SLAC-R-922

April 2009

On decays of B mesons to a strange meson and an η or η' meson at *BABAR*

by

James Francis Hirschauer

B.A., Xavier University, 1999

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Physics

2009

SLAC National Accelerator Laboratory, Menlo Park, Ca 94025

Work supported in part by US Department of Energy contract DE-AC02-76SF00515

This thesis entitled:
On decays of B mesons to a strange meson and an η or η' meson at *BABAR*
written by James Francis Hirschauer
has been approved for the Department of Physics

William T. Ford

James G. Smith

Date _____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Hirschauer, James Francis (Ph.D., Physics)

On decays of B mesons to a strange meson and an η or η' meson at *BABAR*

Thesis directed by Professor William T. Ford

We describe studies of the decays of B mesons to final states $\eta K^*(892)$, $\eta K_0^*(S$-wave), $\eta K_2^*(1430)$, and $\eta' K$ based on data collected with the *BABAR* detector at the PEP-II asymmetric-energy e^+e^- collider at the Stanford Linear Accelerator Center. We measure branching fractions and charge asymmetries for the decays $B \to \eta K^*$, where K^* indicates a spin 0, 1, or 2 $K\pi$ system, making first observations of decays to final states $\eta K_0^{*0}(S$-wave), $\eta K_0^{*+}(S$-wave), and $\eta K_2^{*0}(1430)$. We measure the time-dependent CP violation parameters S and C for the decays $B^0 \to \eta' K^0$, observing CP violation in a charmless B decay with 5σ significance considering both statistical and systematic uncertainties.

"I was determined to know beans."

– Henry Thoreau, *Walden*

For Maura

Acknowledgements

First, I would like to remember Tony Barker. Tony was my good friend during my days of skiing and my advisor during my early days of research. He was always brilliant and always generous. Paradoxically, in his free time, he was susceptible to the lure of easygoing adventure, while in his work he demanded, of himself and his colleagues, diligence and impeccable results. It pleases me when I detect Tony's influence in my thoughts and my words. I am fortunate to have been his friend.

Bill Ford and Jim Smith turned me into a scientist. I came to them with a paucity of computing skills, ignorance of the methods of statistical analysis, a vague understanding of the Standard Model, and a deep curiosity. Through their patient guidance and spirited encouragement not only were my shortcomings remedied, but I have been well prepared for the next step in my career. Thank you both.

I thank my Colorado friends: Mike Wilking, Keith Ulmer, Fred Blanc, Bob Nelson, Pat Toale, Jesper Levinsen, Josh Spitz, and Rune Niclasen. In the end, I think Mike and I probably discussed physics more than baseball or politics, which is saying something; I am glad that my philosophy has been shaped very much by Mike. I may have eventually burst if I had not been able to slowly release my stupid questions on Keith and Fred. I think am most grateful to Bob, Pat, Jesper, and Josh for their off-the-clock influences: beer, wiffle ball, football, foosball, the cube, skiing, and beer. Rune is a great friend; I miss very much our conversations, which, regardless of topic, would eventually be reduced into an analysis of fundamental concepts, and I think Rune *does*

know what that means.

I thank my SLAC friends: Tae Min Hong, Alessandro Gaz, Bryan Fulsom, Tim Piatenko, Jake Anderson, Joe Tuggle, Yanyan Gao, Steve Sekula, Andy Ruland, and Chris Schilling. Each has been a great resource for physics, computing, and statistics knowledge, but more important, each has been a good friend. I will miss our lunch conversations. It has been especially fun to enter parenthood at the same time as Bryan, Tim, and Jake; strangely, the most enthusiasm for our children always seemed to come not from us, but from Tae Min.

I also thank my *BABAR* colleagues on whose advice and expertise I have often called. It was a pleasure to work on the $\eta' K^0$ analysis with our collaborators from Milan: Pietro Biassoni, Alfio Lazzaro, and Fernando Palombo. Through the hard work and cooperation of the Milan group, the Colorado group, and Gabriele Simi and Hassan Jawahery, *BABAR*'s sophisticated analysis of *CP* violation in $b \to s$ penguin modes has been documented thoroughly and concisely; I am proud of our work.

Finally, of course: Maura. It is without doubt that, though she teaches first grade while raising our daughters and running our house with little help from me, I could not have performed this research without her. (Someone should investigate *that* asymmetry!) With Maura I am happy, and that is most important.

That reminds me ... our daughters, Emma and Maggie. I state without reservation that I love them more than myself, but I fear that thanking those who have been more of a hindrance than a help will undermine the many sincere words of gratitude contained in these acknowledgments.

Contents

Chapter

Tables

Table

Figures

Figure

Chapter 1

An historical introduction

Though our inherent tendency to frame an understanding of nature in terms of symmetries has been evident since ancient times, knowledge of the importance of the violation of fundamental symmetries is recent.

In 1955, Gell-Mann and Pais, motivated by the puzzle of the θ and τ particles, introduced what we now call $K^0 - \overline{K}^0$ mixing and the form of charge conjugation C eigenstates of the $K^0 - \overline{K}^0$ system [1]. In 1956, Lee and Yang pointed out that it had not been experimentally tested whether weak processes are symmetric under parity transformations P [2]. Subsequent tests in the β-decay of ^{60}Co showed that in fact the weak interaction does not conserve parity [3]. This discovery called into question the Gell-Mann/Pais framework which relied on C symmetry, but Landau pointed out that CP should be conserved and has the same effect as C in the understanding of the K system [4]. However, in 1964, Cronin and Fitch surprised the field with the discovery of the violation of CP symmetry when they observed the decay $K_L^0 \to \pi^+\pi^-$ [5]. Three years later Sakharov pointed out that the violation of CP symmetry is required for the matter universe that we see [6].

In 1971, Glashow, Weinberg, and Salam (GWS) proposed a robust model for understanding the unified weak and electromagnetic interactions [7] in terms of a spontaneously broken $SU(2) \times U(1)$ gauge symmetry. Shortly after, in 1972, Kobayashi and Maskawa (KM) realized that the method of quark mixing proposed by Cabibbo

in 1963 [8] could be extended to three quark generations such that *CP* violation would be required in the GWS model (barring fine-tuning) [9]. The proposal of six quarks in three generations was a big leap at a time when only three quarks were known and the idea of generations was not established (though a fourth quark had been proposed by Glashow, Iliopoulos, and Maiani (GIM) in 1970 [10]). Through the work of Cabibbo, GWS, KM, and GIM, all the necessary pieces were present, and it was up to experiment to decide how they fit together.

In 1974 the J/ψ particle, understood to be charmonium ($c\bar{c}$), was independently discovered by Ting *et al.* at Brookhaven [11] and Richter *et al.* at SLAC [12]. From 1975 – 1977, the third generation of fundamental particles was established with Perl *et al.* discovering the τ lepton at SLAC [13] and Lederman *et al.* discovering the b quark at Fermilab [14]. Finally, the t quark was discovered in 1994 at Fermilab [15].

It was realized that the long b quark lifetime ($\sim$ 1.5 ps), measured by Ford *et al.* [16] and Jaros, Lockyer, *et al.* [17], along with the long $B^0 - \overline{B}^0$ mixing frequency (~ 0.5 ps^{-1}), made feasible the measurement of mixing-induced *CP* violation in the B system provided sufficient improvement in precision of vertex reconstruction. In 1987, Oddone realized that the existing vertex reconstruction techniques, with resolution of $\sim 100\,\mu$m, were sufficient if experiments could increase the separation of B^0 and $\overline{B}^0$ decay vertices (to $\sim 250\,\mu$m) by boosting the center-of-mass system with respect to the laboratory frame [18]. To that end, asymmetric-energy e^+e^- colliders (the *B-factories*) were constructed at SLAC (USA) and KEK (Japan). Measurements made at the *B*-factories with the *BABAR* and Belle detectors unambiguously confirmed the theory of Kobayashi and Maskawa [19]. In 2008 Kobayashi and Maskawa were awarded the Nobel Prize in Physics (with Nambu) for their theory.

Chapter 2

Theory

2.1 Symmetries

A symmetry is any transformation for which the behavior of a physical system does not change. Noether's theorem, informally stated, says that there is a conserved quantity for every symmetry of nature. For instance, rotations in space leave the laws of physics invariant implying conservation of angular momentum. Rotations in Lorentz space (boosts and rotations in three-space) are continuous transformations that preserve the Minkowski interval. In addition to these continuous transformations, there exist discrete transformations that also preserve the spacetime interval:

- space-reversal or parity (P), which changes $(t, \mathbf{x}) \rightarrow (t, -\mathbf{x})$; spin is even and momentum odd under P;
- time-reversal (T), which changes $(t, \mathbf{x}) \rightarrow (-t, \mathbf{x})$; both spin and momentum are odd under T;
- charge conjugation (C), which changes particles into anti-particles; both spin and momentum are even under C.

2.2 *CP* violation in the Standard Model

In the Standard Model (SM), three generations of fundamental fermions interact via three forces (strong, weak, and electromagnetic) mediated by vector bosons. It is

believed that C, P, and T are good symmetries of the strong and electromagnetic forces, but the vector–axial-vector ($V-A$) form of the weak force has an important effect on its transformation under C and P. Parity is manifestly violated in the weak interaction which acts almost exclusively on left-handed leptons (and right-handed anti-leptons). Also, the absence of left-handed anti-neutrinos in the theory demonstrates very simply the C violation of the weak force, since C changes left-handed neutrinos into left-handed anti-neutrinos. Nevertheless, in general, the $V-A$ form is invariant under the combined transformations of CP, T, and CPT. It should be noted that the CPT theorem states that, assuming the spin-statistics theorem, any Lorentz-invariant, local gauge quantum field theory is invariant under CPT transformations [20, 21].

2.2.1 How the particle got its mass

The SM is represented by $SU(3)_C$ gauge invariance of quark color fields, $SU(2)_L$ of weak isospin, and $U(1)_Y$ of weak hypercharge. The symmetry of the $SU(2)_L \times U(1)_Y$ subgroup is spontaneously broken by the Higgs mechanism [22] giving masses to three vector bosons and all fermions, except the neutrinos, while maintaining gauge invariance.

The Yukawa interaction of the quark fields and the Higgs field is

$$\mathcal{L}_Y = -\lambda^d_{ij}\overline{Q}^L_i \phi d^R_j - \lambda^u_{ij}\overline{Q}^L_i \phi_c u^R_j + \text{h.c.}, \tag{2.1}$$

where i and j index quark flavor, $\lambda^{u,d}$ are 3×3 matrices of dimensionless complex coupling constants, $\overline{Q}^L$ are the left-handed quark doublets, u^R and d^R are right-handed up- and down-type quark singlets, ϕ is the Higgs doublet, and $\phi_c \equiv -i\tau_2\phi^*$ is the Higgs doublet with the form necessary to give mass to the upper members of the quark doublets [23, 24].

In the flavor basis $\lambda^{u,d}$ are not diagonal. The unitary matrices $U^{u,d}$ that diagonalize $\lambda^{u,d}$ transform the quark fields in this way

$$u^{L\prime}_i = U^u_{ij} u^L_i, \qquad d^{L\prime}_i = U^d_{ij} d^L_i, \tag{2.2}$$

where the unprimed states are in the flavor basis. Writing $\mathcal{L}_Y$ in the new basis (denoted with a $'$) and replacing the Higgs field with its vacuum expectation value $\phi_0 = (0, v/\sqrt{2})$,

$$\mathcal{L}'_Y = -\frac{v}{\sqrt{2}}\lambda^{d\prime}_{ii}\overline{d}^{L\prime}_i d^{R\prime}_i\left(1+\frac{h}{v}\right) - \frac{v}{\sqrt{2}}\lambda^{u\prime}_{ii}\overline{u}^{L\prime}_i u^{R\prime}_i\left(1+\frac{h}{v}\right) + \text{h.c.}, \tag{2.3}$$

it is clear that $\mathcal{L}'_Y$ has the standard form of quark mass terms with mass matrices $m^d_{ii} = \frac{v}{\sqrt{2}}\lambda^{d\prime}_{ii}$ and $m^u_{ii} = \frac{v}{\sqrt{2}}\lambda^{u\prime}_{ii}$. Since the mass matrices $m^{u,d}$ are diagonal in this representation, this basis is known as the mass basis.

2.2.2 Flavor changing

The charged current of the weak interaction

$$J^\mu = \frac{1}{\sqrt{2}}\overline{u}^L_i\gamma^\mu d^L_i \tag{2.4}$$

seems to operate only within a single generation i, but transforming from the flavor basis to the mass basis using Eq. 2.2, we can write the charged current in a form that clearly mixes quark generations as

$$J^\mu = \frac{1}{\sqrt{2}}\overline{u}^L_i\gamma^\mu V_{ij} d^L_j. \tag{2.5}$$

In this equation, $V = (U^{u\dagger}U^d)$ is the unitary matrix known as the Cabibbo-Kobayashi-Maskawa (CKM) flavor mixing matrix [8, 9].

2.2.3 The CKM matrix and *CP* violation

The CKM matrix is a 3×3 unitary matrix characterized by three real parameters and one phase. In general, a 3×3 complex matrix has nine real and nine imaginary free parameters, but six constraints from unitarity reduce this to three real parameters and six phases. Five of the six phases are unphysical and can be removed by making phase rotations of the quark fields u^L_i and d^L_i, leaving one irreducible phase. The significance of this remaining phase can be seen by applying the *CP* operator to the term in the

Lagrangian that describes the coupling of the quarks and the charged vector bosons

$$\mathcal{L}_W = g_W \left(\overline{u}_i^L \gamma^\mu W_\mu^+ V_{ij} d_i^L + \overline{d}_i^L \gamma^\mu W_\mu^- V_{ij}^* u_i^L \right), \tag{2.6}$$

where u_L and d_L are the mass eigenstates and g_W is the $SU(2)$ charge. If $V_{ij} = V_{ij}^*$, the operation of CP turns the first term of $\mathcal{L}_W$ into the second term and vice versa, conserving CP symmetry; however, since V contains an irreducible phase $V_{ij} \neq V_{ij}^*$, $\mathcal{L}_W \neq CP\mathcal{L}_W$, and the SM Lagrangian is not invariant under CP.

2.3 The CKM matrix and the unitarity triangle

To be explicit, we write the transformation of the down-type flavor eigenstates (unprimed) into the mass eigenstates (primed) by means of the CKM matrix:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.7}$$

The individual elements of the matrix are determined by experiment. The CKM matrix is commonly written in the Wolfenstein parameterization [25], in which the expansion in powers of the sine of the Cabibbo angle ($\lambda = |V_{us}| \simeq 0.22$) makes the hierarchy of the matrix elements apparent:

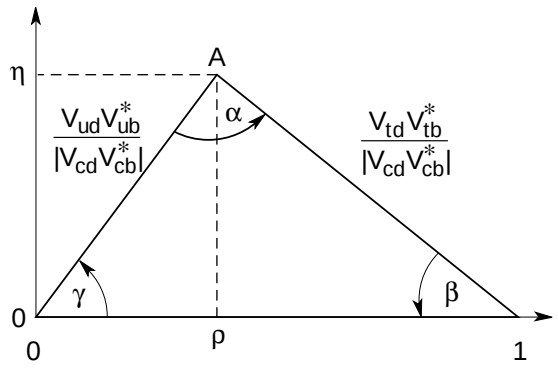
$$V_{CKM} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{2.8}$$

The parameters A, ρ, and η are real and of order unity. Since all parameters are real, the imaginary part of the matrix is contained in the single parameter η, which is attached to the imaginary unit.

The unitarity of the CKM matrix results in nine equations of constraint; six of these equations (for $j = d, s, b$; $k = d, s, b$; and $j \neq k$) are

$$\sum_{i=u,c,t} V_{ij} V_{ik}^* = 0. \tag{2.9}$$

Figure 2.1: The unitarity triangle.



Each of these equations represents a triangle in the complex plane. The most useful of these six triangles, which is relevant for the B system, has sides of roughly equal length and is described as

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \tag{2.10}$$

In Fig. 2.1 we plot this triangle in the complex plane having scaled the sides by $1/|V_{cd}V_{cb}^*|$ such that the bottom of the triangle has unit length. The apex of the triangle can be described by parameters of the Wolfenstein representation of the CKM matrix as $\rho + i\eta$. Since *CP* is violated for $\eta \neq 0$, the area of the triangle corresponds to the magnitude of *CP* violation in the SM. The angles of the triangle are

$$\alpha \equiv \arg\left[-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right],\ \beta \equiv \arg\left[-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right],\ \text{and}\ \gamma \equiv \arg\left[-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right]. \tag{2.11}$$

A primary goal of flavor physics, and the *BABAR* experiment, is to redundantly measure the angles and sides of the unitarity triangle (the elements of the CKM matrix) to determine whether the irreducible CKM phase is the sole cause of *CP* violation in nature.

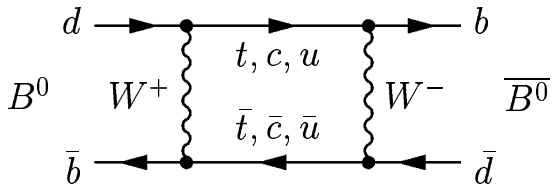


Figure 2.2: Second order weak process (*box* diagram) describing B^0–$\overline{B}^0$ mixing in the Standard Model.

2.4 Neutral B mesons

Neutral B mesons are strongly created in eigenstates of flavor as $|B^0\rangle$ and $|\overline{B}^0\rangle$. These particles oscillate into one another through second order weak processes such as the one in Fig 2.2, implying that the flavor eigenstates are not eigenstates of the Hamiltonian (mass eigenstates). As done by Gell-Mann and Pais in the K system (see Chap. 1), we can write the mass eigenstates as

$$\begin{aligned} |B_L\rangle &= p|B^0\rangle + q|\overline{B}^0\rangle, \\ |B_H\rangle &= p|B^0\rangle - q|\overline{B}^0\rangle, \end{aligned} \tag{2.12}$$

where q and p are complex and normalized such that $|q|^2 + |p|^2 = 1$. The L and H denote the *light* and *heavy* mass states; it will be described later that for neutral B mesons, the mass states have nearly equal lifetimes, but different masses.

We can characterize the mass eigenstates by considering the time evolution of an arbitrary linear combination of flavor eigenstates $a|B^0\rangle + b|\overline{B}^0\rangle$. In the flavor basis, we write the time-dependent Schrodinger equation as

$$i\frac{\partial}{\partial t}\begin{pmatrix} a \\ b \end{pmatrix} = \mathbf{H}\begin{pmatrix} a \\ b \end{pmatrix} = \left(\mathbf{M} - \frac{i}{2}\mathbf{\Gamma}\right)\begin{pmatrix} a \\ b \end{pmatrix}. \tag{2.13}$$

The 2×2 Hermitian matrices that make up the Hamiltonian, $\mathbf{M}$ and $\mathbf{\Gamma}$, are responsible for mixing and decay, respectively. In terms of the masses $m_{H,L}$ and widths $\Gamma_{H,L}$ of the

physical states, the eigenvalues of the Hamiltonian are

$$\begin{aligned} \lambda_H &= m_H - \frac{i}{2}\Gamma_H, \\ \lambda_L &= m_L - \frac{i}{2}\Gamma_L. \end{aligned} \tag{2.14}$$

We can also write the eigenvalues and q/p in terms of the elements of the matrices that compose the Hamiltonian. Because CPT invariance demands that $M_{11} = M_{22}$ and $\Gamma_{11} = \Gamma_{22}$, we can write the eigenvalues of the Hamiltonian in Eq. 2.13 as [26]

$$\lambda_{H,L} = M_{11} - \frac{i}{2}\Gamma_{11} \pm \frac{q}{p}(M_{12} - \frac{i}{2}\Gamma_{12}), \tag{2.15}$$

where

$$\left|\frac{q}{p}\right|^2 = \left|\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}\right|. \tag{2.16}$$

From these relations one can obtain the differences between masses ($\Delta m_B \equiv m_H - m_L$) and widths ($\Delta\Gamma_B \equiv \Gamma_H - \Gamma_L$) of the two physical states

$$\begin{aligned} \Delta m_B &= 2\mathrm{R}e(\frac{q}{p}(M_{12} - \frac{i}{2}\Gamma_{12})), \\ \Delta\Gamma_B &= -4\mathrm{I}m(\frac{q}{p}(M_{12} - \frac{i}{2}\Gamma_{12})), \end{aligned} \tag{2.17}$$

or one can write

$$\begin{aligned} (\Delta m_B)^2 - \frac{1}{4}(\Delta\Gamma_B)^2 &= 4(|M_{12}|^2 - \frac{1}{4}|\Gamma_{12}|^2), \\ \Delta m_B \Delta\Gamma_B &= 4\mathrm{R}e(M_{12}\Gamma_{12}^*). \end{aligned} \tag{2.18}$$

2.4.1 Time evolution of coherent B mesons

At *BABAR* pairs of B^0 and $\overline{B}^0$ mesons are produced coherently in the decay of the $\Upsilon(4S)$. These particles oscillate such that a particle produced as B^0 at time $t = 0$ can be found at time t in the physical state

$$|B^0_{\mathrm{phys}}(t)\rangle \propto e^{-i\lambda_H t}|B_H\rangle + e^{-i\lambda_L t}|B_L\rangle. \tag{2.19}$$

It is most useful to write this relation in terms of flavor states, the masses, and the lifetimes as

$$\begin{aligned} |B^0_{\rm phys}(t)\rangle &= e^{-iMt}e^{-\Gamma t/2}\left[\cos\left(\frac{1}{2}\Delta m_B t\right)|B^0\rangle + i\frac{q}{p}\sin\left(\frac{1}{2}\Delta m_B t\right)|\overline{B}^0\rangle\right], \\ |\overline{B}^0_{\rm phys}(t)\rangle &= e^{-iMt}e^{-\Gamma t/2}\left[i\frac{p}{q}\sin\left(\frac{1}{2}\Delta m_B t\right)|B^0\rangle + \cos\left(\frac{1}{2}\Delta m_B t\right)|\overline{B}^0\rangle\right], \end{aligned} \quad (2.20)$$

where $M = (M_H + M_L)/2$ and $\Gamma = (\Gamma_H + \Gamma_L)/2$. We have made the simplifying assumption that $\Delta\Gamma_B \ll \Delta m_B$. In the kaon system $|\Delta\Gamma_K| \simeq 2|\Delta m_K|$; the difference in the magnitudes of $\Delta\Gamma_K$ and $\Delta\Gamma_B$ arises from the masses of the particles. The relatively large mass of the B allows for large total phase space dominated by flavor-specific decays. The total branching fraction for final states common to B^0 and $\overline{B}^0$, which create the width difference, is of order 10^{-3} in B decays; in K decays the branching fraction for final states common to K^0 and $\overline{K}^0$ dominates.

2.5 Types of *CP* violation in the *B* system

Measurements of interactions of B mesons are sensitive to *CP* violation in three ways:

- *CP* violation in decay, commonly called direct *CP* violation, occurs when the magnitudes of the amplitude for a decay and its *CP* conjugate differ.
- *CP* violation in mixing occurs when the mass eigenstates of a neutral meson system differ from the *CP* eigenstates and is clearly observable in semi-leptonic decays if $|q| \neq |p|$.
- *CP* violation can also occur in the interference between amplitudes for the direct decay of the B meson to a *CP* eigenstate and $B^0\overline{B}^0$ mixing followed by decay to the same final state. This is the primary type of *CP* violation studied at *BABAR*.

2.5.1 *CP* violation in decay

When two (or more) amplitudes contribute to a single decay with different weak *and* strong phases, the magnitudes of the decay amplitude and the *CP* conjugate decay amplitude differ, resulting in different decay rates for the process and its *CP* conjugate. For a decay dominated by two amplitudes, we can write amplitudes for decay to final state f and its *CP* conjugate (neglecting trivial phases) as

$$\begin{aligned} A_f &= A_1 e^{i(\delta_1+\phi_1)} + A_2 e^{i(\delta_2+\phi_2)}, \\ \overline{A}_{\overline{f}} &= A_1 e^{i(\delta_1-\phi_1)} + A_2 e^{i(\delta_2-\phi_2)}. \end{aligned} \tag{2.21}$$

The *weak phases* ϕ_i result from complex terms of the Lagrangian, which come from the irreducible phase of the CKM matrix (Sec. 2.2.3); the weak phases change sign under *CP*. The *strong phases* δ_i come from rescattering in the decay process and do not change sign under *CP*. The difference in decay rates can be written

$$|A_f|^2 - |\overline{A}_{\overline{f}}|^2 \quad \propto \quad A_1 A_2 \sin(\delta_1 - \delta_2)\sin(\phi_1 - \phi_2) \tag{2.22}$$

When $\delta_1 \neq \delta_2$ and $\phi_1 \neq \phi_2$, the difference in decay rates is not zero, and *CP* is violated; i.e., *CP* is violated when the convention-independent quantity $\left|\frac{A_f}{\overline{A}_{\overline{f}}}\right|$ differs from unity. We search for direct *CP* violation in the $B \to \eta K^*$ analysis by measuring the decay rate asymmetry

$$\mathcal{A}_{ch} = \frac{\Gamma^- - \Gamma^+}{\Gamma^- + \Gamma^+}, \tag{2.23}$$

where $\Gamma^\pm \equiv \Gamma(B^\pm \to \eta K^{*\pm})$ is the decay rate (proportional to $|A_f|^2$). In the neutral modes such as $B^0 \to \eta K^{*0}(892)$ where $K^{*0} \to K^\pm\pi^\mp$, the flavor of the B^0 can be determined from the charge of the K, so we define the charge asymmetry in terms of the charge of the K.

2.5.2 *CP* violation in mixing

Since $CP|B^0\rangle = -|\overline{B}^0\rangle$, we can write the *CP* eigenstates of the B meson as

$$\begin{aligned} |B_1\rangle &= \frac{1}{\sqrt{2}}(|B^0\rangle + |\overline{B}^0\rangle), \\ |B_2\rangle &= \frac{1}{\sqrt{2}}(|B^0\rangle - |\overline{B}^0\rangle). \end{aligned} \tag{2.24}$$

Comparing these with the physical eigenstates in Eq. 2.12, we see that the two states are equivalent if the magnitudes of p and q are equal. When $|p/q| \neq 1$, the mass and *CP* eigenstates differ, and *CP* is violated.

This type of *CP* violation is commonly called *indirect CP* violation. Indirect *CP* violation in the kaon system is well known; it was discussed in the introduction that Cronin and Fitch observed indirect *CP* violation in 1964. In the B system, indirect *CP* violation, which is independent of decay channel, is of order 10^{-3} because, as will be discussed later, q/p is very nearly a pure phase.

2.5.3 *CP* violation in interference

The final type of *CP* violation results from interference between the amplitude for simple decay to a final state and the amplitude for $B\overline{B}$ mixing followed by decay to the same final state. Clearly, the effect only occurs for final states that can be reached from both B^0 and $\overline{B}^0$, such as $\eta' K^0$ or $J/\psi K^0$. This mechanism is responsible for *CP* violation in $B^0 \to \eta' K^0$, which is described in this document.

The relevant observable for *CP* violation in interference is similar to that for *CP* violation in decay, but the oscillatory nature of the neutral B mesons requires measurement of the *time-dependent* decay rate asymmetry

$$A_{CP}(t) = \frac{\Gamma(B^0_{\rm phys}(t) \to f_{CP}) - \Gamma(\overline{B}^0_{\rm phys}(t) \to f_{CP})}{\Gamma(B^0_{\rm phys}(t) \to f_{CP}) + \Gamma(\overline{B}^0_{\rm phys}(t) \to f_{CP})}, \tag{2.25}$$

where f_{CP} is the final state discussed above and the state $B^0_{\rm phys}(t)$ ($\overline{B}^0_{\rm phys}(t)$) refers to a B meson that is created in the state $|B^0\rangle$ ($|\overline{B}^0\rangle$) at $t = 0$. Using the expressions for

$B^0_{\rm phys}$ and $\overline{B}^0_{\rm phys}$ from Eq. 2.20, we write the time-dependent decay rates (shortened to f_+ and f_- from $\Gamma(\overline{B}^0_{\rm phys}(t) \to f_{CP})$ and $\Gamma(B^0_{\rm phys}(t) \to f_{CP})$) as

$$f_\pm(t) = \frac{e^{-t/\tau}}{4\tau}\left[1 \pm S_{f_{CP}} \sin(\Delta m_B t) \mp C_{f_{CP}} \cos(\Delta m_B t)\right], \tag{2.26}$$

where τ is the mean B^0 lifetime; the *CP*-violation parameters $S_{f_{CP}}$ and $C_{f_{CP}}$ are

$$S_{f_{CP}} \equiv \frac{2{\rm Im}\lambda_{f_{CP}}}{1+|\lambda_{f_{CP}}|^2}, \qquad C_{f_{CP}} \equiv \frac{1-|\lambda_{f_{CP}}|^2}{1+|\lambda_{f_{CP}}|^2}; \tag{2.27}$$

and $\lambda_{f_{CP}}$ is defined in terms of the *CP* eigenvalue $\eta_{f_{CP}}$ of the final state f_{CP} as

$$\lambda_{f_{CP}} \equiv \frac{q}{p}\frac{\overline{A}_{f_{CP}}}{A_{f_{CP}}} = \eta_{f_{CP}}\frac{q}{p}\frac{\overline{A}_{\overline{f}_{CP}}}{A_{f_{CP}}}. \tag{2.28}$$

As discussed in Sec. 2.5.2, q/p is essentially a pure phase in the B system, so when there is no *CP* violation in decay $|\overline{A}_{\overline{f}_{CP}}/A_{f_{CP}}| = 1$ and $|\lambda_{f_{CP}}| = 1$. Nevertheless, *CP* is still violated if $\lambda_{f_{CP}}$ is complex; i.e., $S_{f_{CP}} \neq 0$.

2.5.3.1 Experimental considerations

In practice, we know neither the flavor at creation nor the flight length of either B meson. In addition, the flavor of the signal B (B_{CP}) at decay into final state f_{CP} is unattainable since f_{CP} is reachable from both B^0 and $\overline{B}^0$. However, we are able to determine (tag) the flavor of the other B meson ($B_{\rm tag}$), as described in Sec. 4.6.3, and we are able to measure the positions of the decay vertices of both B_{CP} and $B_{\rm tag}$, as described in Sec. 4.6.1. From these decay vertex positions, we compute the difference between the proper times of the decays as described in Sec. 4.6.2. Since B_{CP} and $B_{\rm tag}$ are produced in an entangled state in the decay of the $\Upsilon(4S)$, knowledge of the flavor of $B_{\rm tag}$ at decay, along with the mixing frequency and the time difference, allows us to determine the flavor of B_{CP} at decay. We can recast the time-dependent decay rate (Eq. 2.26) in terms of the difference between proper decay times of B_{CP} and $B_{\rm tag}$, $\Delta t \equiv t_{CP} - t_{\rm tag}$, and the flavor of $B_{\rm tag}$ at decay as

$$f_\pm(\Delta t) = \frac{e^{-|\Delta t|/\tau}}{4\tau}\left[1 \pm S_{f_{CP}} \sin(\Delta m_B \Delta t) \mp C_{f_{CP}} \cos(\Delta m_B \Delta t)\right], \tag{2.29}$$

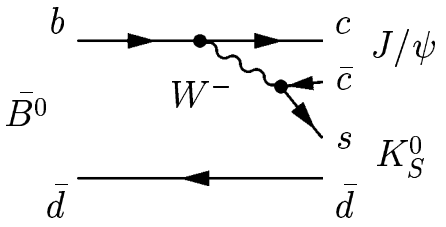


Figure 2.3: Tree diagram for the decay $B^0 \to J/\psi K_S^0$.

where the upper (lower) sign denotes a decay accompanied by a B^0 ($\overline{B}^0$) tag.

2.6 Extracting sin 2β

The importance of *CP* violation in interference becomes apparent when one relates the phase $\lambda_{f_{CP}}$ to the elements of the CKM matrix. The factor of q/p is independent of decay channel. In Sec. 2.4.1, we argued that $\Delta\Gamma_B \ll \Delta m_B$, which along with Eqs. 2.17 and 2.18, implies that $\Gamma_{12} \ll M_{12}$, We can approximate that, to one part in one hundred, q/p is the phase

$$\frac{q}{p} = -\frac{M_{12}^*}{|M_{12}|}. \tag{2.30}$$

With the goal of writing q/p in terms of CKM matrix elements, we examine the vertices in the box diagram (Fig. 2.2), which drives M_{12}. We find $M_{12} \propto (V_{tb}V_{td}^*)^2$. All up-type quarks contribute to the loop in the mixing diagram, but each individual contribution is proportional to the square of the mass of the quark, so the top quark dominates. We can write

$$\frac{q}{p} = \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*}. \tag{2.31}$$

2.7 Tree dominated decays

The decay $B^0 \to J/\psi K_S^0$, which is *CP* odd, occurs almost exclusively through the internal tree diagram in Fig. 2.3 for which $A_{\psi K_S^0} \propto V_{cb}V_{cs}^*$. We also include a diagram for K – $\overline{K}$ mixing for which $(q/p)_K = (V_{cs}V_{cd}^*)/(V_{cs}^*V_{cd})$ [27]. Putting all this together,

we see

$$\begin{aligned}\lambda_{\psi K^0_S} &= \eta_{\psi K^0_S}\left(\frac{q}{p}\right)_B\left(\frac{\overline{A}_{\psi K^0_S}}{A_{\psi K^0_S}}\right)\left(\frac{q}{p}\right)_K, \\ &= -\left(\frac{V^*_{tb}V_{td}}{V_{tb}V^*_{td}}\right)\left(\frac{V_{cb}V^*_{cs}}{V^*_{cb}V_{cs}}\right)\left(\frac{V_{cs}V^*_{cd}}{V^*_{cs}V_{cd}}\right), \\ &= -\left(\frac{V_{td}V^*_{tb}}{V^*_{cb}V_{cd}}\right)\left(\frac{V_{cb}V^*_{cd}}{V^*_{td}V_{tb}}\right). \end{aligned} \tag{2.32}$$

Equation 2.11 implies that $-(V^*_{cb}V_{cd})/(V_{td}V^*_{tb}) = Ae^{i\beta}$ where A is real and β is an angle of the unitarity triangle. So $\lambda_{\psi K^0_S} = -e^{-i2\beta}$, and we can finally write

$$\mathrm{Im}(\lambda_{\psi K^0_S}) = \sin 2\beta. \tag{2.33}$$

Since the decay to $J/\psi K^0_S$ is dominated by a single diagram, we expect that CP violation in decay is negligible, $|\lambda_{\psi K^0_S}| = 1$, implying (from Eq. 2.27)

$$C_{\psi K^0_S} = 0, \qquad S_{\psi K^0_S} = \sin 2\beta. \tag{2.34}$$

In fact, the relationships described above hold for all $b \to c\bar{c}s$ tree-dominated (Fig. 2.3) decays, such as $B^0 \to J/\psi K^{*0}$ and $B^0 \to \eta_c K^0$. The current world averages from all $b \to c\bar{c}s$ decays are $C_{c\bar{c}s} = 0.005 \pm 0.019$ and $S_{c\bar{c}s} = 0.671 \pm 0.024$. These averages come from results from *BABAR* [28] and Belle [29].

2.8 Penguin dominated decays

In this thesis we describe studies of the decay $B^0 \to \eta' K^0$, which is also sensitive to $\sin 2\beta$, but with more theoretical uncertainty because the decay proceeds through both tree and penguin amplitudes (shown in Fig. 2.4). We can write the amplitude for this decay in terms of CKM matrix elements and general contributions from the tree diagram (T) and the penguin diagram with a $q = u, c, t$ quark in the loop (P_q) as

$$A_{\eta' K^0} = P_t V_{tb}V^*_{ts} + P_c V_{cb}V^*_{cs} + (P_u + T)V_{ub}V^*_{us}. \tag{2.35}$$

It is useful to use unitarity to remove the $V_{tb}V^*_{ts}$ factors such that

$$A_{\eta' K^0} = (P_c - P_t)V_{cb}V^*_{cs} + (P_u - P_t + T)V_{ub}V^*_{us}. \tag{2.36}$$

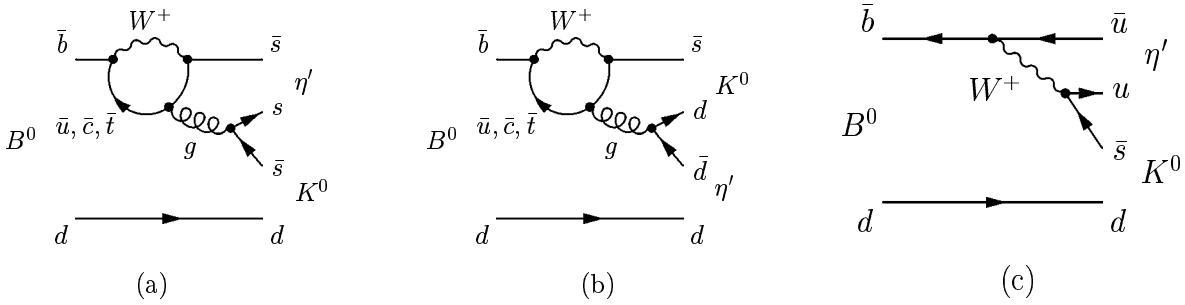


Figure 2.4: Feynman diagrams for the decay $B^0 \to \eta' K^0$(a,b) gluonic penguin and (c) color-suppressed tree.

The first term dominates the decay amplitude since it is of order λ^2 and the second term is of order λ^4, where λ is the sine of the Cabibbo angle from the Wolfenstein parameterization of the CKM matrix. Since the decay $B^0 \to \eta' K^0$ is dominated by a single weak phase, $|\lambda_{\eta' K^0}|$ is approximately unity. The dominant amplitude $V_{cb}V_{cs}^*$ carries the same phase as that in $J/\psi K_S^0$, so the exercise of relating $\lambda_{\eta' K^0}$ to an angle of the CKM matrix proceeds just like the argument for $J/\psi K_S^0$ in Sec. 2.7.

One difference is that here we are discussing the general decay $B^0 \to \eta' K^0$, which includes both $\eta' K_S^0$ and $\eta' K_L^0$ final states. Because we need to correctly treat the differing *CP* eigenvalues ($\eta_{\eta' K_S^0} = -1$, $\eta_{\eta' K_L^0} = +1$) it is most convenient to discuss the product $-\eta_{f_{CP}} S_{f_{CP}}$. So, we conclude that $\mathrm{Im}(\lambda_{\eta' K_S^0}) \simeq \sin 2\beta$ and $\mathrm{Im}(\lambda_{\eta' K_L^0}) \simeq -\sin 2\beta$, and we expect

$$\begin{aligned} -\eta_{\eta' K_S^0} S_{\eta' K_S^0} &\simeq \sin 2\beta, \\ -\eta_{\eta' K_L^0} S_{\eta' K_L^0} &\simeq \sin 2\beta, \\ C_{\eta' K_S^0} &\simeq 0, \\ C_{\eta' K_L^0} &\simeq 0. \end{aligned} \tag{2.37}$$

In the literature, $-\eta S$ measured in a $b \to q\bar{q}s$ penguin decay is typically called an effective $\sin 2\beta$ such that $-\eta_{\eta' K_S^0} S_{\eta' K_S^0} = \sin 2\beta_{\mathrm{eff}}$.

2.9 New physics and ΔS

The penguin dominated $B^0 \to \eta' K^0$ decay is interesting because of its sensitivity to physics beyond the Standard Model. If heavy non-SM particles are present in the loop, the mixing phase can differ from $\sin 2\beta$ [31]. In addition, the CKM-suppressed loop amplitudes and the color-suppressed tree diagram introduce additional weak phases whose contributions may not be negligible [32, 33, 34, 35]. As a consequence, the measured $S_{\eta' K^0}$ may differ from $\sin 2\beta$ even within the SM. This deviation $\Delta S_f \equiv S_{\eta' K^0} - \sin 2\beta$ is estimated in several theoretical approaches. Theorists typically predict ranges of expected ΔS which we list for four theoretical frameworks:

- QCD factorization (QCDF): $(0.00, 0.03)$ [32, 36],
- QCDF with modeled rescattering: $(-0.03, 0.01)$ [37],
- Soft collinear effective theory: $(-0.03, 0.00)$ [38],
- SU(3) symmetry: $(-0.05, 0.09)$ [33, 35, 39].

The larger apparent uncertainties in the estimates from flavor SU(3) result from uncertainties on the input branching fractions (from two-body decays involving η, η', and π^0). All theoretical sectors seem to favor a small value for ΔS, so a large difference between $S_{\eta' K^0}$ and $S_{\psi K^0_S}$ would indicate new physics.

The effects of new physics would come primarily from non-SM particles entering the loop (in the decay amplitude) with a phase different from the CKM phase. New particles in the mixing amplitude would affect $S_{\eta' K^0}$ and $S_{\psi K^0_S}$ in the same way. The non-SM particles are typically taken to be supersymmetric particles including squarks entering with u, c, t and gauginos entering in place of the W.

2.10 The hierarchy of $B \to \eta^{(\prime)} K^{(*)}$ branching fractions

The charged and neutral $B \to \eta' K$ decays first garnered interest when the CLEO collaboration observed their branching fractions (BF) to be much larger than expected [40]. Currently, the $B \to \eta' K$ BF is measured to be ~ 65 times larger than that of $B \to \eta K$, while the $B \to \eta' K^*$ BF is ~ 4 times *smaller* than that of $B \to \eta K^*$. The hierarchy of $B \to \eta^{(\prime)} K^{(*)}$ branching fractions has received considerable attention from theorists. The current world averages for BFs for $\eta' K^0$ and $\eta' K^+$ are already relatively precise at $(64.9 \pm 3.1) \times 10^{-6}$ and $(70.3 \pm 2.5) \times 10^{-6}$, respectively [55]. We contribute to the understanding of the BF hierarchy by improving the BF measurements for $B^0 \to \eta K^{*0}(892)$ and $B^+ \to \eta K^{*+}(892)$, and by searching for decays to $\eta(K\pi)_J$ where the spin $J = 0, 2$.

For decays of interest in this paper, there have been recent calculations from QCD factorization [30, 32, 41] and flavor SU(3) symmetry [42]. As discussed above, charmless B decays to final states with strangeness are expected to be dominated by $b \to s$ penguin amplitudes. The branching fraction for the decay $B \to \eta K^*(892)$ is expected to be larger than most similar decays, such as $B \to \pi^0 K^*$, though not as large as $B \to \eta' K$, due to constructive interference between two penguin amplitudes [43]. While the decay $B \to \eta K^*(892)$ has been seen previously [44, 45], there have been no searches for states with an η meson accompanied by $K^*(1430)$ mesons, and no theoretical predictions exist for these decays. However, it has been noted that the observed hierarchy of branching fractions depends on the spin of the strange meson [32, 43], so the measurement for decays including spin 1 and 2 may help understand the pattern.

In principle, we could measure the *CP* parameter S in any of these $B^0 \to \eta^{(\prime)} K^{(*)0}$ modes. (For $B^0 \to \eta^{(\prime)} K^{*0}$ all K^{*0} final states except $K^0_S \pi^0$ are self-tagging and therefore not suitable for a measurement of mixing-induced *CP* violation.) However, the branching fraction hierarchy, which makes the time-dependent measurement in $B^0 \to \eta' K^0$

so fruitful, makes the same measurements in $B^0 \to \eta K_S^0$ and $B^0 \to \eta' K^{*0}$ unfeasible because of small (or zero) event yields. The measurement in the decay $B^0 \to \eta K^{*0}$ is nearly possible, but the branching fraction is a factor of three smaller than that of $B^0 \to \eta' K^0$ and the efficiency is low for reconstructing the final state $\eta K_S^0 \pi^0$.

2.11 Previous results

In this thesis, we present updates of the measurements of CP-violation parameters S and C in $\eta' K^0$ and branching fractions and charge asymmetries in $B^0 \to \eta K^{*0}(892)$ and $B^+ \to \eta K^{*+}(892)$. There are no previous branching fraction or charge asymmetry results for $B \to \eta K_0^*(S\text{-wave})$ or $B \to \eta K_2^*(1430)$. In Table 2.1, we summarize these previous results [45, 46, 47, 48].

Table 2.1: Summary of previous results for branching fractions (in units of 10^{-6}), charge asymmetries, $-\eta_f S_f$, and C_f.

	BABAR [45][46]	Belle [47][48]
$\mathcal{B}(\times 10^{-6})$		
$\eta K^{*+}(892)$	$25.6 \pm 4.0 \pm 2.4$	$22.8 \pm 3.7 \pm 2.2$
$\eta K^{*0}(892)$	$18.6 \pm 2.3 \pm 1.2$	$19.8 \pm 2.1 \pm 1.4$
$\eta K_0^{*+}(S\text{-wave})$	–	–
$\eta K_0^{*0}(S\text{-wave})$	–	–
$\eta K_2^{*+}(1430)$	–	–
$\eta K_2^{*0}(1430)$	–	–
$\mathcal{A}_{ch}$		
$\eta K^{*+}(892)$	$0.13 \pm 0.14 \pm 0.02$	$-0.09 \pm 0.16 \pm 0.01$
$\eta K^{*0}(892)$	$0.02 \pm 0.11 \pm 0.02$	$-0.04 \pm 0.11 \pm 0.01$
$\eta K_0^{*+}(S\text{-wave})$	–	–
$\eta K_0^{*0}(S\text{-wave})$	–	–
$\eta K_2^{*+}(1430)$	–	–
$\eta K_2^{*0}(1430)$	–	–
$S_{\eta' K^0}$	$0.58 \pm 0.10 \pm 0.03$	$0.64 \pm 0.10 \pm 0.04$
$C_{\eta' K^0}$	$-0.16 \pm 0.07 \pm 0.03$	$-0.01 \pm 0.07 \pm 0.05$
$-\eta_{\eta' K_S^0} S_{\eta' K_S^0}$	0.62 ± 0.11	0.64 ± 0.11
$C_{\eta' K_S^0}$	-0.18 ± 0.07	0.03 ± 0.07
$-\eta_{\eta' K_L^0} S_{\eta' K_L^0}$	0.32 ± 0.28	0.46 ± 0.24
$C_{\eta' K_L^0}$	-0.16 ± 0.07	-0.09 ± 0.16

Chapter 3

The Experiment

3.1 Overview

The *BABAR* experiment operates at the Positron-Electron Project II (PEP-II) asymmetric-energy e^+e^- storage ring [49] of the Stanford Linear Accelerator Center, renamed SLAC National Accelerator Laboratory in October 2008. The primary goal of the experiment is to measure time-dependent *CP*-violation asymmetries in the decays of B mesons to *CP* eigenstates, but the detector is also well suited for measuring the branching fractions of rare B decays.

To maximize the production of B mesons, the PEP-II collider operates at a center-of-mass (CM) energy of 10.58 GeV which corresponds to the mass of the $\Upsilon(4S)$ resonance. The $\Upsilon(4S)$ decays almost exclusively to coherent pairs of B mesons. Approximately 10% of the data are recorded 40 MeV below the $\Upsilon(4S)$ mass for study of $e^+e^- \to q\overline{q}$ continuum events ($q = u, d, s, c$), which are the primary background for these analyses.

To measure the time-dependent *CP* asymmetries, we fully reconstruct one of the B mesons (B_{CP}) in a *CP* eigenstate, determine the flavor of the other B meson ($B_{\rm tag}$) through partial reconstruction, and measure the difference between the proper times of the decays of B_{CP} and $B_{\rm tag}$, which we obtain from the measured distance between their decay vertices. To ensure that the separation of decay vertices is greater than the detector resolution, the energies of the e^+e^- beams are asymmetric, with a 9 GeV

electron beam and a 3.1 GeV positron beam, so that the CM frame is boosted with respect to the laboratory frame. The detector is offset 0.37 m from the interaction point (IP) to maximize acceptance in the boosted CM frame.

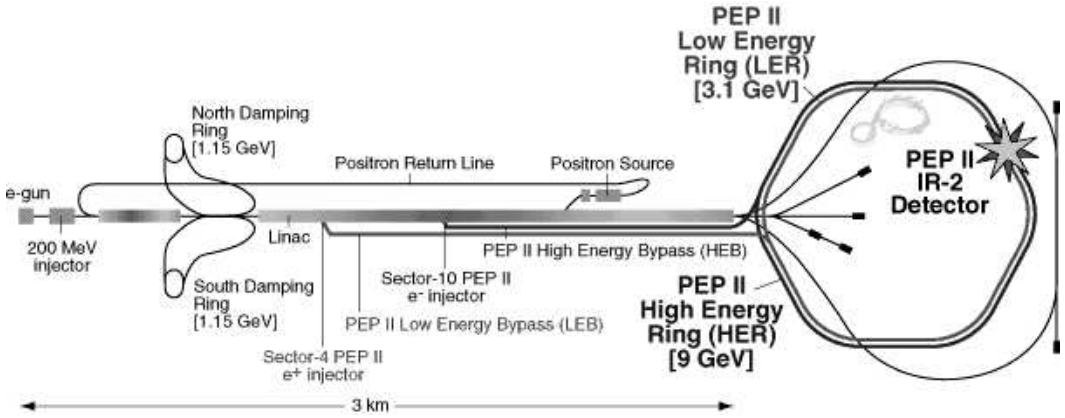


Figure 3.1: Schematic of the linear accelerator and PEP-II.

3.2 The linear accelerator and PEP-II

The linear accelerator (linac) accelerates electrons and positrons to collision energies with radio frequency (RF) electromagnetic waves. Particles from the linac are injected into the PEP-II storage rings where they circulate before colliding at the second interaction region (IR-2) and the *BABAR* detector. The linac and PEP-II are shown schematically in Fig. 3.1.

Electrons are produced from a filament by thermal emission and fed into the linac with a static electric field. The electrons are accelerated to energies of ~ 1 GeV before being redirected to the north damping ring. In the damping ring, the emittance of the beam is reduced via synchrotron radiation and subsequent longitudinal acceleration. (The *emittance* is the spread of the beam in space and momentum.) The damped beam returns to the linac where it is accelerated to collision energy (9.0 GeV).

For the production of positrons, half of the electrons are accelerated almost the

full length of the linac into a tungsten target at energies of 30 GeV. Positrons from the resulting electromagnetic shower are focused and sent via a return line to the source end of the linac for damping in the south ring and acceleration to collision energies (3.1 GeV).

Electrons and positrons are injected into the PEP-II storage rings at collision energies. As they circulate the separate electron and positron rings, they are focused by magnets and accelerated by RF to compensate for energy loss due to synchrotron radiation. As they enter the interaction region the beams are focused and steered into head-on collision; as particles that did not interact depart the interaction region they are steered back into their respective rings.

The PEP-II storage rings each have a circumference of $\sim$ 2200 meters, and the frequency of the accelerating RF is 476 MHz. To achieve high luminosity, approximately every other accelerating phase of the RF is used for a total of $\sim$ 1700 bunches of particles. The bunch spacing is $\sim$ 1.2 m or $\sim$ 4 ns. PEP-II was designed for instantaneous luminosity of $3 \times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, but achieved over four times that goal through collider upgrades.

During typical operation, PEP-II daily delivered several hundred pb^{-1} of integrated luminosity of which more than 95% was recorded by *BABAR*. A total of 426 fb^{-1} of data were recorded at the $\Upsilon(4S)$ energy from October 1999 to October 2007. Since the $B\overline{B}$ production cross-section is about 1.1 nb, the recorded data correspond to 467×10^6 $B\overline{B}$ pairs. For accounting purposes, the eight year running period is divided into six *runs* of $8 - 16$ months. Each run period consists of $20 - 135$ fb^{-1} of recorded data. As described later, the ηK^* analysis uses runs 1–5, and the $\eta' K^0$ analysis uses runs 1–6.

3.3 The *BABAR* detector

The full reconstruction of *CP* eigenstates, with branching fractions as low as 10^{-6}, and the determination of the flavor of $B_{\rm tag}$ require the detection of multiple charged and neutral particles over a wide kinematic range. These considerations indicate the need for a versatile detector and require that the *BABAR* detector have

- acceptance down to small polar angles relative to the boost direction,
- high reconstruction efficiency for charged particles (tracks) with momenta perpendicular to the beam axis (p_T) as low as 60 MeV and for photons with energy as low as 20 MeV,
- resolution of the vertex of fully-reconstructed B decays of 80 μm along the z-axis and 100 μm in the transverse plane,
- resolution of track momentum perpendicular to the beam axis (p_T) of $\sim 0.5\%$,
- excellent energy and angular resolution for photons,
- and efficient and accurate identification of charged particles over a wide momentum range.

The *BABAR* detector, shown in Figs. 3.2 and 3.3, consists of layered subsystems that provide complementary measurements of particle properties and trajectories. Track momenta are measured with a silicon vertex tracker (SVT) of five double-sided layers and a forty-layer drift chamber (DCH) coaxial with a 1.5 T superconducting solenoidal magnet. Neutral cluster (photon) positions and energies are measured with an electromagnetic calorimeter (EMC), which also provides partial K_L^0 reconstruction. Charged hadrons are identified with a detector of internally reflected Cherenkov light (DIRC) and measurements of specific energy loss ($\mathrm{d}E/\mathrm{d}x$) in the DCH and SVT. Finally, the instrumented flux return of the magnet allows discrimination of muons from pions and

additional detection of K_L^0 mesons. A more complete description of the *BABAR* detector can be found elsewhere [50].

3.3.1 *BABAR* coordinate system

The *BABAR* coordinate system is right-handed. The positive z-axis is parallel to the magnetic field of the solenoidal magnet and in the direction of electron beam; the positive y-axis is up; and the positive x-axis is directed horizontally away from the center of the storage ring. The azimuthal angle ϕ is defined as zero on $+x$-axis and increases toward the $+y$-axis. The polar angle θ is zero on the $+z$-axis and π on the $-z$-axis.

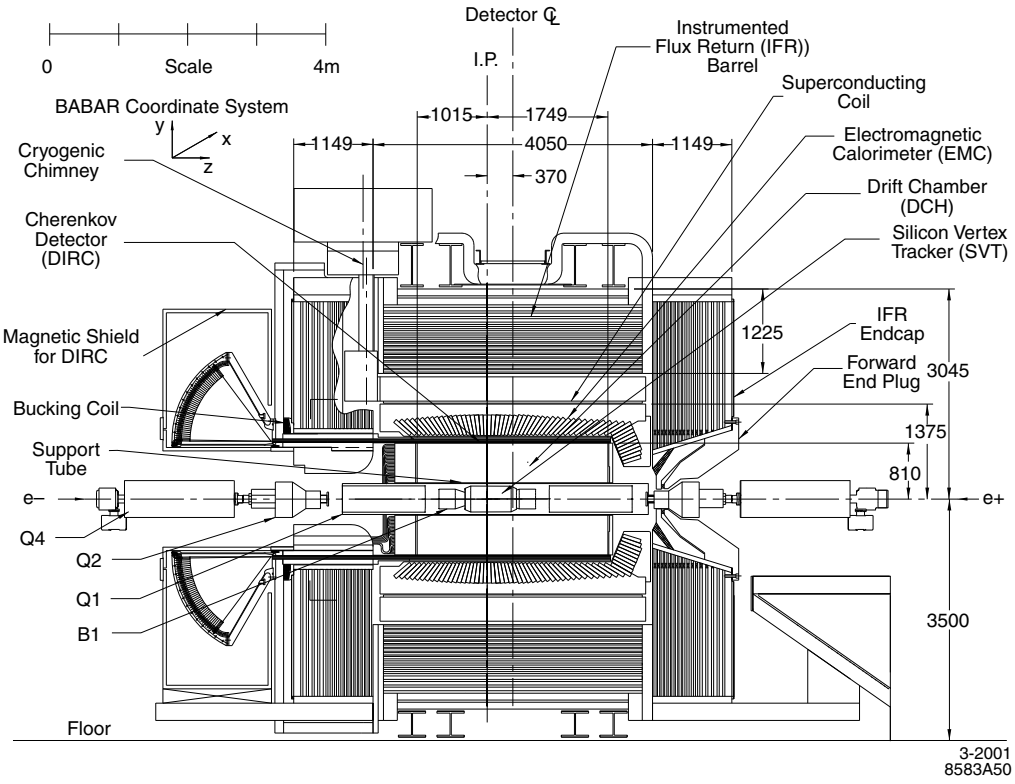


Figure 3.2: Longitudinal cross-section of the *BABAR* detector.

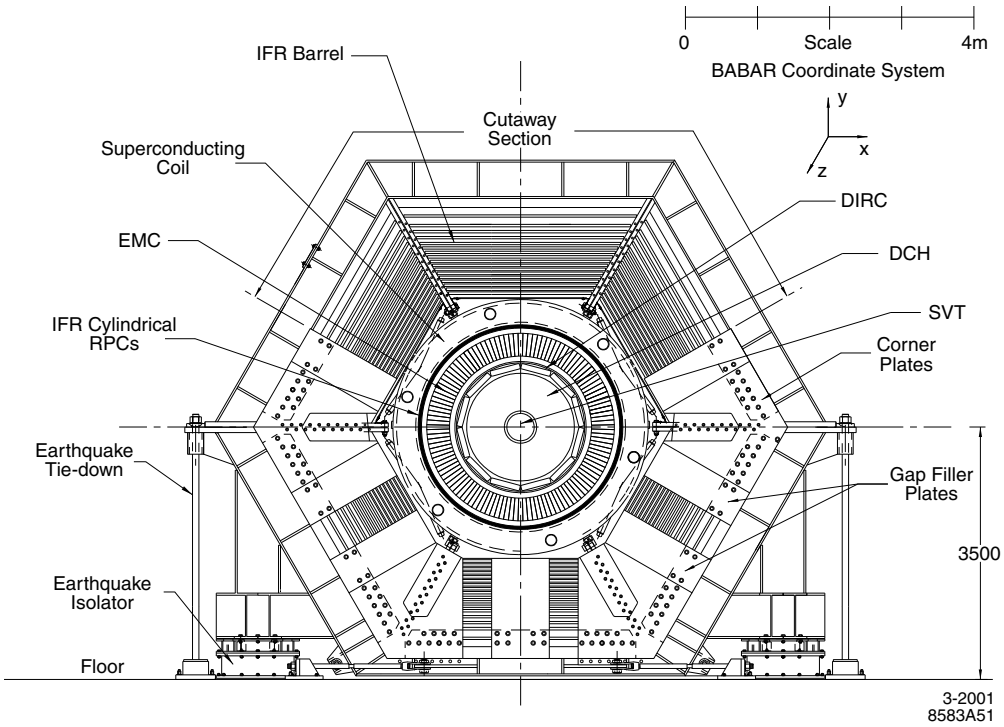


Figure 3.3: Transverse cross-section of the *BABAR* detector.

3.3.2 Silicon vertex tracker

The SVT is designed for high precision measurement of coordinates, both azimuthal and longitudinal, on the trajectories of charged particles close to the e^+e^- interaction region. The SVT comprises five layers of double-sided silicon micro-strip detectors. Layers 1–3 are 3–6 cm from the beampipe with each layer forming a hexagonal prism coaxial with the beampipe; layers 4 and 5 are 12–15 cm from the beampipe with each layer forming a 16- and 18-sided polygonal prism with tapered ends. Figs. 3.4 and 3.5 show transverse and longitudinal views of the SVT.

