et al.*, "Evidence that the $a_0(980)$ and $f_0(980)$ are not elementary particles", *Phys. Lett.* **B 586** (2004) 53.

[18] D. Delepine *et al.*, "Annihilation contribution and $B \to a_0\pi, f_0K$ decays", *Eur. Phys. J.* **C 45** (2006) 693–700.

[19] L. Maiani *et al.*, "A new look at scalar mesons", *Phys. Rev. Lett.* **93** (2004) 212002.

[20] T. Yoshikawa, "Possibility of large electroweak penguin contribution in $B \to K\pi$ modes", *Phys. Rev.* **D 68** (2003) 1–7.

[21] S. Laplace *et al.*, "CP violation and the absence of second class currents in charmless B decays", *Eur. Phys. J.* **C 22** (2001) 431–438.

[22] M. Beneke *et al.*, "QCD factorization for exclusive non-leptonic B-meson decays: General arguments and the case of heavy-light final states", *Nucl. Phys.* **B 591** (2000) 313–418.

[23] K. G. Wilson, “Non-Lagrangian models of current algebra”, *Phys. Rev.* **179** (1969) 1499–1512.

[24] G. P. Lepage and S. J. Brodsky, “Exclusive Processes in Perturbative Quantum Chromodynamics”, *Phys. Rev.* **D 22** (1980) 2157–2198.

[25] R. H. Dalitz, “On the Analysis of τ Meson Data and the Nature of the τ Meson”, *Phil. Mag.* **44** (1953) 1068–1080.

[26] S. M. Flatté, “Coupled-Channel Analysis of the $\pi\eta$ and $K\overline{K}$ Systems near the $K\overline{K}$ Threshold”, *Phys. Lett.* **B 63** (1976) 224–227.

[27] *BABAR* Collaboration, B. Aubert *et al.*, “The *BABAR* detector”, *Nucl. Instrum. Meth.* **A479** (2002) 1–116.

[28] *BABAR* Collaboration, D. Boutigny *et al.*, “*BABAR* technical design report.” SLAC-R-0457.

[29] BELLE Collaboration, A. Bondar *et al.*, “The BELLE detector”, *Nucl. Instrum. Meth.* **A408** (1998) 64–76.

[30] *ed.* M. S. Zisman, “PEP-II: An Asymmetric B Factory. Conceptual Design Report.” SLAC-R-418.

[31] E. Bloom *et al.*, “The PEP-II asymmetric B factory: Design details and R & D results.” presented at 4th European Particle Accelerator Conference (EPAC 94), London, England, 27 Jun - 1 Jul 1994.

[32] W. Kozanecki *et al.*, “Trickle-charge: A new operational mode for PEP-II.” presented at the 9th European Particle Accelerator Conference (EPAC 2004), Lucerne, Switzerland, 5-9 Jul 2004.

[33] J. Seeman *et al.*, “Results and plans of the PEP-II B-factory.” presented at 9th European Particle Accelerator Conference (EPAC 2004), Lucerne, Switzerland, 5-9 Jul 2004.

[34] Spectrosil is a trademark of TSL Group PCL, Wallsend, Tyne on Wear, NE28 6DG, England; sold in the USA by Quartz Products Co., Louisville, KY, USA.

[35] Heavy Flavour Averaging Group (HFAG) Collaboration, E. Barberio *et al.*, "Averages of b-hadron Properties at the End of 2005", `hep-ex/0603003`.

[36] S-2744-08 PIN diode by Hamamatsu Photonics, K. K., Hamamatsu City, Japan.

[37] G. M. Haller and D. R. Freytag, "Analog floating point BiCMOS sampling chip and architecture of the *BABAR* CsI calorimeter front end electronics system at the SLAC B Factory", *IEEE Trans. Nucl. Sci.* **43** (1996) 1610–1614.

[38] D. Best *et al.*, "Crosstalk in the EMC Front End Electronics." *BABAR* Internal Note 550, 2002.

[39] G. P. Taylor, "Study of Charmless Semileptonic B Decays and a Measurement of the CKM Matrix Element $|V_{ub}|$ at *BABAR*",. Ph.D. Thesis, University of London, 2004.

[40] D. D. Altenburg. presentations to *BABAR* calorimeter group 2002-2003.

[41] S. Menke, "Offline Correction of Non-Linearities in the *BABAR* Electromagnetic Calorimeter Electronics." *BABAR* Internal Note 527, 2000.

[42] The Novosibirsk function is defined as $f(x) = A_s \exp(-0.5(\ln^2[1 + \Lambda\tau(x - x_0)]/\tau^2 + \tau))$ where $\Lambda = \sinh(\tau\sqrt{\ln 4})/(\sigma\tau\sqrt{\ln 4})$, the peak is x_0, τ is the tail parameter and A_s is a normalisation factor.

[43] R. Santonico and R. Cardarelli, "Development of Resistive Plate Counters", *Nucl. Instrum. Meth.* **187** (1981) 377–380.

[44] GEANT4 Collaboration, S. Agostinelli *et al.*, "GEANT4: A simulation toolkit", *Nucl. Instrum. Meth.* **A506** (2003) 250–303.

[45] C. Hearty, "Hadronic Event Selection and B-Counting for Inclusive Charmonium Measurements." *BABAR* Internal Analysis Document BAD/30, (2000).

[46] T. Sjostrand, "High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4", *Comput. Phys. Commun.* **82** (1994) 74–90.

[47] W. T. Ford, "Choice of Kinematic Variables in B Meson Reconstruction - Take 3." *BABAR* Internal Analysis Document BAD/53, (2000).

[48] The ARGUS function is defined as $f(x) = Nx[(1-(x/E_{beam}))^2 \exp[p(1-(x/E_{beam}))^2]]^{1/2}$, where N is a normalisation factor.

[49] E. Farhi, "Quantum Chromodynamics Test for Jets", *Phys. Rev. Lett.* **39** (1977) 1587–1588.

[50] R. A. Fisher, "The Use of Multiple Measures in Taxonomic Problems", *Annals of Eugenics* **7** (1936) 179–188.

[51] F. Blanc and others., "Measurements of branching fractions and charge and time-dependent asymmetries for $B \to \eta' K$." *BABAR* Internal Analysis Document BAD/490v7, (2004).

[52] G. Arfken and H. Weber, "Mathematical methods for physicists." Academic Press, 2000.

[53] A. Hoecker, "A Maximum Likelihood Method For DIRC Particle Identification." *BABAR* Internal Analysis Document BAD/23v1, (2000).

[54] R. J. Barlow, "Statistics. A guide to the use of statistical methods in the physical sciences." Wiley, 1997.

[55] W. Verkerke and D. Kirkby, "The RooFit toolkit for data modelling", `physics/0306116`.

[56] F. James and M. Roos, "'Minuit' A System for Function Minimization and Analysis of the Parameter Errors and Correlations", *Comput. Phys. Commun.* **10** (1975) 343–367.

[57] *BABAR* Collaboration, B. Aubert *et al.*, "Measurements of the $B \to X_s\gamma$ branching fraction and photon spectrum from a sum of exclusive final states", *Phys. Rev.* **D72** (2005) 052004.

[58] *BABAR* Collaboration, B. Aubert *et al.*, "B meson decays to ωK^*, $\omega\rho$, $\omega\omega$, $\omega\phi$, and ωf_0", *Phys. Rev.* **D74** (2006) 051102.

[59] M. Beneke and M. Neubert, "QCD factorization in $B \to PP$ and $B \to PV$ decays", *Nucl. Phys.* **B675** (2003) 333–415.

[60] M. Beneke and M. Neubert, "Flavor-singlet B-decay amplitudes in QCD factorization", *Nucl. Phys.* **B651** (2002) 225–248.

[61] K. S. Cranmer, "Kernel estimation in high-energy physics", *Comput. Phys. Commun.* **136** (2001) 198–207.

[62] *BABAR* Tracking Group, 2006. `www.slac.stanford.edu/BFROOT/www/Physics/TrackEfficTaskForce/TrackingTaskForce-recipe.html`.

[63] *BABAR* Neutrals Group, 2006. `www.slac.stanford.edu/BFROOT/www/Physics/Analysis/AWG/Neutrals/validation/recipie18.html`.

[64] P. Clark *et al.*, "Charmless Two-Body B Decays to $\eta'\rho$ and $\eta' K^*$." *BABAR* Internal Analysis Document BAD/960v13, (2006).

[65] LHCb Collaboration, "LHCb reoptimised detector design and performance TDR", `CERN/LHCC 2003-030`.

[66] A. Palano, "Super B Factories", *Nucl. Phys. Proc. Suppl.* **156** (2006) 105–108.

Appendix A

Toy MC Results Plots

In this appendix plots will be presented of the results of all major toy studies undertaken for the analysis.

A.1 SxF Contamination Test Toys

The results of pure toy MC studies to estimate SxF contamination as described in Section 4.4.4.

A.1.1 $\eta \to \gamma\gamma$ Case

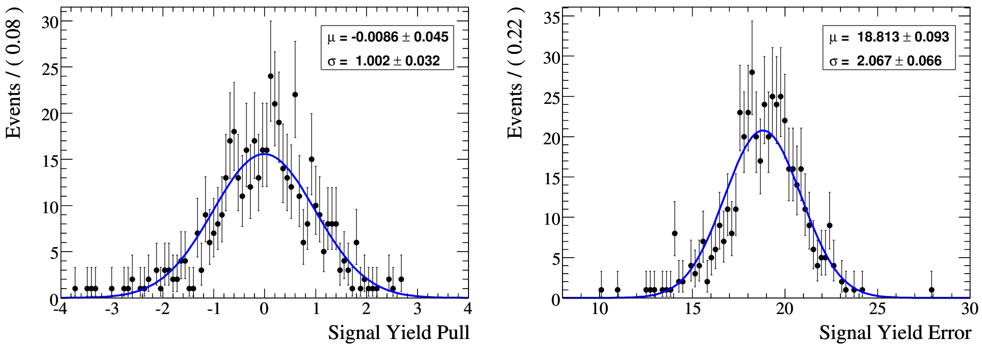

Figure *A.1*: *Toy MC result distributions for the pure toy study to assess potential contamination from SxF events in the $\eta \to \gamma\gamma$ case. The left hand plot shows the pull and the right hand plot the signal yield error. Both are fitted with a Gaussian PDF, overlaid as the blue solid line.*

A.1.2 $\eta \to \pi^+\pi^-\pi^0$ Case

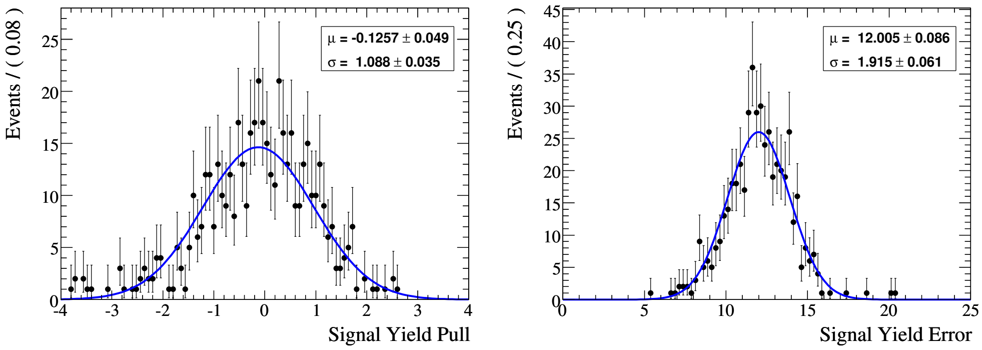

Figure *A.2*: *Toy MC result distributions for the pure toy study to assess potential contamination from SxF events in the $\eta \to \pi^+\pi^-\pi^0$ case. The left hand plot shows the pull and the right hand plot the signal yield error. Both are fitted with a Gaussian PDF, overlaid as the blue solid line.*

A.2 Best Candidate Selection Test Toys

The results of pure toy MC studies to test the need for a best B candidate selection as described in Section 4.6.5.