The five layer design was chosen so that the SVT could provide standalone track reconstruction; three layers are necessary to determine the circular projection of a helix onto the transverse plane; a fourth is used for corroboration; and the fifth layer reduces the impact of hit-inefficiencies on tracking efficiency. Layers 1–3 are located as close to

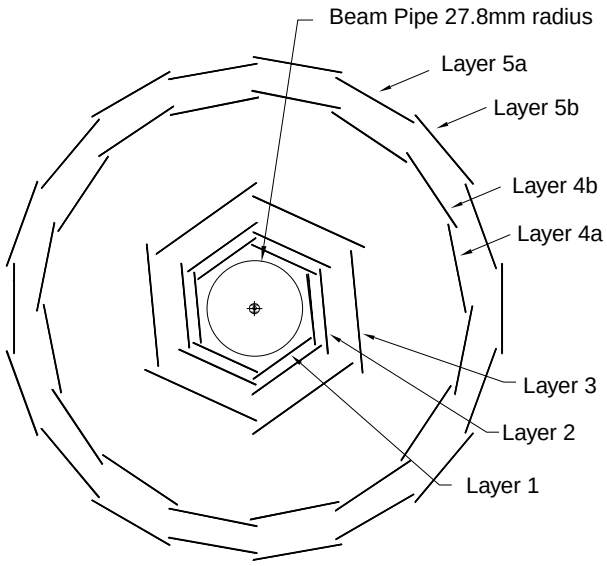


Figure 3.4: Transverse cross-section of the silicon vertex tracker.

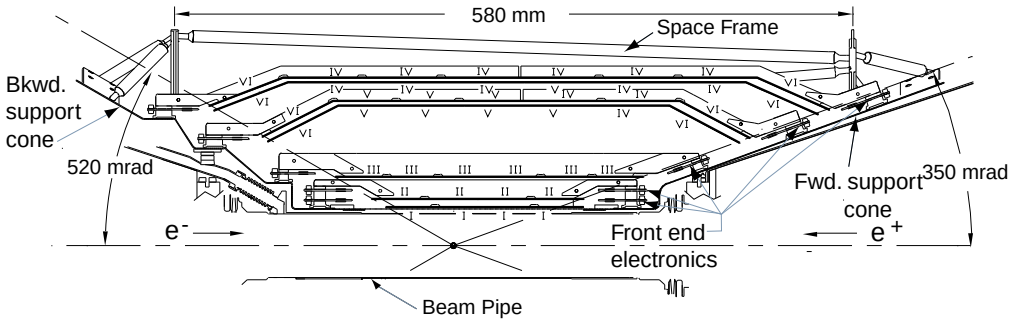


Figure 3.5: Longitudinal cross-section of the silicon vertex tracker.

beampipe as possible for best determination of the trajectories of the decay products. Layers 4 and 5 are located farther from the beampipe to provide position and angle information necessary to link SVT tracks with DCH tracks. Layers 4 and 5 are tapered to provide greater crossing angle for forward and backward tracks.

The single hit precision of the SVT is $10 - 15\,\mu$m in all coordinates in the first three layers and $40\,\mu$m in the outer two layers. The solid-angle coverage of the SVT is 90%. Though not as precise as the DCH dE/dx measurement ($\sim 7\%$ resolution) because there are only five SVT layers, we measure dE/dx in the SVT with 18% resolution.

3.3.3 Drift chamber

The drift chamber (DCH) is the experiment's primary tracker and the experiment's sole tracker for long-lived particles that decay outside the SVT, such as K^0_S and Λ. The DCH (Fig. 3.6) is a cylinder 276.4 cm long with inner radius of 23.6 cm and outer radius of 80.9 cm. To a good approximation, the DCH is coaxial with the ~ 1.5 T quasi-solenoidal magnetic field. Trajectories of charged particles in this field are nearly helical, and their momentum transverse to the z-axis (p_T) is simply related to their radius of curvature. For *BABAR* it is a reasonable approximation to take $p_T = 4.5 \times \rho$ where p_T is measured in MeV and ρ is the radius of the circular x-y projection of a particle's helical trajectory measured in cm [51]. Thus particles with $p_T < 110\,$MeV do not enter the DCH, and particles with $p_T > 365\,$MeV cross all 40 layers of the DCH. The gas filling the DCH is an 80:20 mixture of helium:isobutane. Helium was chosen over argon to minimize multiple scattering, which limits the precision of measured track parameters.

The DCH comprises 7104 hexagonal cells that lie in forty layers grouped in ten superlayers. To measure the z-position of tracks, six *stereo* superlayers are arranged with a slight angular offset with respect to the z-axis ($\pm 45 - \pm 75$ mrad); the remaining four *axial* superlayers are parallel to the z-axis. The axial (A), positive stereo (U), and

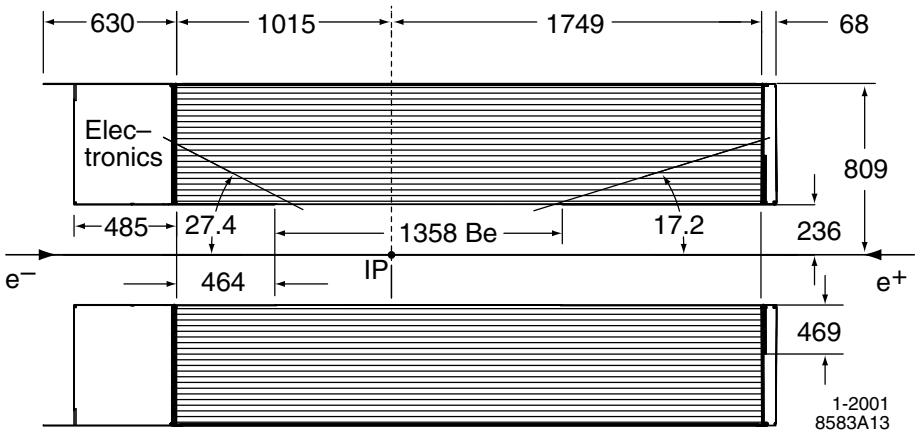


Figure 3.6: Longitudinal cross-section of drift chamber with dimensions (mm). The interaction point (IP) is 370 mm from the center of the chamber.

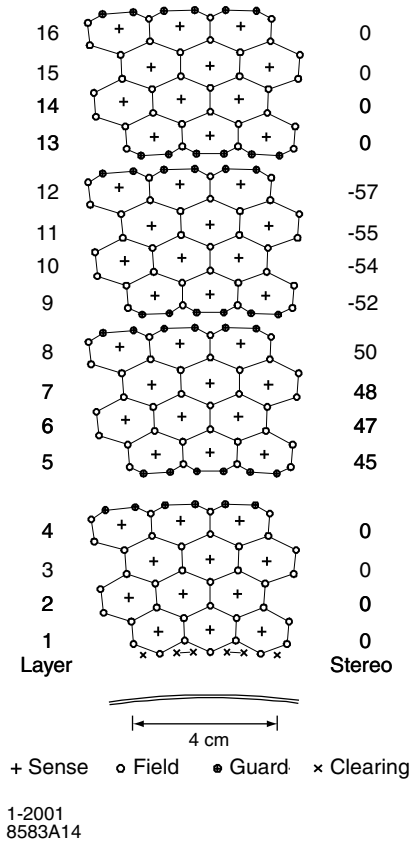


Figure 3.7: Arrangement of drift cells for the four innermost superlayers of the drift chamber. The numbers on the right side give the stereo angles (mrad) of sense wires.

negative stereo (V) superlayers are arranged AUVAUVAUVA. The arrangement of cells in the four innermost superlayers is shown in Fig. 3.7.

Each cell consists of a sense wire surrounded by six field wires, as shown in Fig. 3.8. The sense wire is held at 1930V, and the field wires are at ground. Each cell is ~ 1.2 cm (radial) by ~ 1.9 cm (azimuthal), such that the maximum drift distance is 0.9 cm. The drift velocity of the gas is $\sim 20\,\mu$m/ns which yields a drift time of ~ 500 ns. The time-to-distance calibration yields position resolution of $\sim 100\,\mu$m over most of the drift cell, but $\sim 200\,\mu$m near the edges.

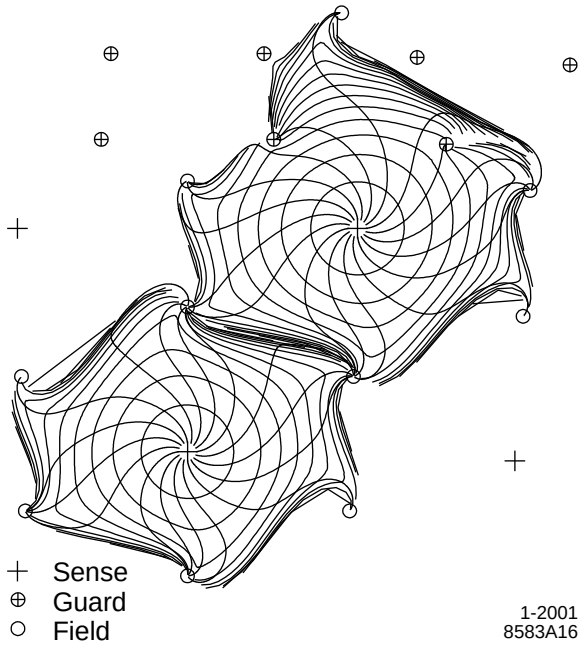


Figure 3.8: Contours of equal drift times in cells of the third and fourth layers of an axial superlayer.

3.3.3.1 *dE/dx*

In addition to coordinates along each track, the DCH provides particle identification (PID). By measuring the total charge deposited in each drift cell, we determine the specific energy loss (dE/dx) of charged particles. The measurements of dE/dx and

particle momentum can be used to determine the type of the particle through comparisons with Bethe-Bloch predictions, see Fig. 3.13. For particles with $p_T < 700\,\mathrm{MeV}$ the DCH is the only source of PID. At these momenta, the DCH yields dE/dx resolution (7.5%) sufficient for good $K - \pi$ separation.

3.3.4 Track reconstruction

We reconstruct tracks from hits in the SVT and DCH using multiple redundant track-finding algorithms and then refit these tracks using Kalman filter fits [52] which account for variations in the magnetic field, scattering, and energy loss [53]. First, segment-based track finders reconstruct DCH tracks of typical momenta. Then specialized track finders reconstruct tracks that originate in the DCH and *loopers*, tracks that do not have sufficient p_T to leave the DCH. After adding additional DCH hits to these tracks, where possible, the tracks are refit with a Kalman filter fit. Next, SVT hits are added to the existing DCH tracks, and the resulting tracks are again refit with a Kalman filter fit. The remaining SVT hits are used to construct SVT-only tracks. Finally, the algorithm attempts to merge these SVT-only tracks with DCH-only tracks allowing for scattering between the SVT and DCH.

These reconstructed tracks are characterized by five variables:

- $\mathbf{d_0}$ the signed closest distance, in the x-y projection, from the z-axis to the trajectory at the point of closest approach of the trajectory to the origin (POCA);
- $\mathbf{z_0}$ the z position of the trajectory at the POCA in the x-y plane;
- ϕ_0 the azimuthal angle corresponding to the track direction in the x-y plane at the POCA;
- ω the signed geometrical curvature;
- $\tan\lambda$ the tangent of the dip angle, $\tan\lambda = \cot\theta$ (polar angle).

The SVT dominates the precision on $\mathbf{z_0}$, $\mathbf{d_0}$, and ϕ_0 for all tracks. Because the precision of z-coordinate measurements is an order of magnitude better for the SVT than for the DCH, the SVT also dominates the determination of the dip angle which is essentially dz/dr. At high momentum, determination of the dip angle is most important for extrapolating trajectories to outer detectors such as the DIRC. The DCH dominates the determination of curvature of the track (p_T), and at low momentum the curvature is most important for the extrapolation to outer detectors.

We determine the resolution on track parameters by comparing the independent reconstructions of the upper and lower halves of cosmic rays ($p_T \simeq 3\,\mathrm{GeV}$) that travel near the IP. We find

- $\sigma_{\mathbf{d_0}} = 23\,\mu\mathrm{m}$,
- $\sigma_{\mathbf{z_0}} = 29\,\mu\mathrm{m}$,
- $\sigma_{\phi_0} = 0.43$ mrad,
- $\sigma_{\tan\lambda} = 0.53 \times 10^{-3}$.

The resolution on the transverse momentum is $\sigma_{p_T}/p_T = (0.5 + 0.1 \cdot p_T)\%$ with p_T in GeV. The first term represents the multiple scattering limit, and the second term arises because the precision of the measurement of the radius of curvature is constant, while the curvature decreases as $1/p_T$.

3.3.5 Detector of internally reflected Cherenkov radiation

An important method for determining the flavor of the tagged B meson uses the charge of a kaon resulting from the cascade decay $b \to c \to s$. Pions and kaons are the only charged daughters of exclusive decays reconstructed in analyses described in this document. For these reasons, it is very important to be able to distinguish between pions and kaons. Measurements of dE/dx in the DCH are responsible for $K-\pi$ separation for

particles with momenta less than 700 MeV; the DIRC provides particle identification (PID) for particles with momenta greater than 700 MeV. To reduce the necessary size and cost of the electromagnetic calorimeter and to reduce the interaction of particles in the DIRC volume, the DIRC was designed to be radially very thin.

The DIRC (Fig. 3.9) consists of 144 bars (17.25 mm thick) of synthetic silica (index of refraction $n = 1.473$) arranged in a 12-sided polygonal prism coaxial with the beampipe. Charged particles with velocity β emit Cherenkov light in a cone of angle θ in a medium with index of refraction n, such that $\cos\theta = 1/n\beta$. The angle of the Cherenkov cone produced by particles travelling through the silica is preserved by internal reflection as the light travels down the silica bars to the rear of the detector where lies an array of photon detectors in the stand off box (SOB). Forward going light is reflected toward the photon detectors in the rear of the detector by mirrors at the front of each silica bar.

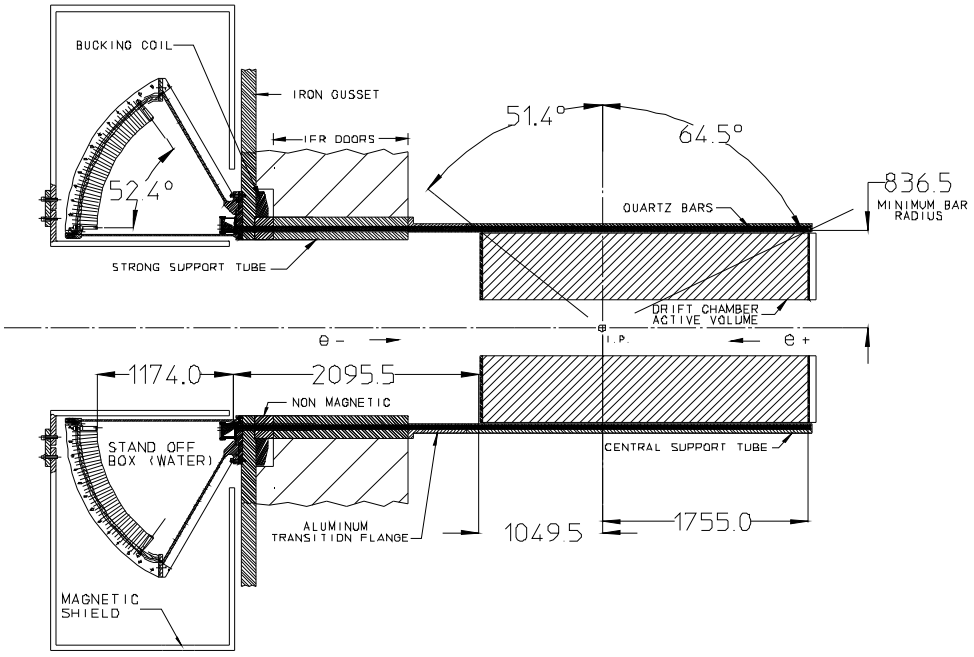


Figure 3.9: Longitudinal cross-section of the DIRC with important dimensions (mm).

Within the SOB (Fig. 3.10) the diameter of the Cherenkov ring increases until it

is detected by an array of photomultiplier tubes (PMTs). Since the indices of refraction of water and silica are similar, the SOB is filled with water to reduce dispersion and signal loss due to total internal reflection at the water/silica interface. At the end of each bar is glued a wedge of silica so that the lower image of the Cherenkov ring is reflected onto the upper image. Finally the photons are detected by an array of closely packed PMTs.

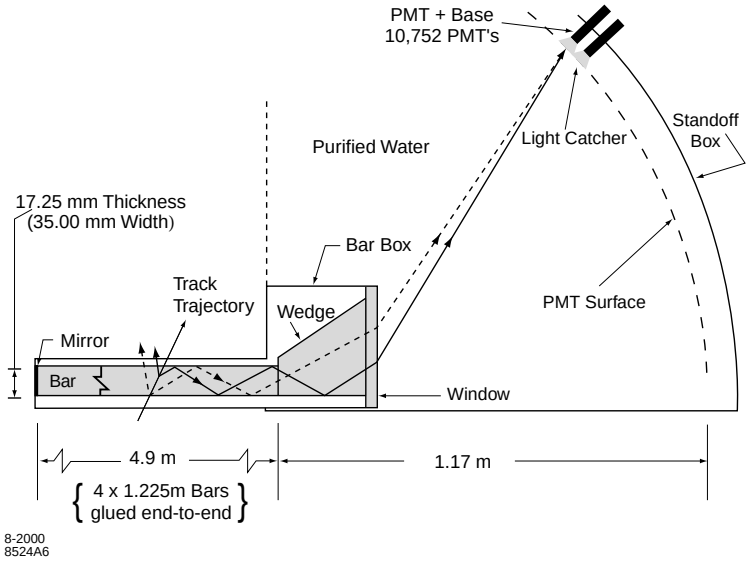


Figure 3.10: Longitudinal cross-section of the top half of the DIRC silica radiator bar and imaging region.

Using position and arrival time information of the Cherenkov light, along with track position, angle, and momentum information from the SVT and DCH, we are able to determine the mass of the particle, and hence type of the particle, that produced each track. With 2.5 mrad resolution on the measured Cherenkov angle, the DIRC provides 4.2σ separation between kaons and pions at 3 GeV. The kaon selection efficiency is about 96% with a pion misidentification rate of about 2%.

3.3.6 Electromagnetic calorimeter

The EMC detects electromagnetic showers of energy 20 MeV to 9 GeV with excellent energy and angular resolution allowing the complete reconstruction of neutral pions and η mesons, via detection of their daughter photons, and the partial reconstruction of K_L^0's (angular information only). The EMC also allows the identification of electrons which is critical for tagging the flavor of neutral B mesons via semi-leptonic decays.

As shown in Fig 3.11, the EMC consists of 56 rings of scintillating crystals arranged in a cylindrical barrel and forward endcap, providing 90% solid angle coverage in the CM frame. The crystals are made of thallium-doped cesium iodide, chosen for a short radiation length (1.85 cm) allowing for shower containment at *BABAR* energies with a compact design. The scintillation light is detected at the rear face of each crystal with two silicon PIN diodes chosen for their good performance in a magnetic field.

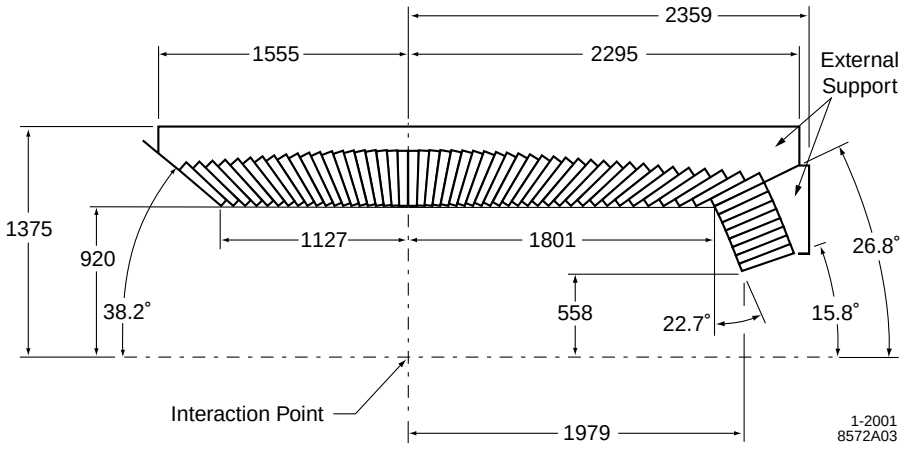


Figure 3.11: Longitudinal cross-section of the top half of the EMC showing the arrangement of the 56 crystal rings with important dimensions (mm).

The EMC is calibrated at low energy (6 MeV) with a radioactive source and at high energy (4-9 GeV) with tracks from e^+e^- scattering (Bhabha scattering), for which the dependence of energy on polar angle is known. The energy resolution as a function

of energy (in GeV) is determined to be

$$\frac{\sigma_E}{E} = \left(\frac{2.3}{E^{(1/4)}} \oplus 1.9 \right) \%,$$

where the sum is in quadrature. The angular resolution is determined to be

$$\sigma_\theta = \sigma_\phi = \left(\frac{3.9}{E^{(1/2)}} \right) \text{mrad}$$

from studies of π^0 and η decays to two photons of approximately equal energies.

The EMC dominates electron identification, which primarily relies on the ratio of shower energy and track momentum (E/p) and the shape of the shower. We also require that dE/dx in the DCH and the Cherenkov angle in the DIRC are consistent with the electron hypothesis.

3.3.7 Instrumented flux return

Surrounding the EMC is a large iron yoke that serves as a flux return for the solenoidal magnet. The flux return is instrumented to provide identification of muons and neutral hadrons, such as K_L^0's. The IFR (Fig. 3.12) consists of a hexagonal barrel section and two end doors. Each section is composed of 19 layers of resistive plate chamber (RPC) detectors or limited streamer tubes (LST) sandwiched between layers of steel. Initially RPC's were used for the entire IFR, but because the RPC performance degraded, the barrel sector RPC's were replaced with LST's in the summers of 2004 and 2006. The efficiency of muon identification of the IFR is at the design goal 90% since the LST upgrade, which is important for B-meson flavor tagging via semi-leptonic decays.

3.3.7.1 K_L^0 Reconstruction : EMC and IFR

The IFR also allows reconstruction of the direction of the K_L^0 momentum; 40% of K_L^0's passing the selection in the $B^0 \to \eta' K_L^0$ analysis are detected with the IFR.

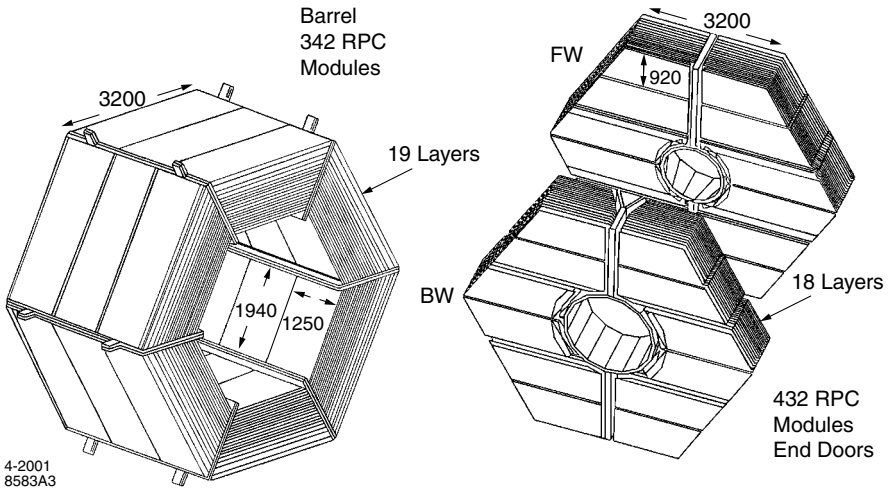


Figure 3.12: Barrel sectors and forward (FW) and backward (BW) end doors of the IFR with important dimensions (mm).

Neutral hadrons are identified as IFR clusters not associated with a charged track. The angular resolution of K_L^0's reconstructed in the IFR is about 60 mrad, the resolution for those detected with the EMC is about 30 mrad. The K_L^0 detection efficiency, for the EMC and IFR combined, is roughly linear with momentum, increasing from 20% at 1 GeV to 40% at 4 GeV.

3.3.8 Particle identification

Though discussed in previous sections, here we provide a qualitative synthesis of the methods of particle identification: we use measured dE/dx from the DCH, Cherenkov angle from the DIRC, E/p from the EMC, and hits in the IFR to distinguish between charged electrons, muons, kaons, pions, and protons. Measurements from these subsystems are used to compute likelihoods for the particle species; requirements are made on ratios of these likelihoods to reject particles of the wrong species.

Essentially, identification of charged hadrons with $p_T < 700\,\mathrm{MeV}$ relies solely on dE/dx; identification of charged hadrons with $p_T > 700\,\mathrm{MeV}$ relies on dE/dx and

information from the DIRC. Plots of dE/dx and Cherenkov angle as functions of track momentum are shown in Fig. 3.13. Electrons are identified as charged tracks with associated EMC clusters with $0.5 < E/p < 1.5$. Muons are detected as hits in the IFR that are associated with tracks in the DCH.

In the $\eta' K^0$ analysis, we make the weak requirement that pions from the η' do not pass the tight selection for kaons, electrons, or protons; the pion fake rate is less than 5% at the relevant momenta. In the ηK^* analysis, we require that pions from the η and K^* are not consistent with kaons, electrons, or protons. We also require that kaons pass the tight kaon selection. The efficiency of the kaon selector is greater than 80% at all momenta, and the pion fake rate is less than 5%.

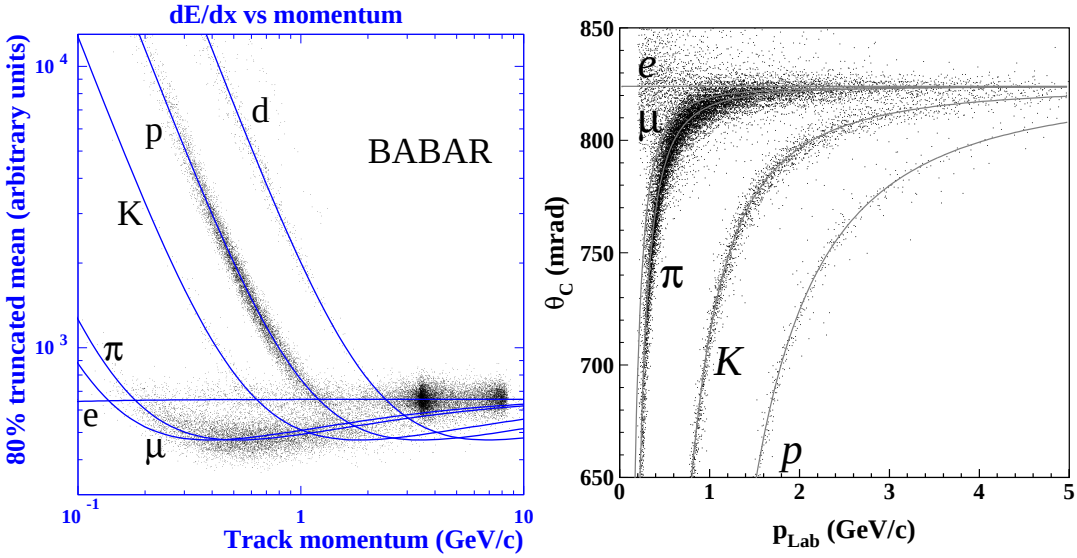


Figure 3.13: Measured dE/dx (in DCH) as function of track momentum with curves showing Bethe-Bloch predictions for several particle species (*left*). The measured Cherenkov angle (in DIRC) as function of track momentum with curves showing predictions for several particle species (*right*).

3.3.9 Trigger and data acquisition

The trigger system is designed to select events of interest with high, stable, and well-understood efficiency (at least 99% for $B\overline{B}$ processes) while keeping the total event

rate manageable for data storage and processing, about 120 Hz. The beam-induced background rates are about 20 kHz for at least one DCH track with $p_T > 120$ MeV or at least one EMC cluster with $E > 100$ MeV. The trigger is a hierarchy of the L1 hardware trigger, which rejects beam-induced backgrounds and has an output rate of less than 2 kHz, followed by the L3 software trigger which selects physics events of interest with an output rate of 120 Hz. Production cross-sections, rates, and L1 trigger rates for important physics processes are shown in Table 3.1.

Front-end electronics (FEE) located on the detector perform low-level signal processing, buffer data, and transfer data to the data acquisition (DAQ) system when triggered. Data from the DCH, EMC, and IFR, such as track p_T, EMC energy deposition, or hits in IFR layers, are used to form L1 *primitives* which are sent to the Global Trigger (GLT). The GLT processes trigger primitives deciding whether to keep an event based on configurable logic and prescaling rates. The maximum response latency determined by the FEE data buffer is 12.8 μs.

Table 3.1: Production cross-sections, rates, and trigger rates at e^+e^- center-of-mass energy of 10.58 GeV. The e^+e^- cross-section is the effective cross-section for events with either the e^+ or e^- interacting with the electromagnetic calorimeter.

Process ($e^+e^- \to$)	Cross-section (nb)	Production rate (Hz)	L1 rate (Hz)
$b\bar{b}$	1.1	3.2	3.2
other $q\bar{q}$	3.4	10.2	10.1
e^+e^-	~ 53	159	156
$\mu^+\mu^-$	1.2	3.5	3.1
$\tau^+\tau^-$	0.9	2.8	2.4

The L3 trigger performs initial event reconstruction, classification, selection, and monitoring. The L3 trigger passes almost all hadronic and leptonic physics events, which make up $\sim 15\%$ of the total output; the rest of the output are prescaled events,

such as Bhabha events used for calibration and luminosity measurements and unfiltered backgrounds (L1 Pass-Throughs) for efficiency calculations.

Events that pass the L3 trigger are sent to the logging manager which writes the data to an *extended tagged container* (XTC) file. Each XTC file contains data from about one hour of running and is called a *run*. The XTC files are sent to online prompt reconstruction (OPR) computing farms. Information on detector conditions is stored in the conditions database for later use during OPR. After the two stages of OPR, prompt calibration (PC) and event reconstruction (ER), reconstructed events are written to *BABAR*'s custom Root-based event store (`kanga`).

Chapter 4

Common Analysis Technique

4.1 Overview

We measure several physical observables – branching fractions, time-integrated decay rate asymmetries, and time-dependent decay rate asymmetries – via decays of B mesons to final states ηK^* and $\eta' K^0$ [54]. These decays are rare processes with branching fractions of $\sim 20 \times 10^{-6}$ and $(64.9 \pm 3.1) \times 10^{-6}$ [55], respectively. (We give the world average branching fraction from the Heavy Flavor Averaging Group for $B^0 \to \eta' K^0$ since we do not report this measurement herein.) We simultaneously isolate the signal events from the immense background and measure the quantities of interest using a maximum likelihood (ML) fit.

We begin with primitive information from detector electronics from which we deduce the energy, momentum, and charge of detected particles. From these basic physical quantities, we reconstruct the quantities related to the B mesons and their decays, such as the decay vertices and related uncertainties, the angular distributions of the decay products, the B flavor ($B^0/\overline{B}^0$), the B charge, and the energy and momenta of the B and other intermediate resonances. We reject the dominant $e^+e^- \to q\bar{q}$ continuum background events, while retaining signal events with high efficiency, by making loose requirements on several of these derived quantities. We use three to six of these derived quantities as input to an ML fit for each analysis.

In general *BABAR* analyses are performed blind; i.e., selection requirements and

analysis methods are determined only with the use of simulated data, data that is outside the signal region (sideband data), and data control samples. However, since these analyses are updates of previous analyses performed on data from runs 1–4 for ηK^* and runs 1–5 for $\eta' K$, we use these datasets, when necessary, for verifying the analysis method.

4.2 Data and Monte Carlo samples

These analyses are based on data collected with the *BABAR* detector at the PEP-II asymmetric-energy e^+e^- collider of the Stanford Linear Accelerator Center. The $\eta' K$ analysis, performed in 2008, is based on the full *BABAR* dataset collected from 1999–2008. The integrated luminosity of the on-resonance data is 426 fb^{-1} corresponding to 467 ± 5 million $B\overline{B}$ pairs. The ηK^* analysis, performed in 2006, is based on data collected from 1999-2006. The integrated luminosity of the on-resonance sample is 312.6 fb^{-1} corresponding to 344 ± 4 million $B\overline{B}$ pairs.

In addition to the data, we use GEANT4 [56] Monte Carlo (MC) simulations to understand the signal decays and rare decays that are backgrounds for these analyses. The MC fully simulates beam conditions, machine backgrounds, the physics of the full interaction ($e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\overline{B} \rightarrow$ final state), interactions with material in the detector, and the detector response. We apply to the MC the same reconstruction and selection that is applied to the data.

4.3 Candidate reconstruction

Associated with each event, as it is recorded in the *BABAR* event store, are lists of reconstructed charged and neutral particles. Loose requirements are placed on members of each list. We use charged particles from three lists: `ChargedTracks`, `GoodTracksVeryLoose`, and `GoodTracksLoose`. `ChargedTracks` are all reconstructed charged particles. Members of `GoodTracksVeryLoose` are tracks from `ChargedTracks`

with a maximum momentum of 10 GeV, a distance of closest approach (DOCA) to the beam in the $x - y$ plane no larger than 1.5 cm, and DOCA in z less than 10 cm. Members of `GoodTracksLoose` are tracks from `GoodTracksVeryLoose` with a minimum transverse momentum of 100 MeV and information from at least 12 layers of the drift chamber. We require photons from an η' to have minimum energy of 100 MeV and all other photons to have minimum energy of 50 MeV.

We reconstruct B candidates by making all possible combinations of neutral particles and charged tracks from these lists with a bottom-up approach. For instance, in the reconstruction of $\eta'_{\rho\gamma}K^0_S$, we build the *decay tree* by combining charged tracks to form ρ^0 and K^0_S meson candidates. An η' candidate, built from photon and ρ^0 candidates, is combined with the K^0_S to form a B candidate. We fit this decay tree, with constraints related to the hypothesis that it is a signal decay, to obtain the four-momenta and decay vertices of the B and all intermediate resonances.

4.3.1 Fitting the decay vertex

We use two different vertex fitting algorithms `Cascade/Geokin` and `TreeFit`. The `Cascade/Geokin` fitter is a generalized least squares, *leaf-by-leaf* fitter in which each stage of the decay tree is fit independently from the bottom-up. A disadvantage of this method is that constraints applied at the top of the tree do not contribute to lower leaf fits. Ideally, one would fit for all vertices simultaneously, but the least squares fit to the entire decay tree involves the inversion of a matrix with dimension of several tens for each candidate, making it computationally unfeasible. The `TreeFit` algorithm [57] solves this problem by using a Kalman filter rather the least squares fit to simultaneously extract the parameters of the decay tree and their uncertainties and correlations. We use the `Cascade/Geokin` method in the ηK^* analysis and the `TreeFit` method in the $\eta' K^0$ analysis. In principle, precise determination of the vertex is more important for the $\eta' K^0$ analysis in which we rely on vertex measurements to determine the difference

in B^0 and $\overline{B}^0$ decay times, which is needed for the time-dependent decay rate. In practice, the difference between `Cascade/Geokin` and `TreeFit` is only substantial in analyses such as $B^0 \to \pi^0 K^0_S$ for which there are no charged tracks present at the B decay vertex.

Regardless of fit algorithm, one can apply various constraints during the vertex fitting. Conservation of four momentum and the geometric constraint, the requirement that a particle production vertex coincides with its parent's decay vertex, are automatically applied. With a mass constraint, one can require that an intermediate resonance have a particular mass; this reduces the free parameters of the four-momentum from four to three. The `Bflight` requirement is a Gaussian constraint that the production vertex of a particle coincide with the e^+e^- interaction point accounting for the uncertainty on this point (10 μm in y, 200 μm in x, and 8 mm in z) and the fact that the B travels before decaying. This constraint is used in the $\eta K^*(892)$ analysis with $\eta \to \gamma\gamma$ and $K^{*+} \to K^0_S \pi^+$ since there is only one charged track present at the B decay vertex.

4.3.2 Kinematic quantities for the B meson

A B meson candidate is characterized kinematically by the invariant energy difference, defined as

$$\Delta E = (2q_0 q_B - s)/(2\sqrt{s}) \tag{4.1}$$

and the beam-energy-substituted mass

$$m_{\rm ES} = \sqrt{(\frac{1}{2}s + \mathbf{p}_0 \cdot \mathbf{p}_B)^2/E_0^2 - p_B^2}; \tag{4.2}$$

where $q_0 = (E_0, \mathbf{p}_0)$ is the four momentum of the beams; $q_B = (E_B, \mathbf{p}_B)$ is the four momentum of the reconstructed signal B; and $s \equiv q_0^2$.

One can also write ΔE in the e^+e^- center-of-mass (CM) frame, denoted with a *, as

$$\Delta E = E^*_B - \frac{1}{2}E^*_0, \tag{4.3}$$

so it is clear that ΔE is the difference between the energy of the signal B and one half of the total energy in the CM. Evaluated in the CM frame, $m_{\rm ES}$ can be seen as the invariant mass of the signal B evaluated with the constraint that $\Delta E = 0$,

$$m_{\rm ES} \;=\; \sqrt{\frac{s}{4} - p_B^{*2}}. \tag{4.4}$$

For correctly reconstructed B mesons, the ΔE distribution peaks at zero and the $m_{\rm ES}$ distribution peaks at the B mass, 5.28 GeV.

4.3.3 Event shape quantities

As mentioned earlier, over 99% of background events are $e^+e^- \rightarrow q\overline{q}$ ($q = u, d, s, c$) continuum events in which no B meson is produced. The topology of these events is very different from that of the signal. We characterize the topology with five event shape quantities all computed in the CM frame. These event shape variables are built from magnitudes and directions of particle momenta, directions of particles with respect to the beam axis, and directions of the thrust axes in the event. The thrust axis of a group of particles is defined as the direction which maximizes the longitudinal momentum of that group.

The five useful variables are

- the thrust angle $\theta_{\rm T}$, which is defined as the angle between the thrust axis of the B candidate daughters and the thrust axis of the other particles in the event,
- the angle with respect to the beam axis of the B momentum θ_B,
- the angle with respect to the beam axis of the B thrust axis θ_{TB},
- the zeroth and second momentum-weighted angular moments L_0 and L_2, defined as

$$L_i \;=\; \sum_j p_j \times \left|\cos\theta_j\right|^i, \tag{4.5}$$

where θ_j is the angle with respect to the B thrust axis of daughter particle j; p_j is its momentum; and the sum excludes the daughters of the B candidate.

The most powerful variable for separating signal and continuum is $|\cos\theta_{\rm T}|$, shown in Fig. 4.1. For signal events, pairs of B mesons are produced nearly at rest and decay isotropically in the CM frame. While the signal B is fully reconstructed, we rarely detect all the daughter particles of the other B. This missing momentum removes all correlation between the directions of the thrust axes of the B mesons because they at rest in the CM, and so the signal $\cos\theta_{\rm T}$ distribution is flat. For continuum events, the daughters of the energetic primary quarks travel in a back-to-back jet-like topology, and the $|\cos\theta_{\rm T}|$ distribution peaks sharply at $+1$. We reduce continuum background by requiring that $|\cos\theta_{\rm T}| < 0.9$ for all modes except those with K^0_L in the final state for which we require that $|\cos\theta_{\rm T}| < 0.8$.

We combine the other four shape variables, $|\cos\theta_{\rm B}|$, $|\cos\theta_{\rm TB}|$, L_0, and L_2, in a linear Fisher discriminant $\mathcal{F}$ which is used as input to the ML fit. The Fisher discriminant is a linear combination of these four variables whose coefficients are chosen to maximize the separation (difference of means divided by quadrature sum of errors) between the signal and continuum background distribution of $\mathcal{F}$. We determine the coefficients from studies of signal MC and off-peak data. We have studied the optimization of $\mathcal{F}$ for a variety of signal modes, and find that a single set of coefficients is nearly optimal for all. We apply a shift to $\mathcal{F}$ based on tagging category information (Sec. 4.6.3) to remove the first order correlation between the mean of the Fisher distribution and the tagging category for $q\overline{q}$ continuum events.

4.4 Backgrounds

The dominant $q\bar{q}$ background events make up $98-100\%$ of the total background, depending on sub-decay channel. The distributions for $m_{\rm ES}$, ΔE, and $\mathcal{F}$ are distinct

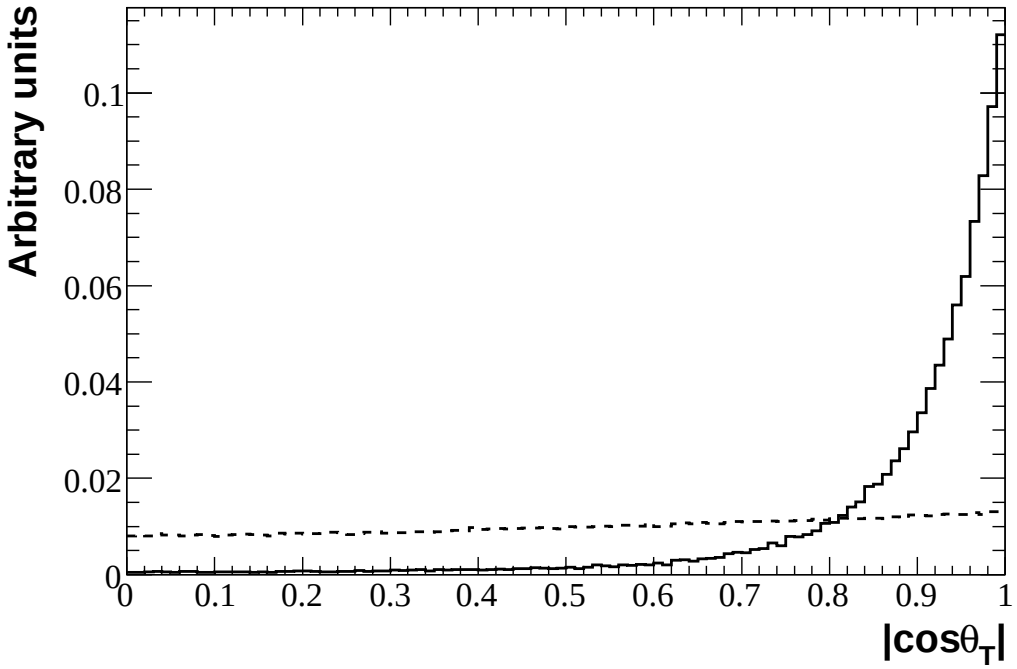


Figure 4.1: Distributions of the angle between thrust axes of the signal B and the rest of the event ($|\cos\theta_{\rm T}|$) for signal (dashed) and $q\bar{q}$ background (solid). The signal distribution is from MC, and the $q\bar{q}$ background distribution is from on-resonance data in sidebands chosen to avoid potential signal events.

for $q\bar{q}$ and signal events, so the ML fit is effective at separating these components. Background from B decays, though only $\sim 2\%$ of the background, can be problematic in some decay channels because their distributions tend to be more signal-like than continuum-like. Without proper treatment in the ML fit, these events would tend to be classified as signal. We identify which sub-decay channels require a careful treatment of the $B\overline{B}$ background with MC studies and design the fit accordingly. Studies related to the $B\overline{B}$ backgrounds are described in detail in Sec. 5.7 for $\eta' K^0$ and Sec. 6.10 for ηK^*. Backgrounds from QED processes such as e^+e^- scattering, heavy lepton pair production, and two photon processes are characterized by two or fewer tracks. We determine with studies of MC that our requirement that an event contain at least four tracks makes this potential background negligible.

4.5 Maximum likelihood fit

We extract parameters of interest from the data with an extended unbinned multivariate maximum likelihood fit. We implement the fit using a RooFit-based [58] software package called `RooRarFit` [59], which ultimately employs MINUIT [60] to maximize the likelihood.

The cuts on the quantities used as input to the ML fit are loose to allow for high efficiency and to provide sufficient sidebands to accurately characterize the background. In Appendix A, we describe studies of the relationship between cut efficiency and expected measurement precision in time-dependent measurements of *CP* parameters S and C in $B^0 \to \eta' K^0_L$. Qualitatively stated, we find that, in the presence of large backgrounds, measurement precision is improved by relying on the ML fit rather than tight cuts for background rejection.

4.5.1 Maximum likelihood method

Consider a set of N measurements of quantities $\mathbf{x} = \{x^1, x^2, ..., x^m\}$ that are randomly distributed according to the *probability density function* (PDF) $\mathcal{P}(\mathbf{x}, \boldsymbol{\alpha})$ where $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, ..., \alpha_n\}$ is a set of parameters. The probability of a single measurement x_i^m falling in the infinitesimal interval $[x_i^m, x_i^m + dx_i^m]$ is $\mathcal{P}(\mathbf{x}_i, \boldsymbol{\alpha}) dx_i^m$. We define the *likelihood* $\mathcal{L}$ for this ensemble of N measurements as

$$\mathcal{L} \equiv \prod_i^N \mathcal{P}(\mathbf{x}_i, \boldsymbol{\alpha}). \tag{4.6}$$

It can be shown that the maximum of $\mathcal{L}$ over parameter space ($\boldsymbol{\alpha}$) gives an unbiased estimate of $\boldsymbol{\alpha}$. To interpret the likelihood in a Bayesian sense, we state that, given the data $\{\mathbf{x}_i\}$, $\mathcal{L}$ is a function that gives the relative probability for different values of $\boldsymbol{\alpha}$. In practice, the parameters $\boldsymbol{\alpha}$ include the number of signal events in the sample and the *CP* parameters S and C; the quantities $\mathbf{x}$ include $m_{\rm ES}$, ΔE, $\mathcal{F}$, and Δt.

One can fix any of the parameters in $\boldsymbol{\alpha}$ if its value is known and estimate the values

of k unknown parameters in $\boldsymbol{\alpha}$ by finding the maximum of $\mathcal{L}$ with the simultaneous solutions of k equations:

$$\frac{\partial \mathcal{L}}{\partial \alpha_j} = 0, \quad j = 1, 2, ..., k. \tag{4.7}$$

For a large number of measurements N, the likelihood as a function of any parameter α_i is Gaussian in the vicinity of the maximum likelihood $\mathcal{L}_{\max} = \mathcal{L}(\hat{\alpha}_i)$,

$$\mathcal{L}(\alpha_i) = \mathcal{L}_{\max} \exp\left[-\frac{(\alpha_i - \hat{\alpha}_i)^2}{2\sigma^2}\right], \tag{4.8}$$

and the statistical uncertainty σ on α_i is

$$\frac{1}{\sigma^2} = \frac{\partial^2 \ln \mathcal{L}}{\partial \alpha_i^2}. \tag{4.9}$$

The important relationships are clear when one considers the function

$$-2 \ln \frac{\mathcal{L}(\alpha_i)}{\mathcal{L}_{\max}} = \frac{(\alpha_i - \hat{\alpha}_i)^2}{\sigma^2}, \tag{4.10}$$

which is a parabola that equals 0 at its minimum $\hat{\alpha}_i$, 1 at $\hat{\alpha}_i \pm \sigma$, 4 at $\hat{\alpha}_i \pm 2\sigma$, etc. In practice the likelihood is not perfectly Gaussian, but we generalize and take the statistical uncertainty σ to be the value such that $-2 \ln \frac{\mathcal{L}(\hat{\alpha}_i + \sigma)}{\mathcal{L}_{\max}} = 1$. We can obtain asymmetric errors in the case that $\mathcal{L}(\hat{\alpha}_i + \sigma) \neq \mathcal{L}(\hat{\alpha}_i - \sigma)$.

We compute the significance of a fit result relative to some other hypothesis, e.g. the zero signal hypothesis, with the likelihood ratio test. We take the significance as the square root of the difference between the value of $\chi^2 \equiv -2 \ln \mathcal{L}$ for zero signal and the value at its minimum. To include systematic uncertainties in the significance, we convolve χ^2 as a function of the parameter of interest with a Gaussian whose width is equal to the systematic uncertainty before finding the difference between the minimum χ^2 and the value for the hypothesis to be tested.

4.5.1.1 Extended maximum likelihood method

If the number of measurements N is also taken to be a random variable, we can construct the *extended* likelihood by multiplying $\mathcal{L}$ by the Poisson probability of making

N measurements when expecting ν measurements:

$$\mathcal{L} = \frac{e^{-\nu}\nu^N}{N!}\prod_i^N \mathcal{P}(\mathbf{x}_i, \boldsymbol{\alpha}). \quad (4.11)$$

Now ν is the overall normalization and a free parameter of the fit.

4.5.1.2 Composite PDFs

We construct the PDF $\mathcal{P}$ as a composite of several functions with each component of the composite corresponding to a component of the data, such as signal or background. In this way, the fraction of each component multiplied by the total normalization ν corresponds to the event *yield* of that component of the data. For instance, in a fit with a signal and background component, we write the extended likelihood as

$$\mathcal{L} = \frac{e^{-\nu}}{N!}\prod_i^N \nu f_{\rm sig}\mathcal{P}_{\rm sig}(\mathbf{x}_i, \boldsymbol{\alpha}) + \nu(1-f_{\rm sig})\mathcal{P}_{\rm bkg}(\mathbf{x}_i, \boldsymbol{\alpha}), \quad (4.12)$$

where $\nu_{\rm sig} \equiv \nu f_{\rm sig}$ and $\nu_{\rm bkg} \equiv \nu(1-f_{\rm sig})$ are the estimators of the number of signal and background events in the data, the parameters of interest in a branching fraction measurement. Notice that the extra degree of freedom (ν) in the extended ML fit reduces the correlation between $\nu_{\rm sig}$ and $\nu_{\rm bkg}$. One can simplify this by writing

$$\mathcal{L} = \frac{e^{-\nu}}{N!}\prod_i^N \sum_j^M \nu_j \mathcal{P}_j(\mathbf{x}_i, \boldsymbol{\alpha}), \quad (4.13)$$

where $j = (\text{sig}, \text{bkg}, ...)$ is one of M components of the data. (We also use components for backgrounds from $B\overline{B}$ events when necessary.)

4.5.1.3 Minimizing $-\ln\mathcal{L}$

In practice, $\mathcal{L}$ is a very small number ($\mathcal{O}$ 10^{-10000}) and computationally difficult to work with. Since the natural logarithm is a monotone operator, it is equivalent and convenient to minimize $-\ln\mathcal{L}$, which is positive and of order 10^5, rather than maximize $\mathcal{L}$. The factor $1/N!$ in Eqs. 4.12 and 4.13 is sometimes omitted since it is an additive constant in $-\ln\mathcal{L}$ and thus has no effect on the parameter fitting.

4.5.1.4 Factorized PDFs and correlations among discriminating variables

We show the correlation coefficients for all pairs of variables used in the fit for on-resonance data (predominantly continuum background) for $\eta' K^0$ and ηK^* in Appendices C and H, respectively. Because the correlations among discriminating variables are low ($< 5\%$), it is a reasonable approximation to factorize the PDF for each fit component:

$$\mathcal{P}_j(\mathbf{x}) \;=\; \mathcal{P}_j(x_1)\mathcal{P}_j(x_2)...\mathcal{P}_j(x_m). \tag{4.14}$$

(These individual functions $\mathcal{P}_j(x_m)$ are typically referred to as *the PDFs.*)

Biases induced by correlations in components other than $q\bar{q}$ background, which are also small, are evaluated in embedded toy studies and accounted for in the analysis as described in Sec. 4.5.4.2. In the $\eta' K^0$ analysis we perform additional studies of biases induced by correlations in the $q\bar{q}$ background. By parameterizing the correlations between several variables (e.g. by allowing the PDF that describes ΔE to be a function of $m_{\rm ES}$), we confirm that these biases are negligible.

4.5.1.5 Floating background parameters

As described earlier, we can float some of the PDF parameters $\boldsymbol{\alpha}$ in the ML fit. By doing this, we include uncertainties in the values of these parameters in the fit statistical error, and the background parameters can be determined by making use of the larger statistics available in the full on-resonance sample rather than restricting the data to on-resonance sidebands, off-peak data, or continuum MC.

We float the most important parameters in the background: the ARGUS parameter, ΔE slope, core $\mathcal{F}$ shape, and in $\eta' K^0$, the Δt resolution model parameters and tagging category fractions. We do not, however, float the double Gaussian component of resonance masses in ηK^*, but determine this from signal MC including appropriate scale factors and shifts as described in Sec. 6.8.4.2. For all parameters floated in our

analyses, we initialize their values to those determined from the $m_{\rm ES}$ or ΔE sidebands. Specific listings of floated parameters, as well as their initial and final values, are given in Appendix C for the $\eta' K^0$ modes and Appendix H for the ηK^* modes. We have tested with toy MC experiments (Sec. 4.5.4.1) that the fitter can handle the number of degrees of freedom we use in our final fits.

4.5.2 Averaging results

Another layer of complication of the ML fit involves measuring the same physical observable in several categories of data. In one analysis, we measure the $B^0 \to \eta K^{*0}(892)$ branching fraction in two sub-decays where the η decays as $\eta \to \gamma\gamma$ or $\eta \to \pi^+\pi^-\pi^0$. In the other analysis, we measure S (and C) in the seven $B^0 \to \eta' K^0$ sub-decays, where within each $\eta' K^0$ sub-decay we measure S (and C) in each of six tagging categories. We can average these individual results in two ways: with a simultaneous fit or by summing scans of $-2\ln\mathcal{L}$ (or multiplying $\mathcal{L}$).

The advantage of the simultaneous fit is that, in cases where multiple parameters are being measured, the simultaneous fit properly accounts for correlations between the parameters. The advantage of summing likelihood curves is that we can properly account for uncorrelated systematic uncertainties. A practical advantage of averaging with likelihood curves is that one works with the simple individual ML fits rather than the grand simultaneous ML fit which converges very slowly.

4.5.2.1 Simultaneous fit

In a simultaneous fit, we form the total likelihood as the product of the individual likelihoods for each category. In the case of tagging categories, we have an individual likelihood $\mathcal{L}_c$ for each of the seven tagging categories (including untagged events) c

$$\mathcal{L}_c \;=\; \frac{e^{-\nu_c}}{N_c!}\prod_i^{N_c}\nu_c\sum_j^{M} f_{j,c}\mathcal{P}_{j,c}(x_i,\boldsymbol{\alpha}), \tag{4.15}$$

where ν_c is the expected number of events in category c; N_c is the number of events to fit in each category c; $f_{j,c}$ is the fraction of the ν_c events in component j with the requirement that $\sum_j^M f_{j,c} = 1$. The product of these likelihoods is the grand likelihood

$$\mathcal{L} = \prod_c^6 \left[\frac{e^{-\nu_c}}{N_c!} \prod_i^{N_c} \nu_c \sum_j^M f_{j,c} \mathcal{P}_{j,c}(x_i, \boldsymbol{\alpha}) \right]. \tag{4.16}$$

In practice, we are interested not in ν_c and $f_{j,c}$, but in ν_j, the total number of events in each component, and $\epsilon_{j,c}$, the fraction of events in component j that are in category c, which is the fraction (efficiency) actually measured by the tagging algorithm. So it is more instructive to write the product $\nu_c f_{j,c}$ in Eq. 4.16 as $\nu_j \epsilon_{j,c}$ making the grand likelihood

$$\mathcal{L} = \prod_c^6 \left[\frac{e^{-\nu_c}}{N_c!} \prod_i^{N_c} \sum_j^M \nu_j \epsilon_{j,c} \mathcal{P}_{j,c}(x_i, \boldsymbol{\alpha}) \right]. \tag{4.17}$$

Buried within the likelihood are six parameters $S_{\mathrm{sig,c}}$. We measure the average S by maximizing the grand likelihood constraining all $S_{\mathrm{sig,c}}$ to be equal.

4.5.2.2 Combining scans of the likelihood

It is useful to consider the averaging of branching fractions measured in two sub-decays when describing the method of likelihood scans. We obtain the likelihood as a function of branching fraction $\mathcal{L}(\mathcal{B})$ by performing the fit many times with the branching fraction fixed to different values. We average the branching fraction measurements by summing the $-2\ln\mathcal{L}(\mathcal{B})$ functions, which is equivalent to taking the product of the likelihood functions, and finding the value of $\mathcal{B}$ at the minimum ($\hat{\mathcal{B}}$) of the resulting function. As discussed in Sec. 4.5.1 the uncertainty on $\hat{\mathcal{B}}$ is the value σ such that $-2\ln\frac{\mathcal{L}(\hat{\mathcal{B}}+\sigma)}{\mathcal{L}_{\max}} = 1$.

For systematic uncertainty (σ_{syst}) that is not correlated between the results to be averaged, we convolve the likelihood for each measurement with a Gaussian of width σ_{syst}, which has the effect of increasing the width of the likelihood from σ to $\sqrt{\sigma^2 + \sigma^2_{\mathrm{syst}}}$.

After summing the corrected curves, we find the value of the branching fraction at the minimum of the resulting function as above.

Correlated systematic uncertainties are treated differently since they affect the results to be averaged with the same sign. To find the correlated systematic uncertainty on the averaged result, we shift the individual likelihood curves one-by-one, summing and finding $\hat{\mathcal{B}}$ for each iteration. The average change between $\hat{\mathcal{B}}$ and the effective $\hat{\mathcal{B}}$ for each iteration is the weighted correlated systematic uncertainty on the averaged result. As above, we convolve the averaged likelihood curve with a Gaussian with a width of this averaged uncertainty to complete the process.

4.5.3 Obtaining PDFs

A primary task in the maximum likelihood method is the determination of the PDFs for the various input quantities in the fit. A PDF must be determined for each discriminating variable ($m_{\rm ES}$, ΔE, $\mathcal{F}$, etc.) for each component (signal, continuum background, etc.). Each PDF is determined with an independent ML fit to the relevant distribution in a well identified data sample. The data samples used for determining the PDFs are exclusive MC for all ηK^*, $\eta' K$, and charmless $B\overline{B}$ components, on-resonance data in sidebands chosen to avoid potential signal events for $q\bar{q}$ background components, and generic $B\overline{B}$ MC for charm $B\overline{B}$ components. As mentioned earlier, the important parameters of the $q\bar{q}$ background PDFs are left free in the final fit to ensure that we correctly characterize the dominant component; therefore, the PDFs determined for $q\bar{q}$ backgrounds are just starting points for the final fit to data.

4.5.3.1 Data control samples

Ideally we would also leave parameters of the signal PDFs free in the fit to data, but relatively small signal yields make such a plan unfeasible. Thus we fix the signal PDF parameters at values obtained from MC, and we confirm that the MC is an accurate

representation of the data with studies of large data control samples. If differences are found, we determine how the MC must be shifted and scaled to match the data.

For $m_{\rm ES}$ and ΔE we study samples of $\sim$ 50,000 $B^- \to D^0\pi^-$ events (with $D^0 \to K^+\pi^-\pi^0$), which are topologically similar to our signal decays. For $\mathcal{F}$ we use samples of $\sim$ 3200 $\eta'_{\rho\gamma}K^+$ events and find that data and MC agree within errors. For η and K^* invariant mass distributions we use the large samples of real resonances that appear in the continuum background to determine the agreement between data and MC. The $\eta_{\gamma\gamma}$ mass requires significantly different shifts for data taking periods Run1, Run2, and Runs3-5 because of EMC calibrations between runs.

The resulting shifts, scales, and related uncertainties are shown in Table 4.1. The results differ for ηK^* and $\eta' K$ analyses because the studies were performed with different processing of the data and analysis configurations. For all parameters, we use the uncertainties from these studies in estimating the final systematic errors.

As described in Sec. 5.3.3, the analysis of $\eta' K^0_L$ differs significantly from that of $\eta' K^0_S$. Because we only know the direction of the K^0_L momentum, we constrain the B and K^0_L masses during vertex fitting for the $\eta' K^0_L$ modes. Because of this unique reconstruction method, the mean and width of the ΔE distribution depends mostly on the angular resolution of the EMC and IFR. The *BABAR* $J/\psi K^0_L$ analysis group have evaluated data/MC agreement for the shape of ΔE for $B^0 \to J/\psi K^0_L$ by reconstructing $J/\psi K^0_S$ events with the K^0_L method. They find that the MC needs to be shifted by -0.40 ± 0.25 MeV and scaled by 1.1 ± 0.1. We find that our results change very little when applying this shift and scale, so we take the total change as a systematic uncertainty.