A.2.1 $\eta \rightarrow \gamma\gamma$ Case

No Selection Case

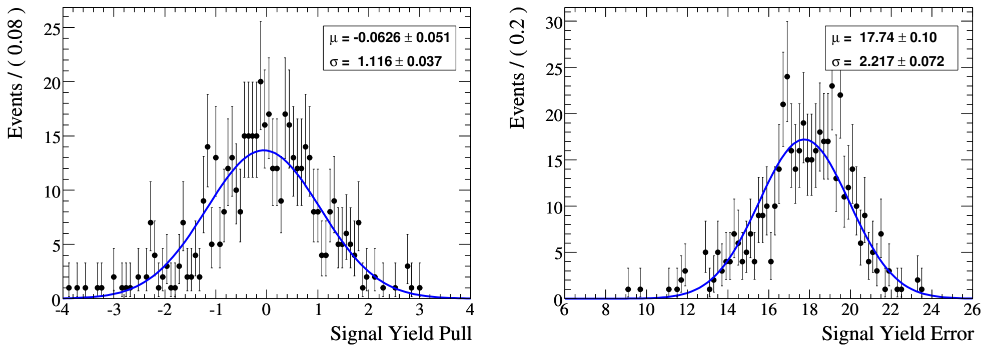

Figure *A.3:* *Toy MC result distributions for the pure toy study for samples with no best candidate selection applied in the $\eta \rightarrow \gamma\gamma$ case. The left hand plot shows the pull and the right hand plot the signal yield error. Both are fitted with a Gaussian PDF, overlaid as the blue solid line.*

$\chi^2_{\eta\pi}$ Selection Case

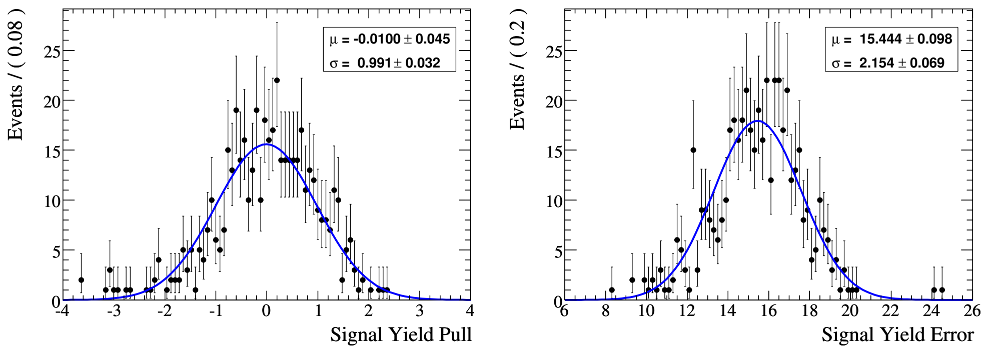

Figure *A.4:* *Toy MC result distributions for the pure toy study for samples with the best candidate selection applied in the $\eta \rightarrow \gamma\gamma$ case. The left hand plot shows the pull and the right hand plot the signal yield error. Both are fitted with a Gaussian PDF, overlaid as the blue solid line.*

A.2.2 $\eta \to \pi^+\pi^-\pi^0$ Case

No Selection Case

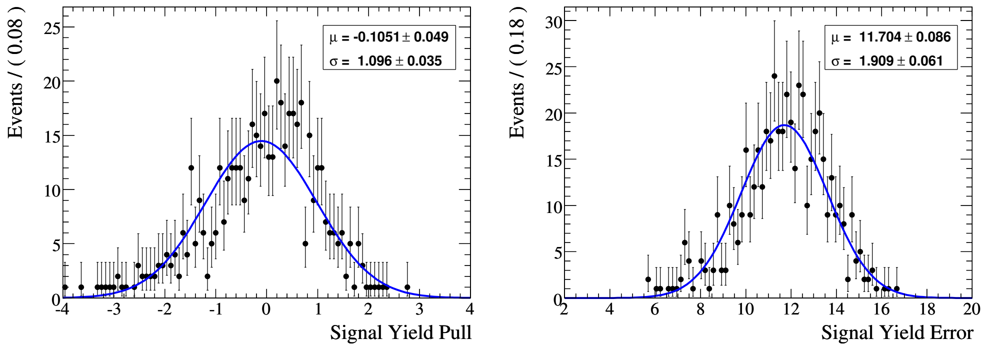

Figure *A.5:* *Toy MC result distributions for the pure toy study for samples with no best candidate selection applied in the $\eta \to \pi^+\pi^-\pi^0$ case. The left hand plot shows the pull and the right hand plot the signal yield error. Both are fitted with a Gaussian PDF, overlaid as the blue solid line.*

$\chi^2_{\eta\pi}$ Selection Case

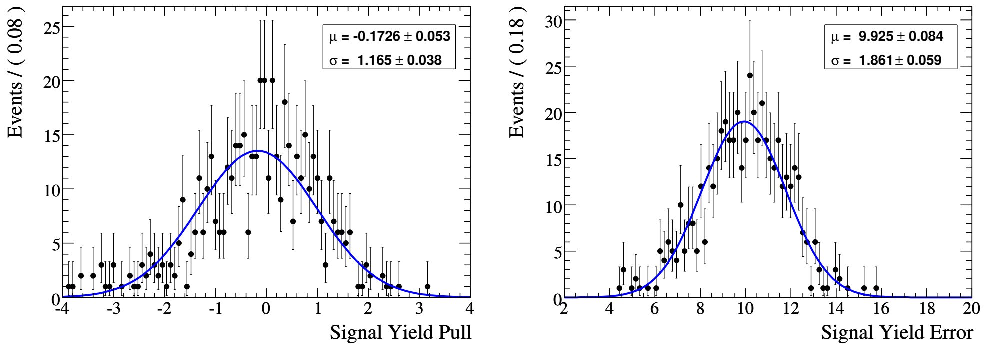

Figure *A.6:* *Toy MC result distributions for the pure toy study for samples with the best candidate selection applied in the $\eta \to \pi^+\pi^-\pi^0$ case. The left hand plot shows the pull and the right hand plot the signal yield error. Both are fitted with a Gaussian PDF, overlaid as the blue solid line.*

A.3 Full Toys

The results of full pure and embedded toy MC studies with the complete fit model as described in Section 4.8.1.

A.3.1 $\eta \to \gamma\gamma$ Case

Pure Toy

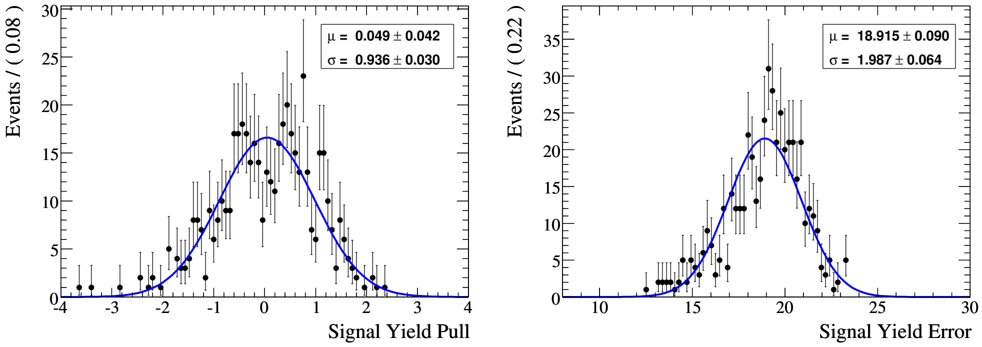

Figure *A.7*: *Toy MC result distributions for the full pure toy study in the $\eta \to \gamma\gamma$ case. The left hand plot shows the pull and the right hand plot the signal yield error. Both are fitted with a Gaussian PDF, overlaid as the blue solid line.*

Embedded Toy

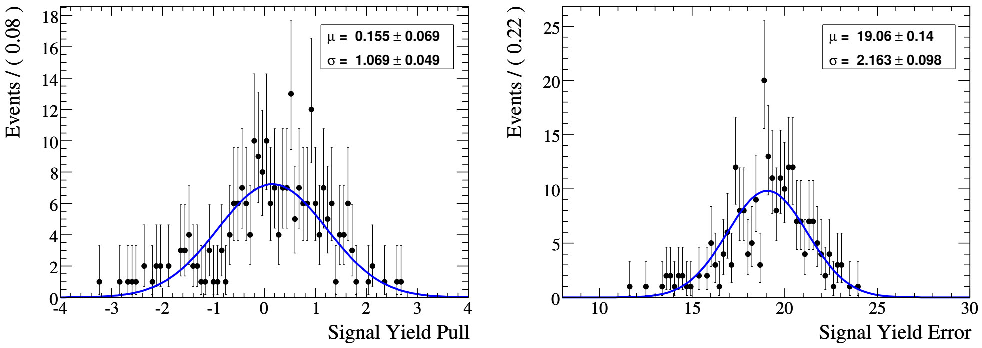

Figure *A.8*: *Toy MC result distributions for the full embedded toy study in the $\eta \to \gamma\gamma$ case. The left hand plot shows the pull and the right hand plot the signal yield error. Both are fitted with a Gaussian PDF, overlaid as the blue solid line.*

A.3.2 $\eta \rightarrow \pi^+\pi^-\pi^0$ Case

Pure Toy

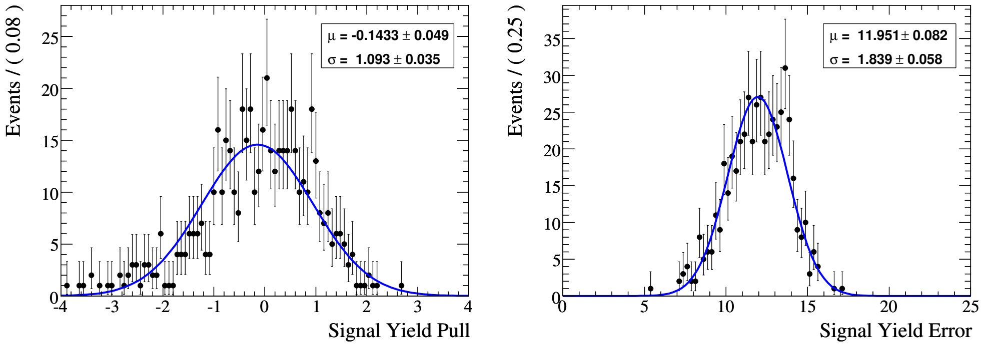

Figure *A.9:* *Toy MC result distributions for the full pure toy study in the $\eta \rightarrow \pi^+\pi^-\pi^0$ case. The left hand plot shows the pull and the right hand plot the signal yield error. Both are fitted with a Gaussian PDF, overlaid as the blue solid line.*

Embedded Toy

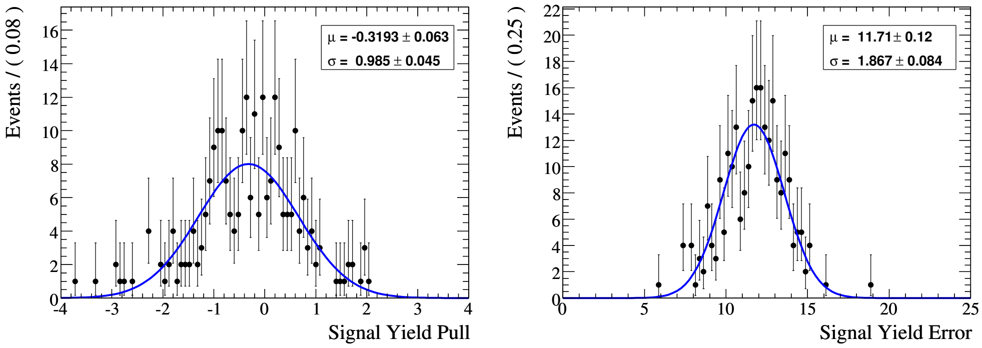

Figure *A.10:* *Toy MC result distributions for the full embedded toy study in the $\eta \rightarrow \pi^+\pi^-\pi^0$ case. The left hand plot shows the pull and the right hand plot the signal yield error. Both are fitted with a Gaussian PDF, overlaid as the blue solid line.*

A.4 Final Fit Bias Measurement Toys

The results of embedded toy MC studies to estimate the bias in the final fit result as described in Section 5.1.

A.4.1 $\eta \to \gamma\gamma$ Case

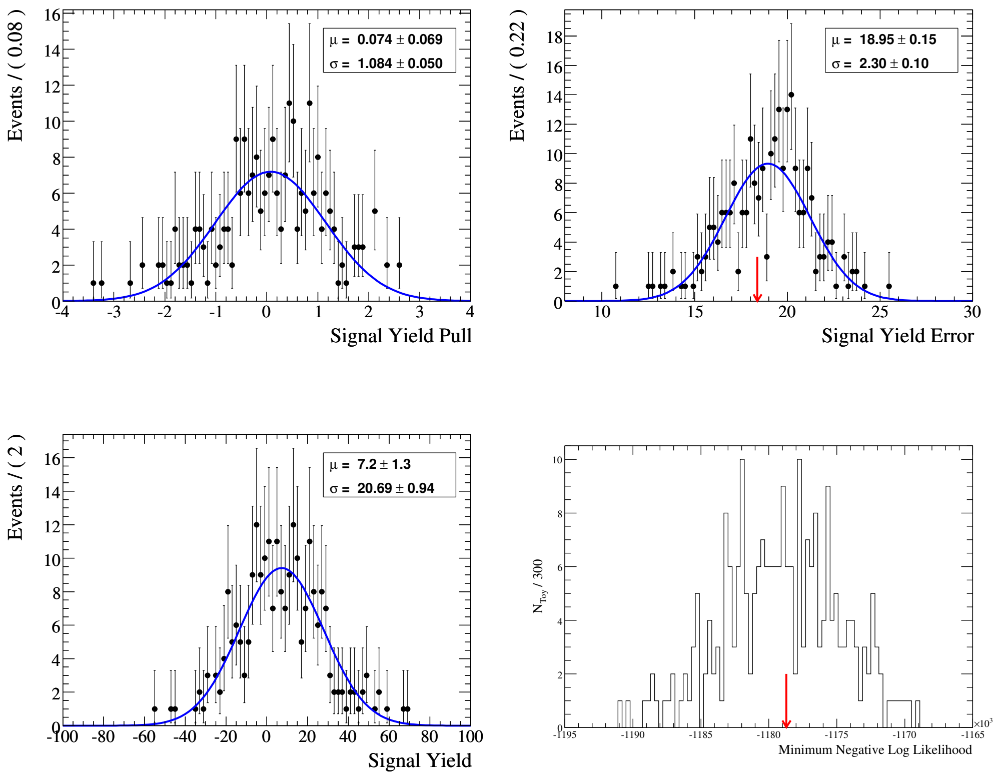

Figure *A.11*: *Toy MC result distributions for the embedded toy study to estimate the fit bias in the $\eta \to \gamma\gamma$ case. The top left hand plot shows the pull and the top right hand plot the signal yield error. The bottom plot shows the signal yield itself. All distributions are fitted with a Gaussian PDF, overlaid as the blue solid line. Finally, the bottom right plot shows the distribution of minimised negative log likelihood results from the toys. Where appropriate the value from the main fit to data is indicated by the red solid arrow.*

A.4.2 $\eta \to \pi^+\pi^-\pi^0$ Case

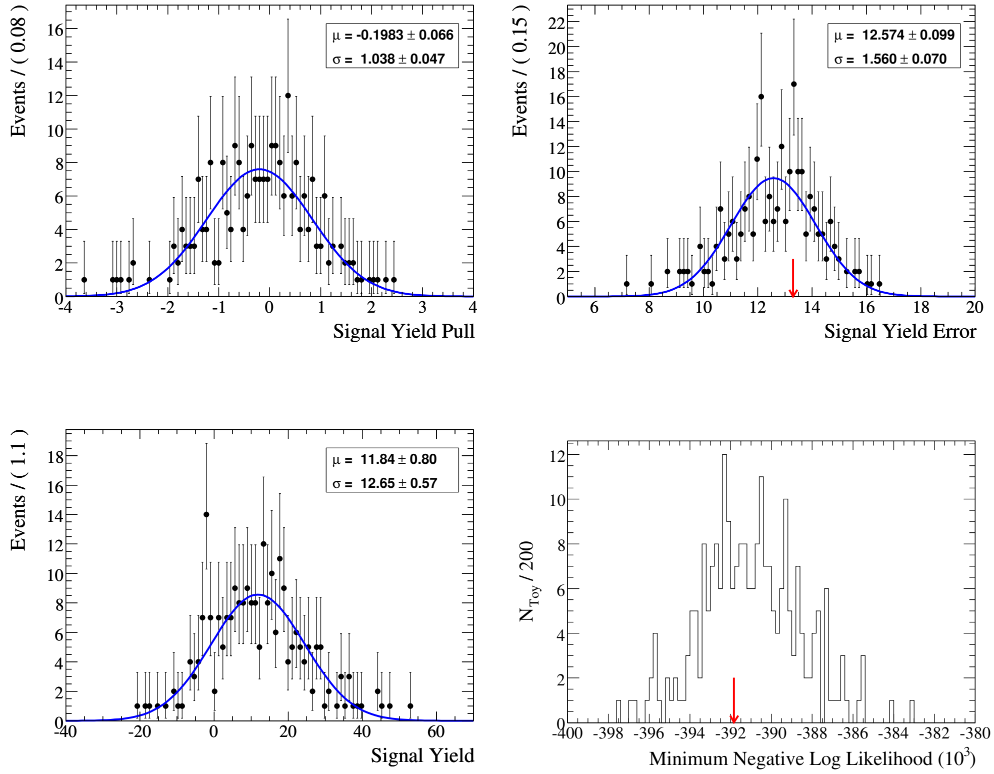

Figure ***A.12:*** *Toy MC result distributions for the embedded toy study to estimate the fit bias in the $\eta \to \pi^+\pi^-\pi^0$ case. The top left hand plot shows the pull and the top right hand plot the signal yield error. The bottom plot shows the signal yield itself. All distributions are fitted with a Gaussian PDF, overlaid as the blue solid line. Finally, the bottom right plot shows the distribution of minimised negative log likelihood results from the toys. Where appropriate the value from the main fit to data is indicated by the red solid arrow.*

Appendix B

Background PDFs

This appendix contains plots showing the MC distributions and PDFs used to model all of the backgrounds accounted for in the analysis.

B.1 $\eta \to \gamma\gamma$ Case

The following plots show the background MC distributions and PDFs for continuum, charmed B and finally charmless B components for the $\eta \to \gamma\gamma$ case. The functional forms used are described in Sections 4.5 and 4.6.

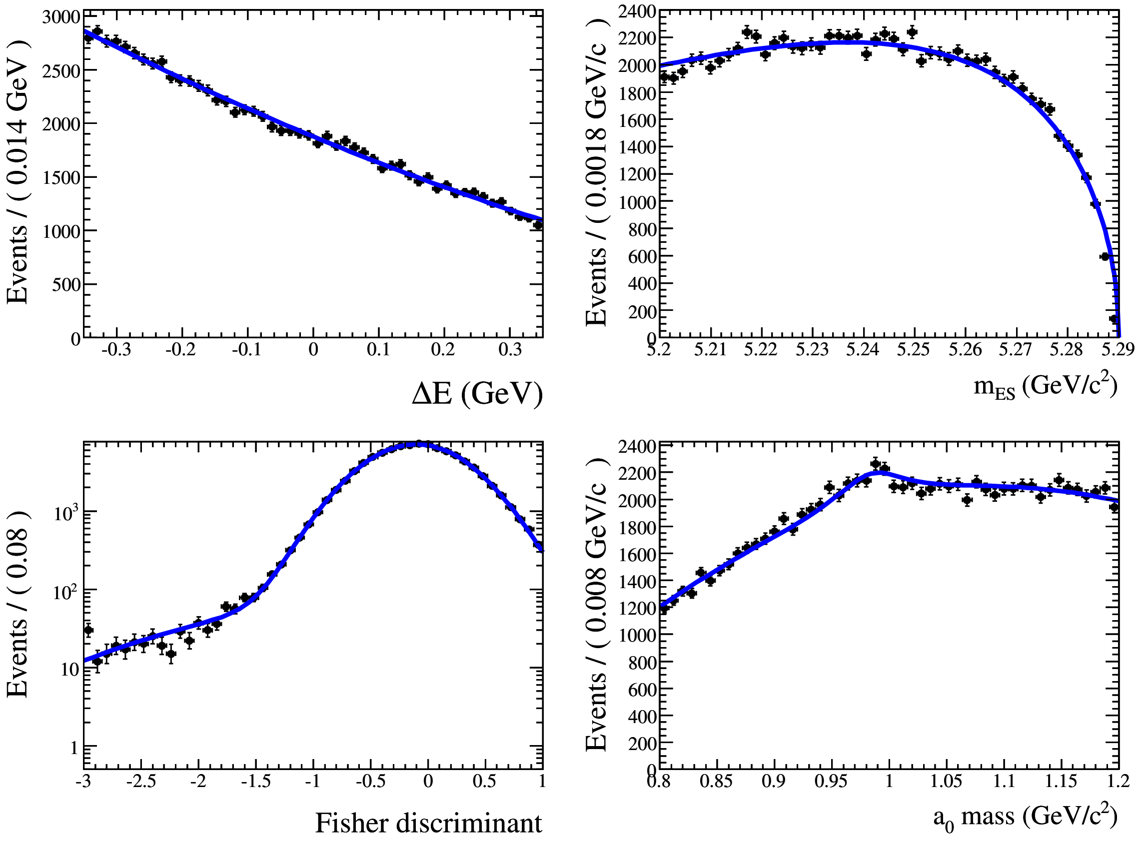

Figure *B.1:* *Continuum background MC distributions and PDFs for the $\eta \to \gamma\gamma$ case. The top left plot shows ΔE, top right m_{ES}, bottom left Fisher and bottom right the a_0 mass. The MC data are represented by the black points while the overall PDF is represented by the blue curve.*