4.5.4 Fit validation

Before fitting the data, we validate the performance of the fit by applying it to simulated data. We check for fit stability, biases on measured parameters, and the reasonableness of reported uncertainties with fits to *toy* MC samples. We perform *pure*

Table 4.1: Results of control sample studies. Scale factors (data/MC) and shift parameters (data–MC) applied to $m_{\rm ES}$, ΔE, and resonance mass distributions to correct for differences between data and Monte Carlo. Corrections for $\eta' K^0_L$ are described in the text.

	ηK^*		$\eta' K^0_S$	
Variable	Scale Factor	Shift (MeV)	Scale Factor	Shift (MeV)
$m_{\rm ES}$	1.00 ± 0.05	0.0 ± 0.2	1.00 ± 0.04	0.0 ± 0.1
ΔE	1.05 ± 0.05	0.0 ± 5.0	1.00 ± 0.02	-2.0 ± 2.0
$\eta \to \gamma\gamma$ mass				
Run1	1.22 ± 0.05	8.7 ± 0.6		
Run2	1.10 ± 0.03	4.5 ± 0.3		
Run3-5	1.03 ± 0.01	6.3 ± 0.2		
$\eta \to \pi^+\pi^-\pi^0$ mass				
Run1-5	1.06 ± 0.02	0.1 ± 0.1		
K^* mass	1.00 ± 0.05	0.0 ± 0.5		

and *embedded* toy MC studies.

4.5.4.1 Pure toy studies

In pure toys, we generate $100 - 500$ toy MC datasets, each with the number of signal and background events expected in the actual data, according to the PDFs used in the fit. Because the simulated data are generated from factorized PDFs, pure toy studies are not sensitive to fit bias due to correlations in the data. Any significant bias observed in pure toys on a parameter of interest is a sign that there are problems with the fit implementation. Pure toys are also useful for confirming that the uncertainties reported by the fit are reasonable.

We fit each dataset and examine the distribution of fit parameters for the ensemble of datasets. Especially useful are the distributions of pulls for each free parameter. The pull is the difference between the input value and the fit value divided by the fit error,

$$\text{pull} \quad = \quad \frac{x_{\rm input} - x_{\rm fit}}{\sigma_{\rm fit}}. \tag{4.18}$$

The expected distribution of pulls is a Gaussian with unit width ($\sigma = 1$) and mean of zero ($\mu = 0$). Significant deviations from unity of the width indicate that uncertainties are over- or under-estimated. Significant deviations from zero of the mean indicate problems with fit implementation.

We also use pure toys to confirm that our fit configuration, specifically the number of free parameters, is stable. When some fraction of pure toy fits do not converge, we can reduce the number of free parameters until all toy fits converge.

4.5.4.2 Embedded toy studies

In embedded toys we construct $100-500$ toy MC datasets, each with the number of signal and background events expected in the actual data. The $q\bar{q}$ background events are generated according to the fit PDFs, but the signal and $B\overline{B}$ background events are embedded from the signal MC. We estimate the fit bias as the difference between the average fitted parameter in embedded toy MC fits and the true value of the parameter, which is known in the toy datasets. Embedded toys are important for evaluating the fit bias, which is due to correlations between fit variables in the signal, self-crossfeed, and $B\overline{B}$ backgrounds. Embedded toys are also sensitive to bias due to residual contributions of $B\overline{B}$ backgrounds to signal parameters, such as yields or CP parameters.

4.5.5 Fit evaluation plots

The output of the ML fit consists of the values of floated parameters, their covariances, and the likelihood. Because it is difficult to understand whether the fit behaved as expected from these numbers alone, we evaluate the fit performance with plots of projections of the total PDF overlaid on the data. We rely heavily on signal-enhanced *projection plots* and *s$\mathcal{P}$lots*.

For projection plots we project the final PDF onto a fit variable, such as $m_{\rm ES}$, ΔE, $\mathcal{F}$, and Δt, for a subset of the data for which the ratio of the likelihood to be

signal and the sum of likelihoods to be signal and background (computed without the variable plotted) exceeds a threshold that optimizes the statistical significance of the plotted signal. Examples of projection plots can be seen in Fig. 5.5 in the $\eta' K^0$ results (Sec. 6.12); the solid lines show the fit result; the dashed lines show the background contributions; and the points are the data.

For *sPlots* we also project the final PDF on the fit variables, but instead of a requirement on the likelihood ratio for each event, we use the fit covariance matrix to accumulate the event-by-event signal (and separately, background) probabilities in bins of the plotted fit variable. The fit is performed without the plotted variable so that the method is unbiased. The advantage of *sPlots* over projection plots is that the *sPlot* method completely isolates fit components without losing events to a requirement on the likelihood ratio; the disadvantage is that the projection method is much more transparent. A detailed, quantitative description of *sPlots* can be found elsewhere [61].

4.6 Time-dependent analyses

The description of the time-evolution of neutral B mesons in Sec. 2.5.3 neglects certain experimental realities. In practice, we modify the ideal time evolution equation (Eq. 2.29) to account for experimental resolution and the performance of the tagging algorithm. We extract the *CP* parameters S and C from the data by using the modified time-evolution equation as a PDF in the ML fit, with S and C free to vary.

One must measure in each event the difference in proper decay times of B_{CP} and $B_{\rm tag}$ (Δt) and the flavor of $B_{\rm tag}$. One must also parameterize the experimental resolution on Δt and the performance of the flavor tagging algorithm and determine the values of the relevant parameters. We obtain Δt by measuring the separation along the beam axis of the B_{CP} and $B_{\rm tag}$ decay vertices (Δz) and converting Δz to Δt using the boost of the $\Upsilon(4S)$ system, the reconstructed four-momentum of the B_{CP}, and the average B lifetime. We determine the flavor of $B_{\rm tag}$ using a neural network based algorithm.

Because the performance of the tagging algorithm is independent of signal mode and the resolution on Δt is dominated by the determination the tag-side decay vertex, we are not limited to our signal events when quantifying the tagging performance or Δt resolution. For this purpose we use a sample of $\sim 100,000$ (cf. ~ 2500 signal events) fully reconstructed, self-tagging, neutral B decays to $D^{(*)-}(\pi^+, \rho^+, a_1^+)$ flavor eigenstates ($B_{\rm flav}$ sample) to characterize the tagging performance and Δt resolution.

4.6.1 Measuring Δz

We determine the decay vertex of the fully reconstructed signal B_{CP} as described in Sec. 4.3, but we also need the position of the decay vertex of $B_{\rm tag}$ to obtain Δz. To maximize efficiency we use an inclusive approach; i.e., we determine the $B_{\rm tag}$ decay vertex using an iterative fit that only requires partial reconstruction of the $B_{\rm tag}$ decay products. As a starting point for this iterative fit, we consider all tracks in the event that are not associated with B_{CP}, ignoring all neutrals. Before fitting these tracks to a vertex, we take steps to ensure that our simple $B_{\rm tag}$ vertex reconstruction is optimized and unbiased. To optimize the reconstruction we use the reconstructed decay vertex and momentum of B_{CP} as additional constraints. Since B_{CP} and $B_{\rm tag}$ are produced back-to-back in the $\Upsilon(4S)$ frame, we can use the reconstructed B_{CP} momentum to point back to the beam spot and help determine the $\Upsilon(4S)$ decay vertex and the $B_{\rm tag}$ direction of flight.

We also remove, from the set of tracks used in $B_{\rm tag}$ vertexing, those tracks that may bias the determination of the vertex. The daughters of long-lived particles, such as K^0_S mesons or Λ baryons, that decay outside the SVT are not used in the fit. However, the daughters of long-lived particles with SVT information are combined into a composite track which is retained in the $B_{\rm tag}$ vertex reconstruction. Tracks from photon conversions are also explicitly removed from the fit. After an initial $B_{\rm tag}$ vertex fit, we recursively remove the track with the largest contribution to the total fit χ^2 until no

tracks contribute more than 6 units to the total χ^2.

The average Δz in these decays is $\sim 260\,\mu$m and the experimental resolution on Δz, which is dominated by the uncertainty on the $B_{\rm tag}$ vertex, is $\sim 130\,\mu$m.

4.6.2 Converting $\Delta z \to \Delta t$

A measurement of Δz is worthless unless we can convert it to Δt. If one assumes that the momentum of the B's in the $\Upsilon(4S)$ frame is negligible, then the boost of the B's and the $\Upsilon(4S)$ are the same in the lab frame. This assumption allows for the simple conversion $\Delta t = \Delta z/\gamma\beta_z c$, where γ is the Lorentz factor of the $\Upsilon(4S)$ in the lab and β_z is the z-component of the $\Upsilon(4S)$ velocity; at *BABAR* $\beta_z\gamma \approx 0.56$. In reality the B's have an average momentum of 340 MeV in the $\Upsilon(4S)$ frame, which implies a γ^* of 1.002 and β^* of 0.06.

Since we fully reconstruct B_{CP}, we improve the Δt resolution by 5% by accounting for the non-zero momenta of the B's in the $\Upsilon(4S)$ frame:

$$\Delta z \;=\; \beta_z\gamma\gamma^*_{CP}c\Delta t + \gamma\beta^*_{CP}\gamma^*_{CP}\cos\theta^*_{CP}c(t_{CP}+t_{\rm tag}), \tag{4.19}$$

where $\beta^*_{CP}, \gamma^*_{CP}$, and $\cos\theta^*_{CP}$ are the velocity, the Lorentz factor, and the cosine of the angle with respect to the beam-axis of B_{CP}. The first term is the simple Δz conversion slightly modified by the factor γ^*_{CP} (~ 1.002). We know $\cos\theta^*_{CP}$ with good precision because B_{CP} is fully reconstructed, but it is difficult to measure $t_{CP}+t_{\rm tag}$. Assuming that $t_{CP}+t_{\rm tag}$ is on average twice the B lifetime ignores the variation with Δt, so we assume that $\langle t_{CP}+t_{\rm tag}\rangle = \tau_B + |\Delta t|$, and we convert Δz to Δt with the relation:

$$\Delta z \;=\; \beta_z\gamma\gamma^*_{CP}c\Delta t + \gamma\beta^*_{CP}\gamma^*_{CP}\cos\theta^*_{CP}c(\tau_B+|\Delta t|). \tag{4.20}$$

The average contribution of the second term ($\sim 40\,\mu$m) is small ($\sim 5\%$ in quadrature) compared to the resolution on Δz from the determination of the $B_{\rm tag}$ vertex ($\sim 130\,\mu$m).

4.6.3 *B* flavor tagging

To determine the flavor of $B_{\rm tag}$, we use a neural network based algorithm that assigns each candidate $B_{\rm tag}$ to one of six hierarchical, mutually exclusive categories, which are defined by the continuous output of the NN and the presence of a lepton $B_{\rm tag}$ daughter for the `Lepton` category. The goal is to determine the flavor of $B_{\rm tag}$ with high efficiency ϵ and low probability w of assigning a wrong flavor to $B_{\rm tag}$. For instance, the probability of misidentifying the flavor of $B_{\rm tag}$ with the charge of the prompt lepton is $\sim 3\%$, but only $\sim 9\%$ of events fall into this tagging category. The tagging categories, in order of increasing mistag rate, are `Lepton`, `Kaon I`, `Kaon II`, `Kaon-Pion`, `Pion`, and `Other`. We are able to tag about three quarters of all events; untagged events help constrain PDF parameters other than S and C.

The figure of merit for the performance of the tagging algorithm is the effective tagging efficiency

$$Q \;=\; \epsilon(1-2w)^2, \tag{4.21}$$

which is approximately related to the statistical uncertainty σ on the coefficients S and C through

$$\sigma \;\propto\; \frac{1}{\sqrt{Q}}. \tag{4.22}$$

The total effective tagging efficiency Q for this algorithm is measured to be $(31.2\pm0.3)\%$. The Q values for each tagging category are shown in Table 4.2.

4.6.4 Mistag probabilities and experimental Δt resolution

As introduced above, we use the $B_{\rm flav}$ sample to measure the average tagging efficiencies (ϵ), average mistag rates (w), and the differences in mistag rates (Δw) and tagging efficiencies (μ) for B^0 and $\overline{B}^0$ tag-side decays in six tagging categories. The results are shown in Table 4.2.

Table 4.2: Efficiencies ϵ, average mistag fractions w, mistag fraction differences $\Delta w \equiv w(B^0) - w(\overline{B}^0)$, tagging efficiency differences $\mu \equiv \epsilon(B^0) - \epsilon(\overline{B}^0)$, and effective tagging efficiency $Q \equiv \epsilon(1-2w)^2$ for each tagging category from the $B_{\rm flav}$ data.

Category	ϵ (%)	w (%)	Δw (%)	μ (%)	Q (%)
`Lepton`	9.0 ± 0.1	2.8 ± 0.3	0.3 ± 0.5	-0.3 ± 0.9	8.0 ± 0.1
`Kaon I`	10.8 ± 0.1	5.3 ± 0.3	-0.1 ± 0.6	0.1 ± 0.9	8.7 ± 0.1
`Kaon II`	17.2 ± 0.1	14.5 ± 0.3	0.4 ± 0.6	0.6 ± 0.8	8.7 ± 0.2
`Kaon-Pion`	13.7 ± 0.1	23.3 ± 0.4	-0.7 ± 0.7	0.2 ± 0.9	3.9 ± 0.1
`Pion`	14.2 ± 0.1	32.5 ± 0.4	5.1 ± 0.7	-2.5 ± 0.9	1.7 ± 0.1
`Other`	9.5 ± 0.1	41.5 ± 0.5	3.8 ± 0.8	1.8 ± 1.0	0.3 ± 0.0
All	74.4 ± 0.1				31.2 ± 0.3

We modify the ideal time evolution equation (Eq. 2.29) to account for the mistag probability of the tagging algorithm by writing the corrected function f' in terms of the ideal function f

$$f'_{\pm} \quad = \quad (1 - w_{\pm}) f_{\pm} + w_{\mp} f_{\mp}, \tag{4.23}$$

where w_+ (w_-) is the probability that a true B^0 ($\overline{B}^0$) meson is tagged as a $\overline{B}^0$ (B^0). The average mistag rate is $w \equiv (w_+ + w_-)/2$ and the difference is $\Delta w \equiv w_+ - w_-$. We can write

$$\begin{aligned} f'(\Delta t) \quad &= \quad \frac{e^{-|\Delta t|/\tau}}{4\tau} \{1 \mp \Delta w \pm \\ &\qquad (1-2w)\left[-\eta_f S_f \sin(\Delta m_d \Delta t) - C_f \cos(\Delta m_d \Delta t)\right]\}, \end{aligned} \tag{4.24}$$

where we have abbreviated $f_{CP} \to f$ in the subscripts, and we explicitly include the CP eigenvalue $-\eta_f$ in the expression. In addition to w and Δw, we also include the tagging efficiency difference μ in the decay rate model, but since μ is consistent with zero (Table 4.2), we omit the complicated formula for simplicity.

Finally, we also modify this equation to account for the resolution on Δt by convolving f' with a resolution function, which is the sum of three Gaussians (called

core, tail, outlier),

$$\begin{aligned}\mathcal{R}(\Delta t, \sigma_{\Delta t}) &= (1 - f_{\rm tail} - f_{\rm out})G(\Delta t, \mu_{\rm core}\sigma_{\Delta t}, \sigma_{\rm core}\sigma_{\Delta t}) \\ &\quad + f_{\rm tail}G(\Delta t, \mu_{\rm tail}\sigma_{\Delta t}, \sigma_{\rm tail}\sigma_{\Delta t}) + f_{\rm out}G(\Delta t, \mu_{\rm out}, \sigma_{\rm out}).\end{aligned} \tag{4.25}$$

To remove to first order the slight dependence of Δt resolution on the signal mode, the means and widths of the core and tail Gaussians are scaled by the uncertainty on Δ_t ($\sigma_{\Delta t}$). The mean and width of outlier Gaussian, which accounts for the few events with incorrectly reconstructed vertices, are fixed at 0 ps and 8 ps, respectively. The other parameters of the triple Gaussian, shown in Table. 4.3, are determined with fits to the $B_{\rm flav}$ sample.

Table 4.3: Summary of Δt resolution function parameters obtained from the $B_{\rm flav}$ sample. We report the mean (μ), width (σ), and fraction of the total function (f) for the core, tail, and outlier Gaussians.

Parameter	Value
Scale `Lepton` (core)	1.0291 ± 0.0420
Scale non-`Lepton`(core)	1.0930 ± 0.0210
Bias `Lepton` (core)	-0.0618 ± 0.0268
Bias non-`Lepton` (core)	-0.1906 ± 0.0125
f (core)	0.8820 ± 0.0080
Scale (tail)	3.0 (fixed)
Bias (tail)	-1.0347 ± 0.1089
f (outlier)	0.0031 ± 0.0005
Width (outlier) (ps)	8.0 (fixed)
Mean (outlier) (ps)	0.0 (fixed)

Chapter 5

$\eta' K^0$ Analysis

5.1 General analysis strategy

In the $B^0 \to \eta' K^0$ analysis we measure the time-dependent *CP* parameters, S and C, with a simultaneous fit to seven sub-decay channels: two *major* modes, $B^0 \to \eta'_{\rho\gamma} K^0_S$ and $B^0 \to \eta'_{\eta\pi\pi} K^0_S$, and five *minor* modes, $B^0 \to \eta'_{\rho\gamma} K^0_{S00}$, $B^0 \to \eta'_{\eta\pi\pi} K^0_{S00}$, $B^0 \to \eta'_{5\pi} K^0_S$, $B^0 \to \eta'_{5\pi} K^0_L$, and $B^0 \to \eta'_{\eta\pi\pi} K^0_L$ where the following decays are implied

- $\eta'_{\rho\gamma} : \eta' \to \rho\gamma$,
- $\eta'_{\eta\pi\pi} : \eta' \to \eta\pi^+\pi^-$ with $\eta \to \gamma\gamma$,
- $\eta'_{5\pi} : \eta' \to \eta\pi^+\pi^-$ with $\eta \to \pi^+\pi^-\pi^0$,
- $K^0_S : K^0_S \to \pi^+\pi^-$,
- $K^0_{S00} : K^0_S \to \pi^0\pi^0$,
- The K^0_L does not decay in our detector.

As a crosscheck, we also measure the *CP* parameters in two self-tagging, charged modes $B^+ \to \eta'_{\rho\gamma} K^+$ and $B^+ \to \eta'_{\eta\pi\pi} K^+$ in which we expect both S and C to be zero.

So that the reader may get a feel for the the signal purity of the data samples and relative importance of each sub-mode, we compare branching fractions, detection efficiencies (determined from MC), the number of events passing selection, and expected yields for all sub-decays in Table 5.1.

Table 5.1: For each sub-decay, we report the sub-decay branching fraction (BF), the detection efficiency ϵ, the number of events passing the total selection (the number of events entering the fit), and the expected yield. We compute the BF's assuming an inclusive branching fraction of 64.9×10^{-6} for $B^0 \to \eta' K^0$ [55].

Mode	BF (10^{-6})	ϵ (%)	# Evts into Fit	Expected Yield
$\eta'_{\eta\pi\pi} K^0_S$	3.9	27	1556	500
$\eta'_{\rho\gamma} K^0_S$	6.6	29	23905	900
$\eta'_{\eta\pi\pi} K^0_{S00}$	1.8	14	1175	110
$\eta'_{\rho\gamma} K^0_{S00}$	3.0	15	28871	200
$\eta'_{5\pi} K^0_S$	2.3	17	546	180
$\eta'_{\eta\pi\pi} K^0_L$	5.7	18	14125	450
$\eta'_{5\pi} K^0_L$	3.3	11	4951	170

5.2 Data and Monte Carlo samples

As mentioned in Sec. 4.2, the $\eta' K^0$ analysis is based on the full *BABAR* on-resonance dataset with an integrated luminosity of 426 fb^{-1} corresponding to 467 million $B\overline{B}$ pairs. For signal MC samples, we use 972,000 generated events for the major modes $\eta'_{\rho\gamma} K^0_S$ and $\eta'_{\eta\pi\pi} K^0_S$. We use 380,000 events for the modes $\eta'_{\rho\gamma} K^0_{S00}$ and $\eta'_{\eta\pi\pi} K^0_L$ and 195000 events for the modes $\eta'_{\eta\pi\pi} K^0_{S00}$, $\eta'_{5\pi} K^0_S$, and $\eta'_{5\pi} K^0_L$. For $B\overline{B}$ background studies, we use ∼650 million generic B^+B^- and ∼650 million generic $B^0\overline{B}^0$ MC events as well as samples of 100-600K events for several dozen individual charmless background modes. For $\eta' K^0_L$ modes we use continuum MC, which is generated with the same integrated luminosity as the real dataset, for training the neural network (Sec. 5.3.3.3).

5.3 Event reconstruction and selection

B candidates are formed by combining an η' candidate with a K^0_S or K^0_L candidate. All tracks are taken from the `GoodTracksVeryLoose` list. (See Sec. 4.3 for description of track lists.) To reduce correlation between ΔE and resonance mass and to improve resolution on the decay vertex, the masses of several resonances whose width

is narrower than detector resolution (η', η, and π^0) are constrained to the nominal value [62] during vertex fitting.

5.3.1 $K^0_S \to \pi^+\pi^-$ reconstruction

Because of its significant lifetime, the K^0_S decay vertex does not coincide with the B decay vertex. We reconstruct $K^0_S \to \pi^+\pi^-$ candidates from two tracks from the `ChargedTracks` list. First we perform a simple fit requiring the K^0_S mass to be between 468 and 528 MeV making a loose requirement on the K^0_S vertex probability to reject combinatorial backgrounds. Then we refit the K^0_S candidate with a `Bflight` constraint, as described in Sec. 4.3.1, and a slightly tighter mass cut (473 – 523 MeV). We find that a mass constraint on the K^0_S does not improve determination of the vertex.

5.3.2 $K^0_S \to \pi^0\pi^0$ reconstruction

Reconstructing the $K^0_S \to \pi^0\pi^0$ vertex is challenging due to the all neutral final state in which we can only measure the energies and interaction positions of the photons. Since the K^0_S travels several cm and the B travels less than one mm, we assume that the K^0_S was produced at the beam-spot when obtaining the vertex. This constraint coupled with mass constraints on the neutral pions and the measured photon energies reduces the number of degrees of freedom to two. These can intuitively be thought of as the distance of the K^0_S vertex from the beam-spot and the opening angle of the pions. With these quantities we can completely characterize the four momentum and the decay vertex of the K^0_S.

5.3.3 K^0_L reconstruction and selection

Candidate K^0_L do not decay in the detector, but deposit part of their energy in the EMC and/or IFR. We reconstruct K^0_L candidates from clusters of energy deposited in the EMC or from hits in the IFR not associated with any charged track in the event

[63]. Of all reconstructed K_L^0 candidates, 60% come from the EMC and the rest from the IFR. If a K_L^0 is detected in both the EMC and IFR, we treat it as an EMC candidate.

Since we cannot measure the K_L^0 candidate's energy, we only measure the momentum direction of the K_L^0, as the direction from the vertex of the η' to the centroid of the EMC (or IFR) cluster. We obtain the K_L^0 four-momentum by constraining the masses of the K_L^0 and the B to their nominal masses during B vertexing with the complete η' four momentum and the K_L^0 direction. As a result of constraining the B mass, the kinematic variables $m_{\rm ES}$ and ΔE are completely correlated, and we use only the ΔE variable in the fit. Additionally, the shape of the ΔE distribution is not the standard Gaussian centered at $\Delta E = 0$. In Fig. 5.1, we compare ΔE distributions from off-resonance data and signal MC.

For the $\eta' K_L^0$ modes, we reduce the background with requirements on three quantities described below: transverse projected missing momentum $P_{\rm miss}^{\rm proj}$, the cosine of the angle between the missing momentum and the beam direction in the lab frame $\cos\theta_{P_{\rm miss}}$, and a neural network of EMC shower shape variables for candidates with a K_L^0 from the EMC. We describe the optimization of these cuts in Appendix A.

5.3.3.1 Transverse projected missing momentum

To obtain the transverse missing momentum of an $\eta' K_L^0$ candidate, we first find the missing momentum $\mathbf{p}_{\rm miss}$ by subtracting the total momentum in the event, excluding the K_L^0 candidate, from the momentum of the beams. We project $\mathbf{p}_{\rm miss}$ onto the direction of the K_L^0 candidate $\hat{\mathbf{p}}_{K_L^0}$ and take the component of the projection that is transverse to the beams. We then subtract from this the transverse K_L^0 momentum obtained from the reconstruction. Missing momentum can also originate from neutrinos in semileptonic decays and leakage in the endcap. We remove the bias from endcap leakage by only considering the transverse momentum, and we remove the bias from missing neutrino momentum by projecting the missing momentum onto the K_L^0 direction. We require

that $P_{\rm miss}^{\rm proj} > -0.7$. In Fig. 5.1, we compare $P_{\rm miss}^{\rm proj}$ distributions from off-resonance data and signal MC.

5.3.3.2 Cosine of missing momentum polar angle

We also characterize K_L^0 candidates by the cosine of the angle between the missing momentum vector and the z-axis, $\cos\theta_{P_{\rm miss}}$. We find an optimum cut on this variable of $\cos\theta_{P_{\rm miss}} < 0.95$. In Fig. 5.1, we compare $\cos\theta_{P_{\rm miss}}$ distributions from off-resonance data and signal MC.

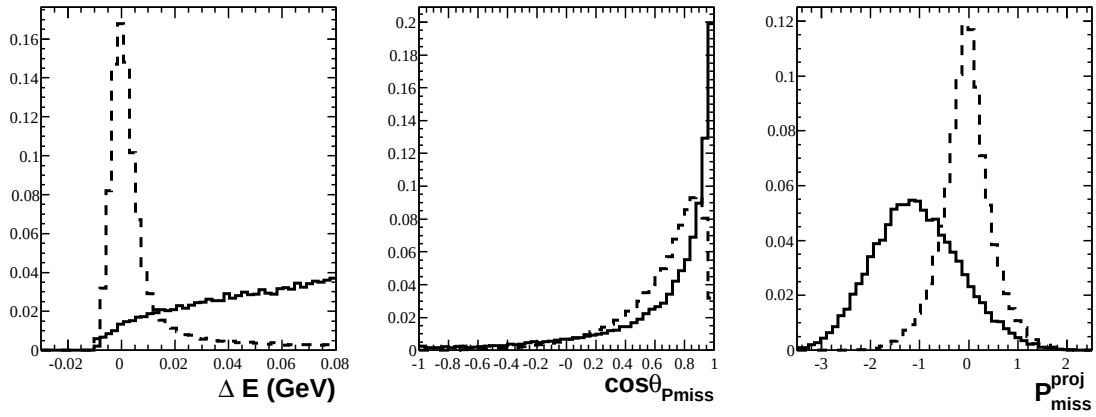


Figure 5.1: Distributions for ΔE, $\cos\theta_{P_{\rm miss}}$, and $P_{\rm miss}^{\rm proj}$ from signal MC (dashed) and off-resonance data (solid) for the mode $\eta'_{\eta\pi\pi}K_L^0$. (Units of vertical axes are arbitrary.)

5.3.3.3 EMC shower shape

For $\eta' K_L^0$ candidates detected in the EMC, we discriminate between signal and background using quantities related to the shape of the electromagnetic shower in the EMC. We use seven shower shape variables as input to a multi-layer perceptron (MLP) artificial neural network (NN), and we cut on the continuous output of the NN. We list the NN input quantities in Appendix B. The NN, which is implemented with the TMVA software package [64], uses the standard sequential back-propagation learning

method and comprises an input layer of seven nodes, two hidden layers with eight and seven nodes, and a single output node. We train the NN on $q\bar{q}$ background MC in 200 training cycles. In Fig. 5.2 we show the NN output for signal and background MC; we require that the NN output be greater than 0.3.

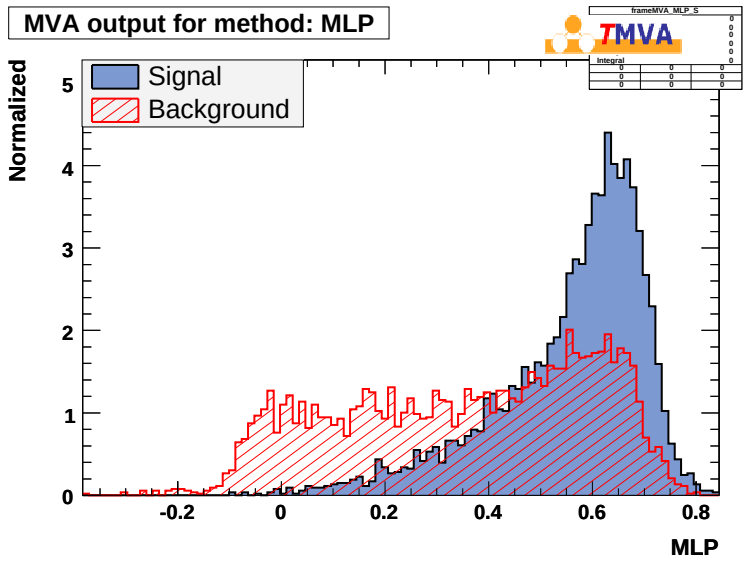


Figure 5.2: Neural net output variable for signal MC and background MC for the mode $\eta'_{\eta\pi\pi}K^0_L$.

5.3.4 Preliminary cuts

A number of preliminary cuts are applied before B candidates are considered for the ML fit. In general, preliminary cuts are loose since the ML fit more efficiently separates signal and background than simple cuts.

- $N_{\rm trks} \geq N_{\rm tracks\,in\,decay\,mode} + 1$ (in order to be able to define a thrust vector for the rest of the event),
- $|\cos\theta_T| \leq 0.9$, ($|\cos\theta_T| \leq 0.8$ for $\eta' K^0_L$),
- $|\Delta E| \leq 0.2$ GeV, ($-0.01 < \Delta E < 0.08$ GeV for $\eta' K^0_L$),

- $5.25 \leq m_{\rm ES} < 5.2893$ GeV, (no cut for $\eta' K^0_L$),
- $-4 < \mathcal{F} < 5$, where $\mathcal{F}$ is the Fisher discriminant,
- $E_\gamma > 30$ MeV for π^0, $E_\gamma > 50$ MeV for $\eta_{\gamma\gamma}$, and $E_\gamma > 100$ MeV in $\eta' \to \rho^0\gamma$,
- $|\cos\theta_\rho| < 0.9$, where θ_ρ is the angle between the direction of a ρ^0-daughter π^+ and the direction of the η' in the ρ^0 rest frame,
- $930 < m_{\eta'(\rho\gamma)} < 980$ MeV,
- $945 < m_{\eta'(\eta\pi\pi)} < 970$ MeV,
- $490 < m_{\eta\to\gamma\gamma} < 600$ MeV,
- $520 < m_{\eta\to 3\pi} < 570$ MeV,
- $470 < m_\rho < 1070$ MeV,
- $120 < m_{\pi^0} < 150$ MeV, ($120 < m_{\pi^0} < 155$ MeV for K^0_{S00}),
- $486 < m^{K_S}_{\pi^+\pi^-} < 510$ MeV,
- $468 < m^{K_S}_{\pi^0\pi^0} < 528$ MeV.
- Particle Identification: The η' daughter charged tracks must NOT satisfy the `Tight` criteria of the electron likelihood selector, the `VeryTight` criteria of the proton likelihood selector, or the `Tight` criteria of the kaon likelihood selector. The efficiency of the selectors in MC is corrected to match the response in real data.

We require the K^0_{S+-} to satisfy:

- fit probability > 0.001,
- flight length significance $> 3\sigma$; the flight length significance is the flight length divided by the uncertainty on the flight length.

For quantities related to the time-dependent analysis we require:

- $|\Delta t| < 20$ ps,
- $\sigma_{\Delta t} < 2.5$ ps.

5.3.5 Multiple candidates per event

In events with more than one B candidate passing all cuts, we choose the candidate with the highest B vertex probability. From MC we determine that, depending on decay mode, we choose a candidate with an incorrectly reconstructed vertex in $7-47\%$ of signal events. We call these events self-crossfeed (SXF). We consider three types of SXF: *total*, *rest of event* (ROE), and *wrong track* (WT).

- Total SXF comprises events which are misreconstructed in any way, e.g. swapping stable daughters between signal B granddaughters. Most of these SXF events affect neither the determination of the B decay vertex nor the distributions of events in $m_{\rm ES}$, ΔE, or $\mathcal{F}$.
- ROE SXF comprises events in which the signal B swaps any stable daughter (charged or neutral) with the rest of the event; it does not include events in which particles are swapped between signal B daughters, granddaughters, etc. Many of these events do not affect vertexing, but all effect $m_{\rm ES}$ and ΔE PDFs.
- Wrong Track SXF includes only events in which a signal B daughter swaps a charged particle with the ROE. Wrong-track misreconstruction affects vertex reconstruction most, but wrong-track SXF comprises only $0.9-3.9\%$ of signal events, depending on sub-mode. The sub-mode with the highest fraction is $\eta'_{5\pi}K^0_L$.

In Table 5.2 we summarize the number of combinations per event in data and signal MC, the total SXF fraction, the ROE-SXF fraction, and the WT-SXF fraction. The

numbers of combinations per event for $\eta'_{5\pi}K^0_S$ and $\eta'_{5\pi}K^0_L$ include combinations in which pions that are daughters of the η' are internally swapped between the η and two bachelor pions. When these combinations are not counted, the numbers of combinations per event in MC are close to ~ 1.3. We describe studies of the effects of misreconstruction on the signal *CP* parameters in Sec. 5.11 with study results in Table 5.20.

Table 5.2: Total self-crossfeed fraction (SXF), rest-of-event SXF fraction, wrong-track SXF fraction, and number of combinations per event from MC and Data.

Mode	Total SXF Fraction (%)	ROE-SXF Fraction (%)	WT-SXF Fraction (%)	Combs/Evt MC	Combs/Evts Data
$\eta'_{\eta\pi\pi}K^0_S$	6.6	5.2	0.9	1.11	1.08
$\eta'_{\rho\gamma}K^0_S$	7.4	6.6	1.8	1.07	1.07
$\eta'_{\eta\pi\pi}K^0_{S00}$	16.8	15.9	0.9	1.24	1.26
$\eta'_{\rho\gamma}K^0_{S00}$	18.6	17.8	1.8	1.21	1.24
$\eta'_{5\pi}K^0_S$	24.3	14.9	2.4	1.83	1.61
$\eta'_{\eta\pi\pi}K^0_L$	31.2	9.6	1.8	1.25	1.18
$\eta'_{5\pi}K^0_L$	47.3	20.6	3.9	2.05	1.79

5.4 $B\overline{B}$ background introduction

While the continuum $q\bar{q}$ events are the dominant background, events from $B\overline{B}$ decays other than signal are present in the data. These $B\overline{B}$ backgrounds include charmless and charmed decays. Accounting for these backgrounds is important because their PDFs are similar to signal, so these events can contribute to the signal in the absence of separate charmless or charm $B\overline{B}$ fit components. We perform several studies to identify these $B\overline{B}$ backgrounds and determine whether dedicated fit components are necessary, but before presenting a detailed description of these studies, it is useful to describe the maximum likelihood fit for the $\eta' K$ analysis.

5.5 Maximum likelihood fit

We perform an unbinned multivariate extended maximum likelihood fit with several discriminating variables and fit components. The configurations are slightly different for each sub-decay channel. For the modes with a K^0_S or K^+, the discriminating variables used in the fits are $m_{\rm ES}$, ΔE, $\mathcal{F}$, Δt, and the tagging category c. For the modes with a K^0_L, we do not use $m_{\rm ES}$. (As described in Sec. 5.3.3, we constrain the mass of the B during vertexing for the $\eta' K^0_L$ analyses , so $m_{\rm ES}$ and ΔE are 100% correlated.)

For the modes $\eta'_{\rho\gamma}K^0_S$, $\eta'_{\rho\gamma}K^0_{S00}$, and $\eta'_{\rho\gamma}K^+$, we use four fit components: signal (sig), continuum $q\bar{q}$ background ($q\bar{q}$), charmless $B\overline{B}$ background (chls), and charm $B\overline{B}$ background (chrm). For the mode $\eta'_{\eta\pi\pi}K^0_L$, we use three fit components: signal, $q\bar{q}$ background, charmless $B\overline{B}$ background. For all other modes we use two components: signal and $q\bar{q}$ background.

As discussed in Sec. 4.6.3 there are six tagging categories plus the untagged events. We allow some parameters of the Δt PDF (parameters of the resolution function and tagging performance) to differ between tagging categories. For each event species j (signal, $q\bar{q}$, etc.) and each tagging category c, we define a PDF for event i as

$$\mathcal{P}^i_{j,c} = \mathcal{P}_j({m_{\rm ES}}^i) \cdot \mathcal{P}_j(\Delta E^i) \cdot \mathcal{P}_j(\mathcal{F}^i) \cdot \mathcal{P}_j(\Delta t^i, \sigma^i_{\Delta t}, c). \tag{5.1}$$

The extended likelihood is constructed, for each sub-decay, from these PDFs as

$$\begin{aligned} \mathcal{L} \;=\; \prod_{c=0}^{6} \frac{e^{-\nu_c}}{N_c!} \prod_i^{N_c} (&\nu_{\rm sig}\epsilon_{{\rm sig,c}}\mathcal{P}^i_{\rm sig,c} + \nu_{q\bar{q}}\epsilon_{q\bar{q},c}\mathcal{P}^i_{q\bar{q}} \\ &+\nu_{\rm chls}\epsilon_{{\rm sig},c}\mathcal{P}^i_{\rm chls} + \nu_{\rm chrm}\epsilon_{{\rm sig},c}\mathcal{P}^i_{\rm chrm}); \end{aligned} \tag{5.2}$$

where ν_j is the number of events in component j; N_c is the number of input events for each category; and $\epsilon_{j,c}$ is the fraction of events in component j that is in category c.

5.6 Probability density functions

As described in Sec. 4.5.3, one PDF must be determined for each discriminating variable for each fit component. We require as many as four variables ($m_{\rm ES}$, ΔE, $\mathcal{F}$, and Δt) for four components (signal, continuum background, and charmless and charm $B\overline{B}$ backgrounds). The data samples used for the determining the PDFs are:

- exclusive MC for all signal PDFs (with data control samples to determine necessary corrections to the MC),
- on-resonance data in sidebands chosen to avoid potential signal events to determine the starting values for all $q\bar{q}$ background PDFs, except for ΔE in $\eta' K^0_L$ modes,
- continuum MC for the $q\bar{q}$ ΔE PDF in $\eta' K^0_L$ modes,
- exclusive MC from charmless $B\overline{B}$ background modes for charmless PDFs,
- inclusive, generic $B\overline{B}$ MC for charm $B\overline{B}$ PDFs.

Appendix C shows PDFs plots and correlations between input variables. In all cases, the best candidate is chosen before the PDFs are determined.

5.6.1 $m_{\rm ES}$

We parameterize the $m_{\rm ES}$ distributions for signal with the sum of two Gaussians. From large control samples of B decays, we determine that the distribution (from MC) must be shifted to match the data because of slight variations of the beam energies, which are used to compute $m_{\rm ES}$. We correct for this effect during the reconstruction of the data.

For the continuum $m_{\rm ES}$ PDF for all modes except $\eta' K^0_L$, we use the on-resonance data in sidebands above and below the ΔE signal region ($100\,{\rm MeV} < |\Delta E| < 200\,{\rm MeV}$)

to obtain the starting value for the parameter of an ARGUS function [65] defined as

$$f(x) = x\sqrt{1-x^2}\exp\left[-\xi(1-x^2)\right], \tag{5.3}$$

with $x \equiv m_{\rm ES}/(\sqrt{s}/2)$ and ξ a free parameter. Since the results for these two samples are consistent, the sum of the two samples is used to determine the PDF parameters. We fix the ARGUS endpoint $\sqrt{s}/2$ at 5.2893 GeV in the PDF for continuum background as determined in fits to control samples. The charm and charmless $B\overline{B}$ components are fit with an ARGUS function and, where needed, an additional Gaussian.

5.6.2 ΔE

We use a double Gaussian shape to parameterize the signal and charmless $B\overline{B}$ distributions. Studies of control samples show that the MC is a reasonable representation of the data, but that the MC must be shifted by -2 MeV. For continuum and charm $B\overline{B}$ background we use a first or second order Chebyshev polynomial.

For $\eta' K^0_L$, we model the continuum background with an ARGUS-like function which has a low, rather than high, cutoff defined as

$$f(x) = x(1-x)^{-2}\exp\left[\xi' x\right], \tag{5.4}$$

where $x \equiv \Delta E - (\Delta E)_{\rm min}$, with $(\Delta E)_{\rm min}$ fixed to -0.01, and ξ' is a free parameter. The $\eta' K^0_L$ modes have this unique ΔE distribution because the B meson is mass-constrained during vertexing. (See Appendix D for more details.)

5.6.3 Fisher discriminant

For all components we parameterize the distribution of $\mathcal{F}$ with an asymmetric Gaussian; i.e., a Gaussian with different widths on the left and right sides of its peak. Modes with high background often exhibit outlier events which require an additional contribution to the continuum background PDF. We model this effect by adding another Gaussian, depending on the amount of background.

5.6.4 Δt

The Δt PDF for all components is the convolution of the time evolution and resolution functions (Eqs. 4.24 and 4.25)

$$\mathcal{P}(\Delta t, \sigma_\Delta t, c) = f'(\Delta t, c) \otimes \mathcal{R}(\Delta t, \sigma_{\Delta t}, c). \tag{5.5}$$

As described in Sec. 5.5, we allow the tagging performance parameters w, Δw, and μ to differ between tagging categories c, and we allow the mean and width of the core Gaussian of the resolution function for events in the `Lepton` category to differ from those in the other five tagging categories.

We take the values of parameters of the signal Δt resolution function from fits to the $B_{\rm flav}$ sample; those values were shown in Table 4.3. We also take signal tagging category efficiencies ϵ, mistag fractions w, mistag differences Δw, and tagging efficiency differences μ for each each tagging category from fits to the $B_{\rm flav}$ sample; those values were listed in Table 4.2. We fix Δm_d and the B lifetimes to the PDG values [62]: $\Delta m_d = 0.507 \pm 0.005$ ps^{-1}, $\tau_{B^+} = 1.638 \pm 0.011$ ps, and $\tau_{B^0} = 1.530 \pm 0.009$ ps in the signal Δt model.

For the $q\bar{q}$ background, we set Δm_d and the B lifetime equal to zero which reduces the time evolution equation to a δ-function. The convolution of the δ-function and $\mathcal{R}$ is just the resolution function. The parameters of $\mathcal{R}$ and the tagging category efficiencies are left free in the fit to data. The *CP* parameters are fixed to zero.

The parameterization of the Δt PDF for the $B\overline{B}$ backgrounds is very similar to that for signal: the parameters of the Δt resolution function and the tagging performance parameters ϵ, w, Δw, and μ are the same as signal, and we fix Δm_d and the B lifetimes to the PDG values [62]. The *CP* parameters, however, are fixed to zero as in the $q\bar{q}$ background. We describe the studies related to this choice in the section regarding systematic uncertainties (Sec. 6.14).

5.7 $B\overline{B}$ backgrounds

After having described the ML fit and the related PDFs, we describe in detail the methods for identifying $B\overline{B}$ backgrounds and for treating these backgrounds in the ML fit.

5.7.1 Charmless $B\overline{B}$ backgrounds

We perform a variety of studies related to backgrounds from B decays using MC samples. For all modes, we apply the full analysis selection to a sample of generic $B\overline{B}$ decays, having removed $b \to c$ and signal events, in order to identify the most troublesome charmless backgrounds. We use these studies to provide a list of modes for further study. We obtain exclusive MC samples of 100-600K events for each of the significant charmless $B\overline{B}$ background samples and mix these events in appropriate proportions according to branching fraction and efficiency. Where branching fraction measurements are unavailable we use use estimates from theory.

For $\eta'_{\eta\pi\pi}K^+$, $\eta'_{\eta\pi\pi}K^0_S$, $\eta'_{\eta\pi\pi}K^0_{S00}$, $\eta'_{5\pi}K^0_S$, and $\eta'_{5\pi}K^0_L$, we find the charmless $B\overline{B}$ backgrounds are negligible ($\lesssim 5$ events entering into the fit). We determine that charmless $B\overline{B}$ backgrounds are significant for the modes $\eta'_{\rho\gamma}K^0_S$, $\eta'_{\rho\gamma}K^0_{S00}$, $\eta'_{\eta\pi\pi}K^0_L$, and $\eta'_{\rho\gamma}K^+$. The contributions of the individual background modes to the total charmless $B\overline{B}$ backgrounds, as determined in this procedure, are shown in Tables 5.3, 5.4, 5.5, and 5.6 for modes $\eta'_{\rho\gamma}K^0_S$, $\eta'_{\rho\gamma}K^0_{S00}$, $\eta'_{\eta\pi\pi}K^0_L$, and $\eta'_{\rho\gamma}K^+$. For each background mode we show the decay channel, an identification number internal to *BABAR*, the detection efficiency, the estimated branching fraction, the daughter branching fractions, the number of events expected in the fit sample, and the number of events used in the cocktail from which we make the input PDFs.

Table 5.3: Potential background for the $B^0 \to \eta'_{\rho\gamma} K^0_S$ mode from exclusive charmless B decays. We show the efficiency for the mode to pass selection cuts, the measured or estimated branching fraction, the appropriate product (daughter) branching fraction, the estimated background normalized to 465.0 million $B\overline{B}$ events, and the number of events we include in the file used for making PDFs. An $*$ denotes an estimated branching fraction.

Bkg. channel	Mode #	MC ϵ (%)	Est. $\mathcal{B}$ (10^{-6})	$\prod \mathcal{B}_i$	Norm. # $B\overline{B}$ Bkg.	# in PDF Bkg. file
$B^+ \to\ a_1^+(\rho^0\pi^+)K^0$	4959	2.25	$34.9^{+6.7}_{-6.7}$	0.231	83.4	5290
$B^0 \to \pi^+\pi^- K^0$ (Dalitz)	6816	1.14	$44.8^{+2.6}_{-2.5}$	0.346	81.2	5145
$B^0 \to\ a_1^0(\rho^-\pi^+)K^0$	4955	2.08	15^*	0.231	33.1	2100
$B^+ \to \rho^+ K^0$	1933	0.82	$8.0^{+1.5}_{-1.4}$	0.500	15	950
$B^0 \to \phi_{3\pi} K^0$	2714	6.32	$8.3^{+1.2}_{-1.0}$	0.053	12.8	809
$B^+ \to\ a_1^+(\rho^+\pi^0)K^0$	4952	0.46	$34.9^{+6.7}_{-6.7}$	0.115	8.5	537
$B^+ \to \rho\ K_0^*(1430)$	6102	0.05	40^*	1.000	8.3	529
$B^+ \to \rho^0 K^{*+}_{K_S\pi^+}(L)$	2357	1.6	$3.6^{+1.9}_{-1.8}$	0.231	6.1	387
$B^0 \to\ K^+K^-K^0$ (Dalitz)	6814	0.13	$24.7^{+2.3}_{-2.3}$	0.346	5	316
$B^0 \to \rho\ K_0^*(1430)$	6104	0.05	20^*	1.000	4.6	292
$B^0 \to \omega\ K^0$	1536	0.61	$5.1^{+0.6}_{-0.6}$	0.308	4.4	279
$B^+ \to \eta'_{\rho\gamma} K^{*+}_{K_S\pi^+}$	2770	2.67	$4.9^{+2.1}_{-1.9}$	0.067	4	255
$B^0 \to \eta'_{\eta\pi\pi} K^0$	1510	0.17	$64.9^{+3.1}_{-3.1}$	0.060	3.1	196
$B^+ \to \pi^+ K^0$	991	0.06	$23.1^{+1.0}_{-1.0}$	0.346	2.3	144
$B^+ \to\ K^{*0}_{K_S\pi^0}\pi^+$	1595	0.23	$10.7^{+0.8}_{-0.8}$	0.167	1.9	121
$B^0 \to \rho^- K^{*+}_{K_S\pi^+}(L)$	2501	0.34	2^*	0.231	0.7	45
$B^0 \to\ K^0K^0$	1774	0.6	$0.96^{+0.21}_{-0.19}$	0.120	0.3	20
$B^+ \to \phi_{3\pi} K^{*+}_{K_S\pi^+}$	3994	0.19	$10.0^{+1.1}_{-1.1}$	0.035	0.3	19
$B^0 \to\ K^{*0}_{K^+\pi^-}\gamma$	1710	< 0.01	$40.1^{+2.0}_{-2.0}$	0.667	0.2	14
$B^0 \to \rho^- K^{*+}_{K_S\pi^+}(T)$	2502	0.03	4^*	0.231	0.1	8
$B^0 \to\ K^+K_SK_S$ $(N.R.)$	3915	0.02	$11.5^{+1.3}_{-1.3}$	0.120	0.1	7
Total					279.4	17463

Table 5.4: Potential background for the $\eta'_{\rho\gamma}K^0_{S00}$ mode from exclusive charmless B decays. Column headers are explained in Table 5.3. An $*$ denotes an estimated branching fraction.

Bkg. channel	Mode #	MC ϵ (%)	Est. $\mathcal{B}$ (10^{-6})	$\prod \mathcal{B}_i$	Norm. # $B\overline{B}$ Bkg.	# in PDF Bkg. file
$B^0 \to \pi^+\pi^- K^0_{\pi^0\pi^0}$ (Dalitz)	8176	0.65	$44.8^{+2.6}_{-2.5}$	0.154	20.4	274
$B^+ \to\ a_1^+ K^0$	4709	0.21	$34.9^{+6.7}_{-6.7}$	0.500	16.4	220
$B^+ \to \rho^+\rho^0(L)$	2390	0.08	$18.3^{+3.4}_{-3.4}$	0.96	6.4	85
$B^+ \to\ a_1^+(\rho^0\pi^+)\pi^0$	4799	0.09	$26.4^{+6.8}_{-6.8}$	0.500	5.2	69
$B^0 \to \rho^+\rho^-(L)$	2498	0.04	$24.2^{+3.1}_{-3.2}$	0.96	4.5	60
$B^+ \to \rho^+ K^0$	1933	0.23	$8.0^{+1.5}_{-1.4}$	0.500	4.3	57
$B^+ \to \eta'_{\rho\gamma}\rho^+$	2775	0.31	$9.1^{+3.7}_{-2.8}$	0.295	3.8	50
$B^0 \to\ a_1^+\pi^-$	4157	0.02	$31.7^{+3.7}_{-3.7}$	1.000	2.2	30
$B^+ \to\ a_1^+(\rho^+\pi^0)\pi^0$	4957	0.03	$26.4^{+6.8}_{-6.8}$	0.500	1.6	21
$B^0 \to \rho^0\pi^0\pi^0$ $(N.R.)$	3590	0.06	5.0^*	1.000	1.3	17
$B^+ \to \pi^+\pi^0\pi^0$ $(N.R.)$	1938	0.01	10^*	1.000	0.5	6
$B^0 \to\ K^{*0}_{K^+\pi^-}K^0$	1956	0.23	$0.95^{+0.95}_{-0.95}$	0.333	0.3	4
$B^0 \to \pi^0\pi^0\ K^{*0}_{K^+\pi^-}$	5297	0.01	10^*	0.667	0.2	3
$B^+ \to \rho^-\pi^0 K^+$	2490	< 0.01	10^*	1.000	0.2	2
$B^+ \to\ K^{*+}_{K_S\pi^+}K^0$	1944	0.14	0.95^*	0.167	0.1	1
$B^+ \to\ b_1^+(\omega\pi^+)\pi^0$	7951	0.02	$1.0^{+0.9}_{-0.9}$	0.891	0.1	0
$B^0 \to \pi^+\pi^- K^0$ (Dalitz)	6816	< 0.01	$44.8^{+2.6}_{-2.5}$	0.346	0	0
Total					68.5	899

Table 5.5: Potential background for the $B^0 \to \eta'_{\eta\pi\pi}K^0_L$ mode from exclusive charmless B decays. Column headers are explained in Table 5.3. An $*$ denotes an estimated branching fraction.

Bkg. channel	Mode #	MC ϵ (%)	Est. $\mathcal{B}$ (10^{-6})	$\prod \mathcal{B}_i$	Norm. # $B\overline{B}$ Bkg.	# in PDF Bkg. file
$B^0 \to \eta'_{\eta\pi\pi}K^0$	1510	0.76	$64.9^{+3.1}_{-3.1}$	0.060	13.6	718
$B^0 \to \eta'_{\eta\pi\pi}K^0_S(\pi^0\pi^0)$	4116	0.59	$64.9^{+3.1}_{-3.1}$	0.027	4.8	251
$B^+ \to \eta'_{\eta\pi\pi}K^+$	1506	0.04	$70.2^{+2.5}_{-2.5}$	0.174	2	107
$B \to \eta'\ell\nu$	4760	< 0.01	80^*	1.000	1.6	84
Total					22	1160

Table 5.6: Potential background for the $B^+ \to \eta'_{\rho\gamma} K^+$ mode from exclusive charmless B decays. Column headers are explained in Table 5.3. An $*$ denotes an estimated branching fraction.

Bkg. channel	Mode #	MC ϵ (%)	Est. $\mathcal{B}$ (10^{-6})	$\prod \mathcal{B}_i$	Norm. # $B\overline{B}$ Bkg.	# in PDF Bkg. file
$B^+ \to\ K^+\pi^-\pi^+$ (Dalitz)	6846	1.17	$54.8^{+2.9}_{-2.9}$	1.000	243.4	6247
$B^+ \to\ a_1^0 K^+$	4874	2.12	20^*	1.000	160.7	4125
$B^0 \to\ a_1^-(\rho^0\pi^-)K^+$	4871	2.27	$16.3^{+3.7}_{-3.7}$	0.500	70.1	1800
$B^0 \to \rho^- K^+$	1044	1.49	$9.9^{+1.6}_{-1.5}$	1.000	56.2	1442
$B^+ \to \rho\ K^*(1430)$	6102	0.15	40^*	1.000	23.1	594
$B^+ \to \phi_{3\pi} K^+$	2713	6.45	$5.4^{+1.2}_{-1.2}$	0.155	20.5	526
$B^+ \to\ K^+K^-K^+$ (Dalitz)	6845	0.13	$32.5^{+1.5}_{-1.5}$	1.000	16.5	423
$B^+ \to \omega\ K^+$	1250	0.62	$6.9^{+0.5}_{-0.5}$	0.891	14.5	372
$B^0 \to\ a_1^-(\rho^-\pi^0)K^+$	4960	0.45	$16.3^{+3.7}_{-3.7}$	0.500	13.8	354
$B^0 \to \rho^0 K^{*0}_{K^+\pi^-}(L)$	2359	1.73	$2.8^{+0.8}_{-0.8}$	0.667	12.3	315
$B^+ \to \rho^0\pi^+$	1220	0.36	$8.7^{+1.0}_{-1.1}$	1.000	12	308
$B^+ \to\ a_1^0\pi^+$	4156	0.15	$20.4^{+5.8}_{-5.8}$	1.000	11.9	305
$B^0 \to \rho\ K^*(1430)$	6104	0.15	20^*	1.000	11.5	296
$B^+ \to \eta'_{\rho\gamma}\pi^+$	1509	2.3	$4.4^{+0.4}_{-0.4}$	0.295	11.3	290
$B^0 \to \pi^- K_0^{*+}(1430)_{K^+\pi^0}$	4697	0.19	$49.7^{+7.8}_{-9.0}$	0.310	10.8	278
$B^0 \to\ a_1^+(\rho^0\pi^+)\pi^-$	1012	0.17	$31.7^{+3.7}_{-3.7}$	0.500	10.1	259
$B^+ \to \rho^0 K^{*+}_{K^+\pi^0}(L)$	2355	1.83	$3.6^{+1.9}_{-1.8}$	0.333	8.3	214
$B^0 \to\ K^+\pi^-$	1028	0.11	$19.4^{+0.6}_{-0.6}$	1.000	8.3	213
$B^0 \to \eta'_{\rho\gamma} K^{*0}_{K^+\pi^-}$	2268	2.89	$3.8^{+1.2}_{-1.2}$	0.197	8.2	210
$B^+ \to \eta'_{\eta\pi\pi} K^+$	1506	0.16	$70.2^{+2.5}_{-2.5}$	0.174	7.6	194
$B^+ \to \rho^+\rho^0(L)$	2390	0.11	$18.3^{+3.4}_{-3.4}$	0.96	7.5	191
$B^0 \to\ K^{*+}_{K^+\pi^0}\pi^-$	1226	0.47	$9.8^{+1.1}_{-1.1}$	0.333	5.8	149
$B^+ \to \eta'_{\rho\gamma} K^{*+}_{K^+\pi^0}$	2773	3.1	$4.9^{+2.1}_{-1.9}$	0.098	5.7	145
$B^+ \to \rho^+\ K^{*0}_{K^+\pi^-}(L)$	2244	0.33	$4.6^{+0.75}_{-0.75}$	0.666	3.8	97
$B^+ \to \pi^+\pi^+\pi^-$ $(N.R.)$	1230	0.27	$3.0^{+3.0}_{-3.0}$	1.000	3	77
$B^0 \to \rho^+\rho^-(L)$	2498	0.03	$24.2^{+3.1}_{-3.2}$	0.96	2.6	67
$B^0 \to \rho^0 K^{*0}_{K^+\pi^-}(T)$	2360	0.28	$2.8^{+0.8}_{-0.8}$	0.667	2	51
$B^0 \to\ f_0 K^{*0}_{K^+\pi^-}$	3359	0.55	$2.15^{+2.15}_{-2.15}$	0.444	2	50
$B^0 \to\ a_1^0 K^{*0}(L)$	5329	0.09	7^*	0.667	1.6	41
$B^+ \to \omega\pi^+$	1248	0.06	$6.7^{+0.6}_{-0.6}$	0.891	1.4	36
$B^+ \to\ K^+\pi^0$	1587	0.03	$12.9^{+0.6}_{-0.6}$	1.000	1.4	36
$B^0 \to \pi^+\pi^- K^0$ (Dalitz)	6816	0.02	$44.8^{+2.6}_{-2.5}$	0.343	0.9	22
$B^0 \to \rho^- K^{*+}_{K^+\pi^0}(L)$	2499	0.35	2^*	0.333	0.9	22
$B^+ \to \rho^-\pi^+\pi^+$	4151	0.02	5^*	1.000	0.4	9
$B^+ \to \rho^+\ K^{*0}_{K^+\pi^-}(T)$	2243	0.03	$4.6^{+0.75}_{-0.75}$	0.666	0.3	8
$B^0 \to \rho^- K^{*+}_{K^+\pi^0}(T)$	2500	0.03	4^*	0.333	0.2	4
Total					770.6	19770

5.7.2 Charm $B\overline{B}$ backgrounds

In earlier *BABAR* analyses, we observed that PDFs for charm $B\overline{B}$ background and continuum events were very similar and that charm and continuum yields were highly correlated. For this reason, we did not use separate charm components in earlier fits, allowing the charm events to be absorbed into the yield of the dominant continuum background. In this analysis, the fit seems to be behaving differently from this expectation. While this does not necessarily bias the results, we attempt to understand this effect.

Because the branching fractions of the charmless $B\overline{B}$ background modes are known with a precision of $10-15\%$ (Tables 5.3–5.6), we are confident in the reasonableness of our estimation of charmless $B\overline{B}$ events entering each fit. Nevertheless, the charmless $B\overline{B}$ yields were observed to float $1.6-5.0$ times higher than expected in fits to the Run1-5 subset of the data. In embedded toy MC studies, we find that the fitted charmless yield is actually smaller than the number of embedded events, revealing that the higher-than-expected charmless yield in fits to Run1-5 data is not a fit bias, but the result of an unaccounted for component of the data, either charm or additional charmless $B\overline{B}$ events.