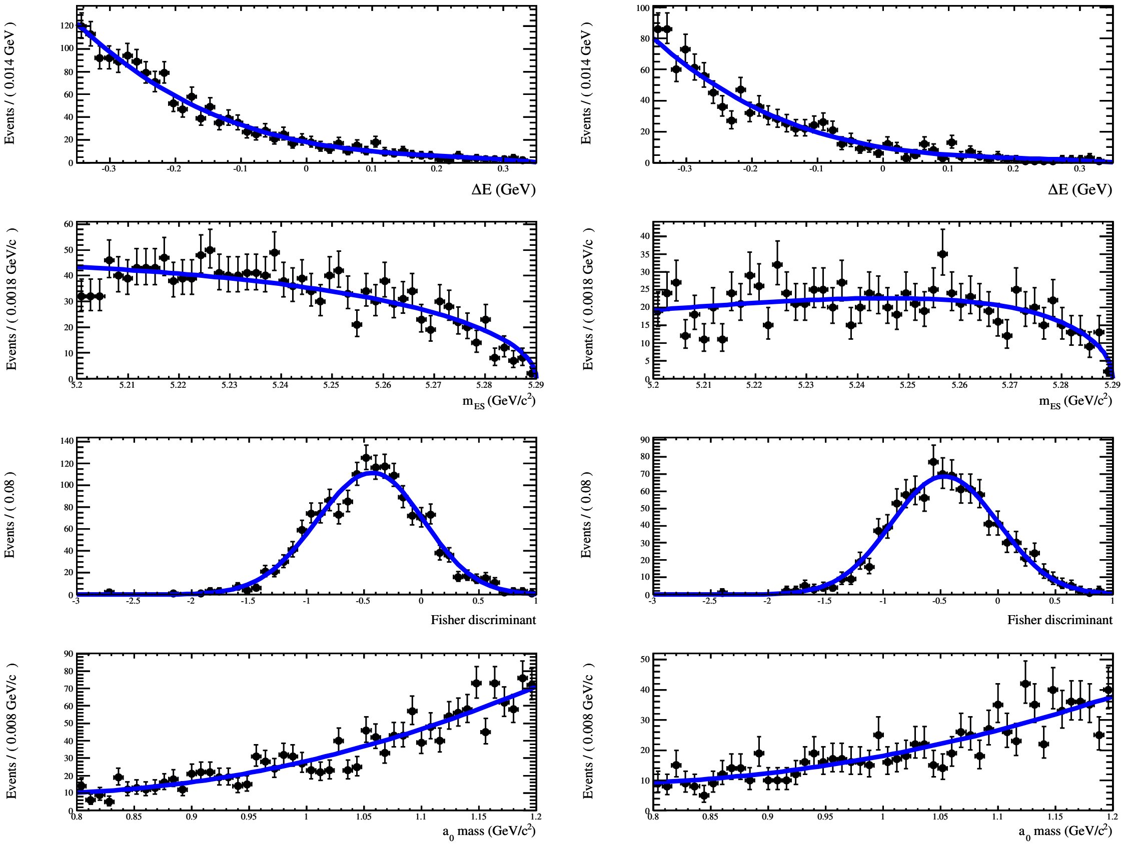

Figure *B.2:* *Charmed B background MC distributions and PDFs for the $\eta \to \gamma\gamma$ case. The left hand column contains the distributions for the charged B case and the right hand column those for the neutral B case. Both columns contain, in descending order: ΔE, m_{ES}, Fisher and finally the a_0 mass. The MC data are represented by the black points while the overall PDF is represented by the blue curve.*

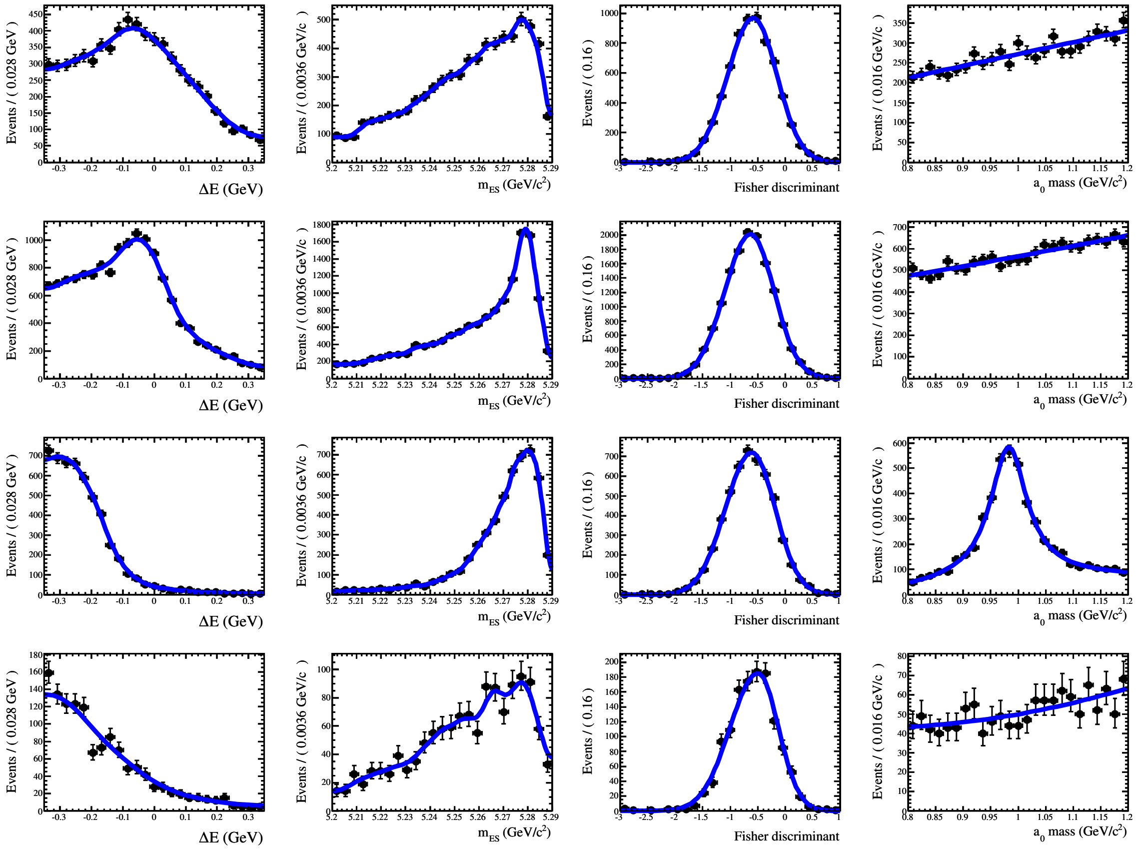

Figure *B.3:* *Charmless B background MC distributions and PDFs for the $\eta \to \gamma\gamma$ case. Starting from the left the first column contains the distributions for the ΔE variable, the second $m_{\rm ES}$, the third Fisher and the fourth the a_0 mass. Row-wise, starting from the top, the modes shown are: $B^{\pm} \to \rho^{\pm}\eta$, $B^{\pm} \to \rho(1450)^{\pm}\eta$, $B^0 \to a_0^{\pm}\rho^{\mp}$ and finally $B^{\pm} \to a_0^0\rho^{\pm}$.*

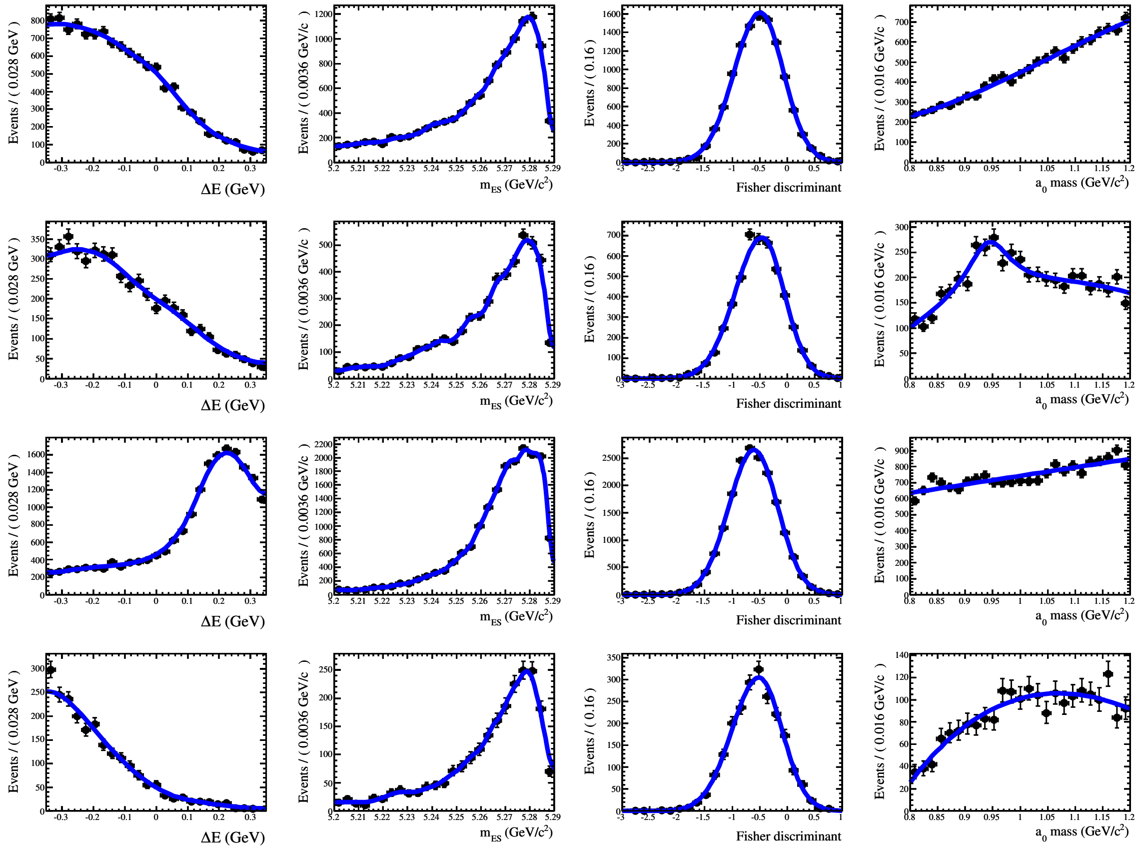

Figure *B.4:* *Charmless B background MC distributions and PDFs for the $\eta \to \gamma\gamma$ case. Starting from the left the first column contains the distributions for the ΔE variable, the second $m_{\rm ES}$, the third Fisher and the fourth the a_0 mass. Row-wise, starting from the top, the modes shown are: non-resonant $B^{\pm} \to \pi^{\pm}\pi^0\pi^0$, $B^{\pm} \to \rho^{\pm}\pi^0$, $B^0 \to \eta\pi^0$ and finally $B^{\pm} \to a_1^{\pm}\pi^0$.*