The crosscheck mode $\eta'_{\rho\gamma}K^+$ with ~ 3200 signal events is useful for investigating this issue. In Fig. 5.3, we show $s\mathcal{P}lots$ of $m_{\rm ES}$ and Δt for the charmless $B\overline{B}$ component from fits with no charm component to Run1-5 data for the mode $\eta'_{\rho\gamma}K^+$. The distribution of data in the charmless $m_{\rm ES}$ $s\mathcal{P}lot$ suggests that non-peaking events (charm or $q\bar{q}$) are being pulled into the charmless component. The broad Δt shape indicates that these non-peaking events are charmed B decays rather than combinatorial continuum events, for which Δt is narrow since there is no particle decaying with a certain lifetime. The $s\mathcal{P}lots$ from $\eta'_{\rho\gamma}K^0_S$ are similar, though with fewer events. Finally, in fits for the yields only (no Δt information), we find that there is no enhancement of the charmless

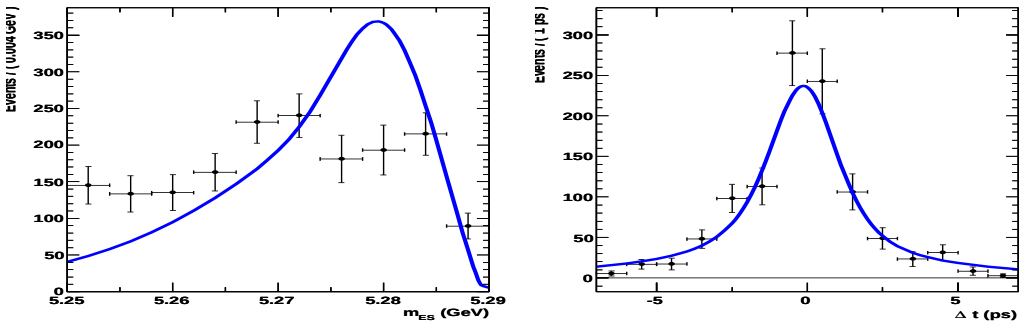


Figure 5.3: $s\mathcal{P}lots$ for $m_{\rm ES}$ (*left*) and Δt (*right*) in the charmless $B\overline{B}$ component from fits (with no charm $B\overline{B}$ component) to Run1-5 data for $\eta'_{\rho\gamma}K^+$.

$B\overline{B}$ yields.

All of this indicates that non-peaking charmed $B\overline{B}$ events are being absorbed into the charmless $B\overline{B}$ component, rather than the $q\bar{q}$ component, because of their broad Δt distribution. Since the charmless $B\overline{B}$ component exhibits higher correlations with the signal than does the $q\bar{q}$ component, we prefer not to let the charm events feed into the charmless component even though we are comfortable allowing them to feed into the $q\bar{q}$ component. Thus, we investigate the use of a charm $B\overline{B}$ fit component for modes in which we observe the enhancement of the charmless yield: $\eta'_{\rho\gamma}K^+$, $\eta'_{\rho\gamma}K^0_{S00}$, and $\eta'_{\rho\gamma}K^0_S$.

5.7.3 Charm $B\overline{B}$ component in the fit

For the $\eta' \to \rho\gamma$ modes, we perform fits (with charmless and charm components) to Run1-5 data fixing and floating the charmless and charm yields. We obtain PDFs (Figs. C.1, C.3, and C.6 in Appendix C) from the charm events in generic $B\overline{B}$ MC that pass the selection for each mode. The final charm MC samples are ~ 1600 $\eta'_{\rho\gamma}K^0_S$ events, ~ 1200 $\eta'_{\rho\gamma}K^0_{S00}$ events, and ~ 6400 $\eta'_{\rho\gamma}K^+$ events; with the number of generated events we can compute the number of charm events in the data that enter the fit.

The results of these fits are in Tables 5.7, 5.8, and 5.9. The expected numbers of charm and charmless events are listed in the captions. We show $m_{\rm ES}$ and Δt $s\mathcal{P}lots$

from the $\eta'_{\rho\gamma}K^+$ fit (with charm) to Run1-5 data in Fig. 5.4 for comparison with *sPlots* from the fit without charm referred to above. In the *sPlots* from the fit with a charm component, the data are in better agreement with the overlays of the PDFs indicating that events are being correctly classified when the charm component is used.

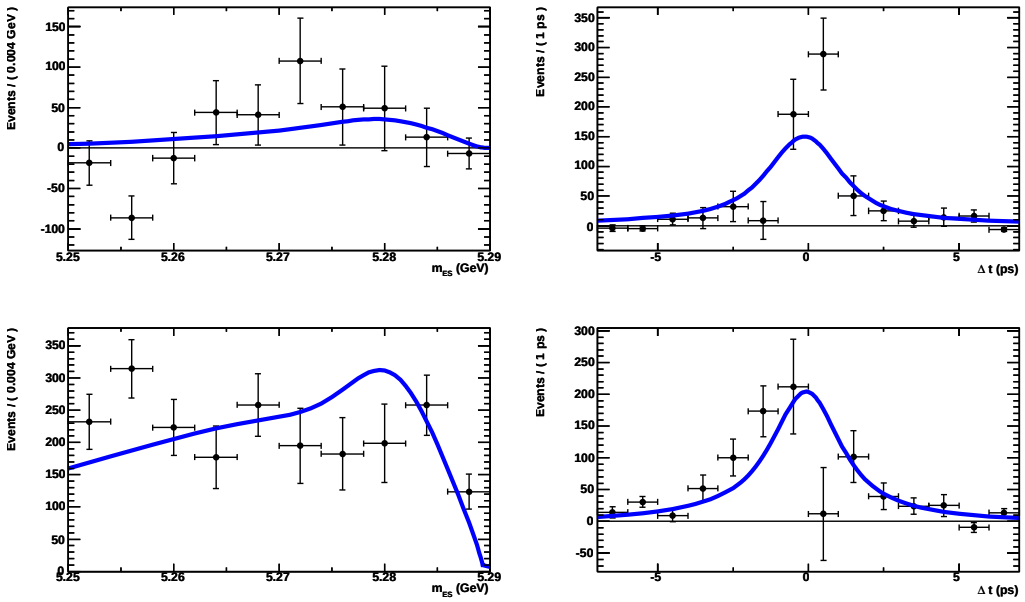


Figure 5.4: *sPlots* for $m_{\rm ES}$ (*left*) and Δt (*right*) in the charmless (*top*) and charm (*bottom*) $B\overline{B}$ components from fits (with a charm $B\overline{B}$ component) to Run1-5 data for $\eta'_{\rho\gamma}K^+$.

5.7.4 Fixing/floating charmless and charm yields

Because the charmless and charm yields float to reasonable values in fits to Run1-5 data and the Run1-5 *sPlots* confirm that the events are being categorized correctly, we use a charm component in the $\eta'_{\rho\gamma}K^0_S$, $\eta'_{\rho\gamma}K^0_{S00}$, and $\eta'_{\rho\gamma}K^+$ fits to the full dataset. Finding zero failed fits in 200 embedded toy experiments, we determine that the fit is stable with both charmless and charm yields floating; however, we choose to use our detailed knowledge of the make-up of the charmless $B\overline{B}$ component to fix the charmless yield at the expected number while allowing the charm yield to float. By using as much

information as available, we will measure S and C as accurately as possible.

Table 5.7: Comparison of Run1-5 results for the mode $\eta'_{\rho\gamma}K^0_S$ from fits with and without a charm component and with different treatments of the charmless (`nChls`) and charm (`nChrm`) yields. We expect 230 charmless and 444 charm events.

ML fit quantity	Flt `nChls` No `nChrm`	Fix `nChls` No `nChrm`	Fix `nChls` Flt `nChrm`	Flt `nChls` Flt `nChrm`
Events into fit	18464	18464	18464	18464
Signal yield	825 ± 36.2	840.1 ± 35.9	823.5 ± 35.9	830.2 ± 36.4
Chls yield	364.4 ± 54.6	230	230	111.1 ± 85.9
Charm yield	–	–	214.7 ± 59.7	319.7 ± 99.5
S	0.45 ± 0.13	0.44 ± 0.13	0.45 ± 0.13	0.45 ± 0.13
C	-0.09 ± 0.10	-0.09 ± 0.10	-0.09 ± 0.1	-0.09 ± 0.10

Table 5.8: Comparison of Run1-5 results for the mode $\eta'_{\rho\gamma}K^0_{S00}$ from fits with and without a charm component and with different treatments of the charmless (`nChls`) and charm (`nChrm`) yields. We expect 57 charmless and 390 charm events.

ML fit quantity	Flt `nChls` No `nChrm`	Fix `nChls` No `nChrm`	Fix `nChls` Flt `nChrm`	Flt `nChls` Flt `nChrm`
Events into fit	21643	21643	21643	21643
Signal yield	151.6 ± 27.7	227.9 ± 25.9	181.2 ± 25.7	164.7 ± 27.9
Chls yield	364.9 ± 61.8	57	57	164.5 ± 77.0
Charm yield	–	–	436.7 ± 77.3	345.1 ± 99.4
S	0.48 ± 0.43	0.28 ± 0.28	0.32 ± 0.36	0.37 ± 0.40
C	0.02 ± 0.32	0.01 ± 0.22	-0.02 ± 0.27	-0.01 ± 0.30

5.8 Fit validation

We use toy MC experiments to check for fit stability, bias on S or C, and the reasonableness of uncertainties reported by the fit. We conclude the fit is stable because we observe no failed fits in pure or embedded toy studies for all modes. In these toy studies, we fix the parameters of the signal Δt PDF to values obtained from fully-simulated MC, rather than the $B_{\rm flav}$ sample, to separate the fit bias from systematic uncertainties related to our use of the $B_{\rm flav}$ sample.

Table 5.9: Comparison of Run1-5 results for the mode $\eta'_{\rho\gamma}K^+$ from fits with and without a charm component and with different treatments of the charmless (`nChls`) and charm (`nChrm`) yields. We expect 771 charmless events and 1770 charm events.

ML fit quantity	Flt `nChls` No `nChrm`	Fix `nChls` No `nChrm`	Fix `nChls` Flt `nChrm`	Flt `nChls` Flt `nChrm`
Events into fit	74166	74166	74164	74164
Signal yield	2561.1 ± 65.8	2696.8 ± 65.5	2556.1 ± 65.1	2590 ± 65.7
Chls yield	1647.3 ± 92.1	771	771	213.3 ± 112
Charm yield	–	–	1568.3 ± 109	2057.7 ± 153
S	-0.07 ± 0.07	-0.07 ± 0.07	-0.06 ± 0.07	-0.06 ± 0.07
C	0.01 ± 0.06	0.01 ± 0.05	0.01 ± 0.06	0.01 ± 0.06

5.8.1 Pure toy experiments

We construct 200 pure toy samples per sub-mode, each with the number of events in the actual data, generating signal, background, and $B\overline{B}$ events according to the PDFs shown in Appendix C with S and C from the values obtained in fits of the Run1-5 data. We summarize the pulls from fits of these samples (defined in Sec. 4.5.4.1) for S, C, and signal yield in Table 5.10.

5.8.2 Embedded toy experiments

We perform 140-550 MC toy experiments per sub-mode depending on the number of available signal MC events. For each sub-mode we perform the maximum number of experiments without oversampling the signal MC, so as not to artificially reduce the uncertainty on the biases. Signal and $B\overline{B}$ events are taken from MC, while background and charm events are generated according to the PDFs shown in Appendix C. We generate 279 charm events for $\eta'_{\rho\gamma}K^0_{S00}$ and 539 charm events for $\eta'_{\rho\gamma}K^0_S$.

We summarize in Table 5.11 the number of embedded signal events, embedded charmless $B\overline{B}$ background events, the bias on the signal yield, the biases on S and C for the individual sub-modes, the weighted average of the sub-modes, and the simultaneous

Table 5.10: Summary of S, C, and yield (nSig) pulls from pure toy experiments for all modes.

$\eta'_{\eta\pi\pi}K^0_S$ (200 Experiments)

	mean	sigma
dtSig_S	0.00 +/- 0.07	1.04 +/- 0.05
dtSig_C	-0.09 +/- 0.07	1.01 +/- 0.05
nSig	-0.01 +/- 0.07	0.92 +/- 0.05

$\eta'_{\rho\gamma}K^0_S$ (200 Experiments)

	mean	sigma
dtSig_S	-0.08 +/- 0.07	1.03 +/- 0.05
dtSig_C	0.00 +/- 0.07	1.04 +/- 0.05
nSig	-0.07 +/- 0.07	0.96 +/- 0.05

$\eta'_{\eta\pi\pi}K^0_{S00}$ (200 Experiments)

	mean	sigma
dtSig_S	-0.09 +/- 0.08	1.09 +/- 0.05
dtSig_C	-0.02 +/- 0.08	1.10 +/- 0.06
nSig	-0.17 +/- 0.07	1.04 +/- 0.05

$\eta'_{\rho\gamma}K^0_{S00}$ (100 Experiments)

	mean	sigma
dtSig_S	0.06 +/- 0.10	0.99 +/- 0.07
dtSig_C	0.01 +/- 0.09	0.91 +/- 0.07
nSig	0.03 +/- 0.10	0.94 +/- 0.07

$\eta'_{5\pi}K^0_S$ (100 Experiments)

	mean	sigma
dtSig_S	0.10 +/- 0.11	1.12 +/- 0.08
dtSig_C	0.08 +/- 0.12	1.25 +/- 0.09
nSig	-0.03 +/- 0.11	1.13 +/- 0.08

$\eta'_{\eta\pi\pi}K^0_L$ (200 Experiments)

	mean	sigma
dtSig_S	-0.03 +/- 0.08	1.08 +/- 0.05
dtSig_C	-0.04 +/- 0.07	0.99 +/- 0.05
nSig	0.04 +/- 0.07	1.02 +/- 0.05

$\eta'_{5\pi}K^0_L$ (200 Experiments)

	mean	sigma
dtSig_S	-0.05 +/- 0.09	1.24 +/- 0.07
dtSig_C	0.03 +/- 0.08	1.08 +/- 0.06
nSig	-0.36 +/- 0.09	1.18 +/- 0.06

fit. The individual sub-mode toys and the simultaneous-fit toys were performed using independent toy datasets, so the uncertainties should be considered when comparing results. In Table 5.12, we report the means and RMS of the S and C error distributions along with the RMS of the S and C distributions for the individual sub-mode fits and the simultaneous fit. We find that the mean of the S (C) error distribution agrees with the RMS of the S (C) distribution, indicating that the errors reported by the fit are reasonable.

5.8.2.1 Averaged sub-mode toy results and small-sample bias

As mentioned above, along with results obtained from simultaneous-fit toy experiments, we present weighted averages of sub-mode toy results in Tables 5.11 and 5.12. One advantage of using the biases obtained from sub-mode toy experiments is that, in principle, one can reduce the width of the combined likelihood function, and thus the statistical uncertainty on the result, by correcting for the bias mode-by-mode before averaging sub-mode results. Another advantage is that one can obtain a more precise estimate of the bias from sub-mode toys because, while one runs into oversampling issues near $200-300$ embedded toy experiments for the simultaneous fit, one can run 550 $\eta'_{\eta\pi\pi}K^0_S$ sub-mode toy experiments, 260 $\eta'_{\rho\gamma}K^0_S$ experiments, etc.

However, in principle, averaging sub-mode toy results also propagates any small-sample bias, potentially present in the minor-modes, into the average bias. Any bias measured in simultaneous-fit toy experiments, with over 2400 events, is not expected to result from small sample effects. To confirm that the small sample contribution to the averaged bias is negligible, we separate each of the 175 simultaneous-fit toy datasets into 7 reduced toy datasets, one for each sub-mode, and perform the appropriate sub-mode fit on each of these 7 sets of 175 toy datasets. We take the weighted average of the sub-mode S and C from these 7 sets of 175 toy experiments. We find that the averaged S and C differ by just 0.002 from the S and C obtained from the simultaneous fit. We

Table 5.11: Biases on *CP* parameters and their errors for 140-550 embedded toy MC experiments. (The number of experiments is the maximum without oversampling the signal MC.) We report the number of experiments, embedded signal, embedded charmless $B\overline{B}$ background, the bias on the signal yield, the bias on S, and the bias on C from individual sub-mode fits. We also report the averaged sub-mode results and the results from 175 toy experiments for the simultaneous fit. The signal yield bias is from time-dependent fits. The embedded MC is generated with the values $S = 0.703$ and $C = 0.000$.

Final state	# toy experiments	# signal input	# $B\overline{B}$ input	Signal Bias	Bias on S	Bias on C
$\eta'_{\eta\pi\pi}K^0_S$	550	470	0	-3.7 ± 0.5	0.010 ± 0.007	-0.012 ± 0.005
$\eta'_{\rho\gamma}K^0_S$	260	970	279	35.1 ± 1.4	0.002 ± 0.007	-0.007 ± 0.006
$\eta'_{\eta\pi\pi}K^0_{S00}$	200	108	0	-3.0 ± 0.6	0.080 ± 0.034	0.022 ± 0.020
$\eta'_{\rho\gamma}K^0_{S00}$	270	199	69	7.3 ± 1.3	0.054 ± 0.022	-0.007 ± 0.016
$\eta'_{5\pi}K^0_S$	190	173	0	-1.8 ± 0.4	0.021 ± 0.020	-0.009 ± 0.014
$\eta'_{\eta\pi\pi}K^0_L$	235	353	22	-25.3 ± 1.7	-0.019 ± 0.018	-0.007 ± 0.013
$\eta'_{5\pi}K^0_L$	140	170	0	-9.5 ± 1.5	0.099 ± 0.035	0.007 ± 0.022
Weighted Avg.					0.013 ± 0.004	-0.007 ± 0.003
Simultaneous $\eta' K^0$ Fit	175				0.006 ± 0.006	-0.008 ± 0.006
Simultaneous $\eta' K^0_S$ Fit	175				0.008 ± 0.006	-0.007 ± 0.005
Simultaneous $\eta' K^0_L$ Fit	175				0.003 ± 0.014	-0.006 ± 0.011

Table 5.12: Mean values of the error distributions for S and C and the RMS of the S and C distributions for each sub-mode from 140-550 embedded toy MC experiments. (The number of experiments is the maximum without oversampling the signal MC.) We also report the averaged sub-mode results and the results from 175 toy experiments for the simultaneous fit. All results are reported in units of 10^{-3}.

Final state	Mean of S Error	RMS of S Error	RMS of S dist.	Mean of C Error	RMS of C Error	RMS of C dist.
$\eta'_{\eta\pi\pi}K^0_S$	156 ± 1	14 ± 1	167 ± 8	115 ± 1	5 ± 1	121 ± 6
$\eta'_{\rho\gamma}K^0_S$	112 ± 1	8 ± 1	114 ± 7	88 ± 1	3 ± 1	97 ± 6
$\eta'_{\eta\pi\pi}K^0_{S00}$	381 ± 5	75 ± 5	481 ± 34	267 ± 2	31 ± 2	280 ± 20
$\eta'_{\rho\gamma}K^0_{S00}$	325 ± 3	53 ± 3	357 ± 22	258 ± 2	36 ± 2	267 ± 16
$\eta'_{5\pi}K^0_S$	246 ± 3	38 ± 3	280 ± 20	180 ± 1	14 ± 1	192 ± 14
$\eta'_{\eta\pi\pi}K^0_L$	271 ± 2	32 ± 2	278 ± 18	199 ± 1	15 ± 1	201 ± 13
$\eta'_{5\pi}K^0_L$	358 ± 6	70 ± 6	416 ± 35	265 ± 3	32 ± 3	263 ± 22
Weighted Average :	76 ± 1	7 ± 1	79 ± 5	57 ± 1	2 ± 1	61 ± 4
Simultaneous Fit:	75 ± 1	3 ± 1	75 ± 6	58 ± 1	2 ± 1	61 ± 6

conclude that the additional bias from small samples is negligible and does not preclude correcting for the bias mode-by-mode.

5.9 Fit results

We present fit results for the $\eta' K^0_S$ modes in Table 5.13, the $\eta' K^0_L$ modes in Table 5.14, and the charged mode crosschecks in Table 5.15. Bias corrections are only applied for the corrected combined results at the bottom of each table. As a crosscheck of the time-dependent (TD) fits, we also measure the branching fractions in *yield-only* fits. Since these yield-only fits are only crosschecks, we make these approximations in converting the yield to a branching fraction: the biases on the yields are obtained from toy studies with the TD fits, and we have applied an estimated relative MC efficiency correction of 0.9 for K^0_L reconstruction.

5.9.1 Signal-enhanced projection plots

In Fig. 5.5 we show $m_{\rm ES}$ and ΔE projection plots for all five $\eta' K^0_S$ sub-modes. In Fig. 5.6 we show ΔE projection plots for the two $\eta' K^0_L$ sub-modes. In Figs. 5.7 and 5.8 we show the Δt projections and asymmetry for the combined $\eta' K^0_S$ sub-modes and combined $\eta' K^0_L$ sub-modes. In all projection plots, the background has been reduced with a cut on the likelihood ratio (computed without the plotted variable), which was optimized separately for each sub-mode and projection, as described in Sec. 4.5.5.

5.9.2 *s$\mathcal{P}$lots*

In Appendix E we show the *s$\mathcal{P}$lots* (Sec. 4.5.5) for all modes. To generate *s$\mathcal{P}$lots* for all fit components in the $\eta'_{\rho\gamma} K^0_S$, $\eta'_{\rho\gamma} K^+$, and $\eta'_{\rho\gamma} K^0_{S00}$ analyses, we float the charmless $B\overline{B}$ yields even though they are fixed in the final fit.

Table 5.13: Results for the $\eta' K^0_S$ sub-modes. We show uncorrected and bias-corrected combined results.

	$\eta'_{\eta\pi\pi}K^0_S$	$\eta'_{\rho\gamma}K^0_S$	$\eta'_{\eta\pi\pi}K^0_{S00}$	$\eta'_{\rho\gamma}K^0_{S00}$	$\eta'_{5\pi}K^0_S$
Yield-only fits:					
Events to fit	1556	23905	1175	28871	546
Signal yield	495.9 ± 24.4	1050.1 ± 42.2	107.7 ± 14.2	197.9 ± 31.9	183.4 ± 14.8
Chls yield	–	279	–	69	–
Chrm yield	–	-100 ± 22.2	–	-100 ± 27	–
MC ϵ (%)	26.8	28.5	13.5	14.7	17.1
Fit bias (evts)	-4.6	32	-3.4	6.6	-1.7
$\mathcal{B}(10^{-6})$	66.5 ± 3.2	75.8 ± 3.1	65.1 ± 8.3	61.4 ± 10.2	66.9 ± 5.4
TD fits:					
Events to fit	1470	22775	1056	27057	513
Signal yield	471.5 ± 23.6	1004.7 ± 39.7	105.3 ± 13	205.8 ± 27.6	170.9 ± 14.1
Chls yield	–	279	–	69	–
Chrm yield	–	253.4 ± 67.3	–	530.3 ± 84.2	–
$-\eta$S	0.70 ± 0.17	0.46 ± 0.12	0.51 ± 0.34	0.26 ± 0.33	0.76 ± 0.26
C	-0.17 ± 0.11	-0.13 ± 0.09	-0.19 ± 0.30	0.04 ± 0.26	0.05 ± 0.20
Combined:					
$-\eta$S (uncorr.)			0.537 ± 0.084		
C (uncorr.)			-0.118 ± 0.062		
$-\eta$S (corr.)			0.529 ± 0.084		
C (corr.)			-0.111 ± 0.062		

Table 5.14: Results for the $\eta' K_L^0$ sub-modes. We show uncorrected and bias-corrected combined results.

	$\eta'_{\eta\pi\pi} K_L^0$	$\eta'_{5\pi} K_L^0$
Yield-only fits:		
Events into fit	14125	4951
Signal yield	345.4 ± 37.9	174.7 ± 24
Chls yield	22	–
Chrm yield	–	–
MC ϵ (%)	17.5	11.1
Fit bias (events)	-27.7	-11.1
$\mathcal{B}(10^{-6})$	52.3 ± 5.3	71.3 ± 9.2
TD fits:		
Events into fit	12217	4586
Signal yield	341 ± 32	158.7 ± 21.6
Chls yield	22	–
Chrm yield	–	–
$-\eta$S	0.65 ± 0.22	0.66 ± 0.46
C	0.07 ± 0.19	0.02 ± 0.26
Combined:		
$-\eta$S (uncorr.)	0.642 ± 0.198	
C (uncorr.)	0.047 ± 0.154	
$-\eta$S (corr.)	0.639 ± 0.198	
C (corr.)	0.053 ± 0.154	

Table 5.15: Results for the charged sub-modes.

	$\eta'_{\eta\pi\pi} K^+$	$\eta'_{\rho\gamma} K^+$
Yield-only fits:		
Events into fit	4386	93807
Signal yield	1465.8 ± 41.6	3215.1 ± 105.0
Chls yield	–	944
Chrm yield	–	82.4 ± 422.0
MC ϵ (%)	26.4	29.6
TD fits:		
Events into fit	4260	90759
Signal yield	1440.8 ± 40.9	3067.6 ± 76.5
Chls yield	–	944
Chrm yield	–	1275.3 ± 115.0
S	-0.05 ± 0.09	-0.09 ± 0.06
C	0.08 ± 0.06	0.01 ± 0.05

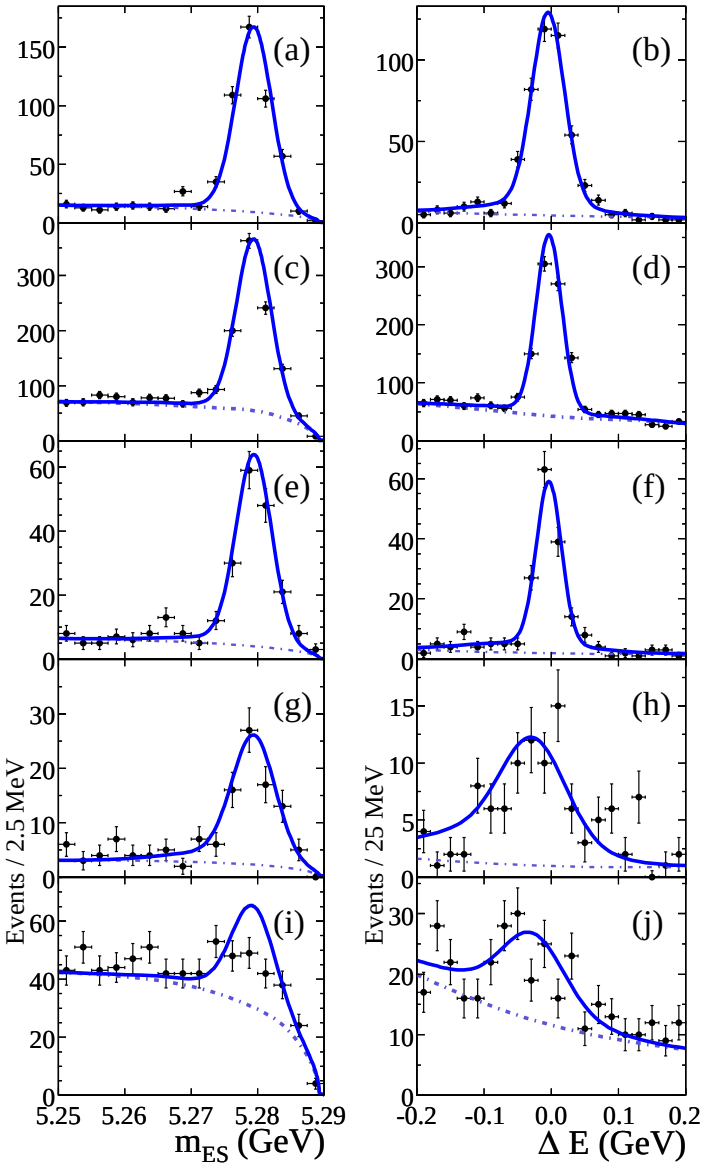


Figure 5.5: Distributions projected onto ($m_{\rm ES}$,ΔE) for $\eta'_{\eta_{(\gamma\gamma)}\pi\pi}K^0_{\pi^+\pi^-}$ (a,b), $\eta'_{\rho\gamma}K^0_{\pi^+\pi^-}$ (c,d), $\eta'_{\eta_{(3\pi)}\pi\pi}K^0_{\pi^+\pi^-}$ (e,f), $\eta'_{\eta_{(\gamma\gamma)}\pi\pi}K^0_{\pi^0\pi^0}$ (g,h), $\eta'_{\rho\gamma}K^0_{\pi^0\pi^0}$ (i,j). The solid lines show the full fit result and the dashed lines show full the background contributions.

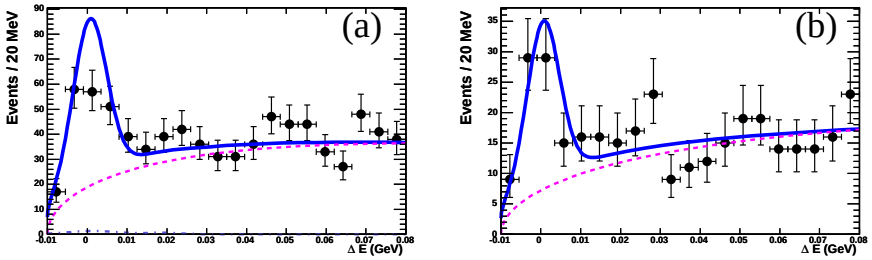


Figure 5.6: Distributions projected onto ΔE for $\eta'_{\eta_{(\gamma\gamma)}\pi\pi}K^0_L$ (a) and $\eta'_{\eta_{(3\pi)}\pi\pi}K^0_L$ (b). The solid lines show the full fit result and the dashed lines show full the background contributions.

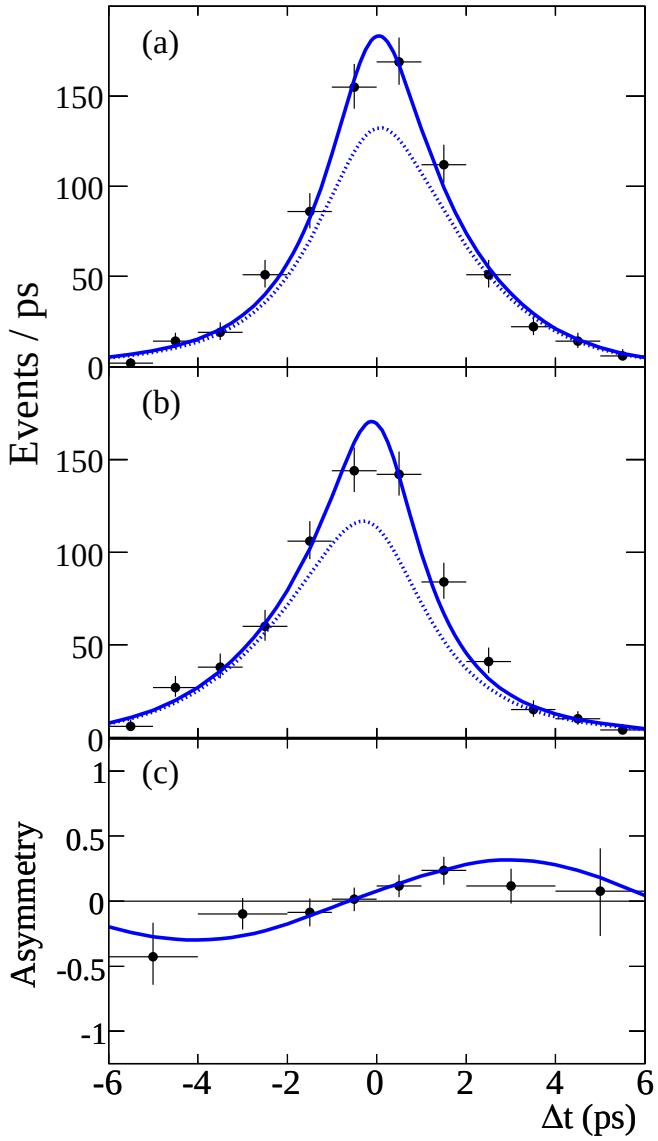


Figure 5.7: Projections onto Δt for data for the five $\eta' K^0_S$ sub-modes combined (points with errors), the fit full function (blue solid line), and the signal only function (blue dashed line). We show (a) B^0 and (b) $\overline{B}^0$ tagged events, and (c) the asymmetry between B^0 and $\overline{B}^0$ tags.

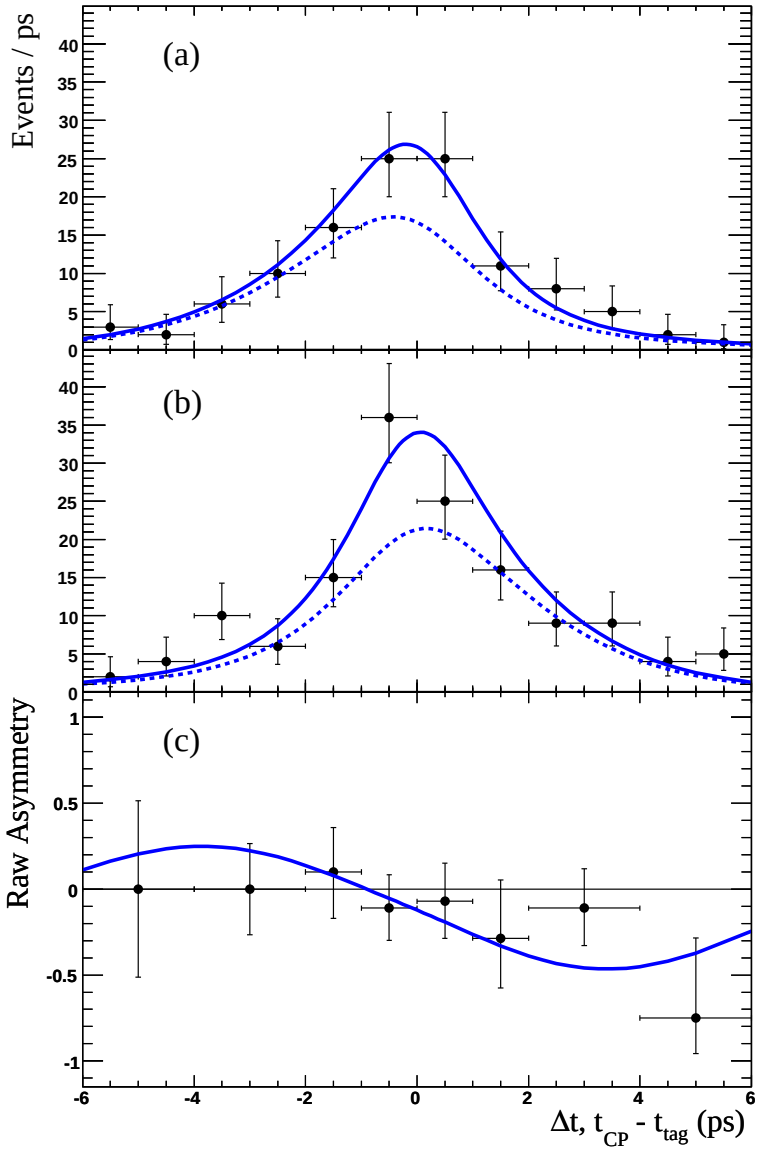


Figure 5.8: Projections onto Δt for data for the two $\eta' K^0_L$ modes combined (points with errors), the fit full function (blue solid line), and the signal only function (blue dashed line). We show (a) B^0 and (b) $\overline{B}^0$ tagged events, and (c) the asymmetry between B^0 and $\overline{B}^0$ tags.

5.10 Averaging sub-mode results

As mentioned in Sec. 4.5.2 and throughout this document, the joint results can be obtained in several ways. One can

- use the simultaneous fit to account for correlations between S and C,
- combine $-2\ln\mathcal{L}$ curves with separate sub-mode systematics to properly account for uncorrelated systematic errors,
- correct for fit biases on S and C mode-by-mode to align the likelihoods and reduce statistical error as mentioned in Sec. 5.8.2.1,
- correct for the fit bias on the combined S and C, to correctly treat small-yield bias for the minor modes (also mentioned in Sec. 5.8.2.1).

We obtain joint results using various combinations of all methods described above. For instance, in Table 5.16, we compare uncorrected results for the sub-mode yields, $S_{\eta' K^0}$, and $C_{\eta' K^0}$ for the simultaneous fit and combination with $-2\ln\mathcal{L}$ curves; in Fig. 5.9, we show the individual $-2\ln\mathcal{L}$ curves and their sum for $S_{\eta' K^0}$ and $C_{\eta' K^0}$. In blind fits, we find that, since correlations and biases are small and statistical uncertainties are four times larger than systematic errors, all combinations of averaging procedures (in bullets above) give almost identical results. We choose to use the simultaneous fit for its elegance and simplicity and to correct for the fit bias on the combined S and C.

5.10.1 Corrected results

We perform three simultaneous fits in which we find S and C for all seven sub-modes, for the five $\eta' K^0_S$ sub-modes, and for the two $\eta' K^0_L$ sub-modes. We show results from the combined $\eta' K^0_S$ fit in Table 5.13, the combined $\eta' K^0_L$ fit in Table 5.14, and the combined $\eta' K^0$ fit in Table 5.16. We summarize bias-corrected results here:

Table 5.16: Uncorrected results for the combined $S_{\eta' K^0}$ and $C_{\eta' K^0}$ and yields for seven sub-modes. The first column are results from the combined fit. The second column contains results from individual fits; combined $S_{\eta' K^0}$ and $C_{\eta' K^0}$ are obtained by combining the $-2\ln\mathcal{L}$ curves.

	Simultaneous Fit (nominal)	Individual Fits ($-2\ln\mathcal{L}$ curves)
$S_{\eta' K^0}$	0.551 ± 0.077	0.554 ± 0.077
$C_{\eta' K^0}$	-0.094 ± 0.058	-0.092 ± 0.058
Signal Yields		
$\eta'_{\eta\pi\pi} K^0_S$	472.1 ± 23.6	471.5 ± 23.6
$\eta'_{\rho\gamma} K^0_S$	1004.4 ± 39.7	1004.7 ± 39.7
$\eta'_{\eta\pi\pi} K^0_{S00}$	105.5 ± 13.0	105.3 ± 13.0
$\eta'_{\rho\gamma} K^0_{S00}$	203.4 ± 27.4	205.8 ± 27.6
$\eta'_{5\pi} K^0_S$	171.2 ± 14.1	170.9 ± 14.1
$\eta'_{\eta\pi\pi} K^0_L$	340.2 ± 31.9	345.4 ± 37.9
$\eta'_{5\pi} K^0_L$	160.6 ± 21.6	174.7 ± 24.0

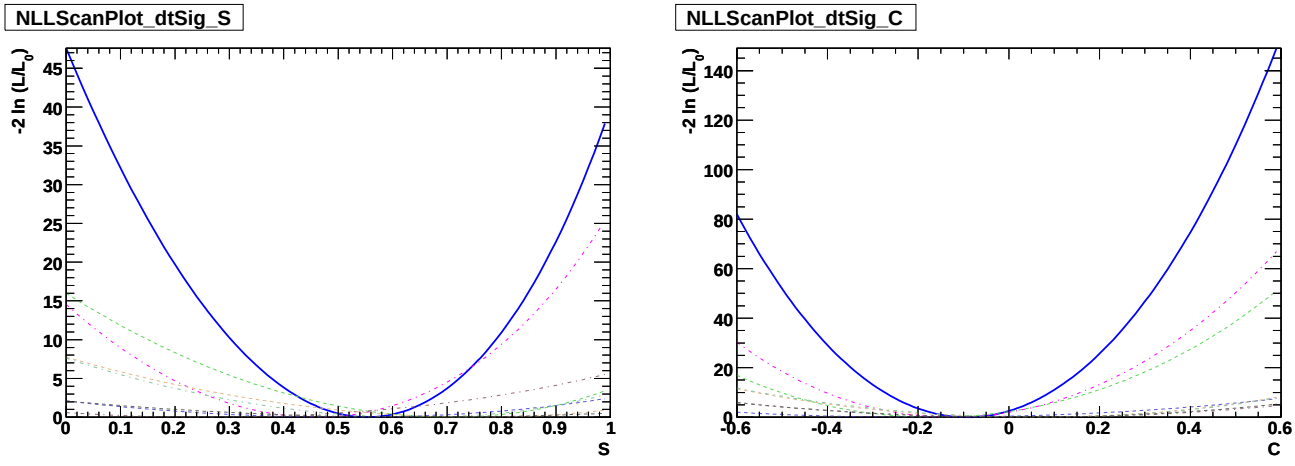


Figure 5.9: Scans of $-2\ln\mathcal{L}/\mathcal{L}_{\max}$ versus $-\eta S$ (left) and C (right). The solid blue line is for combined neutral sub-decays. The dashed lines are for the seven sub-modes $\eta'_{\eta\pi\pi} K^0_S$, $\eta'_{\rho\gamma} K^0_S$, $\eta'_{\eta\pi\pi} K^0_{S00}$, $\eta'_{\rho\gamma} K^0_{S00}$, $\eta'_{5\pi} K^0_S$, $\eta'_{\eta\pi\pi} K^0_L$, and $\eta'_{5\pi} K^0_L$.

$$
\begin{aligned}
S_{\eta' K^0} &= 0.545 \pm 0.077, \\
C_{\eta' K^0} &= -0.086 \pm 0.058, \\
-\eta_{\eta' K^0_S} S_{\eta' K^0_S} &= 0.529 \pm 0.084, \\
C_{\eta' K^0_S} &= -0.111 \pm 0.062, \\
-\eta_{\eta' K^0_L} S_{\eta' K^0_L} &= 0.639 \pm 0.198, \\
C_{\eta' K^0_L} &= 0.053 \pm 0.154. \qquad (5.6)
\end{aligned}
$$

5.11 Systematic uncertainties

Even though the precision of these results is dominated by statistical uncertainties, we perform detailed and careful studies of many sources of systematic uncertainty on S and C. All numbers refer to errors for the seven-mode combined results. In Sec. 5.11.1 we report systematic uncertainties separately for $\eta' K^0_S$ and $\eta' K^0_L$. We sum all systematic errors in quadrature to obtain final uncertainties shown in Table 5.21.

- **PDF parameterization**: We estimate the related errors principally by variation of the fit parameters. In Table 5.17, we summarize all of the variations and the resulting shifts in S and C. All changes are combined in quadrature to obtain an error of 0.004 for S and 0.010 for C.

- ***CP* content in $B\overline{B}$ background** : Due to uncertainty of the *CP* content in the $B\overline{B}$ background we perform fits fixing S and C in the $B\overline{B}$ background to various values depending on our knowledge of the content of the $B\overline{B}$ background. This study is described in Appendix F; we find a systematic on the signal S (C) of 0.008 (0.004). We also fit the data fixing the charmless yields at 80% or 120% of the expected yield and find a change of 0.001 (0.000) in S (C) as described in Appendix G. Combining the yield and *CP* parameter effects in quadrature,

we obtain a total systematic of 0.008 (0.004) for S (C).

- **Fit bias**: Toy studies (Sec. 5.8.2) show that there are net biases on S and C of 0.006 ± 0.006 and -0.008 ± 0.006. We correct the final results for the bias and take the statistical uncertainty on the bias as a systematic uncertainty of 0.006 for S and C.

- **SVT alignment**: We vary the alignment of the SVT in signal MC by the size of misalignment found in the real data. There are five different SVT configurations which are considered. Four configurations simulate time dependent misalignment and one (labeled *boost*) simulates the radius-dependent z shift of entire layers. We fit these MC samples and take the observed shift in S and C as a systematic error. The difference for each configuration between the nominal values of S and C and those from the misaligned configuration are presented in Table 5.18. The nominal value was extracted by running signal MC in refit mode with perfect SVT alignment. We perform this study for the major modes $\eta'_{\rho\gamma}K^0_S$ and $\eta'_{\eta\pi\pi}K^0_S$; we take the larger of the two systematics to obtain uncertainties on S (C) of $+0.002 - 0.001$ ($+0.003 - 0.002$).

- **Beam-spot position**: Just as done for the SVT alignment study, we vary the beam-spot y position and error in signal MC by values considered realistic for real data. The variations in the beam-spot parameters and the changes in S and C are presented in Table 5.19. We take the larger of changes in S and C for the major modes $\eta'_{\rho\gamma}K^0_S$ and $\eta'_{\eta\pi\pi}K^0_S$ as the systematic error to obtain uncertainties on S (C) of 0.002 (0.001).

- **Tag-side interference**: We estimate the effect of interference between the CKM-suppressed $\overline{b} \to \overline{u}c\overline{d}$ amplitude with the favored $b \to c\overline{u}d$ amplitude, which occurs for some tag-side B decays [66]. We find an uncertainty of 0.001

Table 5.17: Results of systematic variations of the PDF shapes. We show the nominal values, the amount that we vary these, the source of this variation amount, and the change of S and C for this amount of variation. We group similar quantities together after combining their variations in quadrature.

Quantity	Nominal	± variation	Source of variation	Change in S	Change in C
Δm_d (ps^{-1})	0.507	$+0.005$	PDG	$+0.0024$	$+0.0015$
		-0.005	PDG	-0.0024	-0.0015
τ_B $((ps)$	1.530	$+0.009$	PDG	$+0.0007$	$+0.0005$
		-0.009	PDG	-0.0008	-0.0005
w	Table 4.2	Table 4.2	Table 4.2	0.0036	0.0053
Δw	Table 4.2	Table 4.2	Table 4.2	0.0007	0.0060
Signal $f_{\rm cat}$	Table 4.2	Table 4.2	Table 4.2	0.0005	0.0002
Signal Δt	Table 4.3	Table 4.3	Table 4.3	0.0024	0.0016
ΔE for K^0_S shift (MeV)	-2.00	2.00	Sec. 5.17	0.0007	0.0012
ΔE for K^0_S scale	1.00	0.02	Sec. 5.17	0.0006	0.0001
ΔE for K^0_L shift (MeV)	-0.40	0.25	Sec. 5.17	0.0003	0.0005
ΔE for K^0_L scale	1.10	0.10	Sec. 5.17	0.0001	0.0001
$m_{\rm ES}$ shift (MeV)	0.00	0.10	Table 4.1	0.0006	0.0002
$\mathcal{F}$	PDF fit	mean 0.02 rms 0.01 asym 0.04	Sec. 5.17	0.0007	0.0009
Total				0.0053	0.0085

Table 5.18: The change in S and C due to different SVT configurations for the modes $\eta'_{\rho\gamma}K^0_S$ and $\eta'_{\eta\pi\pi}K^0_S$. The overall uncertainty is calculated by taking the largest positive and negative contributions of the four time dependent configurations adding them in quadrature to the Boost contribution.

	$\eta'_{\rho\gamma}K^0_S$		$\eta'_{\eta\pi\pi}K^0_S$	
Configuration	ΔS	ΔC	ΔS	ΔC
Time1	0.0005	−0.0005	0.0008	−0.0014
Time2	−0.0006	0.0011	0.0005	0.0002
Time3	0.0014	0.0016	−0.0002	0.0028
Time4	−0.0008	0.0001	−0.0005	−0.0005
Boost	−0.0010	0.0003	0.0015	0.0006
Overall	$^{+0.0014}_{-0.0010}$	$^{+0.0016}_{-0.0011}$	$^{+0.0017}_{-0.0005}$	$^{+0.0028}_{-0.0015}$

Table 5.19: The beam-spot y position was varied by $\pm 20\,\mu m$ and the error on y was independently increased to $20\,\mu m$ in the modes $\eta'_{\rho\gamma}K^0_S$ and $\eta'_{\eta\pi\pi}K^0_S$. The overall contribution to the systematics was evaluated by averaging the absolute contributions from the shift in y and adding that in quadrature to the contribution from the error on y.

	$\eta'_{\rho\gamma}K^0_S$		$\eta'_{\eta\pi\pi}K^0_S$	
Beam-spot	ΔS	ΔC	ΔS	ΔC
$y + 20\,\mu m$	−0.0016	−0.0010	0.0025	0.0005
$y - 20\,\mu m$	−0.0004	0.0002	0.0002	0.0000
$\sigma_y + 20\,\mu m$	0.0014	0.0002	0.0011	0.0004
Overall	±0.0017	±0.0007	±0.0018	±0.0005

for S and 0.015 for C at the 68% confidence level.

- **Signal Δt parameterization**: We estimate systematic uncertainty for the appropriateness of taking $\eta' K^0$ signal Δt parameters (Δt resolution, mistag rates, and tagging category fractions) from B_{flav} data. Since the resolution on the signal side is not completely negligible, the B_{flav} resolution parameters are not entirely appropriate for our decay mode. We remove this effect to first order by making the resolution model a function of $\Delta t/\sigma_{\Delta t}$. We determine a systematic uncertainty for second order effects.

  We estimate the effect of differences between the B_{flav} and true signal resolution models by finding the difference between toy results obtained with Δt parameters from B_{flav} MC (rather than data) and results obtained with parameters from signal MC. When using B_{flav} MC parameters in embedded toys, we find a net fit bias on S (C) of $+0.022 \pm 0.004$ ($+0.001 \pm 0.003$). In nominal toy studies (Table 5.11), we find a bias of $+0.013 \pm 0.004$ (-0.007 ± 0.003) on S (C). These toy fits were performed on the same toy datasets, so the uncertainty on the differences of the biases is very small. We take the differences as a systematic of 0.009 on S and 0.008 on C.

  We also separate the effects of using B_{flav} tagging and resolution model parameters with an additional check: we compare results from two sets of simultaneous-fit toy experiments which differ only in the resolution model parameters. We find S (C) differs by 0.002 (0.010). We conclude that, when evaluating the C systematic of 0.008 as described in the preceding paragraph, there is a cancellation of 0.002. We therefore increase the C systematic to 0.010 and keep the S systematic of 0.009.

- **Self-crossfeed**: We investigate a systematic for the effect of self-crossfeed (SXF) on S and C. Table 5.2 in Section 5.3.5 shows that the fraction of events

in which SXF affects vertexing, the "wrong track" SXF (WT-SXF), is small. To confirm that the effects of SXF are small and to evaluate the related systematic, we perform embedded toy experiments, embedding truth-matched signal MC instead of nominal signal MC, for all sub-modes. For the $\eta' K^0_S$ sub-mode with the highest WT-SXF fraction, $\eta'_{5\pi} K^0_S$, we perform four sets of experiments embedding signal from four different samples of MC: truth MC (0% WT-SXF), nominal MC (2.4% WT-SXF), 1.25×nominal MC (3.0% WT-SXF), and 1.75×nominal MC (4.2% WT-SXF). Full results for all modes are shown in Table 5.20. We find shifts on S and C related to SXF of -0.006 and -0.003, respectively. (The uncertainty on these numbers is negligible because the toy fits were performed on the same toy datasets.) We conclude that, in all sub-modes, the wrong-track SXF fraction is small and will be accounted for in bias corrections, where necessary. We take the differences of 0.006 and 0.003 as the related systematics for S and C.

5.11.1 Separate systematic uncertainties for $\eta' K^0_S$ and $\eta' K^0_L$

We also estimate systematic uncertainties separately for the 5-mode K^0_S and 2-mode K^0_L simultaneous fits. Each systematic is estimated as described in Sec. 5.11 for the 7-mode simultaneous fit. We denote each systematic as correlated (C) or uncorrelated (U) to facilitate combining results. In Table 5.21 we summarize the separate systematic uncertainties for $\eta' K^0_S$ and $\eta' K^0_L$.

5.12 Crosschecks

We perform a number of crosschecks to evaluate the stability and precision of the fitting procedure. These tests are performed blind; we shift the central values of results by the same unknown, random value for all tests while leaving the errors unaffected.

Table 5.20: Results of embedded toy experiments in which we embed signal events from samples with different concentrations of wrong-track self-crossfeed (WT-SXF). See Section 5.3.5 for definition of wrong-track SXF. The parameters S and C are generated at 0.7 and 0.0. A * denotes rows with nominal WT-SXF fractions.

Final State	WT-SXF Frac. (%)	# toy expts	# signal input	Signal Yield	$-\eta S$	C
$\eta'_{\eta\pi\pi}K^0_S$	0.0	400	470	481.1 ± 0.4	0.690 ± 0.008	0.002 ± 0.005
	0.9*	550	470	465.4 ± 0.4	0.700 ± 0.007	0.001 ± 0.005
$\eta'_{\rho\gamma}K^0_S$	0.0	270	970	1061.0 ± 1.5	0.684 ± 0.007	0.012 ± 0.005
	1.8*	300	970	1002.0 ± 1.4	0.683 ± 0.007	0.012 ± 0.006
$\eta'_{\eta\pi\pi}K^0_{S00}$	0.0	200	103	107.2 ± 0.6	0.720 ± 0.029	-0.004 ± 0.020
	0.9*	200	103	99.6 ± 0.5	0.734 ± 0.033	-0.004 ± 0.020
$\eta'_{\rho\gamma}K^0_{S00}$	0.0	140	199	215.0 ± 2.0	0.748 ± 0.033	0.040 ± 0.025
	1.8*	400	199	205.6 ± 1.2	0.724 ± 0.018	0.013 ± 0.014
$\eta'_{5\pi}K^0_S$	0.0	150	173	182.8 ± 0.4	0.690 ± 0.022	-0.013 ± 0.014
	2.4*	190	173	171.3 ± 0.4	0.711 ± 0.021	-0.007 ± 0.014
	3.0	150	173	167.7 ± 0.5	0.712 ± 0.022	-0.040 ± 0.015
	4.2	110	173	160.9 ± 0.7	0.712 ± 0.025	-0.032 ± 0.017
$\eta'_{\eta\pi\pi}K^0_L$	0.0	100	353	341.4 ± 2.7	0.657 ± 0.028	-0.016 ± 0.017
	1.8*	240	353	325.3 ± 1.6	0.678 ± 0.017	-0.013 ± 0.012
$\eta'_{5\pi}K^0_L$	0.0	110	170	174.1 ± 1.5	0.736 ± 0.037	0.020 ± 0.024
	3.9*	140	170	158.9 ± 1.5	0.769 ± 0.035	0.004 ± 0.020
Weighted	0.0				0.689 ± 0.005	0.006 ± 0.003
Averages:	~ 1.8				0.695 ± 0.004	0.003 ± 0.003

Table 5.21: Estimates of systematic errors for the combined $\eta' K^0$ results as well as separate $\eta' K^0_S$ and $\eta' K^0_L$ results. We denote each systematic as correlated (C) or uncorrelated (U) to facilitate combining the $\eta' K^0_S$ and $\eta' K^0_L$ results.

	$\eta' K^0_S$		$\eta' K^0_L$		$\eta' K^0$	
Source of error	$\sigma(S)$	$\sigma(C)$	$\sigma(S)$	$\sigma(C)$	$\sigma(S)$	$\sigma(C)$
PDF shapes (C)	0.006	0.009	0.007	0.014	0.005	0.009
$B\overline{B}$ background (U)	0.009	0.005	–	–	0.008	0.004
Fit bias (U)	0.006	0.005	0.014	0.011	0.006	0.006
SVT alignment (C)	$^{+0.002}_{-0.001}$	$^{+0.003}_{-0.002}$	$^{+0.002}_{-0.001}$	$^{+0.003}_{-0.002}$	$^{+0.002}_{-0.001}$	$^{+0.003}_{-0.002}$
Beam position/size (C)	0.002	0.001	0.002	0.001	0.002	0.001
Tag-side interference (C)	0.001	0.015	0.001	0.015	0.001	0.015
$B_{\rm flav}$ signal shape (C)	0.009	0.015	0.017	0.016	0.009	0.010
Self-crossfeed (U)	0.004	0.001	0.023	0.004	0.006	0.003
Total correlated	0.011	0.023	0.019	0.026	–	–
Total uncorrelated	0.012	0.007	0.027	0.012	–	–
Total	0.016	0.024	0.033	0.029	0.016	0.022

In this way, we can compare central values from different configurations of fits to data without learning the actual fit value and risking biasing our choice of fit configuration.

5.12.1 Adding more discriminating variables to the fit

To understand how additional discriminating variables affect the central values and precision of fit results, we have added PDFs for η' mass, ρ mass (where applicable), and ρ helicity (where applicable) to all sub-mode fits. Blind results from these *all-variable fits* to the full dataset are summarized in Table 5.22. The additional complexity yields an improvement in errors on combined S and C of 1.7% and 1.3%, respectively. These improvements would be lost in the rounding.

5.12.1.1 Study of $\eta'_{\rho\gamma} K^0_{S00}$ six-variable fit

The central values for S and C in the all-variable fit study (Table 5.22) are very stable for all modes except $\eta'_{\rho\gamma} K^0_{S00}$. We investigate this instability with embedded toy

Table 5.22: Comparison of nominal three-variable blind fits and crosscheck fits including all variables. Combined S and C values are simple weighted averages.

	Nominal Fit	All-Variable Fit
$\eta'_{\eta\pi\pi}K^0_S$		
$-\eta S$	-0.622 ± 0.166	-0.629 ± 0.165
C	0.231 ± 0.110	0.217 ± 0.109
$\eta'_{\rho\gamma}K^0_S$		
$-\eta S$	-0.370 ± 0.117	-0.373 ± 0.116
C	0.217 ± 0.090	0.199 ± 0.086
$\eta'_{\eta\pi\pi}K^0_{S00}$		
$-\eta S$	-0.402 ± 0.347	-0.390 ± 0.342
C	0.232 ± 0.301	0.266 ± 0.301
$\eta'_{\rho\gamma}K^0_{S00}$		
$-\eta S$	-0.179 ± 0.331	-0.518 ± 0.336
C	0.029 ± 0.261	0.107 ± 0.263
$\eta'_{5\pi}K^0_S$		
$-\eta S$	-0.650 ± 0.261	-0.671 ± 0.260
C	0.036 ± 0.199	0.039 ± 0.198
$\eta'_{\eta\pi\pi}K^0_L$		
$-\eta S$	-0.684 ± 0.223	-0.692 ± 0.218
C	0.036 ± 0.194	0.033 ± 0.184
combined S	-0.482 ± 0.079	-0.507 ± 0.078
combined C	0.179 ± 0.059	0.171 ± 0.058

studies and unblind fits to Run1-5 data. By adding additional variables one at a time, we learn that the ρ mass variable causes the observed shift in S and C. We confirm that the difference between three- and six-variable fit results is decoupled from our treatment of the yields of the charmless and charm $B\overline{B}$ fit components with studies documented in Table 5.23 that show that the yield shifts regardless of our treatment of $B\overline{B}$ fit components.

We consider using the six-variable fit for $\eta'_{\rho\gamma}K^0_{S00}$ since the ρ mass seems to offer discriminatory power. We check which strategy (three-variable or six-variable) minimizes errors on S and C with embedded toy studies; the means of the S error distributions are identical for these strategies (0.243 ± 0.003); the mean of C error distribution is 3.5% lower for the six-variable fit than for the three-variable fit (0.188 ± 0.001 vs. 0.197 ± 0.002). In addition, if we combine the blind six-variable fit results for $\eta'_{\rho\gamma}K^0_{S00}$ with the nominal blind results from the other five modes, the errors on combined S and C are identical (0.079 and 0.059) to those obtained by combining the blind three-variable fit results for $\eta'_{\rho\gamma}K^0_{S00}$ with the nominal blind results from the other modes. Since the six-variable fit offers no improvement on the three-variable fit, we choose to use the simpler three-variable fit for $\eta'_{\rho\gamma}K^0_{S00}$.

Table 5.23: Comparison of Run1-5 results for the mode $\eta'_{\rho\gamma}K^0_{S00}$ from three-variable fits (3-Var Fit) and six-variable fits (6-Var Fit) in which we float the charmless $B\overline{B}$ yield with and without a charm $B\overline{B}$ component.