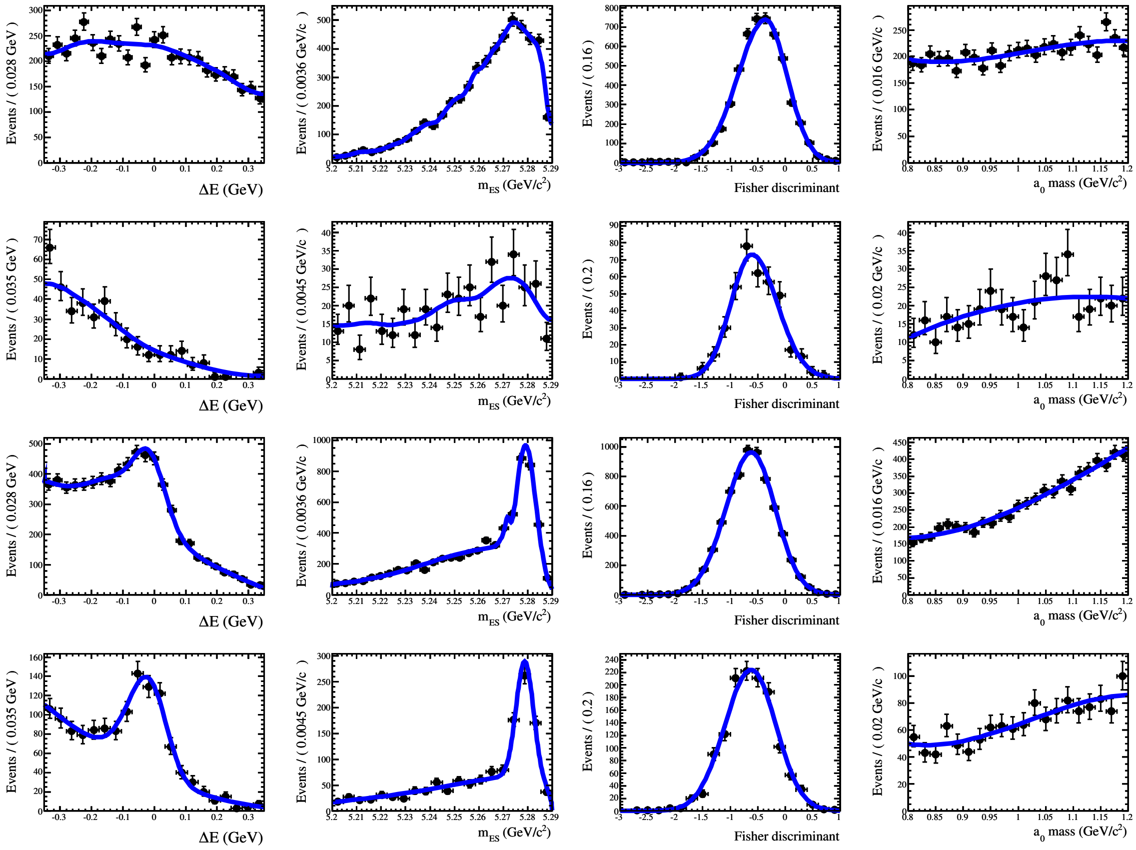

Figure *B.5:* *Charmless B background MC distributions and PDFs for the $\eta \to \gamma\gamma$ case. Starting from the left the first column contains the distributions for the ΔE variable, the second m_{ES}, the third Fisher and the fourth the a_0 mass. Row-wise, starting from the top, the modes shown are: $B^0 \to \pi^0\pi^0$, combined $B \to X_s\gamma$, $B^\pm \to a_0(1450)^\pm\pi^0$ and finally non-resonant $B^\pm \to \eta\pi^\pm\pi^0$.*

B.2 $\eta \to \pi^+\pi^-\pi^0$ Case

The following plots show the background MC distributions and PDFs for continuum, charmed B and finally charmless B components for the $\eta \to \pi^+\pi^-\pi^0$ case. The functional forms used are described in Sections 4.5 and 4.6.

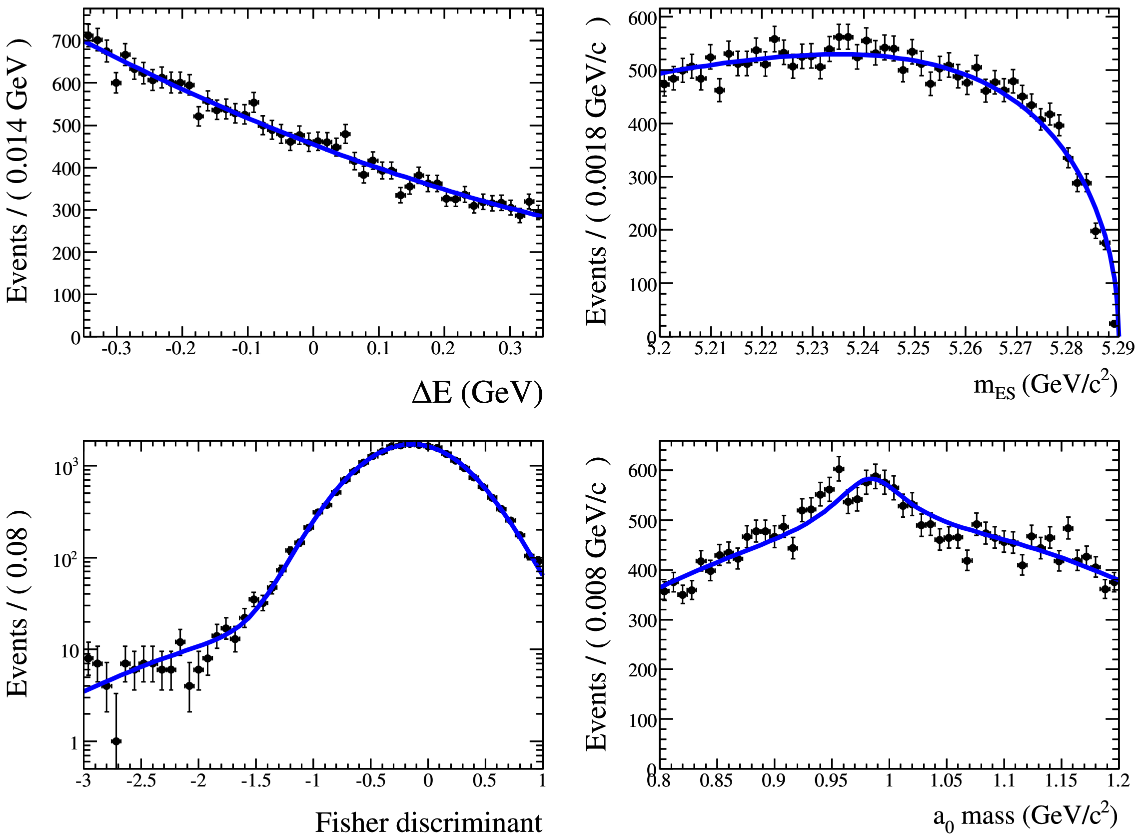

Figure *B.6:* *Continuum background MC distributions and PDFs for the $\eta \to \pi^+\pi^-\pi^0$ case. The top left plot shows ΔE, top right m_{ES}, bottom left Fisher and bottom right the a_0 mass. The MC data are represented by the black points while the overall PDF is represented by the blue curve.*

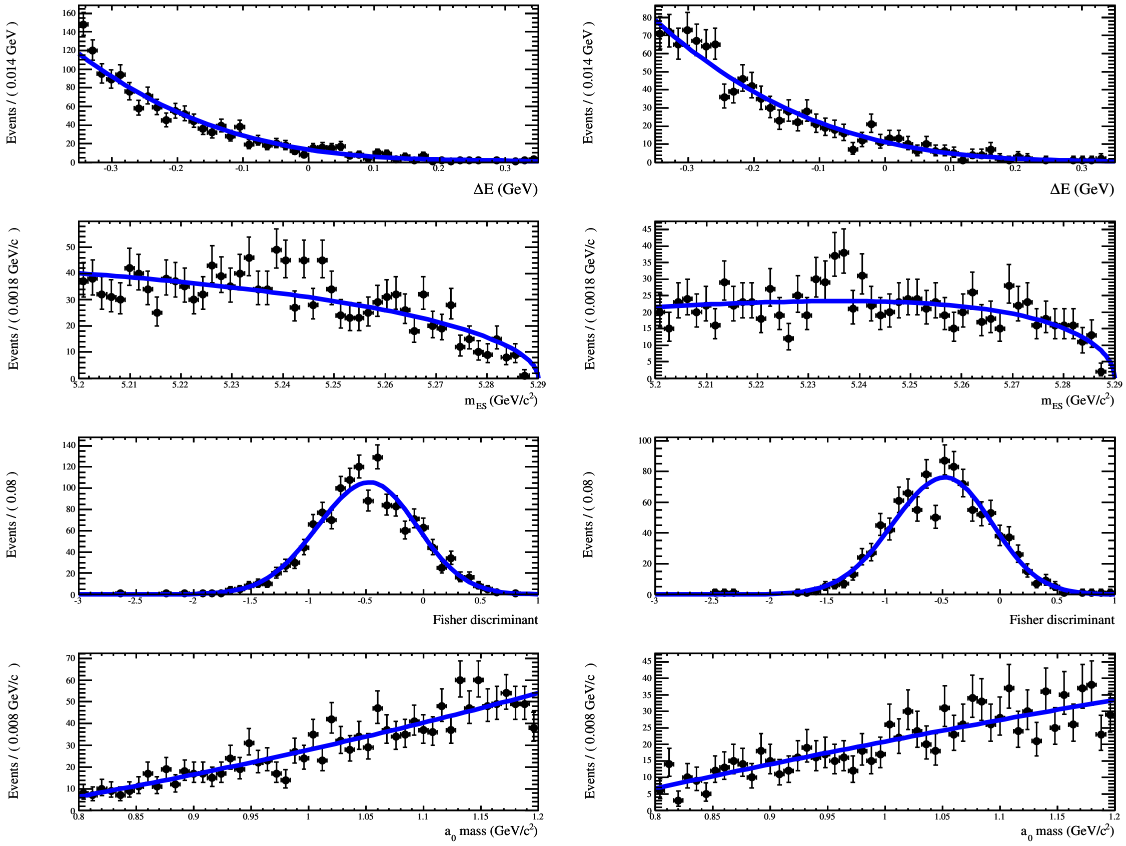

Figure *B.7:* *Charmed B background MC distributions and PDFs for the $\eta \to \pi^+\pi^-\pi^0$ case. The left hand column contains the distributions for the charged B case and the right hand column those for the neutral B case. Both columns contain, in descending order: ΔE, m_{ES}, Fisher and finally the a_0 mass. The MC data are represented by the black points while the overall PDF is represented by the blue curve.*