ML fit quantity	3-Var Fit Flt nChls	6-Var Fit Flt nChls	3-Var Fit No Chrm	6-Var Fit No Chrm
Signal yield	164.6 ± 27.9	149.9 ± 21.6	151.5 ± 27.7	145.6 ± 21.6
Chls yield	164.8 ± 77.0	141.6 ± 59.7	364.8 ± 61.8	320.6 ± 50.5
Charm yield	345.1 ± 99.3	361.0 ± 89.2	0	0
S	0.39 ± 0.40	0.59 ± 0.36	0.48 ± 0.43	0.63 ± 0.37
C	-0.01 ± 0.30	-0.01 ± 0.28	0.02 ± 0.32	0.03 ± 0.29

5.12.1.2 Study of $\eta'_{\rho\gamma}K^0_S$ six-variable fit

Because it carries the most weight in determining combined uncertainties, we have also performed unblind six-variable fits to Run1-5 data for $\eta'_{\rho\gamma}K^0_S$ both fixing and floating the charmless yield. These studies were performed before the charm $B\overline{B}$ component was added the fit, but the results are still useful for diagnostic purposes. The results are listed with the results of the nominal six-variable fit in Table 5.24. We have performed embedded toy experiments in both three- and six-variable configurations to understand how errors on S and C compare. We report the mean of the S and C error distributions in Table 5.25. It is clear that the results are stable between the three- and six-variable fits and that there is little increase in precision for the six-variable fit.

Table 5.24: Comparison of Run1-5 results for $\eta'_{\rho\gamma}K^0_S$ fits with three discriminating variables ($m_{\rm ES}$, ΔE, $\mathcal{F}$, and Δt in TD fits) and six variables ($m_{\rm ES}$, ΔE, $\mathcal{F}$, η' mass, ρ mass, $\mathcal{H}_\rho$, and Δt in TD fits). We performed each fit both floating and fixing the charmless $B\overline{B}$ yield (`nChls`).

$\eta'_{\rho\gamma}K^0_S$ ML fit quantity	Float `nChls` 3 Variables	Fix `nChls` 3 Variables	Float `nChls` 6 Variables	Fix `nChls` 6 Variables
Yield-only fits:				
Events into fit	19356	19356	19356	19356
Signal yield	860.3 ± 38.5	857.5 ± 38.0	852 ± 36.5	856.4 ± 36.4
Chls yield	190.7 ± 90.2	230	335.3 ± 79.8	230
MC ϵ (%)	28.5	28.5	28.5	28.5
Fit bias (events)	33.7	33.7	33.7	33.7
$\mathcal{B}(10^{-6})$	74.7 ± 3.5	74.4 ± 3.4	73.9 ± 3.3	74.3 ± 3.3
TD fits:				
Events into fit	18464	18464	18464	18464
Signal yield	825 ± 36.2	840.1 ± 35.9	816.6 ± 34.6	831.6 ± 34.6
Chls yield	364.4 ± 54.6	230	428.3 ± 53.2	230
S	0.45 ± 0.13	0.44 ± 0.13	0.45 ± 0.13	0.44 ± 0.13
C	-0.09 ± 0.10	-0.09 ± 0.10	-0.08 ± 0.10	-0.08 ± 0.09

Table 5.25: Comparison of errors on S and C from embedded toy studies for the three-variable and six-variable fits for the mode $\eta'_{\rho\gamma}K^0_S$. We report the means of the S and C error distributions.

	3-variable fit	6-variable fit
Mean of S Error Dist	0.120 ± 0.001	0.118 ± 0.001
Mean of C Error Dist	0.0933 ± 0.0003	0.0910 ± 0.0002

5.12.2 Fits with $C = 0$

We perform the simultaneous fit to the full dataset fixing $C = 0$. The fit value of the S parameter shifts -0.002 relative to the nominal fit. The correlation between S and C in the combined fit is 3.1%, so we expect S to be minimally affected by fixing C to zero.

5.12.3 Fits removing variables

We perform the blind simultaneous fit removing one discriminating variable at a time. (We only perform this test for the $\eta' K^0_S$ modes; we do not remove variables from the fits for $\eta' K^0_L$ modes because they already lack an $m_{\rm ES}$ dimension in the nominal fit.) The fit results are shown in Table 5.26. The nominal blind fit results are in the first column. We conclude that the fit is stable even when missing one of the variables.

5.12.4 Splitting S and C by tagging category

We perform the blind simultaneous fit in which the signal CP parameters are not constrained to be equal across tagging categories. The fit values for each category, the weighted averages of the separate category results, and the nominal results are shown in Table 5.27.

Table 5.26: Blind results from the combined fit to the full dataset removing one fit variable at a time. We label each column with the fit variables used in the fit; the nominal blind fit results are in the first column.

Fit vars	$m_{\rm ES}$, ΔE, $\mathcal{F}$, Δt	ΔE, $\mathcal{F}$, Δt	$m_{\rm ES}$, $\mathcal{F}$, Δt	$m_{\rm ES}$,ΔE, Δt
S	-0.481 ± 0.078	-0.471 ± 0.080	-0.512 ± 0.077	-0.535 ± 0.082
C	0.174 ± 0.058	0.159 ± 0.062	0.164 ± 0.059	0.165 ± 0.063
Signal yields:				
$\eta'_{\eta\pi\pi}K^0_S$	468.8 ± 23.5	465.4 ± 26.2	461.8 ± 23.5	475.1 ± 24.9
$\eta'_{\rho\gamma}K^0_S$	999.0 ± 39.6	997.0 ± 47.5	966.7 ± 46.9	1007.2 ± 45.4
$\eta'_{\eta\pi\pi}K^0_{S00}$	104.3 ± 13.0	107.4 ± 15.8	104.3 ± 12.7	111.3 ± 18.0
$\eta'_{\rho\gamma}K^0_{S00}$	201.9 ± 27.4	288.7 ± 44.6	196.2 ± 29.7	201.9 ± 44.9
$\eta'_{5\pi}K^0_S$	170.5 ± 14.1	177.4 ± 15.8	175.9 ± 23.1	172.3 ± 14.9
$\eta'_{\eta\pi\pi}K^0_L$	331.5 ± 31.4	334.9 ± 31.5	334.6 ± 29.1	334.6 ± 31.5
$\eta'_{5\pi}K^0_L$	163.9 ± 21.8	160.9 ± 21.6	160.8 ± 19.8	160.7 ± 21.6

Table 5.27: Blind results of the simultaneous fit in which we allow S and C to float to different values for each tagging category. We also report the weighted average of the separate tagging category results, and the nominal result for comparison.

Tagging category	S	C
`Lepton`	-0.670 ± 0.134	0.122 ± 0.105
`Kaon I`	-0.315 ± 0.146	0.132 ± 0.107
`Kaon II`	-0.569 ± 0.150	0.297 ± 0.117
`Kaon-Pion`	-0.246 ± 0.255	0.080 ± 0.185
`Pion`	-0.205 ± 0.365	0.263 ± 0.257
`Other`	-1.414 ± 1.010	0.624 ± 0.710
Weighted Average	-0.492 ± 0.076	0.174 ± 0.057
Nominal Result	-0.481 ± 0.078	0.174 ± 0.058

5.12.5 Floating signal Δt resolution model parameters in $\eta'_{\rho\gamma}K^0_S$

We perform a blind fit for the $\eta'_{\rho\gamma}K^0_S$ sub-mode in which we fix all PDF parameters and yields at values obtained in the nominal fit and float the important parameters of the signal Δt resolution model: the bias of the tail Gaussian, the fraction of the core Gaussian, the fraction of the outlier Gaussian, and the bias and scale factor of the

core Gaussian for the non-Lepton tags. The other parameters of the signal Δt model, such as the average mistag rates, difference in mistag rates, and difference in tagging efficiencies, are fixed at values from the $B_{\rm flav}$ sample. The resolution model parameters float to reasonable values, and S and C change by only -0.010 and -0.002 with respect to the nominal fit. Because these changes are well within errors, we conclude that our treatment of the Δt resolution model is reasonable.

5.12.6 Run-by-run efficiency and purity in $\eta'_{\eta\pi\pi}K^0_L$

We expect some improvement in K^0_L reconstruction with the installation of all six LST sextants in Run6. In Table 5.28, we report the number of events entering the fit divided by integrated luminosity (an effective background efficiency), the signal efficiency (from MC), and the ratio of the effective background efficiency over the signal efficiency for each run period for the mode $\eta'_{\eta\pi\pi}K^0_L$. One can see that there is slight improvement in Run6.

Table 5.28: The number of events entering the fit divided by the integrated luminosity (effective $\epsilon_{\rm bkg}$), the signal MC efficiency ($\epsilon_{\rm sig}$), and the ratio of the two ($\epsilon_{\rm bkg}/\epsilon_{\rm sig}$).

Run Period	Effective $\epsilon_{\rm bkg}$	$\epsilon_{\rm sig}$	$\epsilon_{\rm bkg}/\epsilon_{\rm sig}$
Run1	0.036 ± 0.001	0.217 ± 0.003	0.166 ± 0.005
Run2	0.033 ± 0.001	0.205 ± 0.002	0.161 ± 0.005
Run3	0.033 ± 0.001	0.200 ± 0.002	0.165 ± 0.005
Run4	0.033 ± 0.001	0.194 ± 0.001	0.170 ± 0.005
Run5	0.033 ± 0.001	0.191 ± 0.001	0.173 ± 0.005
Run6	0.032 ± 0.001	0.210 ± 0.001	0.152 ± 0.005
Run1-6	0.033 ± 0.001	0.199 ± 0.001	0.166 ± 0.005

5.12.7 Conclusions regarding crosschecks

We conclude that the precision of our S and C measurements is optimal in a simple fit that employs $m_{\rm ES}$, ΔE, $\mathcal{F}$, and no additional discriminating variables. We find that measurements of S and C in individual tagging categories agree within errors, supporting our simultaneous fit across tagging categories. Finally, we find additional evidence supporting the important assumption that our use of Δt resolution model parameters from the $B_{\rm flav}$ sample is reasonable.

5.13 Conclusions

We have reconstructed about 2400 $B^0 \to \eta' K^0$ events (2000 $B^0 \to \eta' K^0_S$ events and 400 $B^0 \to \eta' K^0_L$ events), 75% of which are flavor tagged. We find the following raw results from the fit to these data:

$$\begin{aligned} S_{\eta' K^0} &= 0.551 \pm 0.077, \\ C_{\eta' K^0} &= -0.094 \pm 0.058. \end{aligned}$$

We correct for the fit bias (Table 5.11) and add the systematic uncertainties (Table 5.21) as the second error to obtain:

$$\begin{aligned} S_{\eta' K^0} &= 0.545 \pm 0.077 \pm 0.016, \\ C_{\eta' K^0} &= -0.086 \pm 0.058 \pm 0.022. \end{aligned}$$

For the 5-mode $\eta' K^0_S$ combined fit, we find bias-corrected results where the first error is statistical and the second systematic:

$$\begin{aligned} -\eta_{\eta' K^0_S} S_{\eta' K^0_S} &= 0.529 \pm 0.084 \pm 0.016, \\ C_{\eta' K^0_S} &= -0.111 \pm 0.062 \pm 0.024. \end{aligned}$$

For the 2-mode $\eta' K^0_L$ combined fit, we report bias-corrected results where the first error is statistical and the second systematic:

$$\begin{aligned} -\eta_{\eta' K^0_L} S_{\eta' K^0_L} &= 0.639 \pm 0.198 \pm 0.033, \\ C_{\eta' K^0_L} &= 0.053 \pm 0.154 \pm 0.029. \end{aligned}$$

Chapter 6

ηK^* Analysis

6.1 General analysis strategy

In the ηK^* analysis, we measure the branching fractions and time-integrated *CP*-violating charge asymmetries $\mathcal{A}_{ch}$ for the decays $B \to \eta K^*(892)$, $B \to \eta K_0^*(S\text{-wave})$, and $B \to \eta K_2^*(1430)$. We use $K_0^*(S\text{-wave})$ to refer to the coherent superposition of resonant and non-resonant $K\pi$ amplitudes first characterized by the LASS collaboration [67] and described in Sec. 6.3, while $K_0^*(1430)$ refers solely to the resonant contribution to the $K_0^*(S\text{-wave})$.

6.1.1 Sub-decay modes

For each of the three K^* components, $K^*(892)$, $K_0^*(S\text{-wave})$, and $K_2^*(1430)$, we measure the branching fraction (BF) in six sub-modes, two neutral and four charged:

- $B \to \eta K^{*+}$: $\eta_{\gamma\gamma} K^{*+}_{K^+\pi^0}$, $\eta_{\gamma\gamma} K^{*+}_{K^0_S\pi^+}$, $\eta_{3\pi} K^{*+}_{K^+\pi^0}$, $\eta_{3\pi} K^{*+}_{K^0_S\pi^+}$,
- $B \to \eta K^{*0}$: $\eta_{\gamma\gamma} K^{*0}_{K^+\pi^-}$, $\eta_{3\pi} K^{*0}_{K^+\pi^-}$.

We combine the sub-mode BFs to obtain the BF for the neutral and charged modes, $\mathcal{B}(B \to \eta K^{*+})$ and $\mathcal{B}(B \to \eta K^{*0})$, for each of the three K^* components for a total of six BFs:

- $\mathcal{B}(B \to \eta K^{*0}(892))$, $\mathcal{B}(B \to \eta K^{*+}(892))$,

- $\mathcal{B}(B^0 \to \eta K_0^{*0}(S\text{-wave}))$, $\mathcal{B}(B^+ \to \eta K_0^{*+}(S\text{-wave}))$,
- $\mathcal{B}(B^0 \to \eta K_2^{*0}(1430))$, $\mathcal{B}(B^+ \to \eta K_2^{*+}(1430))$.

6.1.2 $K\pi$ mass range

We explored the option of performing a single analysis for all three K^* components over the mass range $0.755 < m_{K\pi} < 1.535$ GeV. Because we observe a large B-candidate multiplicity over this full mass range (1.6 candidates per event), we adopt a conservative approach for improving the precision on the existing $\eta K^*(892)$ measurements while searching for the decays $\eta K_0^*(S\text{-wave})$ and $\eta K_2^*(1430)$. We reduce the maximum candidate multiplicity to 1.3 candidates per event by performing separate fits for the mass regions $0.755 < m_{K\pi} < 1.035$ GeV and $1.035 < m_{K\pi} < 1.535$ GeV. We include $\eta K_0^*(S\text{-wave})$ and $\eta K^*(892)$ components in both fits, but we fix the $\eta K_0^*(S\text{-wave})$ yield and float the $\eta K^*(892)$ yield in the low mass range (LMR); in the high mass range (HMR), we fix the $\eta K^*(892)$ yield and float the $\eta K_0^*(S\text{-wave})$ and $\eta K_2^*(1430)$ yields. In Fig. 6.1 we show the lineshapes as functions of $K\pi$ mass for the three $K\pi$ partial waves, $K_0^*(S\text{-wave})$, $K^*(892)$, and $K_2^*(1430)$, with a vertical line at 1.035 GeV that separates the LMR and HMR.

6.1.3 Sub-mode fits in separate $K\pi$ mass ranges

For the $\eta K^*(892)$ modes, we measure the BFs in six separate fits, one for each sub-mode, and combine the results using $-2\ln\mathcal{L}$ curves as described in Sec. 4.5.2.2. In the HMR the situation is more complicated; we find that the correlation between the $\eta K_0^*(S\text{-wave})$ and $\eta K_2^*(1430)$ BFs is significant ($\sim 40\%$). In order to correctly account for the correlation during the averaging over sub-modes, we combine results using two simultaneous fits (Sec. 4.5.2.1), one for the charged modes and one for neutral modes as described later in this section.

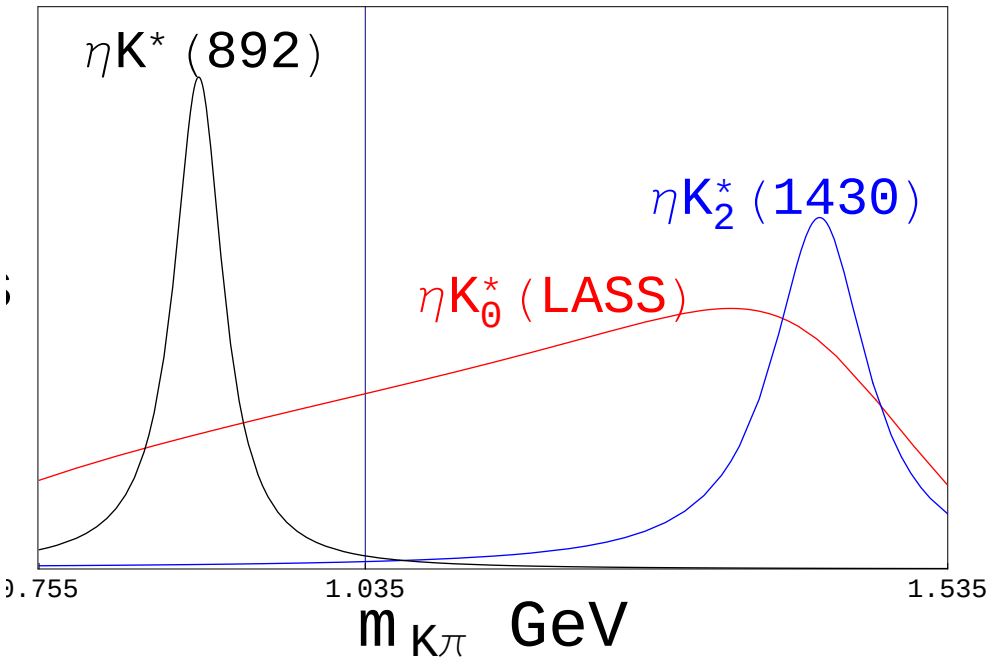


Figure 6.1: Lineshapes for $K_0^*(S\text{-wave})$, $K^*(892)$, and $K_2^*(1430)$. The vertical line denotes the dividing line between the low- and high-mass ranges. The normalizations were chosen solely for illustration.

6.1.4 Fixing yields in LMR and HMR

We fix the $\eta K_0^*(S\text{-wave})$ yield in the LMR fit for the $\eta K^*(892)$ branching fractions, and we fix the $\eta K^*(892)$ yield in the HMR fit for the $\eta K_0^*(S\text{-wave})$ and $\eta K_2^*(1430)$ branching fractions. The number of expected $\eta K_2^*(1430)$ events in the LMR fit ranges from 0 to 5 depending on sub-mode, but is always less than 1% of the $\eta K^*(892)$ signal yield.

We first use the relatively precise, existing branching fraction measurements for the $\eta K^*(892)$ modes to estimate the number of $\eta K^*(892)$ events expected in the HMR. We fix the $\eta K^*(892)$ yield and obtain $\eta K_2^*(1430)$ and $\eta K_0^*(S\text{-wave})$ branching fractions in the HMR fit. We then use these measured branching fractions to refit the $\eta K^*(892)$ branching fractions in the LMR fit. We then iterate one final time to ensure that the

process converges.

To demonstrate that this procedure is reasonable, we show a preview of LMR fit results in Table 6.1 in which the $\eta K_0^*(S\text{-wave})$ yield is fixed as described above and the results of the fit in which the $\eta K_0^*(S\text{-wave})$ yield is allowed to float. The floated $\eta K_0^*(S\text{-wave})$ yields are all within 2σ of the value determined as described above. (The $\eta K^*(892)$ yields are relatively stable, but vary by more than the 2% systematic error quoted in Sec. 6.14 because, in floating the $\eta K_0^*(S\text{-wave})$ yield, we have ignored information from the HMR fit.)

Table 6.1: Comparison of $\eta K_0^*(S\text{-wave})$ and $\eta K^*(892)$ yields from LMR fits in which the $\eta K_0^*(S\text{-wave})$ yield is fixed to a value extrapolated from the HMR fit (nominal) and a fit in which $\eta K_0^*(S\text{-wave})$ yield is allowed to float.

	$\eta K_0^*(S\text{-wave})$ Yield		$\eta K^*(892)$ Yield	
	Nominal	Float Y_{Swave}	Nominal	Float Y_{Swave}
$\eta_{\gamma\gamma} K^{*+}_{K^+\pi^0}$	34.5	1.3 ± 17.0	98.6 ± 16.2	110.1 ± 17.6
$\eta_{3\pi} K^{*+}_{K^+\pi^0}$	14.9	9.3 ± 11.4	56.2 ± 11.4	58.0 ± 12.0
$\eta_{\gamma\gamma} K^{*+}_{K^0_S\pi^+}$	39.8	18.8 ± 20.5	148.5 ± 19.4	157.3 ± 21.4
$\eta_{3\pi} K^{*+}_{K^0_S\pi^+}$	16.5	0.1 ± 12.9	35.8 ± 10.2	41.7 ± 11.6
$\eta_{\gamma\gamma} K^{*0}$	70.4	92.3 ± 30.2	407.0 ± 29.3	396.6 ± 32.2
$\eta_{3\pi} K^{*0}$	27.4	48.2 ± 18.1	110.8 ± 16.2	101.3 ± 17.5

6.2 Data and Monte Carlo samples

As mentioned in Sec. 4.2, the ηK^* analysis, performed in 2006, is based on data collected from 1999-2006. The integrated luminosity of the on-resonance sample is 312.6 fb^{-1} corresponding to 343.5 ± 3.8 million produced $B\overline{B}$ events. For $\eta K^*(892)$ signal MC, we use generated samples of 165,000 events for all modes. For $\eta K_0^*(S\text{-wave})$ signal MC, we use samples of $\sim 150,000$ events, and for $\eta K_2^*(1430)$ we use samples of $\sim 10,000$ events. (The technique for obtaining $\eta K_2^*(1430)$ MC, described in Sec. 6.4,

causes these samples to be smaller than the others.) For $B\overline{B}$ background studies, we use samples of ∼350 million B^+B^- and $B^0\overline{B}^0$ MC events.

6.3 The $K\pi$ S-wave

A generally accepted description of $K\pi$ S-wave production is still undetermined, though several *BABAR* analyses have addressed the issue at least briefly. The $B^\pm \to K^\pm\pi^\mp\pi^\pm$ Dalitz plot analysis measured the $B^+ \to K_0^{*0}(1430)\pi^+$ branching fraction, but was not sensitive to the parameters of the resonance shape due to a lack of statistics [68]. In the $B \to J/\psi K^*$ analysis, *BABAR* was able to extract the relative phase between $K\pi$ S- and P-waves by measuring the forward-backward asymmetry as a function of $K\pi$ mass [69].

For the maximum likelihood fit that we use to extract the $\eta K_0^*(S\text{-wave})$ branching fractions, we need PDFs for the $\eta K_0^*(S\text{-wave})$ component. Since we do not anticipate finding a number of signal events sufficient for determining the PDF parameterization, we rely on measurements made in other experiments.

6.3.1 $K\pi$ S-wave parametrization

In studies of $K\pi$ scattering in the reaction $K^-p \to K^-\pi^+n$, the LASS experiment at SLAC determined a parametrization of the $K\pi$ S-wave amplitude and phase [67, 70]. The total $K\pi$ S-wave scattering amplitude is the sum of non-resonant and resonant components which can be parameterized as

$$S = B\sin(\delta_B+\phi_B)e^{i(\delta_B+\phi_B)} + R\sin(\delta_R+\phi_R)e^{i(\delta_R+\phi_R)}e^{i\theta}, \quad (6.1)$$

where $B, R, \phi_B,$ and ϕ_R are real constants, and δ_B, δ_R, and θ are functions of $K\pi$ mass ($m_{K\pi}$). The first term describes the non-resonant contribution with a typical effective range parametrization,

$$\cot(\delta_B) = \frac{1}{aq} + \frac{rq}{2}, \quad (6.2)$$

where a is the scattering length, r is the effective range, and

$$q(m_{K\pi}) = \sqrt{\frac{(m_{K\pi}^2-(m_K+m_\pi)^2)(m_{K\pi}^2-(m_K-m_\pi)^2)}{4m_{K\pi}^2}}. \quad (6.3)$$

The second term describes the $K_0^*(1430)$ resonant contribution with a Breit-Wigner of mass M_R and width Γ_R,

$$\cot(\delta_R) = \frac{M_R^2-m_{K\pi}^2}{M_R\Gamma(m_{K\pi})}, \quad (6.4)$$

where $q_R = q(m_{K\pi}=M_R)$ and

$$\Gamma(m_{K\pi}) = \Gamma_R\frac{q}{m_{K\pi}}\frac{M_R}{q_R}. \quad (6.5)$$

The LASS data indicate that S-wave scattering is elastic up to $K\pi$ mass of 1.45 GeV where $\eta' K$ production begins. Elasticity and unitarity demand that the scattering amplitude, when viewed as a vector in the complex plane, be described by a circle of radius $\frac{1}{2}$ centered at $(0,\frac{i}{2})$. This is the case when $B=R=1$, $\phi_B=\phi_R=0$, and $\theta=2\delta_B$ in equation (6.1). These requirements on B, R, ϕ_B, and ϕ_R give

$$S = \sin(\delta_B)e^{i\delta_B}+\sin(\delta_R)e^{i\delta_R}e^{i\theta} \quad (6.6)$$

$$= \frac{1}{2i}[e^{2i\delta_B}-e^{i\theta}+e^{i(\delta_R+\frac{\theta}{2})}(e^{i(\delta_R+\frac{\theta}{2})}-e^{-i(\delta_R+\frac{\theta}{2})},)] \quad (6.7)$$

and the requirement on θ yields

$$S = \sin(\delta_B+\delta_R)e^{i(\delta_B+\delta_R)}. \quad (6.8)$$

The above expression demonstrates the unitarity of the scattering amplitude S. The invariant amplitude, T, in this context of S-wave, elastic scattering is proportional to S, $|T|\propto\frac{m_{K\pi}}{q}|S|$. One can write the invariant amplitude as a function of $m_{K\pi}$

$$T = \frac{m_{K\pi}}{q\cot\delta_B-iq}+e^{2i\delta_B}\frac{\Gamma_R\frac{M_R^2}{q_R}}{M_R^2-m_{K\pi}^2-iM_R\Gamma_R\frac{q}{m_{K\pi}}\frac{M_R}{q_R}}. \quad (6.9)$$

6.3.2 MC model of $\eta K_0^*(S\text{-wave})$

We use the standard *BABAR* simulation machinery to produce MC for $B \to \eta K \pi$ decays with a constant matrix element; i.e., flat over the $\eta K \pi$ Dalitz plot. We then use the invariant amplitude T for $m_{K\pi} < 1.8$ GeV and a standard Breit-Wigner for the high mass tail, $m_{K\pi} > 1.8$ GeV, to reweight this non-resonant $B \to \eta K \pi$ MC. We perform the importance sampling using the generated values of the Dalitz variables, $m^2_{K\pi}$ and $m^2_{\eta\pi}$, after skimming, ntuple-generation, pre-selection, and best-candidate selection. In this way, the effects of cuts and detector acceptance are automatically included in the final LASS-shaped $K\pi$ S-wave MC samples. The reweighting is model dependent; values for the four LASS parameters shown in Tab. 6.2, M_R, Γ_R, scattering length, and effective range, are taken from the fits to the LASS data performed by former LASS collaborator Bill Dunwoodie [70].

Table 6.2: Values of LASS parameters for the $K\pi$ S-wave.

Parameter	Value
Resonance Mass (M_R)	1435 ± 5 MeV
Resonance Width (Γ_R)	279 ± 6 MeV
Scattering Length (a)	1.95 ± 0.95
Effective Range (r)	1.76 ± 0.36

6.3.2.1 MC efficiency

We begin with samples of 782000 non-resonant $B \to \eta K \pi$ MC events. After all cuts and reweighting, we are left with approximately 1000 LASS-shaped $B \to \eta K_0^*(S\text{-wave})$ events. The efficiency for the entire process is $\sim 0.1\%$, including the reweighting efficiency, which one needs to remove to determine the final reconstruction

efficiency,

$$\epsilon \;=\; \frac{\#\text{ Events Selected}}{\#\text{ LASS Events Generated}},$$

where

$$\#\text{ LASS Events Generated} \;=\; \epsilon_{\text{reweighting}} \cdot (\#\text{ Non}-\text{res Events Generated}).$$

By reweighting an artificial MC sample generated with one event per 50×50 MeV^2 bin over the Dalitz plot, we find the reweighting efficiency, independent of the reconstruction and selection efficiencies, to be $\epsilon_{\text{reweighting}} = 9.2\%$, with negligible uncertainty. (The true reconstruction efficiency could have also been separated from reweighting efficiency using MC truth information, but this was unavailable to us since we have reconstructed non-resonant $B \to \eta K \pi$ as two-body $B \to \eta K^*$ decays.)

6.4 The $K_2^*(1430)$ resonance

The K_2^* resonance has the typical Breit-Wigner lineshape with world averages for mean and width of 1425.6 ± 1.5 MeV and 98.5 ± 2.7 MeV [62]. As described in Sec. 6.3 for $B \to \eta K_0^*(S\text{-wave})$, we obtain MC for $B \to \eta K_2^*(1430)$ decays by reweighting non-resonant $B \to \eta K \pi$ MC according to the appropriate mass (Breit-Wigner) and helicity ($(3\mathcal{H}^2 - 1)^2$) distributions. Modeling of these distributions is described in Sec. 6.8.

6.5 $K\pi$ helicity

The cosine of the helicity angle ($\mathcal{H} \equiv \cos\theta_{\mathcal{H}}$) of the $K\pi$ system is defined as the cosine of the angle between the direction of the kaon and the B in the $K\pi$ rest frame. In the case of decays of the scalar B meson to a pseudo-scalar and vector ($B \to PV$), a pseudo-scalar and a scalar ($B \to PS$), and a pseudo-scalar and a tensor ($B \to PT$), the vector, scalar, and tensor mesons have well-known helicity distributions. In $B \to PS$ decays, the expected $K_0^*(S\text{-wave})$ helicity distribution is flat. In $B \to PV$ decays, the

$K^*(892)$ helicity distribution is proportional to $P_1(\mathcal{H})^2 = \mathcal{H}^2$. Finally, in the case of $B \to PT$ decays, the tensor meson has a helicity distribution of $P_2(\mathcal{H})^2 = (3\mathcal{H}^2 - 1)^2$.

6.6 Event reconstruction and selection

A B candidate is formed by combining an η candidate with a K^* candidate. To reduce correlation between ΔE and resonance mass and to improve resolution on the decay vertex, the masses of resonances whose width is narrower than detector resolution (η and π^0) are constrained to the nominal value [62] during vertex fitting.

The resonances are reconstructed in the following way, except as noted in subsequent sections. (Particle lists are defined in Sec. 4.3.):

- K^* candidates are reconstructed with charged pions from the `GoodTracksVeryLoose` list, charged kaons from the `GoodTracksLoose` list, and K^0_S whose daughters come from the `ChargedTracks` list.
- $\eta_{\gamma\gamma}$ candidates are made from photons from the `GoodPhotonsLoose` list.
- Two `ChargedTracks` and a π^0 are combined to form an $\eta_{3\pi}$ candidate.
- Event shape quantities are computed with tracks from the `GoodTracksVeryLoose` list.

A number of preliminary cuts are applied before B candidates are considered for the maximum likelihood fit. They are:

- $N_{\rm trks} \geq \max[3, N_{\rm tracks\,in\,signal} + 1]$,
- $120 < m^{\pi^0}_{\gamma\gamma} < 150$ MeV,
- $490 < m^{\eta}_{\gamma\gamma} < 600$ MeV,
- $520 < m^{\eta}_{\pi\pi\pi} < 570$ MeV,

- $m_{K\pi}^{K^*(892)}$: $755 < m_{K\pi}^{K^*} < 1035$ MeV,
- $m_{K\pi}^{K^*(S-wave)}$: $1035 < m_{K\pi}^{K^*} < 1535$ MeV,
- $|\Delta E| \leq 0.2$ GeV,
- $5.25 \leq m_{\rm ES} < 5.2893$ GeV,
- $|\cos\theta_{\rm T}| \leq 0.9$,
- $-3 < \mathcal{F} < 4$.
- Particle identification: We reject events with tracks that satisfy the `Tight` criteria of the electron selector or the `VeryTight` criteria of the proton selector. Charged pion candidates in η candidates must NOT satisfy the `Tight` criteria of the kaon selector.
- For modes with $\eta \to \gamma\gamma$: $E_\gamma > 100\,\mathrm{MeV}$ and $|\cos\theta_{\rm dec}^{\eta}| \leq 0.86$, where $\theta_{\rm dec}^{\eta}$ is the η decay angle, the angle between the photon momentum and the B momentum in the η rest frame. This cut reduces background from modes in which a hard, daughter photon combines with a soft photon to create an η, most notably $B \to K^*\gamma$.
- For modes with $\eta \to \pi^+\pi^-\pi^0$, the π^0 photons satisfy $E_\gamma > 30\,\mathrm{MeV}$.
- We require $E_{\pi^0} > 250\,\mathrm{MeV}$ for all neutral pions.
- For the LMR $\eta K^*(892)$ analysis, helicity cuts are
  - $K^{*+} \to K^+\pi^0$: $-0.7 < \mathcal{H} < 1.0$,
  - $K^{*+} \to K^0\pi^+$: $-0.95 < \mathcal{H} < 1.0$,
  - $K^{*0} \to K^+\pi^-$: $-0.95 < \mathcal{H} < 1.0$.

The cut at $\mathcal{H} > -0.7$ in $K\pi^0$ modes reduces slow π^0 backgrounds. The cut at $\mathcal{H} > -0.95$ in the other modes eliminates the region affected by a drop in efficiency from the slow π^+ momentum cut-off.

- For the HMR analysis, the helicity cut for all modes is $-0.5 < \mathcal{H} < 1.0$. Helicity is flat in the $\eta K_0^*(S\text{-wave})$ modes, and strongly peaked at $\mathcal{H} < -0.5$ for continuum, charmless $B\overline{B}$, and $K^*(892)$ backgrounds. Helicity distributions for the $\eta K_2^*(1430)$ modes are proportional to $(3\mathcal{H}^2 - 1)^2$, but cuts were chosen for the $\eta K_0^*(S\text{-wave})$ modes for which there is a greater chance of measuring a branching fraction.

- For modes in which it is relevant, the K_S^0 lifetime significance $(\tau/\sigma_\tau) > 3$.

6.6.1 Multiple candidates per event

In events with more than one B candidate passing all cuts, we choose the candidate with the η mass closest to the nominal value. From MC we determine that, depending on decay mode, we choose a misreconstructed candidate $14 - 37\%$ of the time. These self-crossfeed events tend to exhibit correlations in fit variables which can bias the fit. In Sec. 6.11.1, we find the bias to be manageable ($\sim 10\%$ of expected signal yield), and we correct for the bias in the final results.

6.7 Maximum likelihood fit

The LMR analysis consists of one unbinned multivariate maximum likelihood fit (Sec. 4.5.1.1) for each of six sub-modes. For each input event, the likelihood ($\mathcal{L}_i$) can be written

$$\mathcal{L}_i = \sum_{j=1}^{m} n_j \mathcal{P}_j(\mathbf{x}_i) \ , \tag{6.10}$$

where $\mathcal{P}_j(\mathbf{x}_i)$ are the probabilities for signal and background parameters evaluated with the observables $\mathbf{x}_i$ of the ith event, and n_j are the numbers of events in the full sample

for each hypothesis. For N input events, the overall likelihood is

$$\mathcal{L} = \frac{e^{-(\sum n_j)}}{N!} \prod_{i=1}^{N} \mathcal{L}_i. \tag{6.11}$$

In the LMR, the components j include signal $\eta K^*(892)$, charmless $B\overline{B}$ background, continuum background, and $\eta K_0^*(S\text{-wave})$ background.

The HMR analysis is similar except the sub-mode likelihoods are combined in a simultaneous ML fit as described in Sec. 4.5.2.1. In the HMR analysis, the $\eta K_0^*(S\text{-wave})$ and $\eta K_2^*(1430)$ components are considered signal, and the background components are $\eta K^*(892)$, continuum, and charmless $B\overline{B}$. A charmless $B\overline{B}$ background component is not used for modes which lack significant contribution to signal ($> 1\%$) from charmless $B\overline{B}$ backgrounds.

The discriminating variables used in the fits are ΔE, $m_{\rm ES}$, the invariant masses of the resonance candidates, m_η and m_{K^*}, an event shape Fisher discriminant ($\mathcal{F}$), and the cosine of the K^* helicity angle $\mathcal{H}$. The helicity variable is described in Sec. 6.6, and the other quantities are described in Chap. 4.

6.8 Probability density functions

As described in Sec. 4.5.3, one PDF must be determined for each discriminating variable for each fit component. We use the six variables listed immediately above for four components: signal $\eta K^*(892)$, continuum background, and charmless $B\overline{B}$ background, and $\eta K_0^*(S\text{-wave})$ background in the LMR and signal $\eta K_0^*(S\text{-wave})$, signal $\eta K_2^*(1430)$, continuum background, and charmless $B\overline{B}$ background in the HMR. Notice that, depending on the fit (LMR or HMR) an ηK^* component might be considered signal or background. When describing the implementation, we make this distinction by referring to *signal-* or *background-*$\eta K^*(892)$. The data samples used for determining the PDFs are

- fully-simulated, exclusive MC for $\eta K^*(892)$ (with data control samples to de-

termine necessary corrections to the MC),

- exclusive MC from non-resonant $\eta K\pi$ reweighted using LASS (Sec. 6.3) or Breit-Wigner parameterizations for $\eta K_0^*(S\text{-wave})$ and $\eta K_2^*(1430)$ PDFs,
- data sidebands for the continuum background,
- exclusive MC from individual modes for the charmless $B\overline{B}$ background.

PDFs for all modes are shown in Appendix H. In all cases, the best candidate is chosen before the PDFs are determined.

6.8.1 $m_{\mathbf{ES}}$

We parametrize each signal- and background-ηK^* $m_{\rm ES}$ distribution with a double Gaussian function. For continuum background we use the ARGUS function as described in Sec. 5.6.1. Each charmless $B\overline{B}$ component is fit with the sum of an ARGUS function and a Gaussian.

6.8.2 ΔE

We use a double Gaussian shape to parameterize all signal- and background-ηK^* as well as charmless $B\overline{B}$ distributions. (Background-ηK^* decays are signal-like since they are correctly reconstructed B decays.) Studies of control samples have shown that the MC is a reasonable representation of the data, but that the MC must be scaled by a factor of 1.05 to have the appropriate width. For continuum background we use a first or second order Chebyshev polynomial, depending on sub-mode.

6.8.3 Fisher discriminant

For all components we parameterize the distribution of $\mathcal{F}$ with an asymmetric Gaussian; i.e., a Gaussian with different widths on the left and right sides of its peak.

Modes with high background often exhibit outlier events which require an additional contribution to the continuum background PDF. We model this effect by adding another Gaussian, depending on the amount of background.

6.8.4 Resonance masses

We obtain the PDFs for the invariant masses of resonances from MC. The term *resonance* refers to η, $K^*(892)$, $K_0^*(S\text{-wave})$, and $K_2^*(1430)$ candidates. (For π^0 candidates, we simply make cuts on the mass distributions and do not fit the mass spectra because we find that the candidates for these particles in the background are dominantly real, so the fit would serve only to de-weight an already small combinatorial background.)

In both signal- and background-$\eta K^*(892)$ components, an η mass shape is fit with a double Gaussian. A signal $K^*(892)$ shape is fit with a Breit-Wigner (or Breit-Wigner plus Gaussian), a signal $K_0^*(S\text{-wave})$ LASS shape with a double Gaussian, and a signal $K_2^*(1430)$ shape with a double Gaussian. For background-$\eta K^*(892)$ components, the $K\pi$ mass distribution is fit with a first or second order Chebyshev polynomial.

6.8.4.1 Resonance components in $q\bar{q}$ and $B\overline{B}$ background

The shape of resonance mass distributions in continuum and charmless $B\overline{B}$ background results from the sum of combinatorial background and real resonance production. We fit these distributions with the sum of the signal shape, to account for the real resonance production, and a first degree Chebyshev polynomial to account for the combinatorial background. The fraction of events with real resonance is floated in the ML fit. The parameters for the true resonance component are fixed to those found for the signal component, after application of the appropriate scale factors and offsets.

6.8.4.2 Resolution measurements

As discussed in Sec. 5.17, we study the invariant mass resolutions in the data using real η and K^* mesons in the continuum background. Where there are differences between MC and data, we determine the appropriate values by which to shift the mean and scale the width of the peaking distribution. For resonances fit with double Gaussians, we apply the shift to both Gaussian components, but scale only the width of the core Gaussian.

We fix the shifts and scale factors to those values listed in Table 4.1. In the case of the $K^*(892)$, the shift and scale factor are determined to be zero and unity within errors. Because the $K_0^*(S\text{-wave})$ lineshape is very wide (300 MeV) and still not well understood, it is difficult to determine whether a shift or scale is needed. Since we employ neither shift nor scale factor for the more well defined $K^*(892)$, we use neither for the $K_0^*(S\text{-wave})$ or $K_2^*(1430)$ shapes.

6.8.5 Resonance helicity

The cosine of the helicity angle of a K^* meson ($\mathcal{H}$) is defined in Sec. 6.5. For continuum and charmless $B\overline{B}$ backgrounds, we expect that $\mathcal{H}$ would have a nearly flat distribution, corresponding to a sum of combinatorial resonance background and background of true resonances from generic production mechanisms. The continuum and charmless background components are modeled by Chebyshev polynomials of varying degree to account for true helicity structure and acceptance effects. In $B \to PS$ decays, the expected flat helicity distribution is modeled with a second order polynomial to account for acceptance effects. In $B \to PV$ decays, the $\mathcal{H}^2$ helicity distribution is modeled with a second or higher order polynomial as is necessary to account for acceptance effects. Finally, in the case of $B \to PT$ decays, the $(3\mathcal{H}^2 - 1)^2$ helicity distribution of the tensor meson is modeled with a fifth order polynomial to account for a small kaon

momentum *roll-off*; i.e., the observed helicity distribution near $\mathcal{H} = 1$ is slightly lower than the theoretical distribution due to a low acceptance of low-momentum K^0_S.

6.9 Efficiency and production rate

The event yield obtained with the ML fit is converted into a branching fraction using the detection efficiency ϵ, the number of produced $B\overline{B}$ pairs $N_{B\bar{B}}$, and the fit bias through the relationship

$$\text{branching fraction} \quad = \quad \frac{\text{yield} - \text{fit bias}}{N_{B\bar{B}} \cdot \epsilon}. \tag{6.12}$$

The uncorrected MC efficiency is the ratio of the number of signal MC events passing selection and the number of generated signal MC events. To obtain the final efficiency, the MC efficiency is corrected for differences between the true efficiency and that of the simulation.

The *BABAR* tracking group studies data and MC for $e^+e^- \to \tau^+\tau^-$ events in which one τ decays via three tracks and the other τ decays with a single hard electron that is used to identify the event as a $\tau^+\tau^-$ event. The tracking group determines that the data and MC agree within errors of 0.54% for `GoodTracksLoose`, 0.45% for `GoodTracksVeryLoose`, and 0.57% for `ChargeTracks` [71]. For the K^0_S efficiency corrections, we follow the recipe provided by the tracking group [72], which accounts for the K^0_S flight length, flight direction, and transverse momentum, and we apply a correction (data/MC) of 0.983 ± 0.019.

The *BABAR* neutral identification and reconstruction group isolates the effects of π^0 reconstruction by comparing rates in data and MC for tagged τ pair events in which one τ decays via $\tau^+ \to \pi^+\nu$ and the rates for $\tau^+ \to \rho^+\nu$ (with $\rho^+ \to \pi^+\pi^0$). We follow the recipe [73] of the neutrals group for correcting the neutral reconstruction efficiency and smearing the MC to make photon energy resolution match the data. Based on our pion selection requirements, we use a correction (data/MC) of 0.968 ± 0.030. There

was no dedicated study of the $\eta \to \gamma\gamma$ reconstruction efficiency but the momentum dependence for the π^0 correction is essentially flat over the relevant range, so we use the correction prescribed for the π selection most similar to our η selection, 0.984 ± 0.036 (data/MC).

As described in Sec. 3.3.8, likelihood-based particle identification selectors are used to veto tracks of unwanted species. During data processing we apply the efficiency corrections as prescribed by the *BABAR* PID group [74]. Finally, the efficiency must be corrected for resonance branching fractions since the sub-decay of interest is forced in generation of signal MC. The relevant efficiency information is summarized in the results tables in Sec. 6.12.

We determine the integrated luminosity using $e^+e^- \to \mu^+\mu^-$, $e^+e^- \to e^+e^-$, and other QED processes with an overall uncertainty of 1.1%; the number of produced $B\overline{B}$ pairs is determined by comparing the ratio of $e^+e^- \to q\overline{q}$ events and $\mu^+\mu^-$ events in on-resonance and off-resonance samples with the assumption that the difference results entirely from $b\overline{b}$ production at the $\Upsilon(4S)$. We assume equal rates of production for charged and neutral B mesons.

6.10 Charmless $B\overline{B}$ backgrounds

We perform a variety of studies related to charmless $B\overline{B}$ backgrounds. We have applied the full analysis selection to the generic $B\overline{B}$ MC sample, removing $b \to c$ and signal decays in order to focus on the most troublesome charmless backgrounds. (We observe that PDFs for charm $B\overline{B}$ and continuum events are very similar and that charm and continuum yields are highly correlated. For this reason, we did not use separate charm components these analyses, and we allow the charm events to be absorbed into the yield of the dominant continuum background.) These studies provide a list of charmless $B\overline{B}$ modes for further study. We obtain exclusive MC samples of $\gtrsim 100$K events for each of the modes and mix these events in appropriate proportions according

to efficiency and branching fraction. Where BF measurements are unavailable we use estimates from theory.

The results of this procedure for the LMR analysis are shown in Table 6.3 for the $\eta K^{*+}_{K^+\pi^0}$ decays, Table 6.4 for the $\eta K^{*+}_{K^0_S\pi^+}$ decays, and Table 6.5 for the $\eta K^{*0}(892)$ decays. For the HMR analysis, results are shown in Table 6.6 for the $\eta K^{*+}_{K^+\pi^0}$ decays, Table 6.7 for the $\eta K^{*+}_{K^0_S\pi^+}$ decays, and Table 6.8 for the $\eta K^{*0}(892)$ decays.

In the LMR analysis, the expected charmless $B\overline{B}$ background is negligible for the $\eta_{3\pi}K^{*+}_{K^+\pi^0}$ (3.2 events) and $\eta_{3\pi}K^{*+}_{K^0_S\pi^+}$ (6.2 events) modes. (See Tables 6.3 and 6.4.) These modes are fit without a charmless $B\overline{B}$ component, but we embed charmless events in toy studies so that the effect of these backgrounds is included in the fit bias. For the HMR analysis, a charmless $B\overline{B}$ component is not necessary for any $\eta_{3\pi}$ modes in which the largest expected contribution is 1.1 events (see Tables 6.6, 6.7, and 6.8).

We also investigate the utility of employing separate fit components for charmless $B\overline{B}$ backgrounds with unique PDFs or especially large contributions to the total expected background. Please see Appendix I for details of the study.

Table 6.3: Potential background for the $\eta K^{*+}_{K^+\pi^0}$ modes, when requiring $0.755 < m_{K\pi} < 1.035$ GeV, from exclusive charmless B decays. We show efficiency for the mode to pass selection cuts, the measured or estimated branching fraction, the appropriate product branching fraction given how the MC was produced, the estimated background normalized to 324.0 million $B\overline{B}$ events and the number of events we include in the file we use for making PDFs. A * denotes that the branching fraction is estimated since no measurement exists.

Signal mode Bkg. channel	Mode Number	MC ϵ (%)	Est. $\mathcal{B}$ (10^{-6})	$\prod \mathcal{B}_i$ (%)	Norm. # $B\overline{B}$ Bkg.	# in PDF Bkg. file
$\eta_{\gamma\gamma} K^{*+}_{K^+\pi^0}(892)$						
$B^+ \to K^{*+}\gamma$	3135	0.24	$40.3^{+2.6}_{-2.6}$	1.000	31.4	502
$B^0 \to \eta_{\gamma\gamma} K^{*0}_{K^+\pi^-}$	1541	0.5	$18.7^{+1.7}_{-1.7}$	0.262	7.9	127
$B^+ \to K_1^{*+}(1273)\gamma$	1453	0.03	$43^{+12.0}_{-12.0}$	1.000	4.3	69
$B^0 \to K_1^{*0}(1273)\gamma$	1452	0.06	16^*	1.000	2.9	46
$B^0 \to K^{*0}(1410)\gamma$	1971	0.06	15.0^*	1.000	2.7	43
$B^0 \to K^{*0}_{K^+\pi^-}\gamma$	1710	0.03	$40.1^{+2.0}_{-2.0}$	0.667	2.6	41
$B^+ \to K^{*+}(1410)\gamma$	1970	0.03	15.0^*	1.000	1.7	26
$B^0 \to \eta\ K^+\pi^-$	5575	0.02	20^*	1.000	1.5	24
$B^0 \to K_2^{*0}(1430)\gamma$	1766	0.03	$12.4^{+2.4}_{-2.4}$	1.000	1.4	22
$B^0 \to K_1^{*0}(1402)\gamma$	1455	0.06	$6^{+6.0}_{-6.0}$	1.000	1.2	19
$B^+ \to K_1^{*+}(1402)\gamma$	1454	0.05	$7.5^{+7.5}_{-7.5}$	1.000	1.1	18
$B^+ \to K_2^{*+}(1430)\gamma$	1765	0.02	$14.5^{+4.3}_{-4.3}$	1.000	0.9	13
$B^+ \to a_0^0(\eta_{\gamma\gamma}\pi^0)\ K^{*+}_{K^+\pi^0}$	6560	0.54	1^*	0.131	0.2	3
$B^0 \to K^{*0}(1680)\gamma$	1973	0.02	2^*	1.000	0.2	2
$B^0 \to \rho^- K^{*+}_{K^+\pi^0}(L)$	2499	0.06	2^*	0.333	0.1	2
$B^0 \to \rho^- K^{*+}_{K^+\pi^0}(T)$	2500	0.01	4^*	0.333	0.1	0
					60.2	957
$\eta_{3\pi} K^{*+}_{K^+\pi^0}(892)$						
$B^0 \to \eta_{3\pi} K^{*0}_{K^+\pi^-}$	1540	0.34	$18.7^{+1.7}_{-1.7}$	0.151	3.1	214
$B^+ \to a_1^0 K^{*+}(K^+\pi^0)(L)$	5327	0.01	10^*	0.333	0.1	4
$B^+ \to \omega\ K^{*+}_{K^+\pi^0}(L, f_L = 1)$	2503	0.03	$0.6^{+1.8}_{-0.6}$	0.297	0	1
					3.2	219

Table 6.4: Potential background for the $\eta K^{*+}_{K^0_S\pi^+}$ modes, when requiring $0.755 < m_{K\pi} < 1.035$ GeV, from exclusive charmless B decays. See Table 6.3 for an explanation of column headers. A * denotes that the branching fraction is estimated since no measurement exists.

Signal mode Bkg. channel	Mode Number	MC ϵ (%)	Est. $\mathcal{B}$ (10^{-6})	$\prod \mathcal{B}_i$ (%)	Norm. # $B\overline{B}$ Bkg.	# in PDF Bkg. file
$\eta_{\gamma\gamma} K^{*+}_{K^0_S\pi^+}(892)$						
$B^+ \to K^{*+}\gamma$	3135	0.21	$40.3^{+2.6}_{-2.6}$	1.000	27.1	433
$B^0 \to \eta_{\gamma\gamma}\ K^{*0}_{K_S\pi^0}$	6564	4.31	$18.7^{+1.7}_{-1.7}$	0.045	11.7	188
$B^+ \to K_1^{*+}(1273)\gamma$	1453	0.04	$43^{+12.0}_{-12.0}$	1.000	5.9	95
$B^0 \to K^{*0}(1410)\gamma$	1971	0.06	15.0*	1.000	2.8	45
$B^0 \to K_1^{*0}(1273)\gamma$	1452	0.05	16*	1.000	2.6	42
$B^0 \to K^{*0}_{K_S\pi^0}\gamma$	5197	0.16	$40.1^{+2.0}_{-2.0}$	0.115	2.4	37
$B^+ \to K^{*+}(1410)\gamma$	1970	0.03	15.0*	1.000	1.5	24
$B^0 \to \eta_{\gamma\gamma} K_S$	2603	2.26	$1.5^{+0.7}_{-0.7}$	0.136	1.5	23
$B^+ \to K_2^{*+}(1430)\gamma$	1765	0.03	$14.5^{+4.3}_{-4.3}$	1.000	1.4	21
$B^0 \to K_2^{*0}(1430)\gamma$	1766	0.03	$12.4^{+2.4}_{-2.4}$	1.000	1.3	20
$B^0 \to K_1^{*0}(1402)\gamma$	1455	0.06	$6^{+6.0}_{-6.0}$	1.000	1.2	19
$B^+ \to K_1^{*+}(1402)\gamma$	1454	0.04	$7.5^{+7.5}_{-7.5}$	1.000	0.9	14
$B^0 \to a_0^0(\eta_{\gamma\gamma}\pi^0)K_S$	6547	1.56	1*	0.136	0.7	11
$B^0 \to \eta_{\gamma\gamma} K^{*0}_{K^+\pi^-}$	1541	0.04	$18.7^{+1.7}_{-1.7}$	0.262	0.7	10
$B^0 \to K^{*0}(1680)\gamma$	1973	0.02	2*	1.000	0.1	2
$B^+ \to K^{*+}(1680)\gamma$	1972	0.02	2*	1.000	0.1	2
					61.9	986
$\eta_{3\pi} K^{*+}_{K^0_S\pi^+}(892)$						
$B^0 \to \eta_{3\pi}\ K^{*0}_{K_S\pi^0}$	6565	2.92	$18.7^{+1.7}_{-1.7}$	0.025	4.4	405
$B^0 \to \eta'_{\eta\pi\pi} K_S$	1510	0.07	$64.9^{+3.5}_{-3.5}$	0.060	0.9	81
$B^0 \to \eta_{3\pi} K_S$	2605	1.62	$1.5^{+0.7}_{-0.7}$	0.078	0.6	56
$B^0 \to \eta_{3\pi} K^{*0}_{K^+\pi^-}$	1540	0.03	$18.7^{+1.7}_{-1.7}$	0.151	0.3	26
					6.2	568

Table 6.5: Potential background for the $\eta K^{*0}_{K^+\pi^-}$ modes, when requiring $0.755 < m_{K\pi} < 1.035$ GeV, from exclusive charmless B decays. See Table 6.3 for an explanation of column headers. A * denotes that the branching fraction is estimated since no measurement exists.

Signal mode Bkg. channel	Mode Number	MC ϵ (%)	Est. $\mathcal{B}$ (10^{-6})	$\prod \mathcal{B}_i$ (%)	Norm. # $B\overline{B}$ Bkg.	# in PDF Bkg. file
$\eta_{\gamma\gamma} K^{*0}_{K^+\pi^-}(892)$						
$B^0 \to K^{*0}_{K^+\pi^-}\gamma$	1710	0.91	$40.1^{+2.0}_{-2.0}$	0.667	78.5	1256
$B^+ \to \eta_{\gamma\gamma} K^{*+}_{K^+\pi^0}$	1539	2.25	$24.3^{+3.0}_{-2.9}$	0.131	23.2	372
$B^+ \to K^{*+}_1(1273)\gamma$	1453	0.1	$43^{+12.0}_{-12.0}$	1.000	14.4	229
$B^+ \to K^{*+}(1410)\gamma$	1970	0.13	15.0*	1.000	6.5	104
$B^0 \to K^{*0}_1(1273)\gamma$	1452	0.11	16*	1.000	5.6	90
$B^0 \to K^{*0}(1410)\gamma$	1971	0.1	15.0*	1.000	4.6	74
$B^+ \to K^{*+}_1(1402)\gamma$	1454	0.16	$7.5^{+7.5}_{-7.5}$	1.000	4	63
$B^+ \to \eta_{\gamma\gamma} K^+$	1513	1.12	$2.5^{+0.3}_{-0.3}$	0.394	3.6	57
$B^0 \to K^{*0}_{K^+\pi^-}\pi^0$	1225	0.75	$1.7^{+0.8}_{-0.8}$	0.667	2.8	44
$B^+ \to K^{*+}_2(1430)\gamma$	1765	0.06	$14.5^{+4.3}_{-4.3}$	1.000	2.7	43
$B^0 \to K^{*0}_2(1430)\gamma$	1766	0.06	$12.4^{+2.4}_{-2.4}$	1.000	2.3	36
$B^0 \to K^{*0}_1(1402)\gamma$	1455	0.1	$6^{+6.0}_{-6.0}$	1.000	2	31
$B^+ \to a^0_0(\eta_{\gamma\gamma}\pi^0)K^+$	6552	0.9	1*	0.394	1.1	18
$B^0 \to a^0_0(\eta_{\gamma\gamma}\pi^0)\ K^{*0}_{K^+\pi^-}$	6562	1.05	1*	0.263	0.9	14
$B^0 \to K^{*0}(1680)\gamma$	1973	0.05	2*	1.000	0.3	5
$B^+ \to K^{*+}(1680)\gamma$	1972	0.05	2*	1.000	0.3	4
$B^0 \to \eta_{\gamma\gamma}\phi_{K^+K^-}$	2904	4.72	0.0*	0.193	0	0
					152.8	2440
$\eta_{3\pi} K^{*0}_{K^+\pi^-}(892)$						
$B^+ \to \eta_{3\pi} K^{*+}_{K^+\pi^0}$	1542	1.56	$24.3^{+3.0}_{-2.9}$	0.075	9.2	209
$B^+ \to \eta_{3\pi} K^+$	1515	0.92	$2.5^{+0.3}_{-0.3}$	0.226	1.7	38
$B^+ \to \eta'_{\eta_{3\pi^0}\pi\pi}\ K^+$	6557	0.05	$69.7^{+2.8}_{-2.7}$	0.144	1.6	37
$B^+ \to \eta'_{\eta\pi\pi} K^+$	1506	0.04	$69.7^{+2.8}_{-2.7}$	0.174	1.4	31
$B^+ \to a^0_0(\eta_{3\pi}\pi^0)K^+$	6553	0.52	1*	0.226	0.4	8
$B^0 \to a^0_0(\eta_{3\pi}\pi^0)\ K^{*0}_{K^+\pi^-}$	6563	0.78	1*	0.151	0.4	8
$B^0 \to a^0_1 K^{*0}(L, f_L = 0.7)$	5329	0.02	10*	0.467	0.4	8
$B^0 \to a^-_0(\eta_{3\pi}\pi^-)K^+$	6546	0.49	1*	0.226	0.4	8
$B^0 \to \omega\ K^{*0}_{K^+\pi^-}(L, f_L = 1)$	2507	0.06	$2.4^{+1.3}_{-1.3}$	0.594	0.3	6
$B^0 \to a^0_1 K^{*0}(T, f_L = 0.7)$	5330	0.02	10*	0.200	0.1	2
$B^+ \to \eta'_{\eta\pi\pi} K^{*+}_{K^+\pi^0}$	1503	0.03	$4.9^{+2.1}_{-1.9}$	0.058	0	0
$B^0 \to \eta_{3\pi}\rho^0$	2314	0.44	0.0*	0.226	0	0
					15.9	355

Table 6.6: Potential background for the $\eta K^{*+}_{K^+\pi^0}$ modes, when requiring $1.035 < m_{K\pi} < 1.535$ GeV, from exclusive charmless B decays. See Table 6.3 for an explanation of column headers. A * denotes that the branching fraction is estimated since no measurement exists.