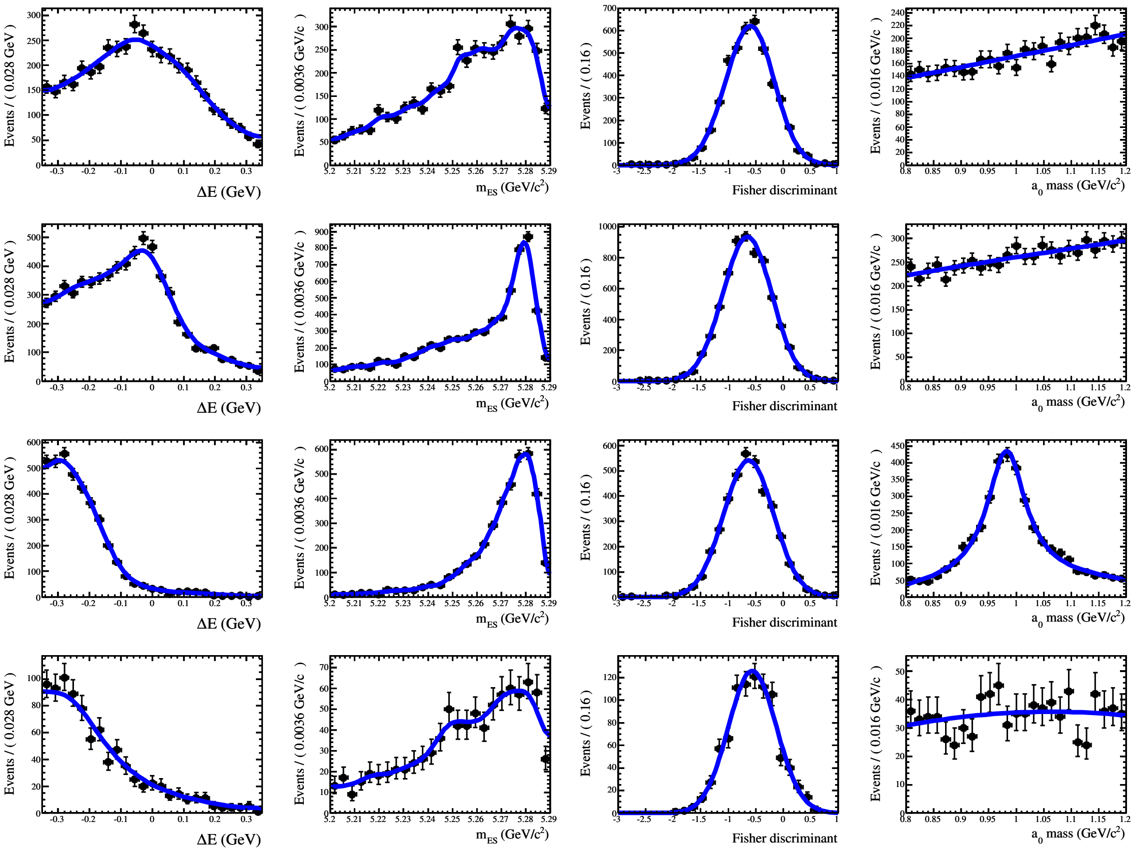

Figure *B.8:* *Charmless B background MC distributions and PDFs for the $\eta \rightarrow \pi^+\pi^-\pi^0$ case. Starting from the left the first column contains the distributions for the ΔE variable, the second m_{ES}, the third Fisher and the fourth the a_0 mass. Row-wise, starting from the top, the modes shown are: $B^\pm \rightarrow \rho^\pm\eta$, $B^\pm \rightarrow \rho(1450)^\pm\eta$, $B^0 \rightarrow a_0^\pm\rho^\mp$ and finally $B^\pm \rightarrow a_0^0\rho^\pm$.*

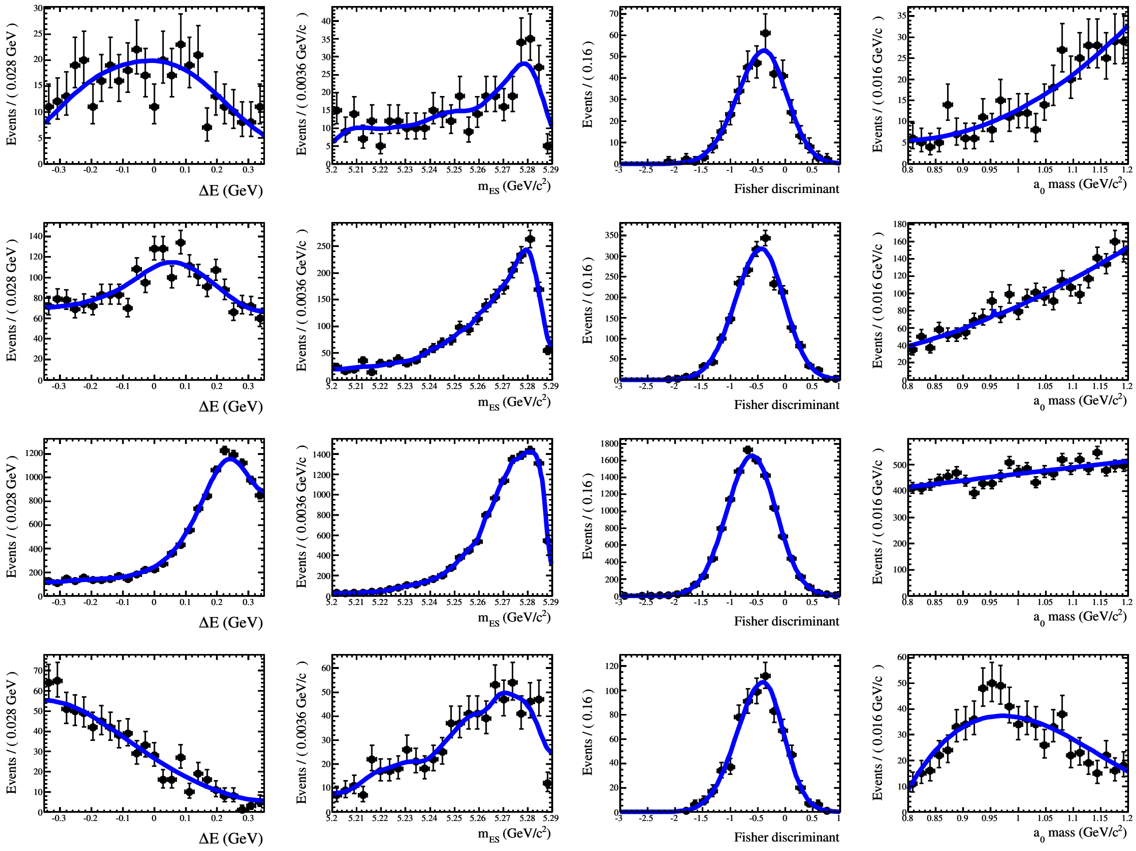

Figure *B.9:* *Charmless B background MC distributions and PDFs for the $\eta \to \pi^+\pi^-\pi^0$ case. Starting from the left the first column contains the distributions for the ΔE variable, the second $m_{\rm ES}$, the third Fisher and the fourth the a_0 mass. Row-wise, starting from the top, the modes shown are: non-resonant $B^\pm \to \pi^\pm\pi^0\pi^0$, $B^\pm \to \rho^\pm\pi^0$, $B^0 \to \eta\pi^0$ and finally $B^\pm \to \rho^\pm\omega$ (longitudinal polarisation).*

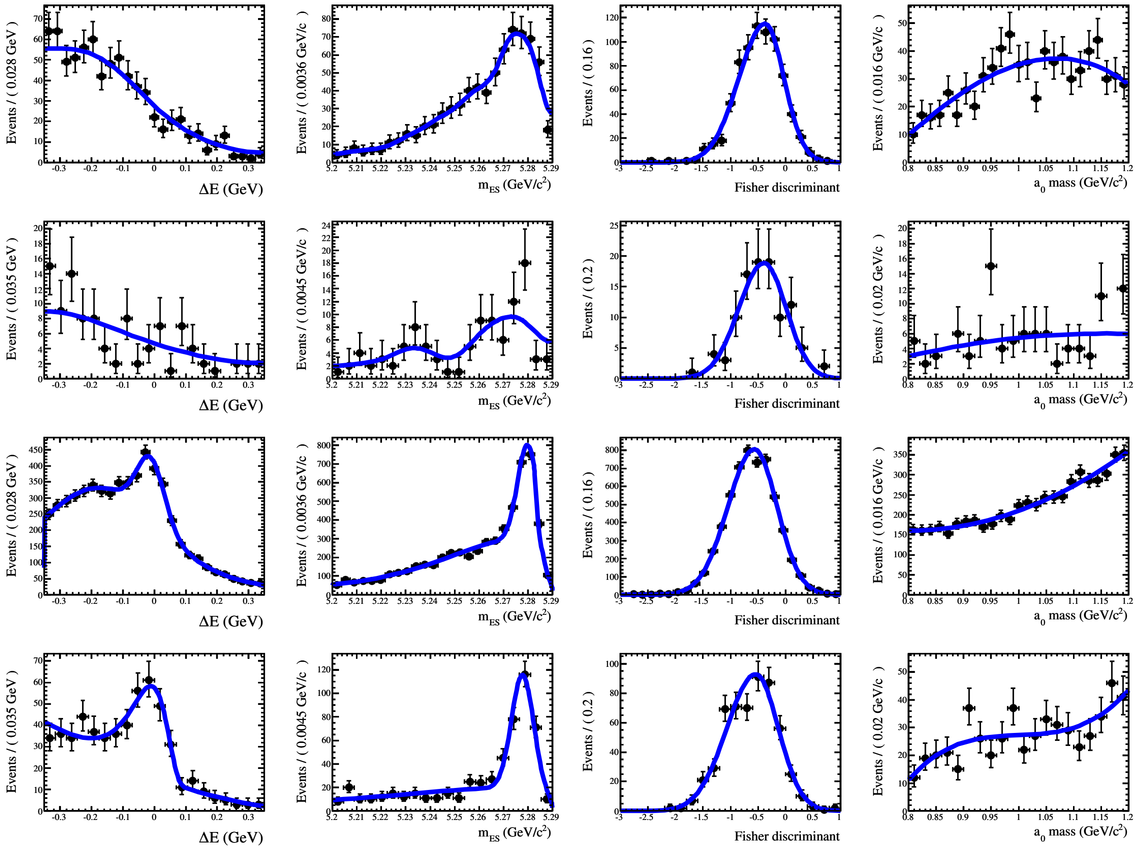

Figure *B.10:* *Charmless B background MC distributions and PDFs for the $\eta \to \pi^+\pi^-\pi^0$ case. Starting from the left the first column contains the distributions for the ΔE variable, the second $m_{\rm ES}$, the third Fisher and the fourth the a_0 mass. Row-wise, starting from the top, the modes shown are: $B^\pm \to b_1^\pm\pi^0$, combined $B \to X_s\gamma$, $B^\pm \to a_0(1450)^\pm\pi^0$ and finally non-resonant $B^\pm \to \eta\pi^\pm\pi^0$.*

Appendix C

DP Fit Result Plots

This appendix contains plots pertaining to the fits to data in regions of the Dalitz plot for the $B^{\pm} \rightarrow a_0(1450)^{\pm}\pi^0$ and non-resonant $B^{\pm} \rightarrow \eta\pi^{\pm}\pi^0$ modes. For each mode experimental data distributions will be presented with the continuum, combined B background and 'signal' PDFs overlaid. The likelihood scan plot, used to calculate the upper limit for each mode, will also be presented.