Signal mode Bkg. channel	Mode Number	MC ϵ (%)	Est. $\mathcal{B}$ (10^{-6})	$\prod \mathcal{B}_i$ (%)	Norm. # $B\overline{B}$ Bkg.	# in PDF Bkg. file
$\eta_{\gamma\gamma} K^{*+}_{K^+\pi^0}(S\text{-wave})$						
$B^+ \to K_2^{*+}(1430)\gamma$	1765	0.09	$14.5^{+4.3}_{-4.3}$	1.000	4.3	69
$B^+ \to K^{*+}\gamma$	3135	0.02	$40.3^{+2.6}_{-2.6}$	1.000	2.6	41
$B^0 \to K^{*0}(1410)\gamma$	1971	0.03	15.0*	1.000	1.4	21
$B^+ \to K^{*+}(1410)\gamma$	1970	0.03	15.0*	1.000	1.4	21
$B^0 \to K_1^{*0}(1273)\gamma$	1452	0.02	16*	1.000	1.2	18
$B^+ \to K_1^{*+}(1273)\gamma$	1453	0.01	$43^{+12.0}_{-12.0}$	1.000	1	16
$B^0 \to K_2^{*0}(1430)\gamma$	1766	0.02	$12.4^{+2.4}_{-2.4}$	1.000	0.8	12
$B^0 \to K_1^{*0}(1402)\gamma$	1455	0.03	$6^{+6.0}_{-6.0}$	1.000	0.5	8
$B^0 \to \eta\, K^+\pi^-$	5575	0.01	20*	1.000	0.5	8
$B^0 \to K^{*0}_{K^+\pi^-}\gamma$	1710	0.01	$40.1^{+2.0}_{-2.0}$	0.667	0.5	7
$B^+ \to K_1^{*+}(1402)\gamma$	1454	0.02	$7.5^{+7.5}_{-7.5}$	1.000	0.4	6
$B^0 \to \eta_{\gamma\gamma} K^{*0}_{K^+\pi^-}$	1541	0.02	$18.7^{+1.7}_{-1.7}$	0.262	0.4	5
$B^0 \to K^{*0}(1680)\gamma$	1973	0.04	2*	1.000	0.3	4
$B^+ \to \pi^+\pi^0\pi^0$	1938	0.01	10*	1.000	0.2	3
$B^+ \to K^{*+}(1680)\gamma$	1972	0.03	2*	1.000	0.2	3
$B^+ \to a_0^0(\eta_{\gamma\gamma}\pi^0)\rho^+$	6678	0.12	1*	0.394	0.1	2
					15.8	244
$\eta_{3\pi} K^{*+}_{K^+\pi^0}(S\text{-wave})$						
$B^0 \to \eta_{3\pi} K^{*0}_{K^+\pi^-}$	1540	0.02	$18.7^{+1.7}_{-1.7}$	0.151	0.2	15
					0.2	15

Table 6.7: Potential background for the $\eta K^{*+}_{K^0_S\pi^+}$ modes, when requiring $1.035 < m_{K\pi} < 1.535$ GeV, from exclusive charmless B decays. See Table 6.3 for an explanation of column headers. A * denotes that the branching fraction is estimated since no measurement exists.

Signal mode Bkg. channel	Mode Number	MC ϵ (%)	Est. $\mathcal{B}$ (10^{-6})	$\prod \mathcal{B}_i$ (%)	Norm. # $B\overline{B}$ Bkg.	# in PDF Bkg. file
$\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}(S\text{-wave})$						
$B^+ \to K_2^{*+}(1430)\gamma$	1765	0.08	$14.5^{+4.3}_{-4.3}$	1.000	3.7	58
$B^+ \to K^{*+}\gamma$	3135	0.02	$40.3^{+2.6}_{-2.6}$	1.000	2.5	39
$B^+ \to K_1^{*+}(1273)\gamma$	1453	0.02	$43^{+12.0}_{-12.0}$	1.000	2.2	35
$B^0 \to K^{*0}(1410)\gamma$	1971	0.03	15.0*	1.000	1.4	22
$B^0 \to K_1^{*0}(1273)\gamma$	1452	0.03	16*	1.000	1.3	20
$B^+ \to K^{*+}(1410)\gamma$	1970	0.02	15.0*	1.000	1.1	17
$B^0 \to K_2^{*0}(1430)\gamma$	1766	0.02	$12.4^{+2.4}_{-2.4}$	1.000	0.8	12
$B^0 \to K_1^{*0}(1402)\gamma$	1455	0.03	$6^{+6.0}_{-6.0}$	1.000	0.6	9
$B^0 \to \eta_{\gamma\gamma}\ K^{*0}_{K_S\pi^0}$	6564	0.19	$18.7^{+1.7}_{-1.7}$	0.045	0.5	8
$B^+ \to K_1^{*+}(1402)\gamma$	1454	0.01	$7.5^{+7.5}_{-7.5}$	1.000	0.3	5
$B^+ \to K^{*+}(1680)\gamma$	1972	0.03	2*	1.000	0.2	3
$B^0 \to K^{*0}(1680)\gamma$	1973	0.02	2*	1.000	0.1	2
$B^0 \to \eta_{\gamma\gamma}K_S$	2603	0.2	$1.5^{+0.7}_{-0.7}$	0.136	0.1	2
$B^+ \to a_0^0(\eta_{\gamma\gamma}\pi^0)\ K^{*+}_{K_S\pi^+}$	6559	0.11	1*	0.090	0	0
$B^0 \to a_0^0(\eta_{\gamma\gamma}\pi^0)K_S$	6547	0.05	1*	0.136	0	0
$B^+ \to a_0^+(\eta_{\gamma\gamma}\pi^+)K_S$	6539	0.03	1*	0.136	0	0
$B^0 \to \eta_{\gamma\gamma}\rho^0$	2313	0.26	0.0*	0.394	0	0
					14.8	232
$\eta_{3\pi}K^{*+}_{K^0_S\pi^+}(S\text{-wave})$						
$B^0 \to \eta_{3\pi}K^{*0}_{K^+\pi^-}$	1540	0.06	$18.7^{+1.7}_{-1.7}$	0.151	0.6	39
$B^0 \to \eta\ K_S\pi^0$	6426	0.02	20*	0.230	0.3	21
$B^0 \to \eta_{3\pi}\ K^{*0}_{K_S\pi^0}$	6565	0.13	$18.7^{+1.7}_{-1.7}$	0.025	0.2	13
$B^0 \to \eta'_{\eta_{3\pi^0}\pi\pi}K_S$	6556	0.01	$64.9^{+3.5}_{-3.5}$	0.050	0.1	4
$B^0 \to \eta_{3\pi}K_S$	2605	0.12	$1.5^{+0.7}_{-0.7}$	0.078	0	3
$B^+ \to a_0^0(\eta_{3\pi}\pi^0)K_S$	6548	0.04	1*	0.078	0	0
					1.2	80

Table 6.8: Potential background for the $\eta K^{*0}_{K^+\pi^-}$ modes, when requiring $1.035 < m_{K\pi} < 1.535$ GeV, from exclusive charmless B decays. See Table 6.3 for an explanation of column headers. A * denotes that the branching fraction is estimated since no measurement exists.

Signal mode Bkg. channel	Mode Number	MC ϵ (%)	Est. $\mathcal{B}$ (10^{-6})	$\prod \mathcal{B}_i$ (%)	Norm. # $B\overline{B}$ Bkg.	# in PDF Bkg. file
$\eta_{\gamma\gamma} K^{*0}_{K^+\pi^-}$ (S-wave)						
$B^0 \to K^{*0}_2(1430)\gamma$	1766	0.24	$12.4^{+2.4}_{-2.4}$	1.000	9.5	151
$B^0 \to K^{*0}_{K^+\pi^-}\gamma$	1710	0.09	$40.1^{+2.0}_{-2.0}$	0.667	7.4	118
$B^+ \to K^{*+}_1(1273)\gamma$	1453	0.03	$43^{+12.0}_{-12.0}$	1.000	4	63
$B^0 \to K^{*0}(1410)\gamma$	1971	0.08	15.0^*	1.000	3.7	58
$B^+ \to \eta_{\gamma\gamma} K^{*+}_{K_S\pi^+}$	1538	0.34	$24.3^{+3.0}_{-2.9}$	0.090	2.4	38
$B^0 \to K^{*0}_1(1273)\gamma$	1452	0.04	16^*	1.000	1.9	30
$B^+ \to \eta_{\gamma\gamma} K^{*+}_{K^+\pi^0}$	1539	0.14	$24.3^{+3.0}_{-2.9}$	0.131	1.4	23
$B^+ \to K^{*+}(1410)\gamma$	1970	0.02	15.0^*	1.000	1.1	17
$B0 \to \eta\pi^+\pi^-$	5573	0.06	5^*	1.000	1	16
$B^+ \to K^{*+}\gamma$	3135	0.01	$40.3^{+2.6}_{-2.6}$	1.000	0.9	14
$B^0 \to K^{*0}_1(1402)\gamma$	1455	0.05	$6^{+6.0}_{-6.0}$	1.000	0.9	14
$B^+ \to K^{*+}_2(1430)\gamma$	1765	0.02	$14.5^{+4.3}_{-4.3}$	1.000	0.8	13
$B^0 \to K^{*0}(1680)\gamma$	1973	0.1	2^*	1.000	0.6	10
$B^+ \to K^{*+}_1(1402)\gamma$	1454	0.02	$7.5^{+7.5}_{-7.5}$	1.000	0.5	8
$B^+ \to K^{*+}(1680)\gamma$	1972	0.03	2^*	1.000	0.2	3
$B^+ \to a^0_0(\eta_{\gamma\gamma}\pi^0)K^+$	6552	0.02	1^*	0.394	0	0
$B^+ \to \eta_{\gamma\gamma} K^+$	1513	0.01	$2.5^{+0.3}_{-0.3}$	0.394	0	0
					36.3	576
$\eta_{3\pi} K^{*0}_{K^+\pi^-}$ (S-wave)						
$B^+ \to \eta_{3\pi} K^{*+}_{K^+\pi^0}$	1542	0.1	$24.3^{+3.0}_{-2.9}$	0.075	0.6	13
$B0 \to \eta\pi^+\pi^-$	5573	0.03	5^*	1.000	0.5	11
$B^+ \to \eta'_{\eta_{3\pi^0}\pi\pi}\ K^+$	6557	0.01	$69.7^{+2.8}_{-2.7}$	0.144	0.2	4
$B^+ \to \eta_{3\pi} K^+$	1515	0.01	$2.5^{+0.3}_{-0.3}$	0.226	0	0
$B^0 \to \eta_{3\pi}\rho^0$	2314	0.37	0.0^*	0.226	0	0
					1.3	28

6.11 Fit validation

We check for fit stability and biases on the yields due to residual correlations in signal, self-crossfeed, and charmless $B\overline{B}$ backgrounds using embedded MC toy studies.

6.11.1 Embedded toy experiments

In embedded toys studies we construct and fit $200 - 500$ toy MC datasets, each with the total number of events expected in the actual data, but with various assumptions about the data content. The $q\bar{q}$ background events are generated according to the fit PDFs, and all other components ($\eta K^*(892)$, $\eta K_0^*(S\text{-wave})$, $\eta K_2^*(1430)$, and charmless $B\overline{B}$) are embedded from MC. We observe no failed fits in embedded toy studies for all modes.

We expect that the fit will have some bias. The main bias comes from the residual correlations among the fit variables. We also expect bias due to the fitting of charmless $B\overline{B}$ background events as signal, though this type of bias is mostly eliminated with the use of charmless fit components where necessary. To ensure that bias from this source is not neglected, we embed the expected number of $\eta K^*(892)$, $\eta K_0^*(S\text{-wave})$, $\eta K_2^*(1430)$, and charmless events in all toy studies regardless of whether all four fit components are used in the fit. For instance, though we do not include the $\eta K_2^*(1430)$ as a component in the LMR $\eta K^*(892)$ fit, we have embedded the expected number of $\eta K_2^*(1430)$ events in toy studies.

Before unblinding the fit results, we confirm that the fit bias is stable for a variety of possible fit outcomes by embedding different numbers of $\eta K^*(892)$, $\eta K_0^*(S\text{-wave})$, $\eta K_2^*(1430)$, and charmless $B\overline{B}$ events. In Tables 6.9 and 6.10 we show results from embedded toy studies for the LMR and HMR analyses. These studies pre-date the inclusion of the $\eta K_2^*(1430)$ mode in the analysis, but the results are still useful for understanding the effects of backgrounds on the fit biases. We determine the numbers

to embed by extrapolating the results of previous analyses to this larger dataset.

After unblinding results we recompute bias numbers adjusting the number of events we embed to produce mean fit yields close to the result from the final fit to data, using linear interpolation to close in on the final value. The results of these embedded toy studies are given in Tables 6.11 and 6.12. In most cases the fit bias is less than 10% of the signal yield. In cases of low signal yield, such as $\eta_{3\pi}K^{*0}_{K^+\pi^-}$($S$-wave), the bias can be as much as 16% of the signal yield. We correct results to account for these biases, and we assign appropriate systematic uncertainties.

Table 6.9: Summary of results from embedded toy MC studies for the LMR fits with 200-500 toy experiments. We show the numbers of $\eta K^*(892)$ signal, charmless $B\overline{B}$ background, and $\eta K_0^*(S\text{-wave})$ events embedded in continuum background generated from the PDFs, along with the bias on the $\eta K^*(892)$ yield. These toy studies pre-date the inclusion of $\eta K_2^*(1430)$ component in the fit, but the results are still useful for understanding the effect of backgrounds on the fit bias. All numbers are in events.

Final state	$\eta K^*(892)$ input	$B\overline{B}$ input	ηK_0^* input	$\eta K^*(892)$ Fit bias
$\eta_{\gamma\gamma} K^{*+}_{K^+\pi^0}$	88	0	31	4.4 ± 0.8
	88	4	31	2.7 ± 0.7
	88	10	31	4.1 ± 0.8
	88	44	0	4.1 ± 0.7
	88	0	46	4.0 ± 0.8
$\eta_{\gamma\gamma} K^{*+}_{K^0_S\pi^+}$	99	0	35	3.5 ± 0.8
	99	179	35	6.8 ± 0.8
	99	269	35	8.6 ± 0.9
	99	179	0	7.0 ± 0.8
$\eta_{\gamma\gamma} K^{*0}_{K^+\pi^-}$	297	0	56	6.4 ± 1.3
	297	94	56	6.3 ± 1.2
	297	141	56	6.3 ± 1.3
	297	94	0	7.5 ± 1.3
$\eta_{3\pi} K^{*+}_{K^+\pi^0}$	34	0	0	2.8 ± 0.5
	34	0	13	3.5 ± 0.6
	34	10	13	3.8 ± 0.5
	34	0	26	3.1 ± 0.6
$\eta_{3\pi} K^{*+}_{K^0_S\pi^+}$	29	0	15	3.0 ± 0.5
	29	10	15	5.1 ± 0.4
	29	20	15	6.5 ± 0.5
	29	0	0	2.3 ± 0.4
	29	0	30	4.3 ± 0.5
$\eta_{3\pi} K^{*0}_{K^+\pi^-}$	86	0	23	4.5 ± 0.7
	86	114	23	10.7 ± 0.7
	86	171	23	12.6 ± 0.9
	86	114	0	10.9 ± 0.7

Table 6.10: Summary of results from embedded toy MC studies for the HMR fit with 200-500 toy experiments. We report the number of ηK_0^*(S-wave) signal, charmless $B\overline{B}$ background, and $\eta K^*(892)$ events embedded in continuum background generated from the PDFs, along with the bias on the ηK_0^*(S-wave) yield. These studies pre-date the inclusion of $K_2^*(1430)$ in the fit, but the results are still useful for understanding the effect of backgrounds on the bias. All numbers are in events.

Final state	ηK_0^* input	$B\overline{B}$ input	$\eta K^*(892)$ input	ηK_0^* Fit bias
$\eta_{\gamma\gamma} K^{*+}_{K^+\pi^0}$	110	0	7	7.1 ± 0.8
	110	68	7	7.1 ± 0.9
	110	102	7	8.2 ± 0.9
	110	68	14	8.5 ± 0.9
$\eta_{\gamma\gamma} K^{*+}_{K^0_S\pi^+}$	22	0	6	2.4 ± 0.6
	22	12	6	2.8 ± 0.6
	22	24	6	1.8 ± 0.6
	22	12	12	2.2 ± 0.6
$\eta_{\gamma\gamma} K^{*0}_{K^+\pi^-}$	142	0	16	9.2 ± 1.0
	142	131	16	9.7 ± 1.1
	142	195	16	9.2 ± 1.0
	142	131	24	12.9 ± 1.1
$\eta_{3\pi} K^{*+}_{K^+\pi^0}$	27	0	0	1.8 ± 0.5
	27	0	3	2.9 ± 0.5
	27	10	3	2.3 ± 0.5
	27	0	6	2.3 ± 0.5
$\eta_{3\pi} K^{*+}_{K^0_S\pi^+}$	48	0	0	3.3 ± 0.5
	48	0	3	3.3 ± 0.5
	48	10	3	3.9 ± 0.5
	48	0	6	4.4 ± 0.5
$\eta_{3\pi} K^{*0}_{K^+\pi^-}$	65	0	0	3.4 ± 0.6
	65	0	6	5.3 ± 0.6
	65	10	6	6.6 ± 0.7
	65	0	12	5.7 ± 0.7

Table 6.11: Summary of final, unblind results from 200–500 embedded toy experiments for the LMR analysis in which we estimate bias on $\eta K^*(892)$ yield. We report the number of signal [ηK^*], charmless $B\overline{B}$ background, $\eta K_0^*(S\text{-wave})$ [ηK_0^*], and $\eta K_2^*(1430)$ [ηK_2^*] events embedded in continuum background generated from the PDFs, along with the bias. For each fit component, we embed a number of events such that the mean of toy yield distribution equals the fit yield from the nominal fit. All numbers are in events.

Final state	ηK^* input	$B\overline{B}$ input	ηK_0^* input	ηK_2^* input	$\eta K^*(892)$ Fit bias
$\eta_{\gamma\gamma} K^{*+}_{K^+\pi^0}$	96	6	37	1	7.0 ± 0.7
$\eta_{\gamma\gamma} K^{*+}_{K^0_S\pi^+}$	140	151	43	4	12.1 ± 0.7
$\eta_{\gamma\gamma} K^{*0}_{K^+\pi^-}$	399	60	67	4	15.4 ± 1.0
$\eta_{3\pi} K^{*+}_{K^+\pi^0}$	53	0	16	4	3.8 ± 0.5
$\eta_{3\pi} K^{*+}_{K^0_S\pi^+}$	32	0	18	4	4.6 ± 0.4
$\eta_{3\pi} K^{*0}_{K^+\pi^-}$	98	76	26	1	12.6 ± 0.6

Table 6.12: Summary final, unblind results from 200–500 embedded toy experiments for the HMR analysis in which we estimate biases on $\eta K_0^*(S\text{-wave})$ and $\eta K_2^*(1430)$ yields. We report the number of $\eta K^*(892)$ [ηK^*], charmless $B\overline{B}$ background, $\eta K_0^*(S\text{-wave})$ [ηK_0^*], and $\eta K_2^*(1430)$ [ηK_2^*] events embedded in continuum background generated from the PDFs, along with the biases. For each fit component, we embed a number of events such that the mean of toy yield distribution equals the fit yield from the nominal fit. All numbers are in events.

Final state	ηK_0^* input	ηK_2^* input	$B\overline{B}$ input	ηK_1^* input	ηK_0^* Fit bias	ηK_2^* Fit bias
$\eta_{\gamma\gamma} K^{*+}_{K^+\pi^0}$	87	27	0	8	9.3 ± 0.9	-0.6 ± 0.5
$\eta_{\gamma\gamma} K^{*+}_{K^0_S\pi^+}$	48	17	0	7	5.1 ± 0.6	-0.6 ± 0.4
$\eta_{\gamma\gamma} K^{*0}_{K^+\pi^-}$	156	74	115	20	17.1 ± 1.1	-1.0 ± 0.7
$\eta_{3\pi} K^{*+}_{K^+\pi^0}$	39	17	0	3	5.7 ± 0.6	-0.9 ± 0.4
$\eta_{3\pi} K^{*+}_{K^0_S\pi^+}$	45	3	0	3	3.2 ± 0.5	1.1 ± 0.3
$\eta_{3\pi} K^{*0}_{K^+\pi^-}$	65	43	0	7	8.9 ± 0.7	-0.8 ± 0.6

6.12 Results

We show $B \to \eta K^*(892)$ branching fraction and charge asymmetry ($\mathcal{A}_{ch}$) results from the LMR analysis in Table 6.13. We show results from the HMR analysis for the $B \to \eta K_0^*(S\text{-wave})$ and $B \to \eta K_2^*(1430)$ branching fractions and charge asymmetries in Tables 6.14 and 6.15. For each decay, we show

- the number of combinations per event for data (before choosing the best candidate),
- the number of events that enter the fit,
- the fit yields with statistical errors (no uncertainties means that the value was held constant),
- the fit bias as determined in embedded toy studies,
- the detection efficiency as determined from MC,
- corrections to the efficiency due to data/MC disagreement for K_S^0 and π^0 detection efficiency,
- the corrected efficiency,
- the product of B daughter branching fractions (for final states where a specific decay chain was required in generation of MC),
- the statistical significance of the yield,
- the branching fraction,
- the charge asymmetry,
- the combined branching fractions, with systematic errors (Sec. 6.14), for neutral and charged decays,

- the combined significance considering both statistical and systematic uncertainties,
- the combined charge asymmetry with systematic uncertainties.

6.12.1 Signal-enhanced projection plots and *s$\mathcal{P}$lots*

In Fig. 6.2 we show projections onto $m_{\rm ES}$ of data sub-samples enriched with a threshold requirement on the signal likelihood (computed without the variable plotted) that optimizes the sensitivity. These projections are for the combined results: $\eta K^{*0}(892)$ and $\eta K^{*+}(892)$ in the LMR and neutral and charged $\eta K_0^*(S\text{-wave}) + \eta K_2^*(1430)$ in the HMR. For the HMR, separation of the $\eta K_0^*(S\text{-wave})$ and $\eta K_2^*(1430)$ signals is afforded mainly by the $K\pi$ mass and helicity shapes; projections of these distributions are shown in Fig. 6.3. We show individual projection plots for all sub-modes and fit variables in Appendix K. We show individual *s$\mathcal{P}$lots* for all sub-modes in Appendix J.

6.12.2 Average over sub-modes

For the $\eta K^*(892)$ modes, we combine the branching fractions and charge asymmetries ($\mathcal{A}_{ch}$) for the four charged B decays ($\eta_{\gamma\gamma}K^{*+}_{K^+\pi^0}$, $\eta_{3\pi}K^{*+}_{K^+\pi^0}$, $\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}$, and $\eta_{3\pi}K^{*+}_{K^0_S\pi^+}$) as well as the neutral B decays ($\eta_{\gamma\gamma}K^{*0}$ and $\eta_{3\pi}K^{*0}$) using scans of the likelihood as described in Sec. 4.5.2.2. The $-2\ln\mathcal{L}$ curves are shown, for the branching fractions and charge asymmetries, in Figs. 6.4 and 6.5 for each sub-mode and for the sum of sub-modes. We perform a simultaneous fit for the $\eta K_0^*(S\text{-wave})$ and $\eta K_2^*(1430)$ results, so there is no need to average over sub-modes after the fit. As mentioned above, the combined results are shown in Tables 6.13, 6.14, and 6.15.

Table 6.13: Branching fraction and charge asymmetry ($\mathcal{A}_{ch}$) results for $\eta K^*(892)$ decays from the low mass range fit.

ML fit quantity	$\eta_{\gamma\gamma}K^{*+}_{K^+\pi^0}$	$\eta_{3\pi}K^{*+}_{K^+\pi^0}$	$\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}$	$\eta_{3\pi}K^{*+}_{K^0_S\pi^+}$	$\eta_{\gamma\gamma}K^{*0}$	$\eta_{3\pi}K^{*0}$
#Data combs/event	1.13	1.28	1.16	1.30	1.08	1.23
Events to fit	10041	5568	7665	4084	18691	9906
Fit $K^*(892)$ yield	$98.6^{+16.2}_{-16.2}$	$56.2^{+11.4}_{-11.4}$	$148.5^{+19.4}_{-19.4}$	$35.8^{+10.2}_{-10.2}$	$407.0^{+29.3}_{-29.3}$	$110.8^{+16.2}_{-16.2}$
Fit $B\overline{B}$ yield	-2 ± 30	–	42 ± 45	–	52 ± 43	72 ± 25
Fix $K_0^*(S$-wave) yield	34.5	14.9	39.8	16.5	70.4	27.4
ML-fit bias (events)	7.0 ± 0.7	3.8 ± 0.5	12.1 ± 0.7	4.6 ± 0.4	15.4 ± 1.0	12.6 ± 0.6
MC ϵ (%)	11.9	8.5	22.3	15.4	24.2	16.4
K^0_S corr. (%)	–	–	98.1	98.1	–	–
π^0 corr. (%)	95.3	93.8	98.3	96.8	98.3	96.8
Corr. ϵ (%)	11.3	8.0	21.5	14.6	23.8	15.9
$\prod \mathcal{B}_i$ (%)	13.1	7.5	9.0	5.2	26.3	15.1
$\mathcal{S}$	6.9	6.1	8.6	3.8	17.6	6.3
$\mathcal{B}(10^{-6})$	18.0 ± 3.2	25.4 ± 5.5	20.5 ± 2.9	11.9 ± 3.9	18.2 ± 1.4	10.9 ± 2.0
$\mathcal{A}_{ch}$	0.19 ± 0.16	-0.05 ± 0.20	-0.03 ± 0.13	-0.23 ± 0.28	0.24 ± 0.07	0.12 ± 0.14
Combined results						
$\mathcal{B}(10^{-6})$		$18.9 \pm 1.8 \pm 1.3$			$16.5 \pm 1.1 \pm 0.8$	
$\mathcal{S}$		13.0			18.8	
$\mathcal{A}_{ch}$		$0.01 \pm 0.08 \pm 0.02$			$0.21 \pm 0.06 \pm 0.02$	

Table 6.14: Branching fraction and charge asymmetry ($\mathcal{A}_{ch}$) results for $\eta K_0^*(1430)$ decays from the high mass range fit, the same fit as the results in Table 6.15.

ML fit quantity	$\eta_{\gamma\gamma}K^{*+}_{K^+\pi^0}$	$\eta_{3\pi}K^{*+}_{K^+\pi^0}$	$\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}$	$\eta_{3\pi}K^{*+}_{K^0_S\pi^+}$	$\eta_{\gamma\gamma}K^{*0}$	$\eta_{3\pi}K^{*0}$
#Data combs/event	1.09	1.21	1.03	1.17	1.03	1.15
Events to fit	6568	3840	5181	2849	13122	7431
Fit $K_0^*(S$-wave) yield	92.9 ± 19.6	39.0 ± 11.7	55.3 ± 15.7	48.9 ± 10.9	162.6 ± 25.4	69.0 ± 17.1
Fit $B\overline{B}$ yield	-43 ± 28	–	-17 ± 39	–	74 ± 63	–
Fix $K^*(892)$ yield	8.0	3.3	6.9	2.7	21.2	8.0
ML-fit bias (events)	9.3 ± 0.9	5.7 ± 0.6	5.1 ± 0.6	3.2 ± 0.5	17.1 ± 1.1	8.9 ± 0.7
MC ϵ (%)	10.2	7.7	12.6	9.6	15.2	10.5
K^0_S corr. (%)	–	–	98.1	98.1	–	–
π^0 corr. (%)	95.3	93.8	98.3	96.8	98.3	96.8
Corr. ϵ (%)	9.7	7.2	12.2	9.1	14.9	10.2
$\prod \mathcal{B}_i$ (%)	13.1	7.5	9.0	5.2	26.3	15.1
$\mathcal{S}$	4.3	3.4	3.0	4.4	5.3	3.6
$\mathcal{B}(10^{-6})$	19.2 ± 4.5	18.0 ± 6.3	13.3 ± 4.2	28.1 ± 6.7	10.8 ± 1.9	11.4 ± 3.2
$\mathcal{A}_{ch}$	-0.05 ± 0.21	0.03 ± 0.29	0.13 ± 0.25	0.18 ± 0.22	0.14 ± 0.15	-0.18 ± 0.25
Combined results						
$\mathcal{B}(10^{-6})$		$18.2 \pm 2.6 \pm 2.6$			$11.0 \pm 1.6 \pm 1.5$	
$\mathcal{S}$		5.9			5.7	
$\mathcal{A}_{ch}$		$0.05 \pm 0.13 \pm 0.02$			$0.06 \pm 0.13 \pm 0.02$	

Table 6.15: Branching fraction and charge asymmetry ($\mathcal{A}_{ch}$) results for $\eta K_2^*(1430)$ decays from the high mass range fit, the same fit as the results in Table 6.14.

ML fit quantity	$\eta_{\gamma\gamma}K^{*+}_{K^+\pi^0}$	$\eta_{3\pi}K^{*+}_{K^+\pi^0}$	$\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}$	$\eta_{3\pi}K^{*+}_{K^0_S\pi^+}$	$\eta_{\gamma\gamma}K^{*0}$	$\eta_{3\pi}K^{*0}$
#Data combs/event	1.09	1.21	1.03	1.17	1.03	1.15
Events to fit	6568	3840	5181	2849	13122	7431
Fit $K_2^*(1430)$ yield	25.6 ± 11.6	19.8 ± 8.1	12.4 ± 9.6	1.9 ± 4.7	71.9 ± 16.6	39.8 ± 13.2
Fit $B\overline{B}$ yield	-43 ± 28	–	-17 ± 39	–	74 ± 63	–
Fix $K^*(892)$ yield	8.0	3.3	6.9	2.7	21.2	8.0
ML-fit bias (events)	-0.6 ± 0.5	-0.9 ± 0.4	-0.6 ± 0.4	1.1 ± 0.3	-1.0 ± 0.7	-0.8 ± 0.6
MC ϵ (%)	13.3	9.5	13.6	10.8	18.7	12.8
K^0_S corr. (%)	–	–	98.1	98.1	–	–
π^0 corr. (%)	95.3	93.8	98.3	96.8	98.3	96.8
Corr. ϵ (%)	12.7	8.9	13.1	10.3	18.4	12.4
$\prod \mathcal{B}_i$ (%)	6.6	3.8	4.5	2.6	13.7	7.6
$\mathcal{S}$	2.3	2.6	1.8	0.2	4.7	3.4
$\mathcal{B}(10^{-6})$	9.1 ± 4.0	17.8 ± 7.2	6.4 ± 4.7	0.9 ± 5.1	8.4 ± 1.9	12.5 ± 4.1
$\mathcal{A}_{ch}$	-0.16 ± 0.41	-0.82 ± 0.47	0.05 ± 0.58	-1.00 ± 1.56	-0.20 ± 0.23	0.23 ± 0.31
Combined results						
$\mathcal{B}(10^{-6})$		$9.1 \pm 2.7 \pm 1.4$			$9.6 \pm 1.8 \pm 1.1$	
$\mathcal{S}$		3.5			5.3	
$\mathcal{A}_{ch}$		$-0.45 \pm 0.30 \pm 0.02$			$-0.07 \pm 0.19 \pm 0.02$	

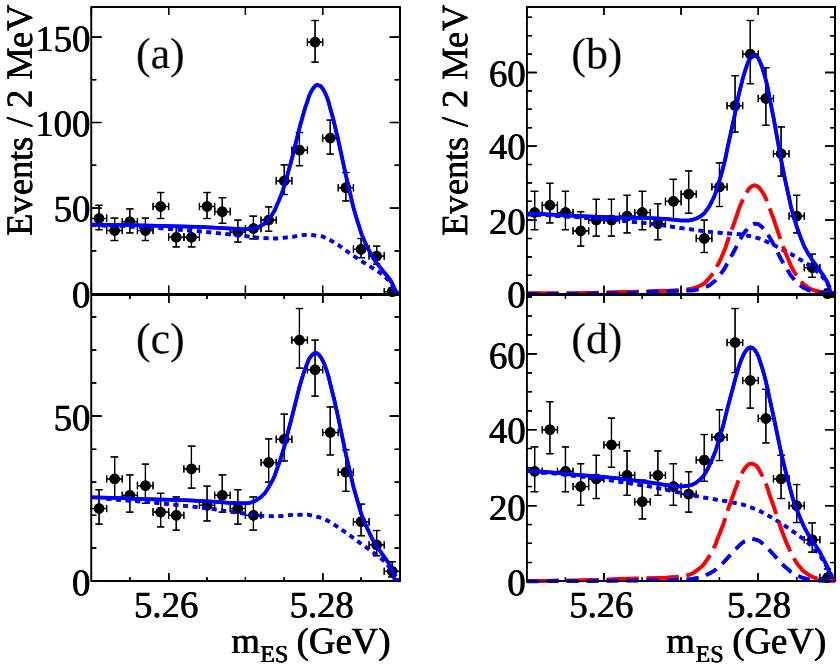


Figure 6.2: Projections onto $m_{\rm ES}$ obtained with a cut on the signal likelihood (see text) for (a) $B^0 \to \eta K^{*0}(892)$, (b) $B^0 \to \eta K_0^{*0}(S\text{-wave})$ (long-dashed, red) plus $B^0 \to \eta K_2^{*0}(1430)$ (short-dashed, blue), (c) $B^+ \to \eta K^{*+}(892)$, and (d) $B^+ \to \eta K_0^{*+}(S\text{-wave})$ (long-dashed, red) plus $B^0 \to \eta K_2^{*0}(1430)$ (short-dashed, blue). Points with uncertainties represent the data, solid curves the full fit functions, and dotted curves the full background functions.

6.12.3 Correlation between $\eta K_0^*(S\text{-wave})$ and $\eta K_2^*(1430)$ yields

As described, in Sec 6.1.1 we use the simultaneous fits in the HMR analysis to properly account for the correlation between $\eta K_0^*(S\text{-wave})$ and $\eta K_2^*(1430)$ branching fractions. In Fig. 6.6 we show the likelihood projected on the planes of $\eta K_0^*(S\text{-wave})$ and $\eta K_2^*(1430)$ branching fractions for neutral and charged modes. The solid dots show the final fit value. The contours denote 1-sigma steps ($\sqrt{\Delta\chi^2} = 1$) about the central value. The tilt angle of the ellipse shows the correlations which are both -0.42.

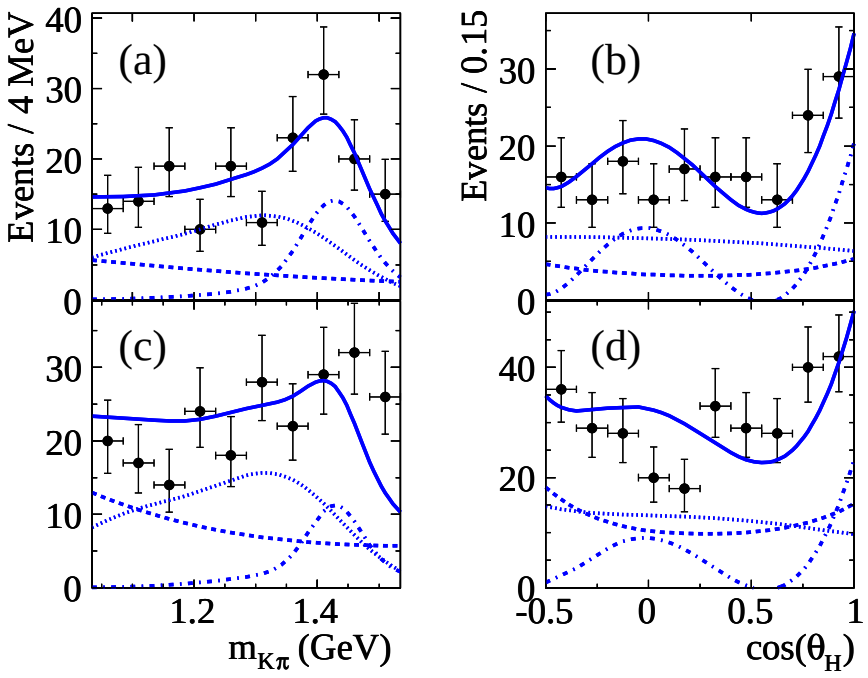


Figure 6.3: Projection of the signals in the HMR, obtained with a cut on the signal likelihood (see text): $K\pi$ mass for (a) B^0, and (c) B^+ channels; $\mathcal{H}$ for (b) B^0, and (d) B^+ channels. Points with uncertainties represent the data, solid curves the full fit functions, dotted curves the $K_0^*(1430)$ portion, dot-dashed curves the $K_2^*(1430)$ portion, and dashed curves the full background functions.

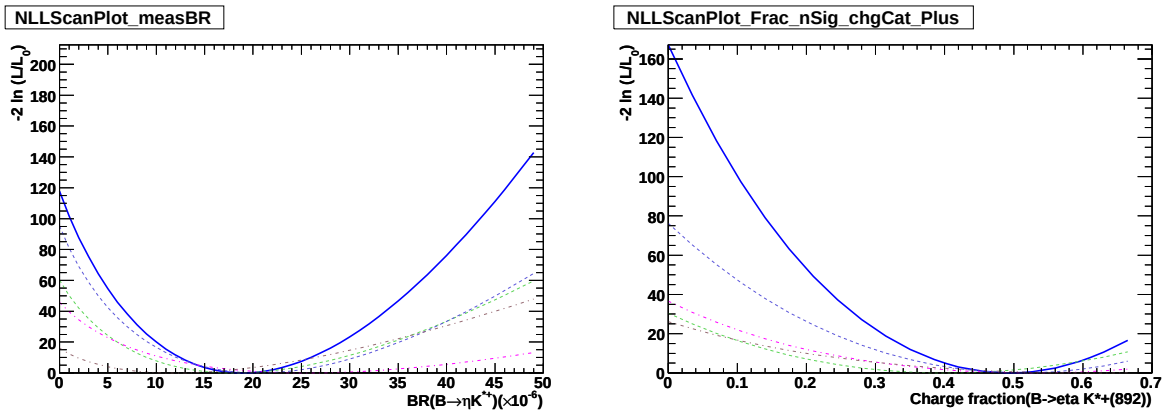


Figure 6.4: Plots of individual (dashed) and combined (solid, blue) $-2\ln\mathcal{L}$ for branching fraction fits (left) and charge fraction (right) are shown for the decay $B^+ \to \eta K^{*+}(892)$.

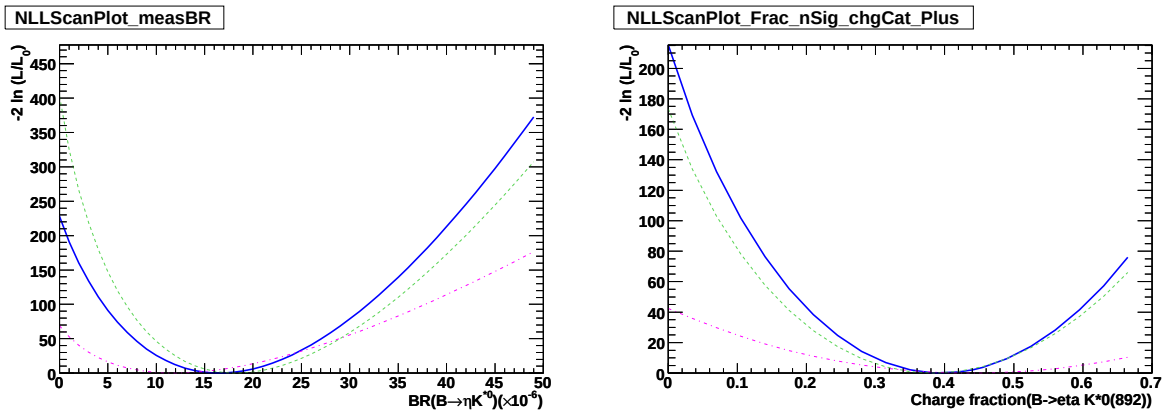


Figure 6.5: Plots of individual (dashed) and combined (solid, blue) $-2\ln\mathcal{L}$ for branching fraction fits (left) and charge fraction (right) are shown for the decay $B^0 \to \eta K^{*0}(892)$.

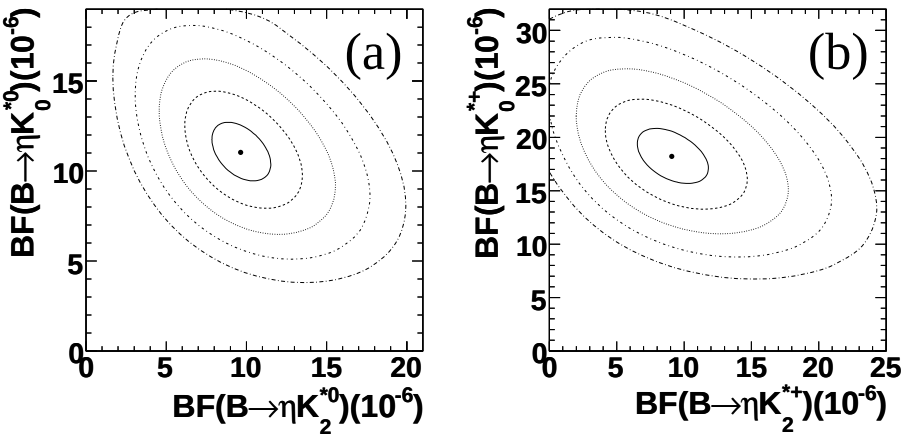


Figure 6.6: Likelihood projected onto the $\mathcal{B}(K_0^*(S\text{-wave}))$–$\mathcal{B}(K_2^*(1430))$ plane for (a) B^0 and (b) B^+ decays. Contours denote 1-sigma steps ($\sqrt{\Delta\chi^2} = 1$) about the central value.

6.13 Partial wave interference

The $B \to \eta K^*$ analyses involve potential interference between $K\pi$ S-wave and P-wave amplitudes in the LMR and S-wave and D-wave amplitudes in the HMR. The P-wave and D-wave are sufficiently separated in $K\pi$ mass that their interference is negligible. The amplitudes as functions of $K\pi$ mass and helicity are

$$A_D(m_{K\pi}, \mathcal{H}) \propto A_{BW_2}(m_{K\pi}) \cdot (3\mathcal{H}^2 - 1), \tag{6.13}$$

$$A_P(m_{K\pi}, \mathcal{H}) \propto A_{BW_1}(m_{K\pi}) \cdot \mathcal{H}, \tag{6.14}$$

$$A_S(m_{K\pi}, \mathcal{H}) \propto A_{LASS}(m_{K\pi}); \tag{6.15}$$

where $\mathcal{H}$ is $\cos\theta_{\mathcal{H}}$; A_{BW_J} is the mass-dependent J-spin relativistic Breit-Wigner lineshape; and A_{LASS} is the LASS lineshape. Very generally, one can write the LMR and HMR decay rates in terms of these amplitudes as

$$\begin{aligned}
\frac{d^2\Gamma_{\rm LMR}}{d(\cos\theta_H)dm_{K\pi}} &= |A_P|^2 + |A_S|^2 + 2{\rm Re}[A_P A_S^* e^{-i\delta_0}], \\
&= |A_{BW_1}|^2\mathcal{H}^2 + |A_{LASS}|^2 \\
&\qquad + 2\ \mathcal{H}\ {\rm Re}[A_{BW_1} A_{LASS}^* e^{-i\delta_0}], \qquad (6.16) \\
\frac{d^2\Gamma_{\rm HMR}}{d(\cos\theta_H)dm_{K\pi}} &= |A_D|^2 + |A_S|^2 + 2{\rm Re}[A_D A_S^* e^{-i\delta_0'}], \\
&= |A_{BW_2}|^2(3\mathcal{H}^2 - 1)^2 + |A_{LASS}|^2 \\
&\qquad + 2\ (3\mathcal{H}^2 - 1)\ {\rm Re}[A_{BW_1} A_{LASS}^* e^{-i\delta_0'}], \qquad (6.17)
\end{aligned}$$

where $\delta_0^{(\prime)}$ are process-dependent arbitrary phase differences between the S-wave and the P- and D-waves.

The interference between the $K\pi$ S-wave and P-wave amplitudes in the LMR fit vanishes over a symmetric helicity range since one is integrating an odd function, the product of the zeroth and first Legendre polynomials. Deviation from a symmetric helicity range will have the largest effect for the $K^* \to K^+\pi^0$ modes in the LMR $\eta K^*(892)$ analysis, in which we use an asymmetric helicity range $-0.7 < \mathcal{H} < 1.0$ to

reduce background from soft pions. For the other LMR modes we expect the interference to be minimal since we use the cut $-0.95 < \mathcal{H} < 1.0$.

The interference between S-wave and D-wave in the HMR only vanishes over the full helicity range, $-1.0 < \mathcal{H} < 1.0$; i.e., the general inner product defining orthogonality of Legendre polynomials is over the full range. ($S + P$ interference vanished due to the odd-ness of $P_0(x)P_1(x)$; the $S+D$ interference term is even, $P_0(x)P_2(x)$.) For all modes in the HMR analysis, we use an asymmetric helicity range $-0.5 < \mathcal{H} < 1.0$ to reduce background from soft pions and $\eta K^*(892)$.

As described in Sec. 6.14, we estimate a systematic error for the contribution of interference to the decay rates (for publication) by integrating Eqs. 6.16 and 6.17 over the appropriate mass and helicity ranges. We know the magnitude and phase of each partial wave amplitude as a function of $K\pi$ mass; we obtain the acceptance as functions of $K\pi$ mass and helicity from MC; and we know the relative normalizations since we have measured the branching fractions. We do not know the values of the phase shifts δ_0 and δ_0', so we vary them from $0 \to 2\pi$ taking the average interference contribution as the systematic uncertainty.

We estimate the systematic uncertainty related to interference to be 4.7% for $\eta K^*(892)$ modes in which $K^* \to K^+\pi^0$, 0.9% for other $\eta K^*(892)$ modes, and 10.5% for $\eta K_0^*(S\text{-wave})$ and $\eta K_2^*(1430)$ modes. In Appendix L we describe extensive studies, which were performed after publication, of partial wave interference in this analysis. We measure δ_0 and δ_0' in the data and confirm that these estimates of the effects of interference are reasonable.

6.14 Systematic errors

We show estimates of systematic errors from various sources for the LMR and HMR analyses in Tables 6.16, 6.17, and 6.18. Even though the precision of these results is dominated by statistical uncertainties, we perform meticulous studies of many sources

of systematic error. In the tables we distinguish which systematics are correlated and uncorrelated among different sub-decay channels of the same mode with a "C" or a "U." The distinction is relevant to the combining of sub-mode results as described in Sec. 4.5.2.2.

The systematic errors related to the fit yield, fit bias, LASS-shape, fixed yield, and charmless $B\overline{B}$ background are additive, while the other errors are primarily multiplicative. Only the additive errors decrease the significance of the result; multiplicative systematic errors simply scale the measured branching fraction. For the high mass range fits, all errors are taken to be correlated. We describe the sources of systematic error:

- Interference: As outlined in Sec. 6.13, we perform calculations to quantify the contribution that S-wave/P-wave (S-wave/D-wave) interference contributes to the LMR (HMR) yields.

- ML fit yield: We would like to float signal PDF parameters in the fit. For these modes, which have a small number of signal events, this is not practical. As an alternative, we use control samples (Sec. 5.17) to study how well the MC models the data in ΔE, $m_{\rm ES}$, $\mathcal{F}$, and resonance masses. We determine nominal shift and scale factors to apply to our signal PDFs. From these studies we also determine appropriate values by which to vary our signal parameters for estimating systematic uncertainties. We apply these variations, one at a time, to our signal PDFs and re-run the ML fit. (We do not specifically vary the helicity PDF parameters in this procedure since studies with the $\eta K^{*0}(892)$ sample show that the effect of this variation is negligible.)

- $\eta K_0^*(S\text{-wave})$ LASS reweighting: As discussed in Sec. 6.3, we make $\eta K_0^*(S\text{-wave})$ MC samples by reweighting non-resonant $\eta K\pi$ MC according to the LASS distribution. To determine the LASS-related systematic uncertainty for each sub-mode, we make two additional $\eta K_0^*(S\text{-wave})$ MC samples by

Table 6.16: Estimates of systematic errors for the $\eta K^*(892)$ analyses. Multiplicative errors are given in percent, additive ones in events.

Quantity	$\eta_{\gamma\gamma}K^{*+}_{K^+\pi^0}$	$\eta_{3\pi}K^{*+}_{K^+\pi^0}$	$\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}$	$\eta_{3\pi}K^{*+}_{K^0_S\pi^+}$	$\eta_{\gamma\gamma}K^{*0}$	$\eta_{3\pi}K^{*0}$
Multiplicative errors (%)						
Track multiplicity (C)	1.0	1.0	1.0	1.0	1.0	1.0
Tracking eff/qual (C)	0.5	1.4	0.4	1.3	0.9	1.8
π^0/ γ eff (C)	6.0	6.0	3.0	3.0	3.0	3.0
K^0_S efficiency (C)	–	–	1.9	1.9	–	–
Number $B\overline{B}$ (C)	1.1	1.1	1.1	1.1	1.1	1.1
Branching fractions (U)	0.7	1.8	0.7	1.8	0.7	1.8
MC statistics (U)	0.7	0.9	0.5	0.6	0.6	0.5
$\cos\theta_{\rm T}$ (C)	0.5	0.5	0.5	0.5	0.5	0.5
Total multiplicative (%)	6.3	6.4	4.0	4.5	3.6	4.3
Additive errors (events)						
Fit yield (U)	1.9	0.8	1.7	0.7	4.4	1.0
Fixed yield (U)	2.8	1.4	3.9	1.6	7.7	3.7
LASS shape (U)	0.3	0.2	0.1	0.2	2.2	0.1
Fit bias (U)	3.5	1.9	6.1	2.3	7.7	6.3
$B\overline{B}$ Background (U)	0.1	1.0	0.1	1.0	0.7	0.9
Interference (C)	4.7	2.6	1.3	0.3	3.7	1.0
Total additive (events)	6.8	3.7	7.6	3.1	12.5	7.5
Total errors [$\mathcal{B}(10^{-6})$]						
Additive	1.2	1.8	1.1	1.2	0.6	0.9
Uncorrelated	0.9	1.4	1.1	1.2	0.6	0.9
Correlated	1.4	2.0	0.8	0.5	0.7	0.5

Table 6.17: Estimates of systematic errors for the $\eta K_0^*(S\text{-wave})$ analyses. Multiplicative errors are given in percent, additive ones in events.

Quantity	$\eta_{\gamma\gamma}K^{*+}_{K^+\pi^0}$	$\eta_{3\pi}K^{*+}_{K^+\pi^0}$	$\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}$	$\eta_{3\pi}K^{*+}_{K^0_S\pi^+}$	$\eta_{\gamma\gamma}K^{*0}$	$\eta_{3\pi}K^{*0}$
Multiplicative errors (%)						
Track multiplicity (C)	1.0	1.0	1.0	1.0	1.0	1.0
Tracking eff/qual (C)	0.5	1.4	0.4	1.3	0.9	1.8
π^0/ γ eff (C)	6.0	6.0	3.0	3.0	3.0	3.0
K^0_S efficiency (C)	–	–	1.9	1.9	–	–
Number $B\overline{B}$ (C)	1.1	1.1	1.1	1.1	1.1	1.1
Branching fractions (C)	0.7	1.8	0.7	1.8	0.7	1.8
MC statistics (C)	0.4	0.6	0.3	0.5	0.3	0.5
$\cos\theta_{\rm T}$ (C)	0.5	0.5	0.5	0.5	0.5	0.5
Total multiplicative (%)	6.3	6.7	4.0	4.5	3.6	4.2
Additive errors (events)						
Fit yield (C)	1.8	0.7	0.8	1.4	3.3	2.0
Fixed yield (C)	1.8	1.1	1.7	0.9	3.3	1.9
LASS shape (C)	1.7	0.6	1.3	1.0	2.2	1.9
Fit bias (C)	4.2	2.9	2.6	1.6	8.6	4.5
$B\overline{B}$ Background (C)	0.6	1.0	0.9	1.0	0.2	1.0
Interference (C)	9.8	4.1	5.9	5.2	17.2	7.3
Total additive (events)	11.1	5.3	6.9	5.9	19.9	9.3
Total errors [$\mathcal{B}(10^{-6})$]						
Additive	2.3	2.6	1.7	3.4	1.3	1.6
Multiplicative	1.2	1.4	0.5	1.3	0.4	0.5
Total Correlated	2.6	2.9	1.8	3.6	1.4	1.7
Weighted averages [$\mathcal{B}(10^{-6})$]						
Additive		2.3			1.4	
Multiplicative		1.0			0.4	
Total Correlated		2.5			1.5	

Table 6.18: Estimates of systematic errors for the $\eta K_2^*(1430)$ analyses. Multiplicative errors are given in percent, additive ones in events.

Quantity	$\eta_{\gamma\gamma}K^{*+}_{K^+\pi^0}$	$\eta_{3\pi}K^{*+}_{K^+\pi^0}$	$\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}$	$\eta_{3\pi}K^{*+}_{K^0_S\pi^+}$	$\eta_{\gamma\gamma}K^{*0}$	$\eta_{3\pi}K^{*0}$
Multiplicative errors (%)						
Track multiplicity (C)	1.0	1.0	1.0	1.0	1.0	1.0
Tracking eff/qual (C)	0.5	1.4	0.4	1.3	0.9	1.8
π^0/ γ eff (C)	6.0	6.0	3.0	3.0	3.0	3.0
K^0_S efficiency (C)	–	–	1.9	1.9	–	–
Number $B\overline{B}$ (C)	1.1	1.1	1.1	1.1	1.1	1.1
Branching fractions (C)	2.5	3.0	2.5	3.0	2.5	3.0
MC statistics (C)	0.3	0.5	0.3	0.5	0.3	0.4
$\cos\theta_{\rm T}$ (C)	0.5	0.5	0.5	0.5	0.5	0.5
Total multiplicative (%)	6.7	7.0	4.7	5.1	4.3	4.8
Additive errors (events)						
Fit yield (C)	1.3	0.4	1.0	0.6	1.7	1.9
Fixed yield (C)	0.2	0.0	0.0	0.0	0.2	0.2
Fit bias (C)	0.4	0.5	0.3	0.6	0.5	0.4
LASS shape (C)	0.9	0.5	0.7	0.7	0.9	0.9
$B\overline{B}$ Background (C)	0.1	1.0	0.2	1.0	0.5	1.0
Interference (C)	2.7	1.1	1.3	0.2	7.6	4.2
Total additive (events)	3.1	2.5	1.8	1.5	7.9	4.8
Total errors [$\mathcal{B}(10^{-6})$]						
Additive	1.1	2.2	0.9	1.6	0.9	1.5
Multiplicative	0.6	1.2	0.3	0.1	0.4	0.6
Total Correlated	1.3	2.4	0.9	1.6	0.9	1.6
Weighted averages [$\mathcal{B}(10^{-6})$]						
Additive			1.3		1.0	
Multiplicative			0.5		0.4	
Total Correlated			1.4		1.1	

reweighting according the LASS distribution with parameters shifted by plus and minus one standard deviation. We then perform the fit using the "plus-one-standard-deviation" sample and again using the "minus-one-standard-deviation" sample, each time finding the difference between the yield and the nominal yield. We sum these plus and minus yield differences in quadrature and count this as an additive systematic.

- Fixed yield: As discussed in Sec. 6.1.4, we determine the values at which to fix the $\eta K_0^*(S\text{-wave})$ yield in the LMR fit and the $\eta K^*(892)$ yield in the HMR fit using the appropriate branching fractions and the corrected efficiencies. We vary the value of the fixed yields by the quadrature sum of the propagated branching fraction error and a Poisson fluctuation ($\sqrt{Y}_{\rm fixed}$). We assign a systematic uncertainty of the quadrature sum of the differences between the nominal yield and each of yields measured when varying the fixed yield by plus and minus one sigma.

- ML fit bias: We assign a systematic uncertainty of one-half the fit bias (see Sec. 6.11.1).

- Trigger efficiency: The *BABAR* B-counting group measured the trigger efficiency for multi-hadron events to be in excess of 0.9993. We neglect the systematic error on the tiny inefficiency from this source.

- Track multiplicity: As described in Sec. 6.6, we make a requirement on the minimum number of tracks in the event to guarantee the reconstruction of at least one track from the other B decay so that we can compute event shape quantities. The signal MC inefficiency for this cut is a few percent. We assign an uncertainty of 1%.

- Track finding/efficiency: As described in Sec. 6.9, the *BABAR* tracking group

has determined that there is no efficiency correction and provides systematic errors related to track finding by run period and track list. Using the fraction of events in each run period as determined in our fits, we average these systematics and apply a 0.47% and 0.40% systematic uncertainty for tracks from the `GoodTracksLoose` and `GoodTracksVeryLoose` lists.

- π^0, $\eta_{\gamma\gamma}$ finding: As described in Sec. 6.9, we use the recipe provided by the *BABAR* neutrals group to determine the efficiency correction and systematic uncertainty associated with π^0 and $\eta_{\gamma\gamma}$ reconstruction. We assign a systematic of 3% per π^0 and η, added linearly.

- Number $B\overline{B}$: Also described in Sec. 6.9, the error on B counting is estimated to be 1.1%.

- Branching fractions of daughters: We take the uncertainty on the branching fraction from the PDG [62] as a systematic error.

- MC statistics: Analyses use about 165K events. With a typical efficiency of 20%, this gives a relative error of 0.7% from binomial statistics. Actual values are used for each mode.

- Event shape cuts: We make a requirement on the event shape variable $\cos\theta_{\rm T}$. The expectation is that the $\cos\theta_{\rm T}$ distribution for signal should be nearly flat. We take the systematic uncertainty to be half of the difference between the observed signal MC efficiency of the $\cos\theta_{\rm T}$ cut used for each analysis and the expected efficiency given a flat distribution.

- Particle ID: PID efficiency corrections were applied during data processing for all analyses. Efficiency is determined from tables produced by the *BABAR* PID group where data control samples determine efficiencies with an uncertainty

of ∼0.1% when integrated over our samples. Thus we take the PID efficiency systematic error to be negligible.

- Charmless $B\overline{B}$ background: As charmless $B\overline{B}$ background was found to be small for all $\eta_{3\pi}$ modes except $\eta_{3\pi}K^{*0}_{K^+\pi^-}(892)$, we include no correction to our yields and assign a conservative uncertainty of one event. For all $\eta_{\gamma\gamma}$ modes, a charmless component is included in the fit, and, in principle, the uncertainty due to these backgrounds is included in the statistical error on the signal yields. However, we assign an additional uncertainty to account for modeling of the charmless $B\overline{B}$ background, derived from differences in the signal yields when such background is included or not or when we use the $b \to c$ $B\overline{B}$ background sample rather than the nominal one.

6.14.1 Charge asymmetry systematics

We expect some charge asymmetry in the detection of tracks because our detector is made of matter. Studies of $\sim 5 \times 10^8$ charged tracks with requirements to remove low momentum protons from beam backgrounds and high momentum electrons from QED backgrounds put a stringent bound on detector charge asymmetry effects for kaons and pions, at all momenta, of 1%. We assign a 2% systematic uncertainty for $\mathcal{A}_{ch}$. In addition, we see that the charge asymmetry of the background is consistent with zero in all cases.

6.15 Conclusion

We analyze 344 million $B\overline{B}$ pairs. We improve the branching fraction and charge asymmetry measurements for the decays $B \to \eta K^*(892)$ and make first measurements for the decays $B \to \eta K_0^*(S\text{-wave})$ and $B \to \eta K_2^*(1430)$, considering these three decays as mutual backgrounds. For $B \to \eta K^*(892)$, we find

$$\begin{aligned}
\mathcal{B}(B \to \eta K^{*+}(892)) &= (18.9 \pm 1.8 \pm 1.3) \times 10^{-6}, \\
\mathcal{A}_{ch} &= 0.01 \pm 0.08 \pm 0.02, \\
\mathcal{B}(B \to \eta K^{*0}(892)) &= (16.5 \pm 1.1 \pm 0.8) \times 10^{-6}, \\
\mathcal{A}_{ch} &= 0.21 \pm 0.06 \pm 0.02.
\end{aligned}$$

For $B \to \eta K_0^*(S\text{-wave})$, we find

$$\begin{aligned}
\mathcal{B}(B^+ \to \eta K_0^{*+}(S\text{-wave})) &= (18.2 \pm 2.6 \pm 2.6) \times 10^{-6}, \\
\mathcal{A}_{ch} &= 0.05 \pm 0.13 \pm 0.02, \\
\mathcal{B}(B^0 \to \eta K_0^{*0}(S\text{-wave})) &= (11.0 \pm 1.6 \pm 1.5) \times 10^{-6}, \\
\mathcal{A}_{ch} &= 0.06 \pm 0.13 \pm 0.02.
\end{aligned}$$

And for $B \to \eta K_2^*(1430)$, we find

$$\begin{aligned}
\mathcal{B}(B^+ \to \eta K_2^{*+}(1430)) &= (9.1 \pm 2.7 \pm 1.4) \times 10^{-6}, \\
\mathcal{A}_{ch} &= 0.01 \pm 0.08 \pm 0.02, \\
\mathcal{B}(B^0 \to \eta K_2^{*0}(1430)) &= (9.6 \pm 1.8 \pm 1.1) \times 10^{-6}, \\
\mathcal{A}_{ch} &= -0.07 \pm 0.19 \pm 0.02.
\end{aligned}$$

We also calculate the branching fraction for the resonant decays to $\eta K_0^*(1430)$ using the composition of $K_0^*(S\text{-wave})$ described in Sec. 6.3.1. (For this model, 9.6% of the $\eta K_0^*(S\text{-wave})$ branching fraction is in the region above the HMR where the model

is most uncertain.) We find

$$\begin{aligned}\mathcal{B}(B^0 \to \eta K_0^{*0}(1430)) &= (9.6 \pm 1.4 \pm 0.7 \pm 1.1) \times 10^{-6}, \\ \mathcal{B}(B^+ \to \eta K_0^{*+}(1430)) &= (15.8 \pm 2.2 \pm 1.4 \pm 1.7) \times 10^{-6},\end{aligned}$$

where the third errors arise from the uncertainties on the branching fraction $K_0^*(1430) \to K\pi$ [62] and the resonant fraction of $K_0^*(S\text{-wave})$.

Chapter 7

Discussion

We report measurements of the time-dependent *CP* parameters S and C in the decay $B^0 \to \eta' K^0$ obtained in five $\eta' K^0_S$ and two $\eta' K^0_L$ sub-decay modes; results are summarized in Table 7.1. We also report measurements of branching fractions and time-integrated charge asymmetries for decays of charged and neutral B mesons to final states $\eta K^{*0}(892)$, $\eta K^{*+}(892)$, $\eta K^{*0}_0(S\text{-wave})$, $\eta K^{*+}_0(S\text{-wave})$, $\eta K^{*0}_2(1430)$, and $\eta K^{*+}_2(1430)$; results are summarized in Table 7.2. We use the full *BABAR* dataset (467 million $B\overline{B}$ pairs) for the $\eta' K^0$ results and the Run1-5 *BABAR* dataset (344 million $B\overline{B}$ pairs) for the ηK^* results.

Table 7.1: Summary of results for $-\eta_f S_f$ and C_f from decays $B^0 \to \eta' K^0$.

	$-\eta_f S_f$	C_f
$\eta' K^0$	$0.55 \pm 0.08 \pm 0.02$	$-0.09 \pm 0.06 \pm 0.02$
$\eta' K^0_S$	$0.53 \pm 0.08 \pm 0.02$	$-0.11 \pm 0.06 \pm 0.02$
$\eta' K^0_L$	$0.64 \pm 0.20 \pm 0.03$	$0.05 \pm 0.15 \pm 0.03$

Significant changes to the previous *BABAR* analysis for $\eta' K^0$ [46] include 20% more data, improved track reconstruction, improved K^0_L selection, and the addition of the $\eta'_{5\pi} K^0_L$ decay channel. Despite the modest increase in data, the uncertainties on $S_{\eta' K^0}$ and $C_{\eta' K^0}$ decrease by 20% and 25% due to the other improvements. Our measurement of $S_{\eta' K^0}$ is more than 5σ from zero, confirming *CP* violation in a charmless B decay;

Table 7.2: Summary of results for branching fractions (in units of 10^{-6}) and charge asymmetries for decays $B \to \eta K^*$.

	$\mathcal{B}(10^{-6})$	$\mathcal{A}_{ch}$
$\eta K^{*+}(892)$	$18.9 \pm 1.8 \pm 1.3$	$0.01 \pm 0.08 \pm 0.02$
$\eta K^{*0}(892)$	$16.5 \pm 1.1 \pm 0.8$	$0.21 \pm 0.06 \pm 0.02$
$\eta K_0^{*+}(S\text{-wave})$	$18.2 \pm 2.6 \pm 2.6$	$0.05 \pm 0.13 \pm 0.02$
$\eta K_0^{*0}(S\text{-wave})$	$11.0 \pm 1.6 \pm 1.5$	$0.06 \pm 0.13 \pm 0.02$
$\eta K_2^{*+}(1430)$	$9.1 \pm 2.7 \pm 1.4$	$-0.45 \pm 0.30 \pm 0.02$
$\eta K_2^{*0}(1430)$	$9.6 \pm 1.8 \pm 1.1$	$-0.07 \pm 0.19 \pm 0.02$

the measurement of $C_{\eta' K^0}$ is consistent with the Standard Model (SM) expectation of zero. We combined our $\eta' K_S^0$ results with the $\eta' K_L^0$ results ($-\eta_f S_f$) from a different *BABAR* analysis for publication in Physical Review D [75]. (The published $\eta' K_L^0$ results are in good agreement with our results.)

Significant changes to the previous *BABAR* analysis for ηK^* [45] include a factor of ~ 4 increase in data and the search for decays of B mesons to high-mass $K\pi$ resonances in final states $\eta K_0^*(S\text{-wave})$ and $\eta K_2^*(1430)$. We improve the uncertainties on $\eta K^*(892)$ branching fraction results by a factor of ~ 2. We make first observations of the decays $B^+ \to \eta K_0^{*+}(S\text{-wave})$, $B^0 \to \eta K_0^{*0}(S\text{-wave})$, and $B^0 \to \eta K_2^{*0}(1430)$, and we find evidence for $B^+ \to \eta K_2^{*+}(1430)$. We observe the charge asymmetry in $\eta K^{*0}(892)$ to be just over 3σ from zero. This is not predicted by the theory and is assumed to be a statistical fluctuation. These results were published in Physical Review Letters [76].

7.1 ΔS

Measurements of $\sin 2\beta_{\rm eff}$ in $B^0 \to \eta' K^0$ and other decay modes dominated by a $b \to q\bar{q}s$ penguin amplitude (Fig. 2.4), including $B^0 \to \phi K^0, K^0 K^0 K^0$, $\pi^0 K^0$, $\rho^0 K^0$, ωK^0, and $K^+ K^- K^0$, are important because of their sensitivity to new physics. As described in Sec. 2.10, heavy, non-SM particles can appear in the loop (with a different

weak phase) causing $\sin 2\beta_{\rm eff}$ to differ from $\sin 2\beta$ from $b \to c\bar{c}s$ decays. To complicate matters, additional, non-negligible weak phases can be introduced by CKM-suppressed loop amplitudes and color-suppressed tree diagrams, so that ΔS is not expected to be zero in the SM. The expected value of ΔS in the SM is channel dependent.

Estimates of ΔS from QCD factorization are in the ranges $(0.0, 0.2)$, $(-0.03, 0.03)$, and $(0.01, 0.12)$ for ωK^0_S, $\eta' K^0$, and $\pi^0 K^0_S$, respectively [36, 38, 37]; SU(3) symmetry provides bounds of $(-0.05, 0.09)$ for $\eta' K^0$ and $(-0.06, 0.12)$ for $\pi^0 K^0_S$ [39]. Predictions that use isospin symmetry to relate several amplitudes, including the $I = \frac{3}{2}$ $B \to K\pi$ amplitude, give an expected value for $S_{\pi^0 K^0_S}$ near 1.0 instead of $\sin 2\beta$ [77]. The modification of the *CP* asymmetry due to the presence of suppressed tree amplitudes in $B^0 \to \phi(K^+K^-)K^0$ is of order 0.01 [36, 78], while at higher K^+K^- masses a larger contribution at of order 0.1 is possible [37].

After the 2005 winter conferences, the field was excited by a 3.7σ discrepancy between the naive average of $\sin 2\beta_{\rm eff}$ measured in $b \to q\bar{q}s$ penguin modes (0.43 ± 0.07) and the average in $b \to c\bar{c}s$ modes (0.73 ± 0.04). By the 2008 summer conferences, the discrepancy had decreased significantly: the naive $b \to q\bar{q}s$ average (0.64 ± 0.04) is now only 0.7σ from the $b \to c\bar{c}s$ average (0.67 ± 0.02). The latest *BABAR*, Belle, and averaged results from $b \to q\bar{q}s$ penguins are compared with the $b \to c\bar{c}s$ average in Fig. 7.1. (The measurements of C in $q\bar{q}s$ modes (Fig. 7.2) are also consistent with the SM expectation of zero.)

This apparent agreement notwithstanding, discussions at the 34th International Conference on High Energy Physics in Philadelphia, PA, made it clear that some theorists are not willing to concede that these measurements are consistent with the SM. It is commonly noted that the measured $\sin 2\beta_{\rm eff}$ is lower than the $c\bar{c}s$ measurement in six of eight $q\bar{q}s$ penguin modes, and theorists working in diverse frameworks of QCD factorization [36], QCD factorization with modeled rescattering [37], and soft collinear effective theory [38] expect ΔS to be positive for most $q\bar{q}s$ penguin modes.

Unfortunately, with predictions for ΔS of ~ 0.03 and statistical uncertainties in the most promising $q\bar{q}s$ mode ($\eta' K^0$) of 0.07, a precise measurement of ΔS will come from future experiments. Since the LHCb experiment at CERN's Large Hadron Collider will match the precision of current measurements with 2 fb^{-1} of data, they will only improve precision by a factor of 2.2 with 10 fb^{-1} [79]. A precise determination of ΔS will have to wait for the Super B-factories, which plan for $50 - 100$ ab^{-1} of data and a factor of 10 improvement in precision.

7.1.1 Minimal flavor violation

Despite the dramatic successes of the SM, there are several reasons to believe that the SM is only an effective low-energy theory of nature. In addition to the obvious issues of neutrino mass and gravity, the hierarchy problem, the strong *CP* problem, and the large baryon asymmetry of the universe need explanation. The hierarchy problem states that, to prevent the Higgs mass from being very large, new physics must appear at an energy scale of $\Lambda_{\rm NP} \sim 1$ TeV [80]. The energy scale of NP is also bounded by flavor physics measurements including those described in this thesis and others from neutral meson mixing; for flavor-generic NP, the measured value of Δm_K implies $\Lambda_{\rm NP} \geq 10^4$ TeV and Δm_B implies $\Lambda_{\rm NP} \geq 10^3$ TeV [81].

The tension between these bounds leads to the seemingly unnatural requirement that NP have highly non-generic flavor structure; e.g., that NP conserve flavor. This is unnatural because the SM, the effective low-energy description of NP, does not respect flavor. One popular method for elegantly resolving this apparent contradiction is the principle of *minimal flavor violation* (MFV). Qualitatively, MFV requires that all flavor- and *CP*-violating processes in the new theory (SM+NP) come from the Yukawa couplings (Eq. 2.3) of the SM [82].

At the LHC, one could attempt to exclude MFV by looking for enhanced flavor-changing neutral currents (FCNC) in NP processes. Reference [83] speculates that given

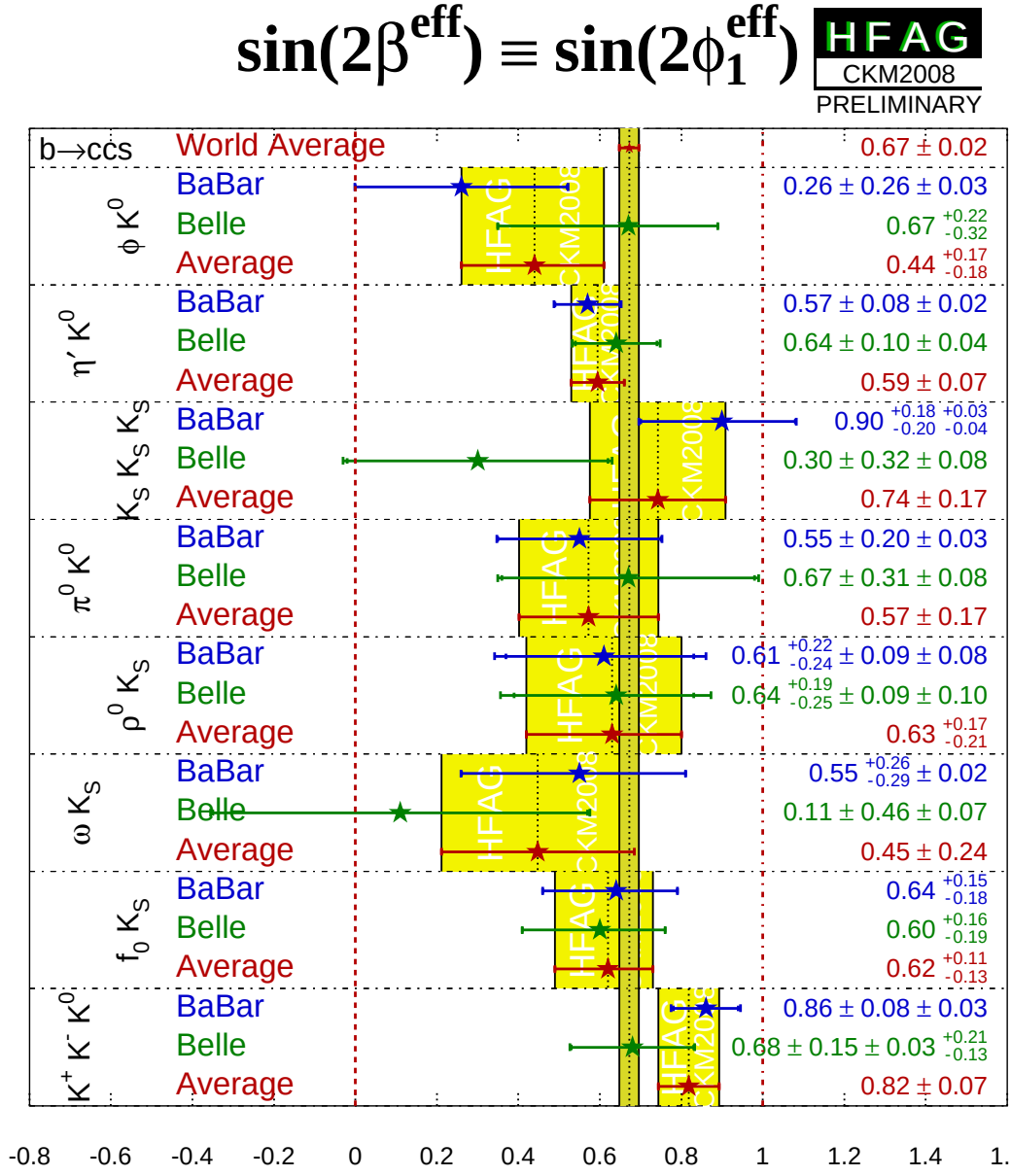


Figure 7.1: Results for $\sin 2\beta_{\rm eff}$ from $b \to q\bar{q}s$ penguin decays and $\sin 2\beta$ from $b \to c\bar{c}s$ decays from the Heavy Flavor Averaging Group [84].

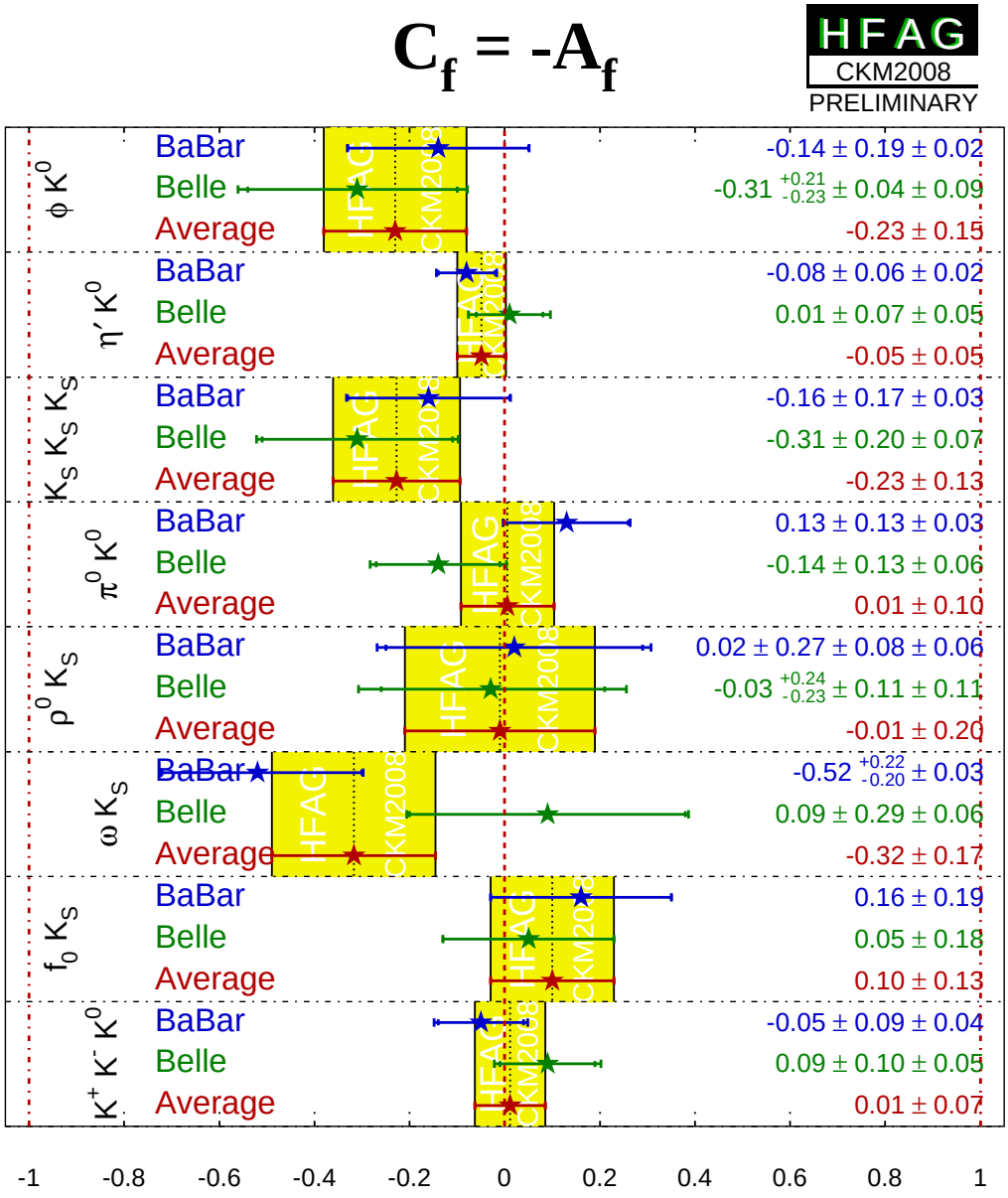


Figure 7.2: Comparison of results for the direct CP-violation parameter C from $b \to q\bar{q}s$ penguin decays from the Heavy Flavor Averaging Group [84].

the fortunate scenario that NP is supersymmetry (SUSY), the stop squark is the next-to-lightest SUSY particle, and the decay of the stop to the bottom quark ($\tilde{t} \to \chi_1^+ b$) is kinematically forbidden, one can measure the decay length of the stop in the decay to the charm quark ($\tilde{t} \to \chi_1^0 c$). A short stop decay length would indicate an enhanced rate for this FCNC and would, therefore, disfavor the principle of minimal flavor violation.

7.2 $B \to \eta^{(\prime)} K^{(*)}$ branching fraction hierarchy

As introduced in Sec. 2.10, decays of B mesons to final states that include flavor-singlet states η and η' are of considerable interest to theorists. The hierarchy of $B \to \eta^{(\prime)} K^{(*)}$ branching fractions has commanded interest for almost two decades. As early as 1991, Lipkin suggested that interference between penguin diagrams where the spectator quark follows either the $\eta^{(\prime)}$ or $K^{(*)}$ can result in differing branching fractions for the four $\eta^{(\prime)} K^{(*)}$ final states. In recent work (2003), Beneke and Neubert confirmed Lipkin's explanation and found that radiative corrections calculated with QCD factorization improve the agreement between data and theory [30].

Unfortunately, theory errors are large and the only measurements with adequate precision for testing the theories are from the modes $B \to \eta' K$ and $B \to \eta K^*$. Since precise measurements in $B \to \eta K$ and $B \to \eta' K^*$ will not be made until the Super B-factories and the effects of strange-meson spin on the hierarchy are still not well understood, we hope that our measurements in $B \to \eta K_0^*(S\text{-wave})$ and $B \to \eta K_2^*(1430)$ will motivate and facilitate theoretical and experimental work in the near future and beyond.

Bibliography

[1] M. Gell-Mann and A. Pais, Phys. Rev. **97**, 1387 (1955).

[2] T.D. Lee and C.N. Yang, Phys. Rev. **104**, 254 (1956).

[3] C.S. Wu *et al.*, Phys. Rev. **105**, 1413 (1957); R.L. Garwin *et al.*, Phys. Rev. **105**, 1415 (1957); J.I. Friedman and V.L. Telegdi, Phys. Rev. **105**, 1681 (1957).

[4] L.D. Landau, Nucl. Phys. **3**, 127 (1957).

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

[6] A. Sakharov, Pisma Zh. Eksp. Teor. Fiz. **5**, 32 (1967).

[7] S.L. Glashow, Nucl. Phys. **22**, 579 (1961); S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967); A. Salam, in *Proceedings of the 8th Nobel Symposium*, ed. N. Swartholm, Almquist, and Wiksells, Stockholm (1968).

[8] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963).

[9] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).

[10] S.L. Glashow, J. Iliopoulos, and L. Maiani, Phys. Rev. D **2**, 1285 (1970).

[11] J.J. Aubert *et al.*, Phys. Rev. Lett. **33**, 1404 (1974).

[12] J.E. Augustine *et al.*, Phys. Rev. Lett. **33**, 1406 (1974).

[13] M.L. Perl *et al.*, Phys. Rev. Lett. **35**, 1489 (1975).

[14] S.W. Herb *et al.*, Phys. Rev. Lett. **39**, 252 (1977).

[15] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **74**, 2626 (1995).

[16] MAC Collaboration, E. Fernandez *et al.*, Phys. Rev. Lett. **51**, 1022 (1983).

[17] Mark-II Collaboration, N. S. Lockyer *et al.*, Phys. Rev. Lett. **51**, 1316 (1983).

[18] P. Oddone, in *Proceedings of the UCLA Workshop: Linear Collider $B\overline{B}$ Factory Conceptual Design*, ed. D. Stork, World Scientific, 243 (1987).

[19] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **89**, 201802 (2002); Belle Collaboration, K. Abe *et al.*, Phys. Rev. D **66**, 071102 (2002).

[20] T.D. Lee, *Particle Physics and Introduction to Field Theory*, Updated First Edition, Routledge U.K. (1990); page 320.

[21] G. Luders, Kong. Danks. Vid. Selskab, Mat.-Fys. Medd. **28**, 5 (1954); W. Pauli, Nuovo Cimento **6**, 204 (1957).

[22] F. Englert and R. Brout, Phys. Rev. Lett. **13**, 321 (1964); P. W. Higgs, Phys. Rev. Lett. **13**, 508 (1964); G. S. Guralnik, C. R. Hagen, and T. W. Kibble, Phys. Rev. Lett. **13**, 585 (1964).

[23] M. Peskin and D. Schroeder, *Introduction to Quantum Field Theory*, Westfield Press (1995).

[24] F. Halzen and A. Martin, *Quarks and Leptons*, Wiley, (1984).

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

[26] I. Bigi and A. Sanda, *CP Violation*, University Press, Cambridge, England (2000).

[27] P. F. Harrison and H. R. Quinn, editors, *The BABAR Physics Book: Physics at an Asymmetric B Factory*, Report of the *BABAR* Physics Workshop, SLAC Report 504 (1998).

[28] *BABAR* Collaboration, B. Aubert *et al.*, arXiv:0808.1903v1 [hep-ex].

[29] Belle Collaboration, K.F. Chen *et al.*, Phys. Rev. Lett. **98**, 031802 (2007).

[30] M. Beneke and M. Neubert, Nucl. Phys. B **651**, 225 (2003).

[31] Y. Grossman and M. P. Worah, Phys. Lett. B **395**, 241 (1997); D. Atwood and A. Soni, Phys. Lett. B **405**, 150 (1997); M. Ciuchini *et al.*, Phys. Rev. Lett. **79**, 978 (1997).

[32] M. Beneke and M. Neubert, Nucl. Phys. B **675**, 333 (2003).

[33] C.-W. Chiang, M. Gronau, and J. L. Rosner, Phys. Rev. D **68**, 074012 (2003);

[34] D. London and A. Soni, Phys. Lett. B **407**, 61 (1997).

[35] Y. Grossman, Z. Ligeti, Y. Nir, and H. Quinn, Phys. Rev. D **68**, 015004 (2003).

[36] M. Beneke, Phys. Lett. B **620**, 143 (2005).

[37] H. Y. Cheng, C-K. Chua, and A. Soni, Phys. Rev. D **72**, 014006 (2005), Phys. Rev. D **71**, 014030 (2005); S. Fajfer, T. N. Pham, and A. Prapotnik-Brdnik Phys. Rev. D **72**, 114001 (2005).

[38] A. R. Williamson and J. Zupan, Phys. Rev. D **74**, 014003 (2006).

[39] M. Gronau, J. L. Rosner, and J. Zupan, Phys. Rev. D **74**, 093003 (2006).

[40] CLEO Collaboration, B.H. Behrens *et al.*, Phys. Rev. Lett. **80**, 3710 (1998).

[41] M.-Z. Yang, Y.-D. Yang, Nucl. Phys. B **609**, 469 (2001); M. Beneke, M. Neubert, Nucl. Phys. B **651**, 225 (2003).

[42] C.-W. Chiang, M. Gronau, and J. L. Rosner, Phys. Rev. D **68**, 074012 (2003); C.-W. Chiang *et al.*, Phys. Rev. D **69**, 034001 (2004).

[43] H. J. Lipkin, Phys. Lett. B **254**, 247 (1991).

[44] CLEO Collaboration, S.J. Richichi *et al.*, Phys. Rev. Lett. **85**, 520 (2000).

[45] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. D **70**, 032006 (2004).

[46] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **98**, 031801 (2007).

[47] Belle Collaboration, K. Abe *et al.*, Phys. Rev. Lett. **98**, 031802 (2007).

[48] Belle Collaboration, K. Abe. *et al.*, BELLE-CONF-0408 (2004).

[49] PEP-II Conceptual Design Report, SLAC Report No. SLAC-R-418 (1993).

[50] *BABAR* Collaboration, B. Aubert *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 1 (2002).

[51] J. Jackson, *Classical Electrodynamics*, Third Edition, John Wiley and Sons, Inc., (1999); page 586, equation 12.42.

[52] P. Billoir, Nucl. Instr. Methods A225 (1984) 225.

[53] D.N. Brown *et al.*, The *BABAR* Track Fitting Algorithm, in *Proceedings of CHEP 2000*, Padova, Italy (200).

[54] Except as noted explicitly, we use a particle name to denote either member of a charge conjugate pair.

[55] Heavy Flavor Averaging Group, world average $B^0 \to \eta' K^0$ branching fraction,

`http://www.slac.stanford.edu/xorg/hfag/rare/ichep08/charmless/OUTPUT/HTML/charmless_table3.html`

[56] S. Agostinelli *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **506**, 250 (2003).

[57] W.D. Hulsbergen, Nucl. Instrum. Meth. A552 (2005) 566-575.

[58] `http://roofit.sourceforge.net`

[59] `http://rarfit.sourceforge.net`

[60] `http://wwwasdoc.web.cern.ch/wwwasdoc/minuit/minmain.html`

[61] M. Pivk and F. R. Le Diberder, Nucl. Instrum. Methods Phys. Res., Sect. A **555**, 356 (2005).

[62] Y.-M. Yao et al., J. Phys. **G33**, 1 (2006).

[63] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. D **66**, 032003 (2002).

[64] `http://tmva.sourceforge.net`

[65] ARGUS Collaboration, H. Albrecht *et al.*, Phys. Lett. B **241**, 278 (1990).

[66] Owen Long, Max Baak, Robert N. Cahn, and David Kirkby, Phys. Rev. D **68**, 034010 (2003).

[67] D. Aston *et al.*, Nucl. Phys. B **296** (1988) 493.

[68] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. D **78**, 012004 (2008).

[69] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. D **76**, 031102 (2007).

[70] W. M. Dunwoodie, private communication.

[71] `http://www.slac.stanford.edu/BFROOT/www/Physics/TrackEfficTaskForce/TrackingTaskForce-2006.html`

[72] `http://www.slac.stanford.edu/BFROOT/www/Physics/TrackEfficTaskForce/TrackingTaskForce-2007.html#KsRecipe`

[73] `http://www.slac.stanford.edu/BFROOT/www/Physics/Analysis/AWG/Neutrals/validation/recipie18.html#mozTocId170488`

[74] `http://www.slac.stanford.edu/BFROOT/www/Physics/Tools/Pid/PidOnMc/pidonmc.html`

[75] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. D **79**, 052003 (2009).

[76] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **97**, 201802 (2006).

[77] A. J. Buras, R. Fleischer, S. Recksiegel, and F. Schwab, Phys. Rev. Lett. **92**, (2004) 101804; R. Fleischer, S. Jager, D. Pirjol, and J. Zupan, arXiv:0806.2900 [hep-ph]; M. Gronau and J. L. Rosner arXiv:0807.3080 [hep-ph].

[78] G. Buchalla, G. Hiller, Y. Nir, and G. Raz, JHEP **0509**, 074 (2005).

[79] Y. Xie *et al.*, LHCb Collaboration, CERN-LHCb-2007-130 (2007); S. Amato *et al.*, LHCb Collaboration, CERN-LHCb-2007-045 (2007).

[80] M. Schmaltz, Plenary talk at ICHEP 2002, arXiv:hep-ph/0210415 (2002).

[81] Y. Grossman, arXiv:hep-ph/0310229 (2003).

[82] G. D'Ambrosio, G.F. Giudice, G. Isidori, A. Strumia, arXiv:hep-ph/0207036 (2002).

[83] G. Hiller and Y. Nir, arXiv:0802.0916 (2008).

[84] `http://www.slac.stanford.edu/xorg/hfag/triangle/summer2008/index.shtml#qqs`

[85] Y. Gao and A. Gritsan, *BABAR* analysis document 1428;

Appendix A

Cut optimization for $\eta' K_L^0$

For the $\eta' K_L^0$ modes, we make cuts on the projected transverse missing momentum ($P_{\rm miss}^{\rm proj}$), the cosine of the angle between the missing momentum and the beam direction in the lab frame ($\cos\theta_{P_{\rm miss}}$), and the output of a neural network (NN) of shower shape variables for candidates with a K_L^0 from the EMC.

For each of these quantities, we start by finding the set of cut values that maximizes the signal significance $S/\sqrt{S+B}$; i.e., the signal yield divided by the square root of number of events entering the fit ("Very Tight" cuts in Table A.1). This method yields a good starting point, but only optimizes errors for a cut-and-count branching fraction analysis. Since we perform an ML fit for S and C, we expect that looser cuts would be optimal to minimize the error on S and C. We incrementally relax the cuts obtained by signal significance optimization and obtain the expected errors on S and C from embedded toy MC experiments (Sec. 6.11). In the actual analysis, we use the set of cuts that minimizes the expected errors on S and C. We report results from these studies in Table A.1 for $\eta'_{\eta\pi\pi} K_L^0$ and Table A.2 for $\eta'_{5\pi} K_L^0$. We find the optimal cuts, those in the *Loose* column to be the same for both modes.

Table A.1: Results from cut optimization study for $\eta'_{\eta\pi\pi}K^0_L$. The "Very Tight" cut values maximize signal significance ($S/\sqrt{S+B}$). The "Loose" cut values minimize the errors on S and C from toy MC studies and are used in this analysis. We report cut values, events entering the Run1-6 fit, signal efficiency, expected Run1-6 signal yield, mean of the S and C error distributions for embedded toys with Run1-6 statistics, and blind fit values to Run1-6 data as a final crosscheck.

	Very Tight	Tight	Loose	Very Loose
NN output Cut	0.50	0.40	0.30	0.20
$P^{\text{proj}}_{\text{miss}}$ Cut	-0.46	-0.60	-0.70	-0.80
$\cos\theta_{P_{\text{miss}}}$ Cut	0.93	0.94	0.95	0.96
Events to Fit	6253	8826	12085	14992
MC ϵ (%)	15.2	17.6	19.8	21.0
Expected nSig	249	310	353	375
S error	0.310 ± 0.004	0.273 ± 0.003	0.257 ± 0.003	0.262 ± 0.003
C error	0.222 ± 0.002	0.198 ± 0.002	0.191 ± 0.002	0.190 ± 0.001
blind S	-0.76 ± 0.26	-0.68 ± 0.24	-0.68 ± 0.22	-0.59 ± 0.22
blind C	0.09 ± 0.22	0.05 ± 0.21	0.04 ± 0.19	0.05 ± 0.19

Table A.2: Results from cut optimization study for $\eta'_{5\pi}K^0_L$. The "Very Tight" cut values maximize signal significance ($S/\sqrt{S+B}$) in the $\eta'_{\eta\pi\pi}K^0_L$ mode. We report cut values, events entering the Run1-6 fit, signal efficiency, expected Run1-6 signal yield, mean of the S and C error distributions for embedded toys with Run1-6 statistics, and blind fit values to Run1-6 data as a final crosscheck.

	Very Tight	Tight	Loose	Very Loose
NN output Cut	0.50	0.40	0.30	0.20
$P^{\text{proj}}_{\text{miss}}$ Cut	-0.46	-0.60	-0.70	-0.80
$\cos\theta_{P_{\text{miss}}}$ Cut	0.93	0.94	0.95	0.96
Events to Fit	2497	3602	4901	6144
MC ϵ (%)	10.9	12.1	12.8	13.5
Expected nSig	130	145	155	165
S error	0.460 ± 0.010	0.410 ± 0.009	0.390 ± 0.006	0.391 ± 0.008
C error	0.318 ± 0.005	0.289 ± 0.004	0.283 ± 0.003	0.287 ± 0.003
blind S	-1.21 ± 0.52	-0.89 ± 0.49	-0.68 ± 0.46	-0.77 ± 0.45
blind C	0.03 ± 0.31	0.08 ± 0.28	0.04 ± 0.26	0.16 ± 0.26

Appendix B

EMC shower shape quantities

The seven variables used as input to the neural network in the $\eta' K_L^0$ analyses are listed below and shown in Fig B.1.

- **Number of EMC crystals.**

- **Second moment:**

$$\frac{\sum E_i \cdot r_i^2}{\sum E_i},$$

where E_i is the energy of crystal i, and r_i is the distance of crystal i to the cluster center.

- **Lateral moment:**

$$\frac{\sum_{i=2,n} E_i \cdot r_i^2}{(\sum_{i=2,n} E_i \cdot r_i^2) + 25(E_0 + E_1)},$$

with the crystals in descending energy order.

- **S1/S9:** The energy of the most energetic crystal (S1) divided by the energy sum of the 3x3 crystal block (S9) with the most energetic crystal in its center.

- **S9/S25:** The energy sum of the 3x3 crystal block (S9) with the most energetic crystal in it's center, divided by the energy sum of the 5x5 crystal block (S25) with the most energetic crystal in it's center.

- **Zernike moments $|Z_{20}|$, $|Z_{42}|$:** The spatial energy distribution of an EMC cluster can be expressed as a series of Zernike polynomials (ζ): $E(x, y) =$

$\sum_{n,m} Z_{n,m} \cdot \zeta_{n,m}(r, \phi)$; where (x, y) are the Cartesian coordinates in the plane of the calorimeter; (r, ϕ) are the polar coordinates of the Zernike polynomials $(0 \leq r \leq 1)$; and n, m are non-negative integers.

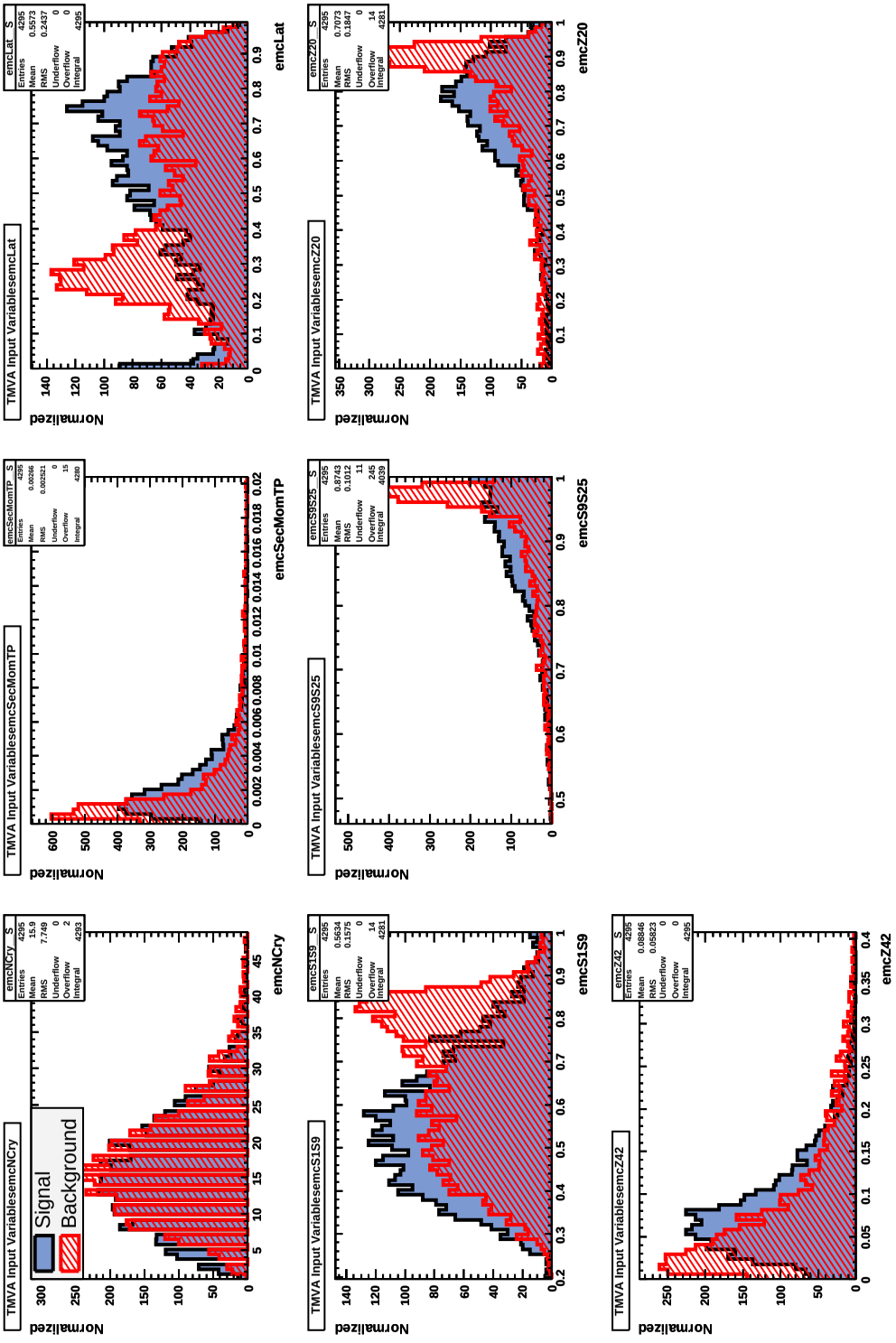


Figure B.1: Comparison of the seven shower shape variables for signal MC (solid blue) and continuum background MC (hatched red) for $\eta'_{\eta\pi\pi}K^0_L$. Rotating clockwise by 90 degrees, the plots are the number of EMC crystals (top left), second moment (top middle), lateral moment (top right), $S1/S9$ (middle left), $S9/S25$ (middle middle), $|Z_{20}|$ (middle right), $|Z_{42}|$ (bottom left).

Appendix C

$\eta' K^0$: PDFs, fit input correlations, and final free parameter values

We show here for each decay mode the signal and background PDFs used in ML fits, the correlation coefficients between the input variables used in the ML fits, and the initial and final values of the PDF parameters that are free in the fit. Correlations are less than 7% between all variables. The 33% correlation between $\mathcal{F}$ and $m_{\rm ES}$ for data in the modes with $\eta' \to \eta\pi^+\pi^-$ is spurious; it is due to signal which is a large fraction of the data for these modes. Bias due to correlations in the signal is accounted for in embedded toy studies. Signal and $B\overline{B}$ background PDFs are determined from MC; for continuum background PDFs we use on-peak sidebands.

C.1 $\eta'_{\rho\gamma}K^0_S$

Correlation matrix for onData (22775 events):

	de	mes	fisher	deltaT
mes	0.0099			
fisher	-0.0342	-0.0734		
deltaT	0.0228	-0.0037	0.0219	
dtErr	0.0336	-0.0228	-0.0438	0.0413

The $\eta'_{\rho\gamma}K^0_S$ fit finds the following values for the parameters which were allowed to float:

Floating Parameter	InitialValue	FinalValue +/- Error	GblCorr.
Frac_nBkg_tagCat_04T1	9.7610e-03	4.5888e-03 +/- 5.29e-04	0.230561
Frac_nBkg_tagCat_04T2	8.3922e-02	8.1936e-02 +/- 1.92e-03	0.328561
Frac_nBkg_tagCat_04T3	1.6215e-01	1.6061e-01 +/- 2.56e-03	0.407178
Frac_nBkg_tagCat_04T4	1.1882e-01	1.1788e-01 +/- 2.25e-03	0.370811
Frac_nBkg_tagCat_04T5	1.4183e-01	1.4046e-01 +/- 2.43e-03	0.391618
Frac_nBkg_tagCat_04T6	1.1325e-01	1.1454e-01 +/- 2.22e-03	0.366261
deBkg_P01	-1.6984e+00	-1.6322e+00 +/- 6.06e-02	0.119880
dtBkgBiasC	5.5338e-02	6.5050e-02 +/- 1.23e-02	0.427125
dtBkgBiasT	1.4082e-01	1.4414e-01 +/- 6.00e-02	0.411307
dtBkgFracC	7.6755e-01	7.9942e-01 +/- 2.12e-02	0.955283
dtBkgFracO	1.7775e-02	1.3576e-02 +/- 1.44e-03	0.464447
dtBkgScfaC	1.1903e+00	1.2016e+00 +/- 1.93e-02	0.883904
dtBkgScfaT	2.5522e+00	2.6238e+00 +/- 1.14e-01	0.926759
dtSig_C	0.0000e+00	-1.2577e-01 +/- 8.94e-02	0.036799
dtSig_S	7.0000e-01	4.5927e-01 +/- 1.17e-01	0.048574
fisBkgC_asym	3.6003e-02	7.5271e-02 +/- 1.18e-02	0.225838
fisBkgC_mean	4.2247e-01	4.3956e-01 +/- 4.46e-03	0.401732
fisBkgC_rms	5.7080e-01	5.6597e-01 +/- 3.17e-03	0.317902
mesBkg_c	-2.5608e+01	-1.9867e+01 +/- 1.80e+00	0.118454
nBkg	1.0000e+04	2.1242e+04 +/- 1.59e+02	0.362095
nChrm	5.3600e+02	2.5336e+02 +/- 6.73e+01	0.634903
nSig	6.6000e+02	1.0047e+03 +/- 3.97e+01	0.197468

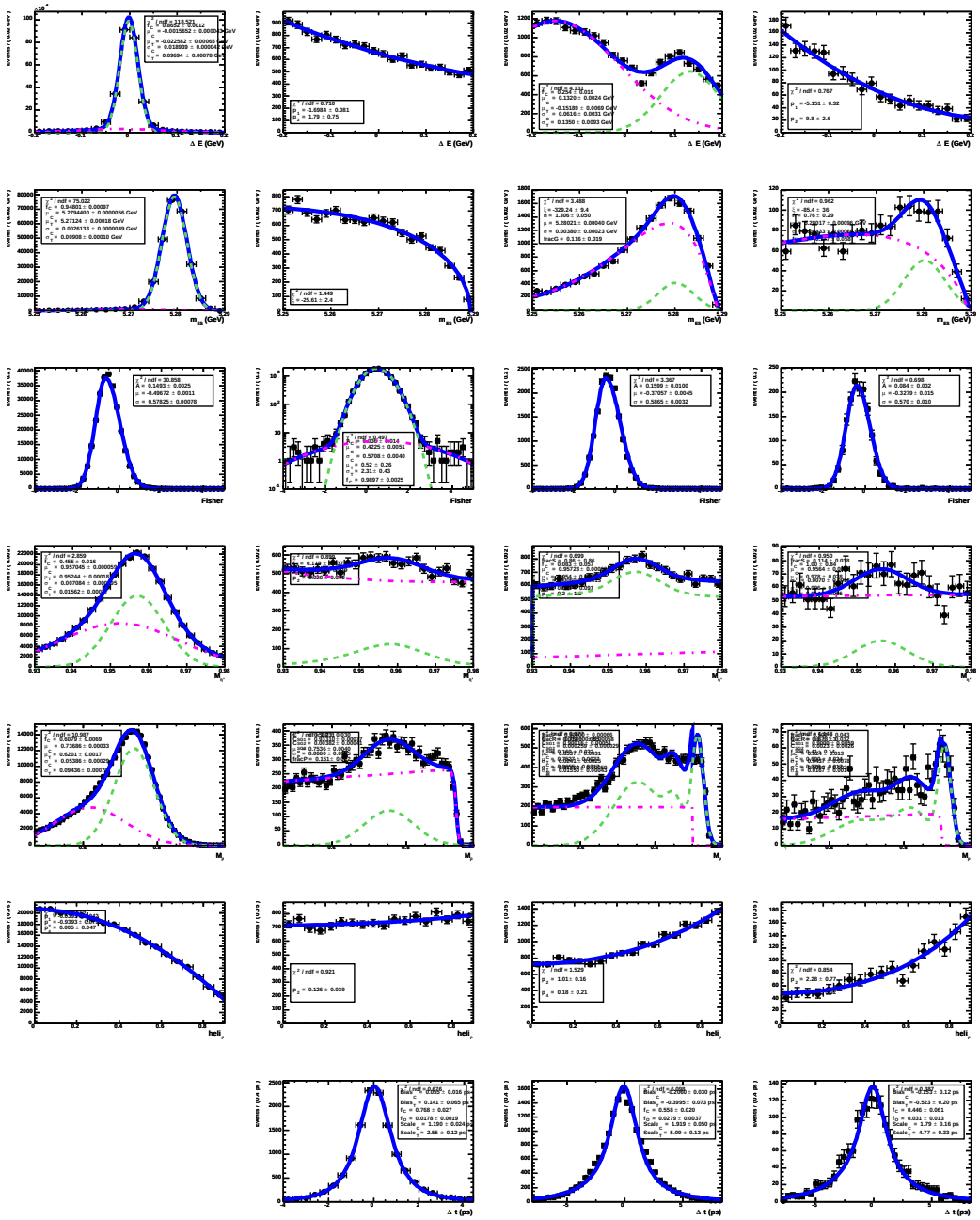


Figure C.1: PDFs for $\eta'_{\rho\gamma}K^0_S$ (from top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η' mass, ρ mass, ρ helicity, and Δt. Signal MC (left), on-peak sidebands (left-center), charmless $B\overline{B}$ (right-center), and $b \to c$ (right).

C.2 $\eta'_{\eta\pi\pi}K^0_S$

```
Correlation matrix for onData (1470 events):
               de       mes    fisher    deltaT
    mes    0.0613
 fisher   -0.0763   -0.3217
 deltaT   -0.0079   -0.0536    0.0135
  dtErr    0.0263   -0.0713    0.0384    0.0103
```

The $\eta'_{\eta\pi\pi}K^0_S$ fit finds the following values for the parameters which were allowed to float:

Floating Parameter	InitialValue	FinalValue +/- Error	GblCorr.
Frac_nBkg_tagCat_04T1	3.7088e-02	1.2118e-02 +/- 3.80e-03	0.150683
Frac_nBkg_tagCat_04T2	8.6538e-02	6.6504e-02 +/- 8.19e-03	0.289262
Frac_nBkg_tagCat_04T3	1.4629e-01	1.3513e-01 +/- 1.11e-02	0.369361
Frac_nBkg_tagCat_04T4	1.0646e-01	1.0434e-01 +/- 1.01e-02	0.342391
Frac_nBkg_tagCat_04T5	1.6140e-01	1.6409e-01 +/- 1.21e-02	0.391656
Frac_nBkg_tagCat_04T6	1.0508e-01	1.1405e-01 +/- 1.04e-02	0.349927
deBkg_P01	-2.0214e+00	-1.9501e+00 +/- 2.75e-01	0.015490
dtBkgBiasC	2.8455e-02	1.7217e-02 +/- 6.08e-02	0.364934
dtBkgBiasT	-1.2390e-01	-2.3695e-01 +/- 2.76e-01	0.394154
dtBkgFracC	8.4129e-01	7.8988e-01 +/- 6.09e-02	0.892320
dtBkgFracO	0.0000e+00	5.6884e-03 +/- 4.66e-03	0.205584
dtBkgScfaC	1.3604e+00	1.2910e+00 +/- 7.21e-02	0.774063
dtBkgScfaT	3.7481e+00	2.9650e+00 +/- 3.21e-01	0.825189
dtSig_C	0.0000e+00	-1.6553e-01 +/- 1.09e-01	0.068914
dtSig_S	7.0000e-01	7.0477e-01 +/- 1.66e-01	0.063050
fisBkgC_asym	4.6816e-02	2.5116e-02 +/- 5.63e-02	0.106149
fisBkgC_mean	3.8793e-01	3.9733e-01 +/- 1.98e-02	0.138429
fisBkgC_rms	6.0198e-01	5.7331e-01 +/- 1.47e-02	0.068787
mesBkg_c	-2.0170e+01	-1.7598e+01 +/- 8.44e+00	0.147285
nBkg	6.0000e+02	9.9846e+02 +/- 3.29e+01	0.143662
nSig	2.7500e+02	4.7153e+02 +/- 2.36e+01	0.176899

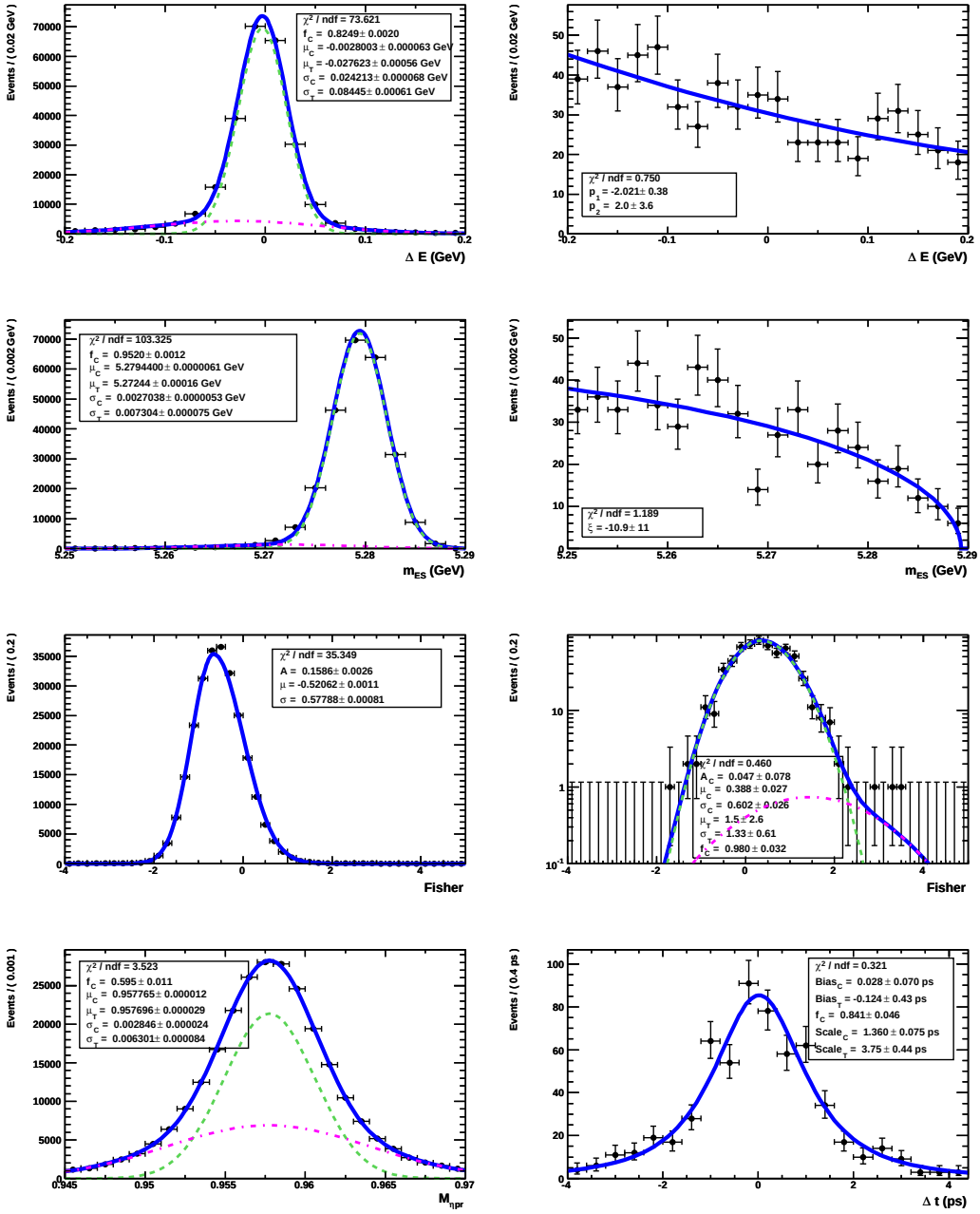


Figure C.2: PDFs for $\eta'_{\eta\pi\pi}K^0_S$ (from top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η' mass, and Δt. Signal MC (left) and on-peak sidebands (right).

C.3 $\eta'_{\rho\gamma}K^+$

```
Correlation matrix for onData (74164 events):
                de       mes    fisher    deltaT
     mes    0.0002
  fisher   -0.0360   -0.0609
  deltaT    0.0390   -0.0108    0.0164
   dtErr    0.0240   -0.0322   -0.0720    0.0040
```

The $\eta'_{\rho\gamma}K^+$ fit finds the following values for the parameters which were allowed to float:

```
  Floating Parameter  InitialValue    FinalValue +/-  Error     GblCorr.
--------------------  ------------  --------------------------  --------
Frac_nBkg_tagCat_04T1    8.1980e-03    3.4255e-03 +/-  2.53e-04  0.164035
Frac_nBkg_tagCat_04T2    9.1122e-02    8.8061e-02 +/-  1.10e-03  0.339724
Frac_nBkg_tagCat_04T3    1.6345e-01    1.6153e-01 +/-  1.42e-03  0.411235
Frac_nBkg_tagCat_04T4    1.2021e-01    1.1887e-01 +/-  1.25e-03  0.375262
Frac_nBkg_tagCat_04T5    1.3663e-01    1.3628e-01 +/-  1.33e-03  0.391314
Frac_nBkg_tagCat_04T6    1.1279e-01    1.1456e-01 +/-  1.23e-03  0.369186
           deBkg_P01   -1.3610e+00   -1.3064e+00 +/-  3.45e-02  0.110743
          dtBkgBiasC    5.0163e-03   -5.0812e-03 +/-  1.36e-02  0.603423
          dtBkgBiasT   -1.3547e-01    1.8434e-03 +/-  1.04e-02  0.603035
          dtBkgFracC    9.0714e-01    5.0000e-01 +/-  9.48e-04  0.079346
          dtBkgFracO    5.3891e-03    9.3073e-03 +/-  6.27e-04  0.399800
          dtBkgScfaC    1.2513e+00    1.6979e+00 +/-  9.74e-03  0.501043
          dtBkgScfaT    3.4955e+00    9.6305e-01 +/-  7.11e-03  0.377984
             dtSig_C    0.0000e+00    6.4849e-03 +/-  5.57e-02  0.055011
             dtSig_S    7.0000e-01   -6.4819e-02 +/-  6.85e-02  0.054524
        fisBkgC_asym    3.1409e-02    6.3497e-02 +/-  6.55e-03  0.204364
        fisBkgC_mean    4.1613e-01    4.3978e-01 +/-  2.37e-03  0.320130
         fisBkgC_rms    5.6825e-01    5.6031e-01 +/-  1.71e-03  0.255301
            mesBkg_c   -2.0381e+01   -1.8881e+01 +/-  9.93e-01  0.107010
                nBkg    1.0000e+04    6.9295e+04 +/-  2.80e+02  0.300165
               nChrm    1.7000e+03    1.5683e+03 +/-  1.09e+02  0.568917
                nSig    2.6000e+03    2.5561e+03 +/-  6.51e+01  0.205200
```

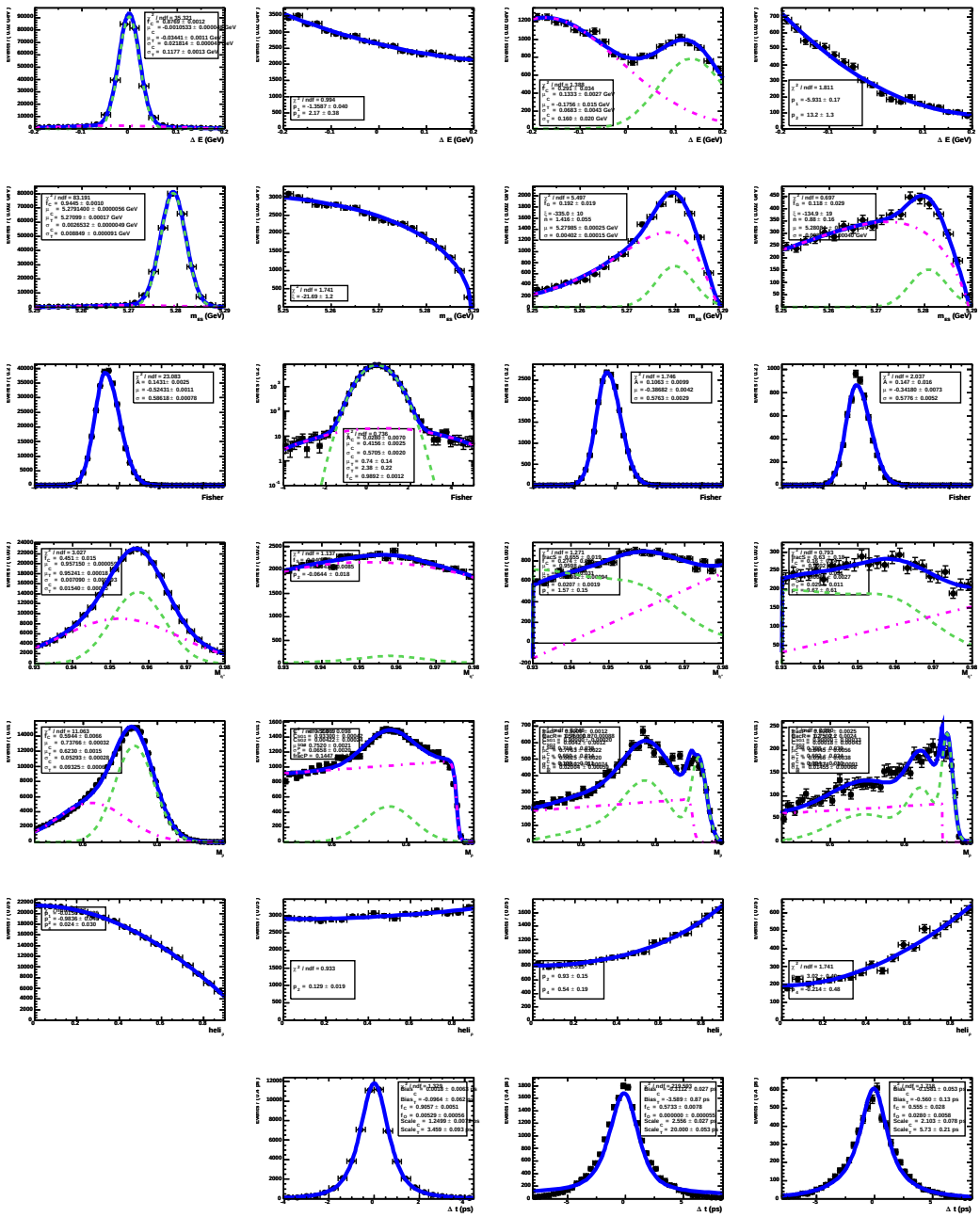


Figure C.3: PDFs for $\eta'_{\rho\gamma}K^+$ (from top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η' mass, ρ mass, ρ helicity, and Δt. Signal MC (left), on-peak sidebands (left-center), charmless $B\overline{B}$ (right-center), and $b \to c$ (right).