C.1 $B^{\pm} \rightarrow a_0(1450)^{\pm}\pi^0$ Fit

Fit result plots for the $B^{\pm} \rightarrow a_0(1450)^{\pm}\pi^0$ fit showing the $\eta \rightarrow \gamma\gamma$ and $\eta \rightarrow \pi^+\pi^-\pi^0$ cases. Finally the likelihood scan for the separate and combined modes is presented.

C.1.1 $\eta \rightarrow \gamma\gamma$ Case

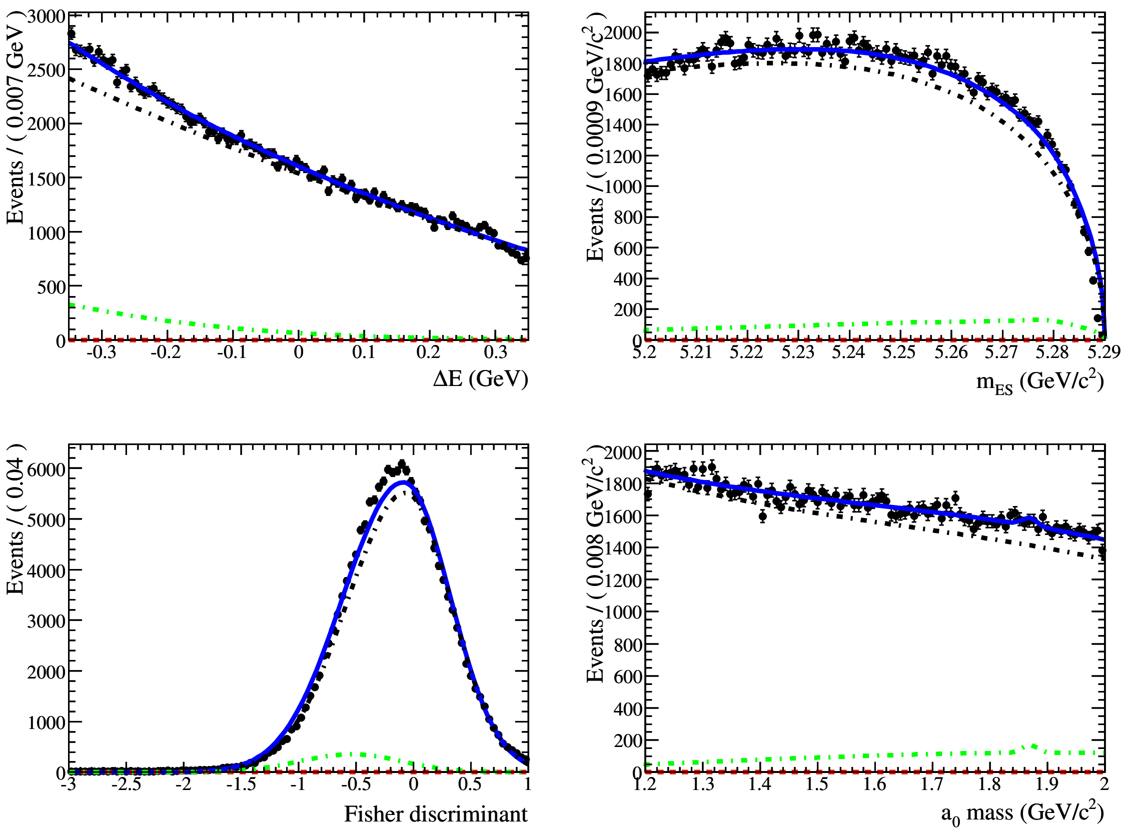

Figure *C.1*: *Experimental data distribution and PDFs for the $B^{\pm} \rightarrow a_0(1450)^{\pm}\pi^0$ DP fit in the $\eta \rightarrow \gamma\gamma$ case. The top left plot shows the ΔE variable, top right m_{ES}, bottom left Fisher and bottom right the a_0 mass. The data are represented by the black points while the overall PDF is represented by the blue curve. The continuum background is represented by the black dot-dashed curve while the combined B background is represented by the Green dot-dashed curve. The signal is represented by the red dashed curve.*

C.1.2 $\eta \to \pi^+\pi^-\pi^0$ Case

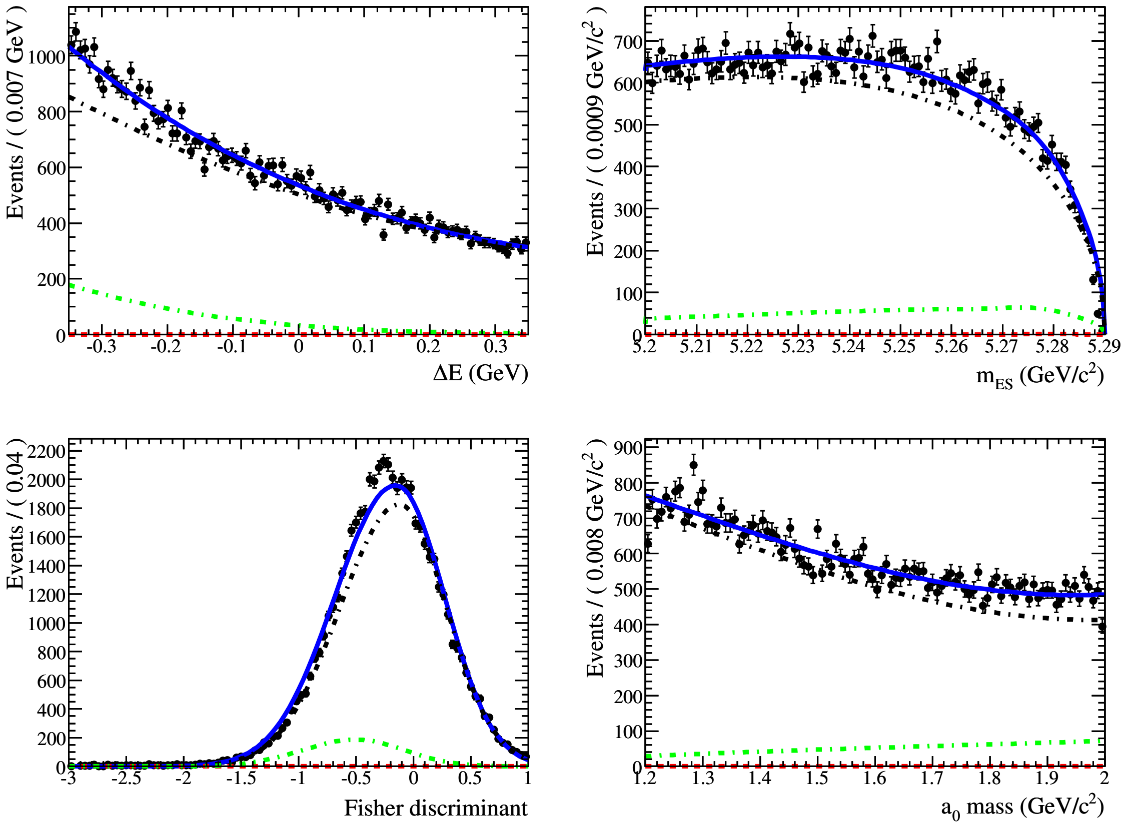

Figure ***C.2:*** *Experimental data distribution and PDFs for the $B^\pm \to a_0(1450)^\pm\pi^0$ DP fit in the $\eta \to \pi^+\pi^-\pi^0$ case. The top left plot shows the ΔE variable, top right $m_{\rm ES}$, bottom left Fisher and bottom right the a_0 mass. The data are represented by the black points while the overall PDF is represented by the blue curve. The continuum background is represented by the black dot-dashed curve while the combined B background is represented by the Green dot-dashed curve. The signal is represented by the red dashed curve.*

C.1.3 Combined Likelihoods

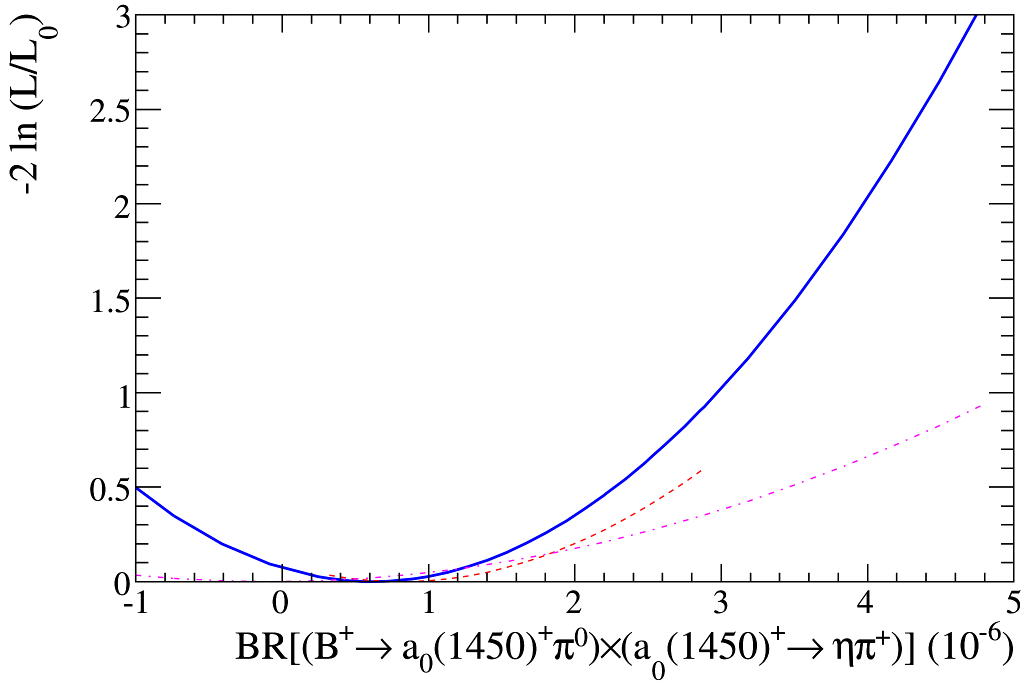

Figure *C.3:* *Likelihood scan for a range of $B^{\pm} \to a_0(1450)^{\pm}\pi^0$ ($a_0(1450)^{\pm} \to \eta\pi^{\pm}$) branching fractions for both η decay modes. The red dashed curve represents the $\eta \to \gamma\gamma$ contribution with the $\eta \to \pi^+\pi^-\pi^0$ represented by the pink dot-dashed curve. The blue solid curve represents the combined likelihood for the two modes. The upper limit on the branching fraction is set using the blue curve.*

C.2 $B^{\pm} \rightarrow \eta\pi^{\pm}\pi^{0}$ (Non-resonant) Fit

Fit result plots for the $B^{\pm} \rightarrow \eta\pi^{\pm}\pi^{0}$ (non-resonant) fit showing the $\eta \rightarrow \gamma\gamma$ and $\eta \rightarrow \pi^{+}\pi^{-}\pi^{0}$ cases. Finally the likelihood scan for the separate and combined modes is presented.

C.2.1 $\eta \rightarrow \gamma\gamma$ Case

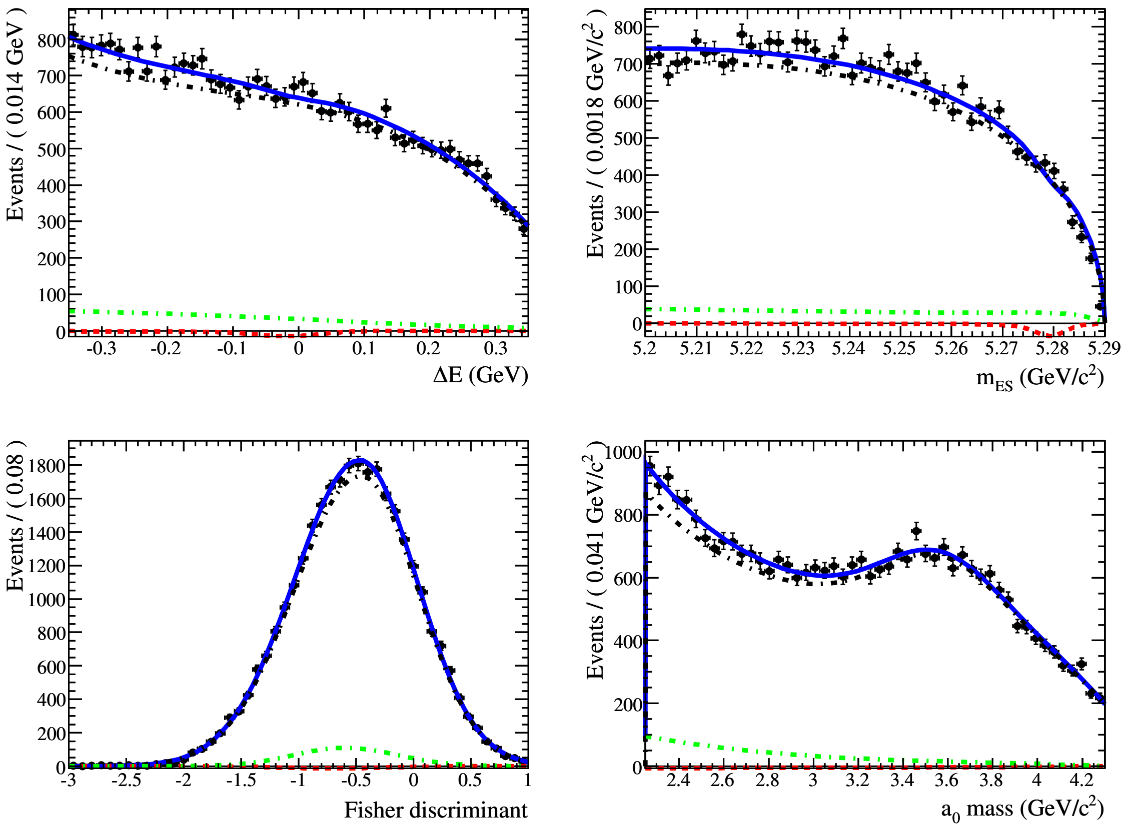

Figure ***C.4:*** *Experimental data distribution and PDFs for the $B^{\pm} \rightarrow \eta\pi^{\pm}\pi^{0}$ (non-resonant) DP fit in the $\eta \rightarrow \gamma\gamma$ case. The top left plot shows the ΔE variable, top right m_{ES}, bottom left Fisher and bottom right the a_0 mass. The data are represented by the black points while the overall PDF is represented by the blue curve. The continuum background is represented by the black dot-dashed curve while the combined B background is represented by the Green dot-dashed curve. The signal is represented by the red dashed curve.*

C.2.2 $\eta \to \pi^+\pi^-\pi^0$ Case

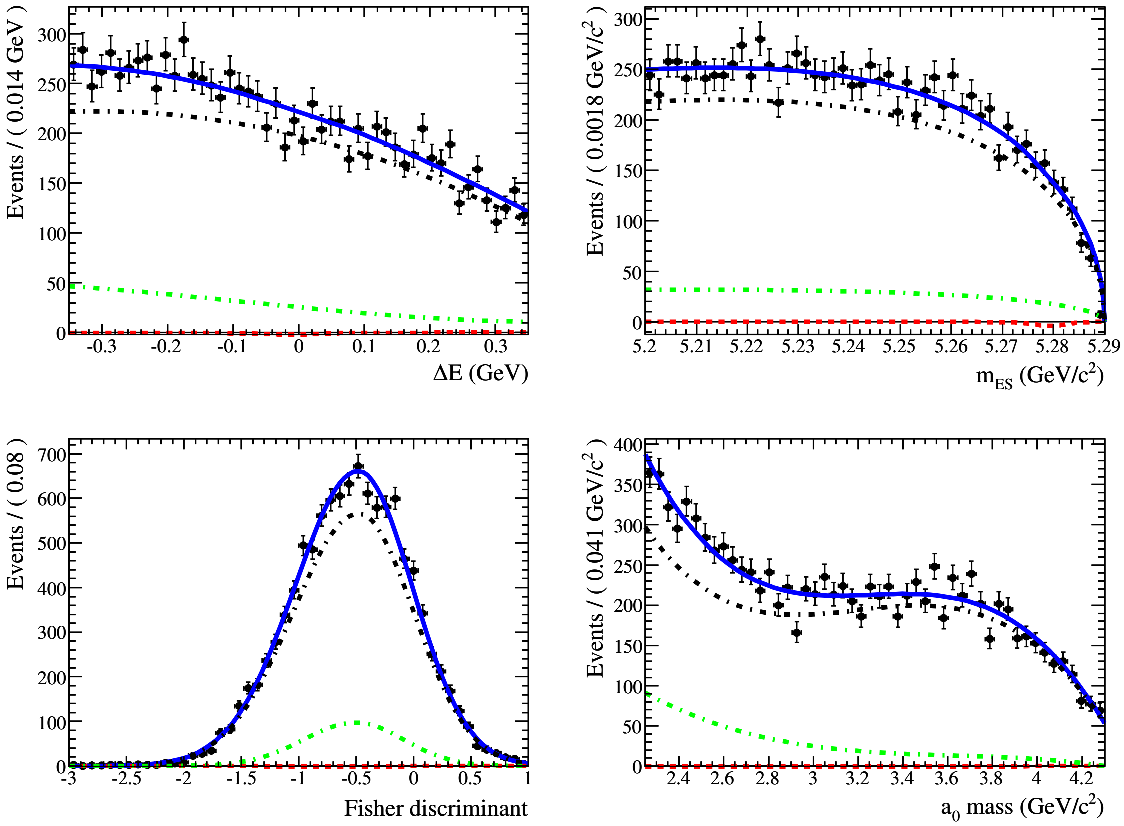

Figure ***C.5:*** *Experimental data distribution and PDFs for the $B^{\pm} \to \eta\pi^{\pm}\pi^0$ (non-resonant) DP fit in the $\eta \to \pi^+\pi^-\pi^0$ case. The top left plot shows the ΔE variable, top right m_{ES}, bottom left Fisher and bottom right the a_0 mass. The data are represented by the black points while the overall PDF is represented by the blue curve. The continuum background is represented by the black dot-dashed curve while the combined B background is represented by the Green dot-dashed curve. The signal is represented by the red dashed curve.*

C.2.3 Combined Likelihoods

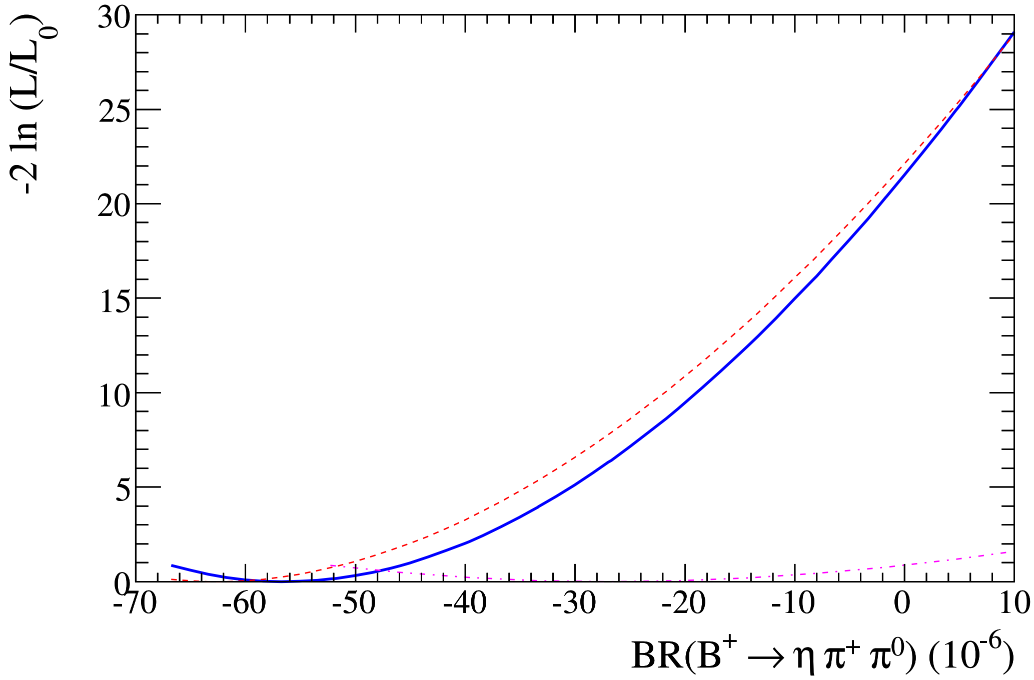

Figure *C.6:* *Likelihood scan for a range of* $B^{\pm} \to \eta\pi^{\pm}\pi^{0}$ *(non-resonant) branching fractions for both* η *decay modes. The red dashed curve represents the* $\eta \to \gamma\gamma$ *contribution with the* $\eta \to \pi^{+}\pi^{-}\pi^{0}$ *represented by the pink dot-dashed curve. The blue solid curve represents the combined likelihood for the two modes. The upper limit on the branching fraction is set using the blue curve.*