C.4 $\eta'_{\eta\pi\pi}K^+$

```
Correlation matrix for onData (3534 events):
                de       mes    fisher    deltaT
     mes    0.0270
  fisher   -0.0785   -0.3277
  deltaT    0.0326   -0.0205    0.0142
   dtErr   -0.0129   -0.0845    0.0626   -0.0377
```

The $\eta'_{\eta\pi\pi}K^+$ fit finds the following values for the parameters which were allowed to float:

```
  Floating Parameter  InitialValue    FinalValue +/-  Error     GblCorr.
--------------------  ------------  --------------------------  --------
Frac_nBkg_chgCat_Plus    5.0000e-01    4.9363e-01 +/-  1.08e-02  0.098221
Frac_nBkg_tagCat_04T1    3.3956e-02    6.7219e-03 +/-  1.89e-03  0.112705
Frac_nBkg_tagCat_04T2    9.8755e-02    7.9878e-02 +/-  5.85e-03  0.311961
Frac_nBkg_tagCat_04T3    1.5959e-01    1.4488e-01 +/-  7.55e-03  0.380443
Frac_nBkg_tagCat_04T4    1.1969e-01    1.1300e-01 +/-  6.78e-03  0.351106
Frac_nBkg_tagCat_04T5    1.4261e-01    1.4422e-01 +/-  7.52e-03  0.379102
Frac_nBkg_tagCat_04T6    1.0555e-01    1.1402e-01 +/-  6.78e-03  0.351351
Frac_nSig_chgCat_Plus    5.0000e-01    5.1212e-01 +/-  1.53e-02  0.105660
           deBkg_P01   -1.3945e+00   -1.6097e+00 +/-  1.95e-01  0.006408
          dtBkgBiasC    2.3112e-02    5.2414e-03 +/-  3.20e-02  0.257706
          dtBkgBiasT   -5.1271e-01   -6.0920e-01 +/-  3.86e-01  0.412189
          dtBkgFracC    9.0955e-01    9.1155e-01 +/-  2.04e-02  0.817211
          dtBkgFracO    1.0919e-02    8.9629e-03 +/-  3.68e-03  0.309387
          dtBkgScfaC    1.2740e+00    1.2677e+00 +/-  3.17e-02  0.648539
          dtBkgScfaT    3.6141e+00    3.4970e+00 +/-  3.94e-01  0.743575
             dtSig_C    0.0000e+00    6.1717e-02 +/-  6.86e-02  0.073521
             dtSig_S    7.0000e-01   -7.0309e-02 +/-  9.36e-02  0.030565
        fisBkgC_asym    5.8204e-02    8.8145e-02 +/-  3.65e-02  0.149136
        fisBkgC_mean    3.5933e-01    3.7496e-01 +/-  1.24e-02  0.143860
         fisBkgC_rms    5.4182e-01    5.3786e-01 +/-  9.20e-03  0.106163
                nBkg    2.0000e+03    2.3170e+03 +/-  4.99e+01  0.109206
                nSig    8.5000e+02    1.2172e+03 +/-  3.73e+01  0.123177
```

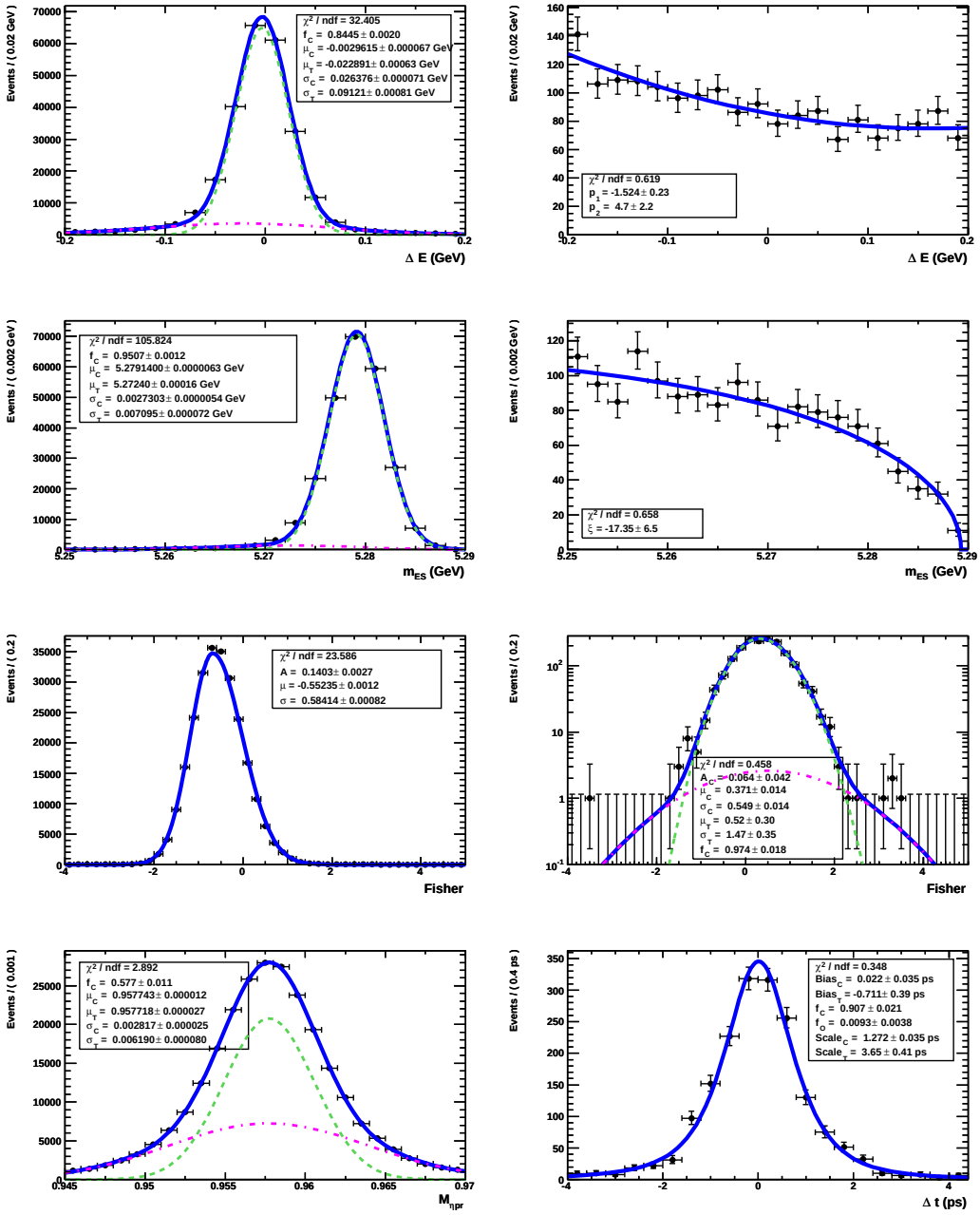


Figure C.4: PDFs for $\eta'_{\eta\pi\pi}K^+$ (from top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η' mass, and Δt. Signal MC (left) and on-peak sidebands (right).

C.5 $\eta'_{\eta\pi\pi}K^0_{s00}$

```
Correlation matrix for onData (1056 events):
                de       mes    fisher    deltaT
     mes    0.0793
  fisher   -0.0371   -0.1363
  deltaT   -0.0103    0.0267    0.0341
   dtErr    0.0555    0.0006   -0.0380   -0.0141
```

The $\eta'_{\eta\pi\pi}K^0_{s00}$ fit finds the following values for the parameters which were allowed to float:

```
  Floating Parameter  InitialValue    FinalValue +/-  Error     GblCorr.
--------------------  ------------  --------------------------  --------
Frac_nBkg_tagCat_04T1    1.0476e-02    2.1759e-03 +/-  1.54e-03  0.053127
Frac_nBkg_tagCat_04T2    5.3333e-02    4.8480e-02 +/-  7.17e-03  0.238969
Frac_nBkg_tagCat_04T3    1.3238e-01    1.2264e-01 +/-  1.09e-02  0.335905
Frac_nBkg_tagCat_04T4    1.1048e-01    1.0849e-01 +/-  1.03e-02  0.319745
Frac_nBkg_tagCat_04T5    1.4476e-01    1.4788e-01 +/-  1.17e-02  0.351393
Frac_nBkg_tagCat_04T6    1.2571e-01    1.2789e-01 +/-  1.10e-02  0.336454
           deBkg_P01   -2.3579e+00   -1.8366e+00 +/-  3.18e-01  0.018273
          dtBkgBiasC    1.9152e-02    5.1068e-02 +/-  5.11e-02  0.131092
          dtBkgBiasT   -1.1332e+00   -2.0471e+00 +/-  1.07e+00  0.488166
          dtBkgFracC    9.3064e-01    9.2688e-01 +/-  1.74e-02  0.678028
          dtBkgScfaC    1.3954e+00    1.3784e+00 +/-  4.93e-02  0.561894
          dtBkgScfaT    5.0000e+00    6.1516e+00 +/-  9.44e-01  0.592485
             dtSig_C    0.0000e+00   -1.9316e-01 +/-  2.96e-01  0.110068
             dtSig_S    7.0000e-01    5.0954e-01 +/-  3.39e-01  0.096802
        fisBkgC_asym    1.2866e-01    2.9410e-02 +/-  5.97e-02  0.112705
        fisBkgC_mean    4.2706e-01    4.3033e-01 +/-  1.86e-02  0.197662
         fisBkgC_rms    5.3615e-01    5.2915e-01 +/-  1.37e-02  0.144102
            mesBkg_c   -1.7360e+01   -3.5872e+01 +/-  8.46e+00  0.165908
                nBkg    6.0000e+02    9.5069e+02 +/-  3.19e+01  0.169650
                nSig    2.7500e+02    1.0530e+02 +/-  1.30e+01  0.295803
```

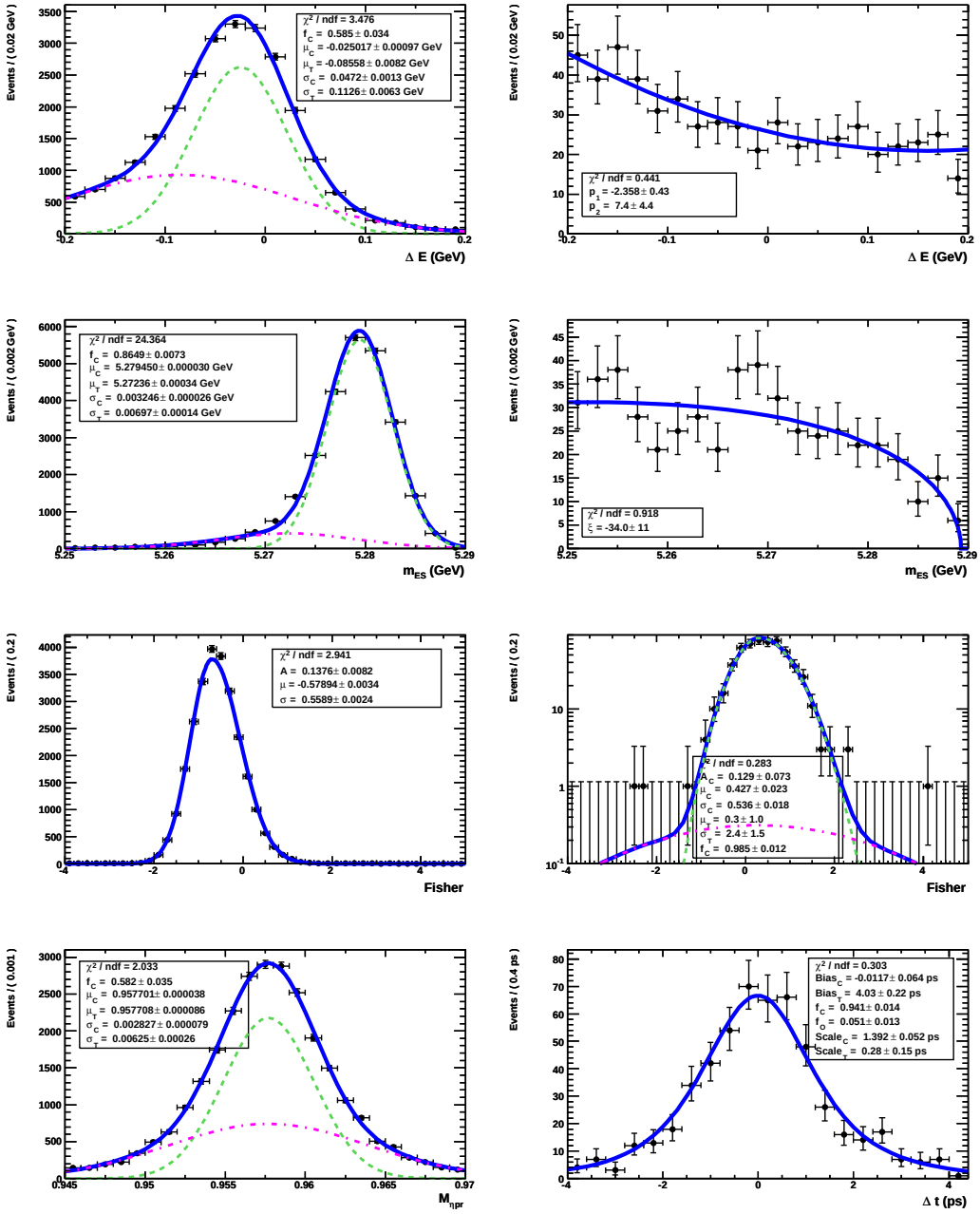


Figure C.5: PDFs for $\eta'_{\eta\pi\pi}K^0_{S00}$ (from top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η' mass, and Δt. Signal MC (left) and on-peak sidebands (right).

C.6 $\eta'_{\rho\gamma}K^0_{s00}$

```
Correlation matrix for onData (27057 events):
                de       mes    fisher    deltaT
     mes   -0.0095
  fisher   -0.0453    0.0048
  deltaT    0.0257   -0.0151    0.0019
   dtErr    0.0246   -0.0207   -0.0494    0.0586
```

The $\eta'_{\rho\gamma}K^0_{s00}$ fit finds the following values for the parameters which were allowed to float:

Floating Parameter	InitialValue	FinalValue +/- Error	GblCorr.
Frac_nBkg_tagCat_04T1	3.7210e-03	1.6482e-03 +/- 3.08e-04	0.252397
Frac_nBkg_tagCat_04T2	5.0910e-02	4.8346e-02 +/- 1.37e-03	0.279427
Frac_nBkg_tagCat_04T3	1.2742e-01	1.2543e-01 +/- 2.08e-03	0.351220
Frac_nBkg_tagCat_04T4	1.1686e-01	1.1610e-01 +/- 2.01e-03	0.338131
Frac_nBkg_tagCat_04T5	1.4421e-01	1.4456e-01 +/- 2.20e-03	0.357407
Frac_nBkg_tagCat_04T6	1.3211e-01	1.3335e-01 +/- 2.12e-03	0.350237
deBkg_P01	-1.8169e+00	-1.8289e+00 +/- 5.36e-02	0.127598
dtBkgBiasC	4.6024e-02	4.7696e-02 +/- 9.85e-03	0.365159
dtBkgBiasT	4.9437e-01	4.2760e-01 +/- 7.20e-02	0.466791
dtBkgFracC	7.4756e-01	8.4892e-01 +/- 1.21e-02	0.929147
dtBkgFracO	1.4792e-02	1.3162e-02 +/- 1.45e-03	0.543209
dtBkgScfaC	1.2144e+00	1.1866e+00 +/- 1.30e-02	0.825596
dtBkgScfaT	3.0702e+00	2.9493e+00 +/- 1.21e-01	0.901146
dtSig_C	0.0000e+00	3.8985e-02 +/- 2.61e-01	0.112608
dtSig_S	4.0000e-01	2.5793e-01 +/- 3.32e-01	0.125628
fisBkgC_asym	3.3376e-02	4.9351e-02 +/- 1.10e-02	0.287759
fisBkgC_mean	4.5140e-01	4.7755e-01 +/- 4.06e-03	0.470505
fisBkgC_rms	5.6411e-01	5.5508e-01 +/- 2.89e-03	0.391116
mesBkg_c	-2.1421e+01	-1.9060e+01 +/- 1.60e+00	0.101165
nBkg	1.0000e+04	2.6251e+04 +/- 1.79e+02	0.404420
nChrm	3.9300e+02	5.3033e+02 +/- 8.42e+01	0.704531
nSig	6.6000e+02	2.0581e+02 +/- 2.76e+01	0.365266

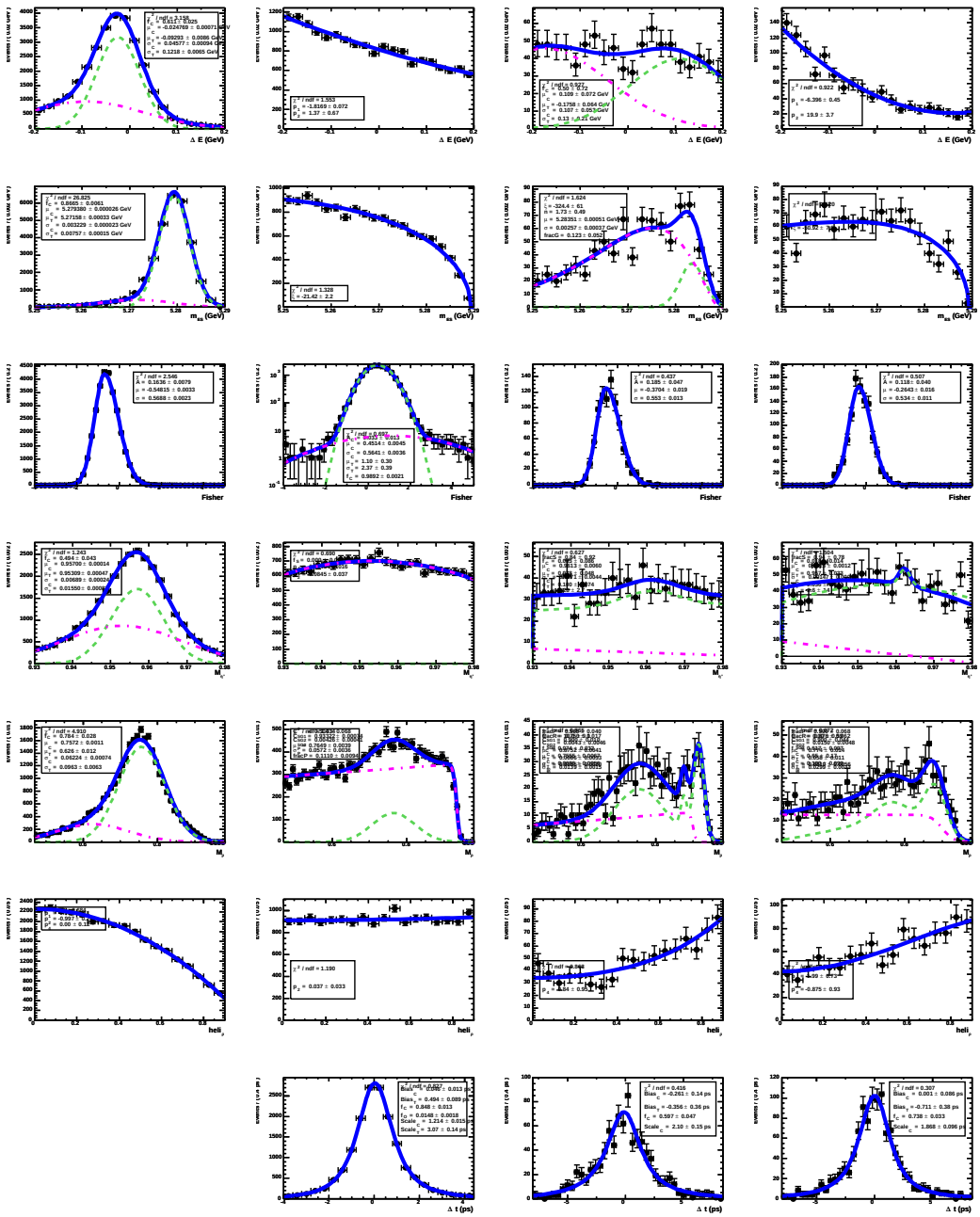


Figure C.6: PDFs for $\eta'_{\rho\gamma}K^0_{S00}$ (from top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η' mass, ρ mass, ρ helicity, and Δt. Signal MC (left), on-peak sidebands (left-center), charmless $B\overline{B}$ (right-center), and $b \to c$ (right).

C.7 $\eta'_{5\pi}K^0_S$

```
Correlation matrix for onData (513 events):
               de       mes    fisher    deltaT
    mes    0.0201
 fisher   -0.0939   -0.3345
 deltaT   -0.0035    0.0364   -0.0555
  dtErr    0.0668    0.0024    0.0524   -0.0165
```

The $\eta'_{5\pi}K^0_S$ fit finds the following values for the parameters which were allowed to float:

Floating Parameter	InitialValue	FinalValue +/- Error	GblCorr.
Frac_nBkg_tagCat_04T1	3.1434e-02	4.8646e-03 +/- 4.34e-03	0.216354
Frac_nBkg_tagCat_04T2	7.6621e-02	7.3156e-02 +/- 1.46e-02	0.300190
Frac_nBkg_tagCat_04T3	1.5914e-01	1.5824e-01 +/- 2.02e-02	0.385378
Frac_nBkg_tagCat_04T4	1.1788e-01	9.4403e-02 +/- 1.63e-02	0.328415
Frac_nBkg_tagCat_04T5	1.5324e-01	1.3603e-01 +/- 1.92e-02	0.372333
Frac_nBkg_tagCat_04T6	1.1395e-01	1.3010e-01 +/- 1.87e-02	0.364487
deBkg_P01	-2.0959e+00	-1.8701e+00 +/- 5.04e-01	0.025418
dtBkgBiasC	-1.7224e-01	-6.9467e-02 +/- 1.01e-01	0.292376
dtBkgBiasT	-5.6099e+00	-5.6285e+00 +/- 9.57e-01	0.746894
dtBkgFracC	9.6091e-01	9.5968e-01 +/- 1.77e-02	0.820812
dtBkgFracO	7.7409e-03	1.8389e-02 +/- 1.26e-02	0.752459
dtBkgScfaC	1.6217e+00	1.6294e+00 +/- 9.20e-02	0.509378
dtBkgScfaT	1.4041e+00	1.2792e+00 +/- 6.46e-01	0.671039
dtSig_C	0.0000e+00	5.1998e-02 +/- 1.98e-01	0.127634
dtSig_S	7.0000e-01	7.5704e-01 +/- 2.58e-01	0.182663
fisBkgC_asym	5.0965e-02	1.2670e-01 +/- 1.01e-01	0.033996
fisBkgC_mean	3.5288e-01	3.6376e-01 +/- 3.31e-02	0.153675
fisBkgC_rms	5.4632e-01	5.7072e-01 +/- 2.37e-02	0.117309
mesBkg_c	-2.0170e+01	-1.8321e+01 +/- 1.41e+01	0.156459
nBkg	6.0000e+02	3.4213e+02 +/- 1.92e+01	0.142790
nSig	2.7500e+02	1.7086e+02 +/- 1.41e+01	0.176318

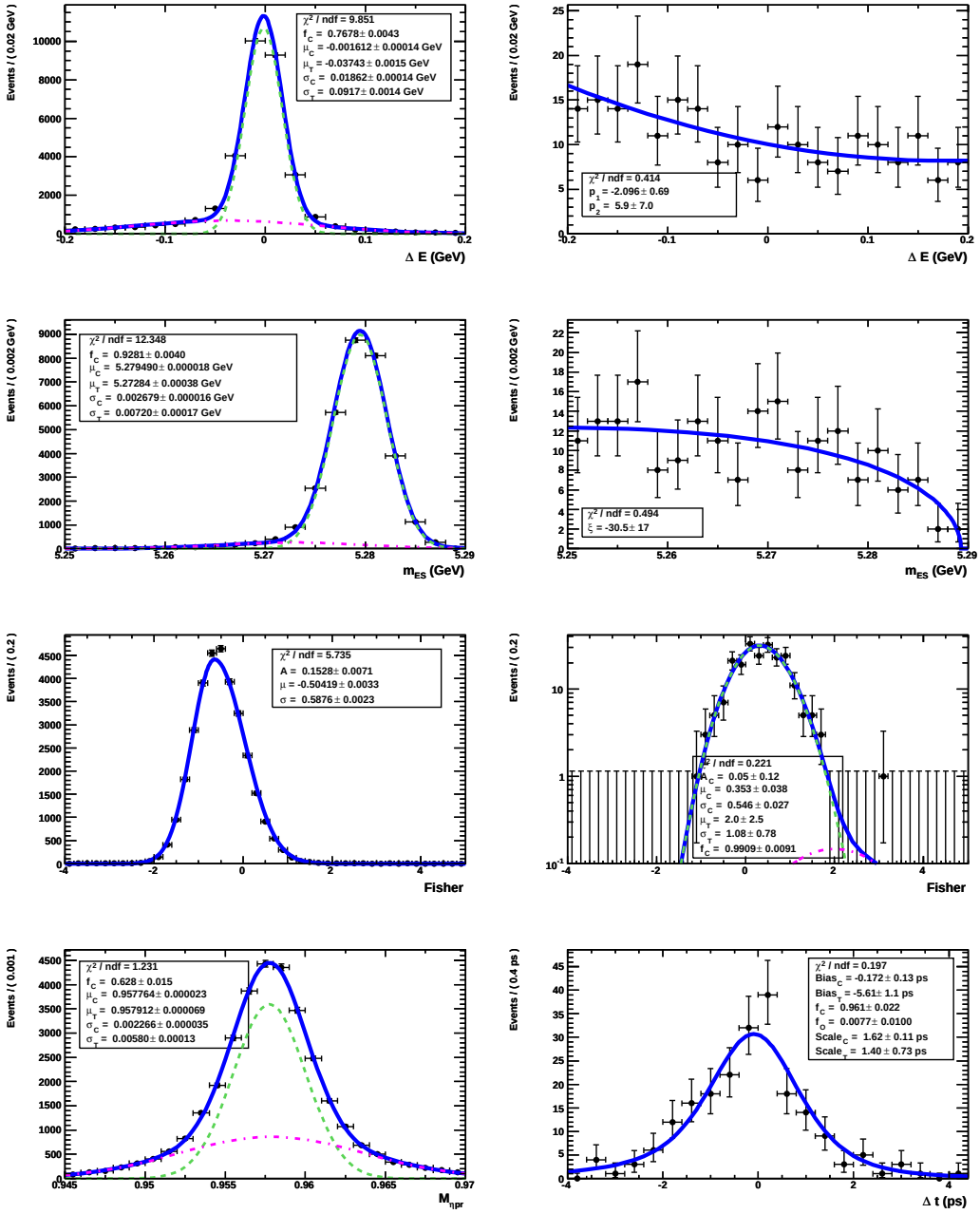


Figure C.7: PDFs for $\eta'_{5\pi}K^0_S$ (from top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η' mass, and Δt. Signal MC (left) and on-peak sidebands (right).

C.8 $\eta'_{\eta\pi\pi}K^0_L$

```
Correlation matrix for onData (12217 events):
                de    fisher    deltaT
  fisher    0.0487
  deltaT   -0.0016   -0.0041
   dtErr    0.0253    0.0302    0.0237
```

The $\eta'_{\eta\pi\pi}K^0_L$ fit finds the following values for the parameters which were allowed to float:

Floating Parameter	InitialValue	FinalValue +/- Error	GblCorr.
Frac_nBkg_tagCat_04T1	1.0426e-02	7.8762e-03 +/- 8.50e-04	0.133473
Frac_nBkg_tagCat_04T2	6.3384e-02	6.1065e-02 +/- 2.24e-03	0.291768
Frac_nBkg_tagCat_04T3	1.4084e-01	1.3986e-01 +/- 3.22e-03	0.386606
Frac_nBkg_tagCat_04T4	1.2065e-01	1.1943e-01 +/- 3.01e-03	0.367659
Frac_nBkg_tagCat_04T5	1.6218e-01	1.6329e-01 +/- 3.43e-03	0.403344
Frac_nBkg_tagCat_04T6	1.2106e-01	1.2259e-01 +/- 3.04e-03	0.369948
deBkg_c	-9.1516e+00	-7.8574e+00 +/- 4.03e-01	0.241539
dtBkgBiasC	1.3170e-03	2.9246e-03 +/- 1.66e-02	0.345070
dtBkgBiasT	-2.4871e-01	-7.4765e-02 +/- 1.00e-01	0.398035
dtBkgFracC	8.9992e-01	8.0874e-01 +/- 2.07e-02	0.934798
dtBkgFracO	2.4000e-02	2.4741e-02 +/- 3.73e-03	0.623593
dtBkgScfaC	1.4044e+00	1.2915e+00 +/- 2.35e-02	0.841679
dtBkgScfaT	5.2648e+00	3.1398e+00 +/- 1.86e-01	0.914210
dtSig_C	0.0000e+00	6.9601e-02 +/- 1.93e-01	0.042170
dtSig_negS	7.0000e-01	6.4514e-01 +/- 2.20e-01	0.076511
fisBkgC_asym	1.2811e-01	1.3931e-01 +/- 1.52e-02	0.128343
fisBkgC_mean	8.4241e-02	9.0187e-02 +/- 5.26e-03	0.257486
fisBkgC_rms	5.3314e-01	5.2750e-01 +/- 3.84e-03	0.180913
nBkg	4.0000e+03	1.1854e+04 +/- 1.12e+02	0.191251
nSig	2.0000e+02	3.4100e+02 +/- 3.20e+01	0.374348

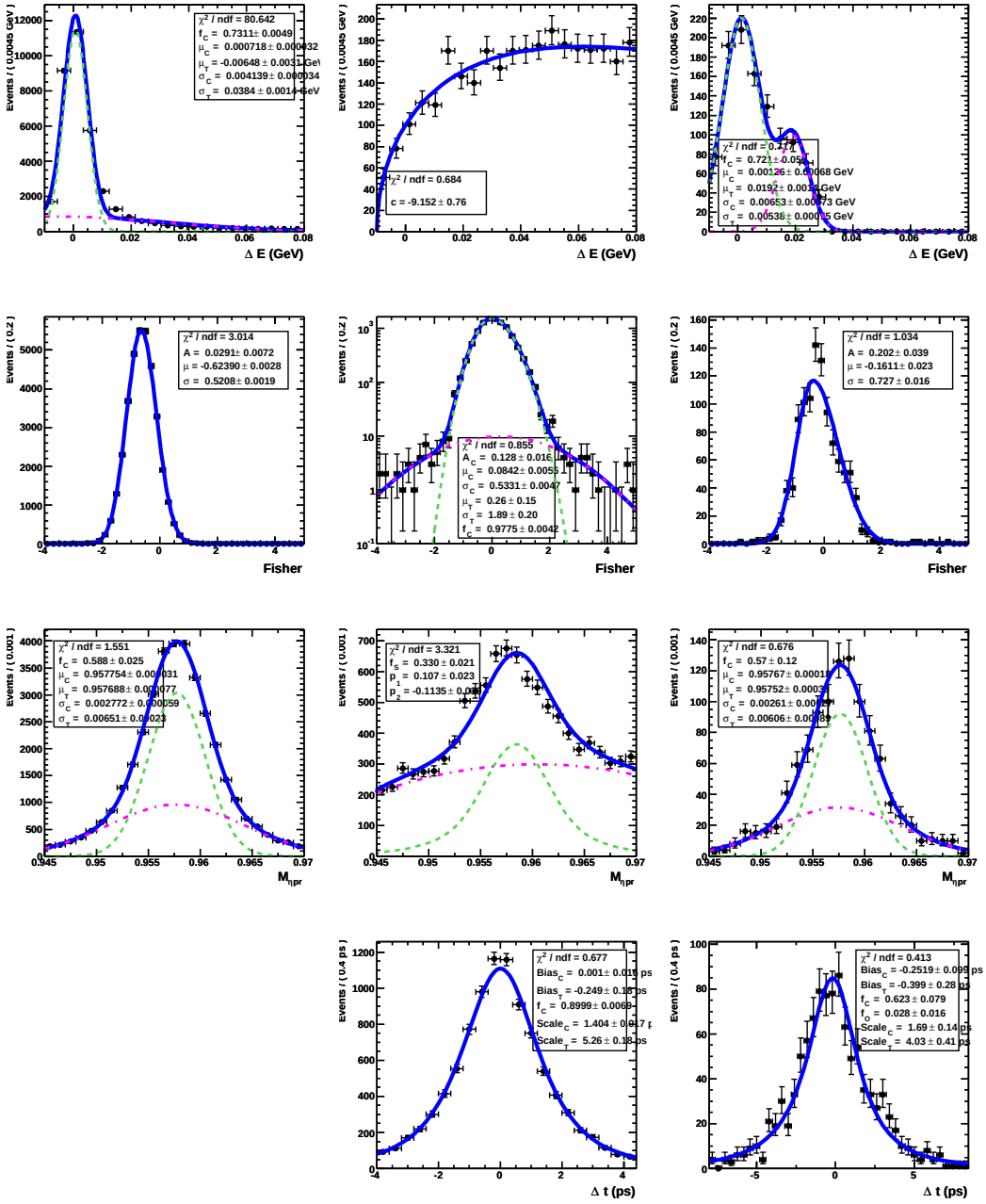


Figure C.8: PDFs for $\eta'_{\eta\pi\pi}K^0_L$ (from top to bottom) ΔE, $\mathcal{F}$, η' mass, and Δt for signal (left), continuum background (middle), and charmless $B\overline{B}$ background (right). The continuum ΔE PDF is obtained from *udsc* MC; all other continuum PDFs are obtained from ΔE sidebands.

C.9 $\eta'_{5\pi}K^0_L$

```
Correlation matrix for onData (4586 events):
               de    fisher    deltaT
  fisher    0.0869
  deltaT   -0.0106    0.0192
   dtErr    0.0493    0.0303    0.0482
```

The $\eta'_{5\pi}K^0_L$ fit finds the following values for the parameters which were allowed to float:

Floating Parameter	InitialValue	FinalValue +/- Error	GblCorr.
Frac_nBkg_tagCat_04T1	1.1238e-02	8.5706e-03 +/- 1.46e-03	0.136895
Frac_nBkg_tagCat_04T2	6.5668e-02	6.4444e-02 +/- 3.76e-03	0.295656
Frac_nBkg_tagCat_04T3	1.4874e-01	1.4941e-01 +/- 5.42e-03	0.393951
Frac_nBkg_tagCat_04T4	1.2561e-01	1.2465e-01 +/- 5.02e-03	0.372193
Frac_nBkg_tagCat_04T5	1.5800e-01	1.5776e-01 +/- 5.54e-03	0.399234
Frac_nBkg_tagCat_04T6	1.0974e-01	1.1055e-01 +/- 4.76e-03	0.357566
deBkg_chi	-4.8974e+00	-5.0627e+00 +/- 6.75e-01	0.279117
dtBkgBiasC	-2.8701e-02	-1.4106e-02 +/- 3.49e-02	0.361255
dtBkgBiasT	-9.9090e-02	-6.7400e-02 +/- 1.36e-01	0.375255
dtBkgFracC	7.3257e-01	7.3182e-01 +/- 2.91e-02	0.902257
dtBkgScfaC	1.4722e+00	1.4698e+00 +/- 4.75e-02	0.818469
dtBkgScfaT	3.8640e+00	3.8416e+00 +/- 1.78e-01	0.832529
dtSig_C	0.0000e+00	2.1117e-02 +/- 2.57e-01	0.038189
dtSig_negS	7.0000e-01	6.5959e-01 +/- 4.61e-01	0.099934
fisBkgC_asym	1.3366e-01	1.3723e-01 +/- 2.50e-02	0.200287
fisBkgC_mean	7.5921e-02	7.8163e-02 +/- 8.75e-03	0.289068
fisBkgC_rms	5.2826e-01	5.1973e-01 +/- 6.47e-03	0.191620
nBkg	4.0000e+03	4.4275e+03 +/- 6.88e+01	0.211902
nSig	6.2000e+01	1.5869e+02 +/- 2.16e+01	0.410331

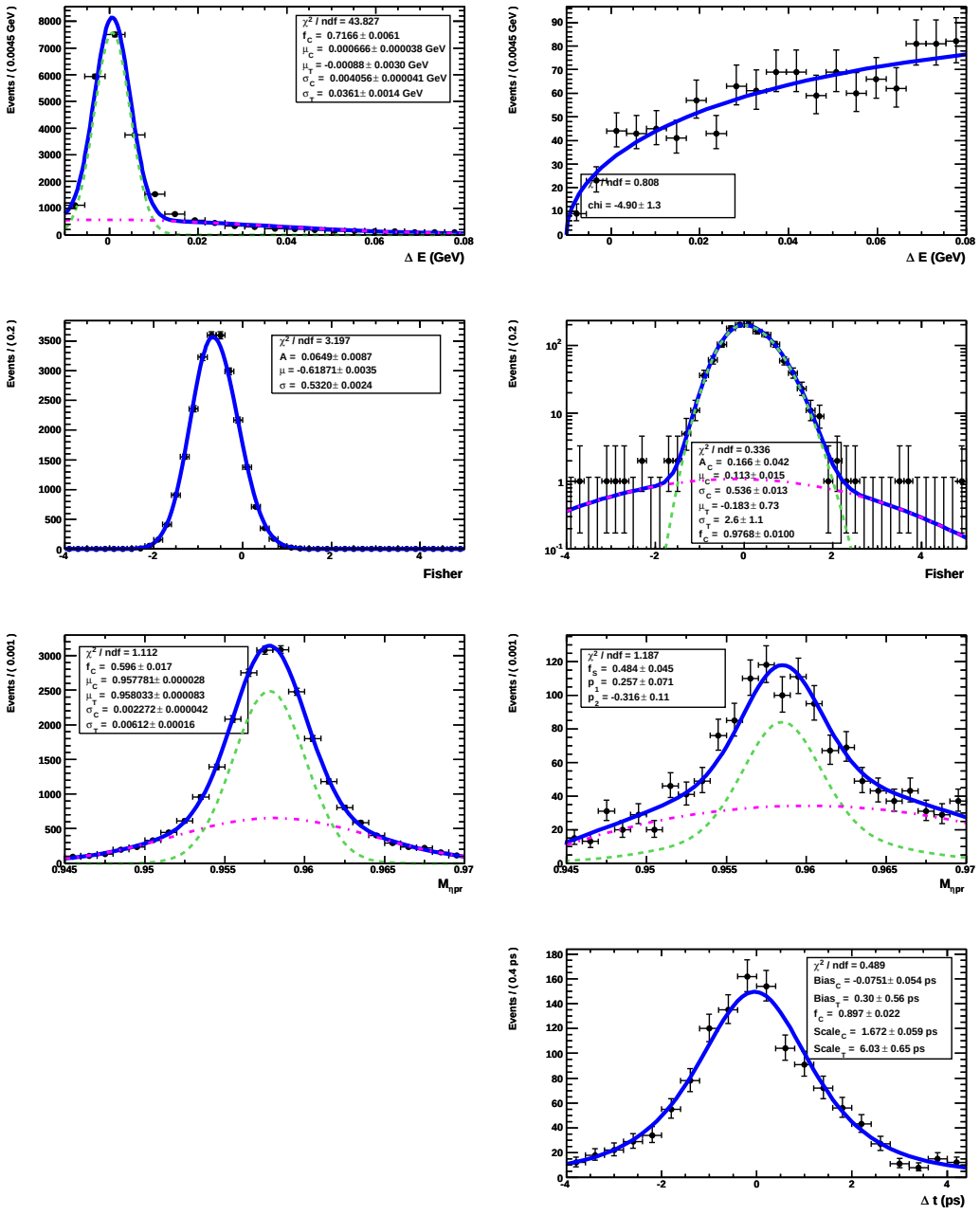


Figure C.9: PDFs for $\eta'_{5\pi}K^0_L$ (from top to bottom) ΔE, $\mathcal{F}$, η' mass, and Δt for signal (left) and continuum background (right). The continuum ΔE PDF is obtained from *udsc* MC; all other continuum PDFs are obtained from ΔE sidebands.

Appendix D

ΔE Distribution in $\eta' K_L^0$

In the modes $\eta'_{\eta\pi\pi} K_L^0$ and $\eta'_{5\pi} K_L^0$, we constrain the mass of the B to the PDG value during vertexing. This causes the continuum and signal ΔE distributions to have the unique shape, with a low-side cut-off, seen in Figs. C.8 and C.9. The low-side endpoint‘ of these shapes is related to the familiar upper endpoint of the $m_{\rm ES}$ Argus shapes in modes with no B mass constraint, such as in Fig. C.1. We can see this relationship by writing $m_{\rm ES}$ and ΔE in the center-of-mass frame

$$m_{\rm ES}{}^2 = E^{*2}_{\rm beam} - p^{*2}_{\rm B}, \tag{D.1}$$

$$\Delta E = E^*_{\rm B} - E^*_{\rm beam}. \tag{D.2}$$

With the B mass constraint, $m_{\rm ES}$ becomes

$$m_{\rm ES}{}^2 = E^{*2}_{\rm beam} - (E^{*2}_{\rm B} - m^2_{\rm B}) \tag{D.3}$$

$$= (E^*_{\rm beam} - E^*_{\rm B})(E^*_{\rm beam} + E^*_{\rm B}) + m^2_{\rm B} \tag{D.4}$$

$$= -\Delta E(E^*_{\rm beam} + E^*_{\rm B}) + m^2_{\rm B}. \tag{D.5}$$

Solving for ΔE, we obtain

$$\Delta E = \frac{m^2_{\rm B} - m_{\rm ES}{}^2}{E^*_{\rm beam} + E^*_{\rm B}} \tag{D.6}$$

$$= \frac{m^2_{\rm B} - m_{\rm ES}{}^2}{\frac{1}{2}\sqrt{s} + E^*_{\rm B}}. \tag{D.7}$$

It is clear that ΔE will be a minimum when $m_{\rm ES}{}^2$ is a maximum and $E^*_{\rm B}$ is a minimum. Theoretically, the smallest that $E^*_{\rm B}$ can be is half the mass of the $\Upsilon(4S)$ and the largest

that $m_{\rm ES}$ can be is 5.2893 GeV. Inserting these values into Equation D.7, we find the lower limit to be

$$\Delta E \;\; \geq \;\; -0.01029 \text{ GeV}. \tag{D.8}$$

We fix the minima of the continuum ΔE PDFs in Figs. C.8 and C.9 at -0.01 GeV using the Argus-like function

$$\frac{\sqrt{x - x_0}}{(1 - (x + x_0))^2} \; \mathrm{e}^{\chi(x+x_0)} \, . \tag{D.9}$$

In Equation D.9, x is the independent variable ΔE, the parameter x_0 is the low-side endpoint which we fix at -0.01, and the Argus shape parameter is χ. In the $\eta'_{\eta\pi\pi}K^0_L$ PDF fit we find $\chi = -7.6 \pm 0.7$.

Appendix E

$\eta' K^0$: $s\mathcal{P}lot$ libraries

We show here the $s\mathcal{P}lots$ for the two charged modes and the seven neutral modes. In the $\eta'_{\rho\gamma}K^0_S$, $\eta'_{\rho\gamma}K^+$, and $\eta'_{\rho\gamma}K^0_{S00}$ $s\mathcal{P}lots$, so that we can view $s\mathcal{P}lots$ for all components, we float the charmless yield even though it is fixed to the expected value in the final fit.

E.1 $\eta'_{\rho\gamma}K^0_S$

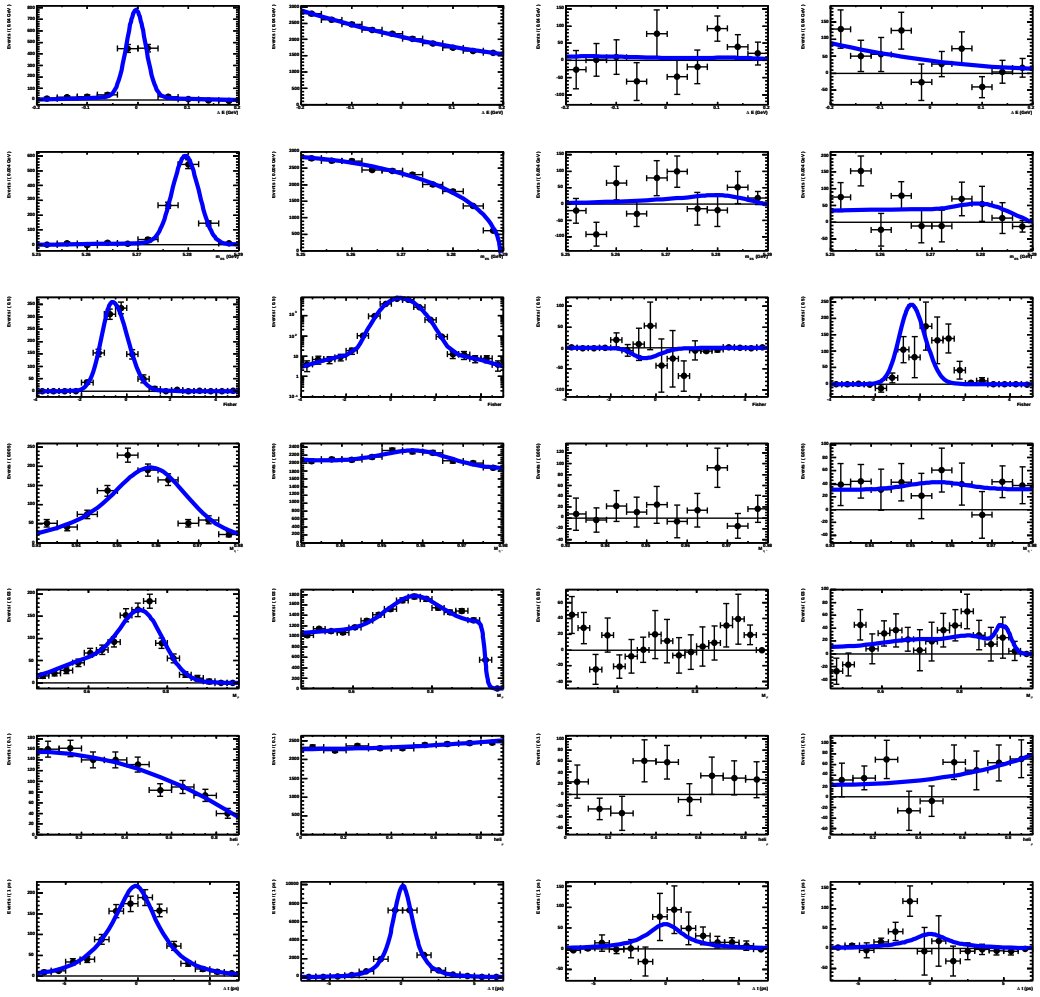


Figure E.1: $s\mathcal{P}lots$ for $\eta'_{\rho\gamma}K^0_S$ (from top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η' mass, ρ mass, ρ helicity, and Δt. Signal MC (left), continuum background (left-center), charmless $B\overline{B}$ (right-center), and $b \to c$ (right).

E.2 $\eta'_{\eta\pi\pi}K^0_S$

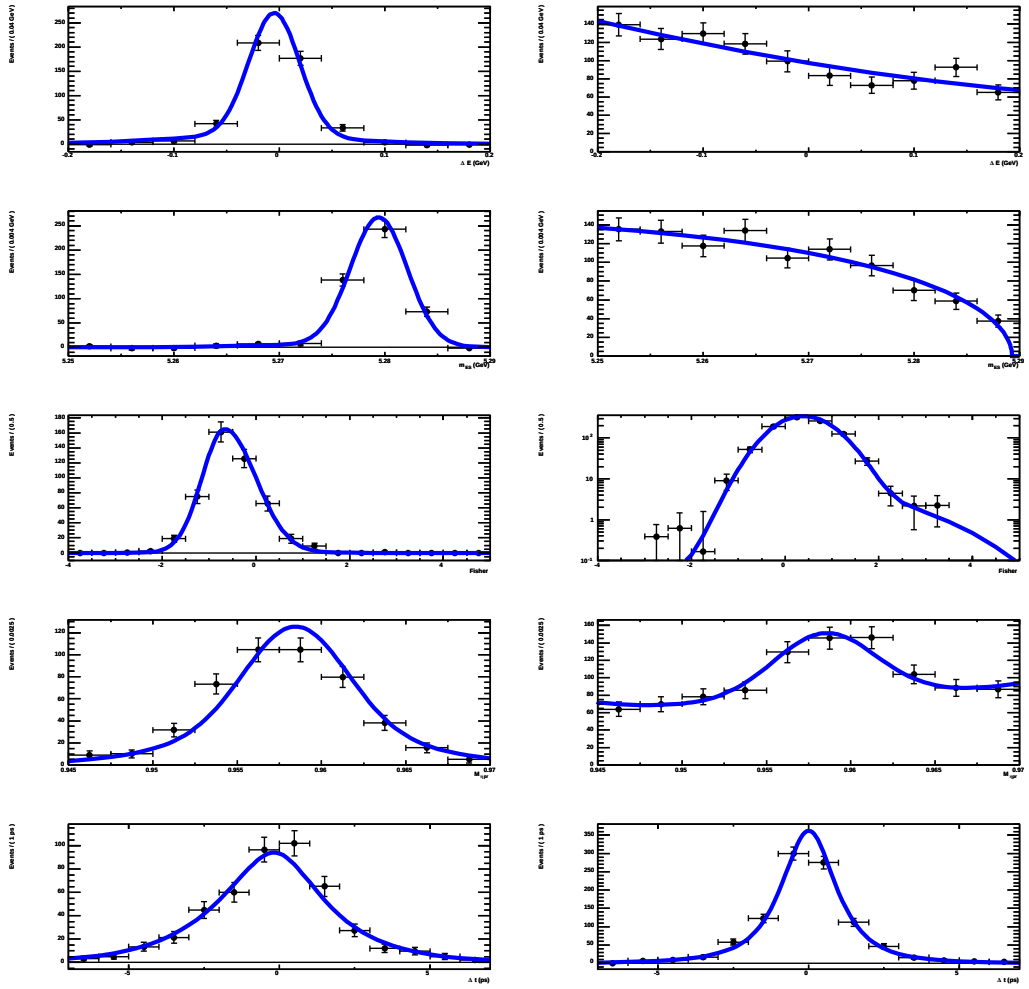


Figure E.2: *sPlots* for $\eta'_{\eta\pi\pi}K^0_S$ (from top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η' mass, and Δt. Signal MC (left) and continuum background (right).

E.3 $\eta'_{\rho\gamma}K^+$

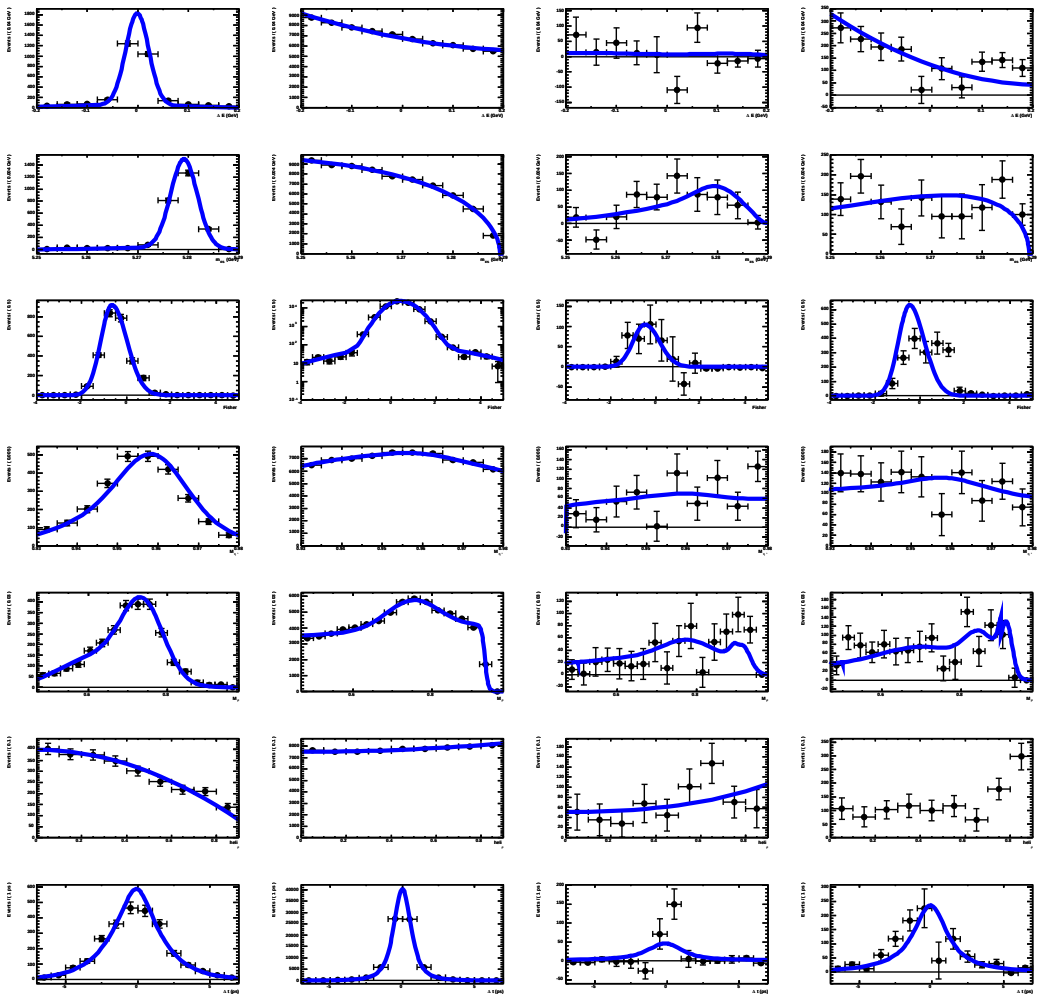


Figure E.3: *sPlots* for $\eta'_{\rho\gamma}K^0_S$ (from top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η' mass, ρ mass, ρ helicity, and Δt. Signal MC (left), continuum background (left-center), charmless $B\overline{B}$ (right-center), and $b \to c$ (right).

E.4 $\eta'_{\eta\pi\pi}K^+$

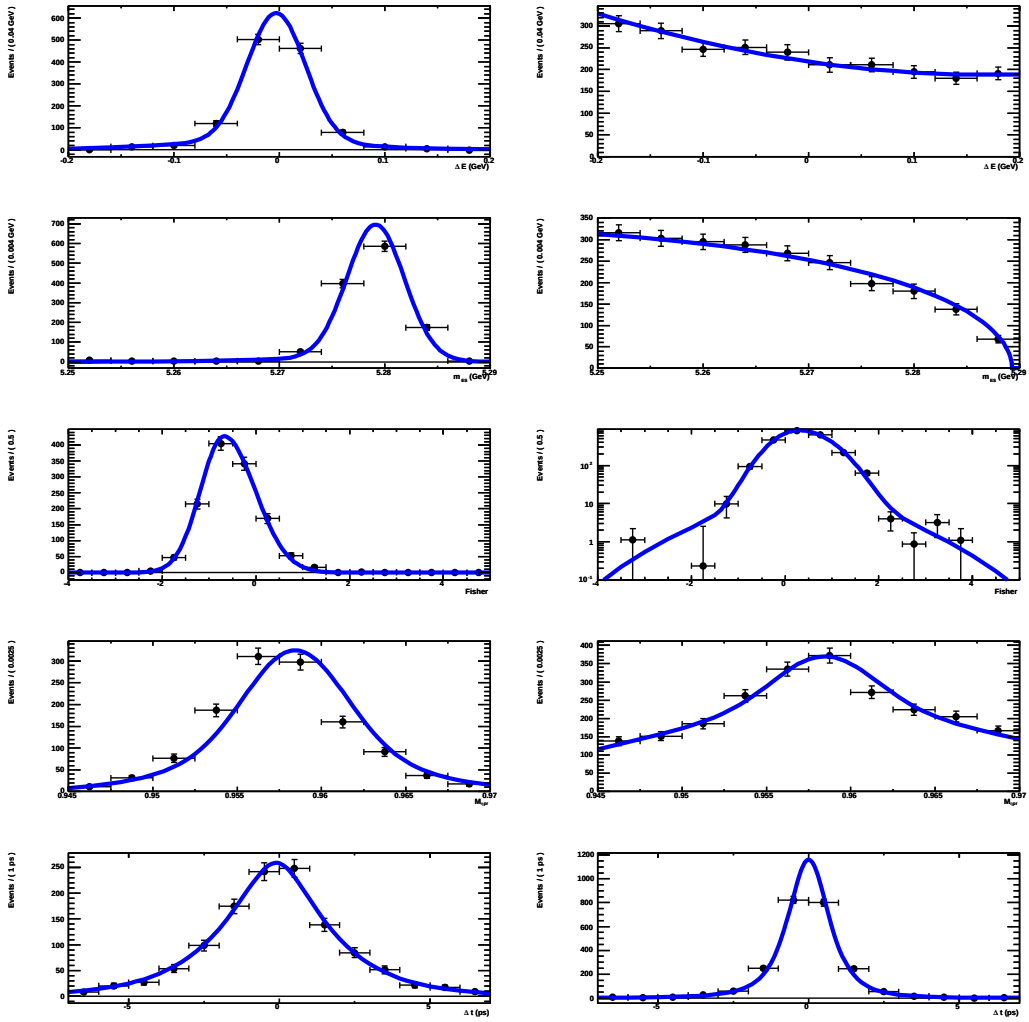


Figure E.4: *sPlots* for $\eta'_{\eta\pi\pi}K^+$ (from top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η' mass, and Δt. Signal MC (left) and continuum background (right).

E.5 $\eta'_{\eta\pi\pi}K^0_{s00}$

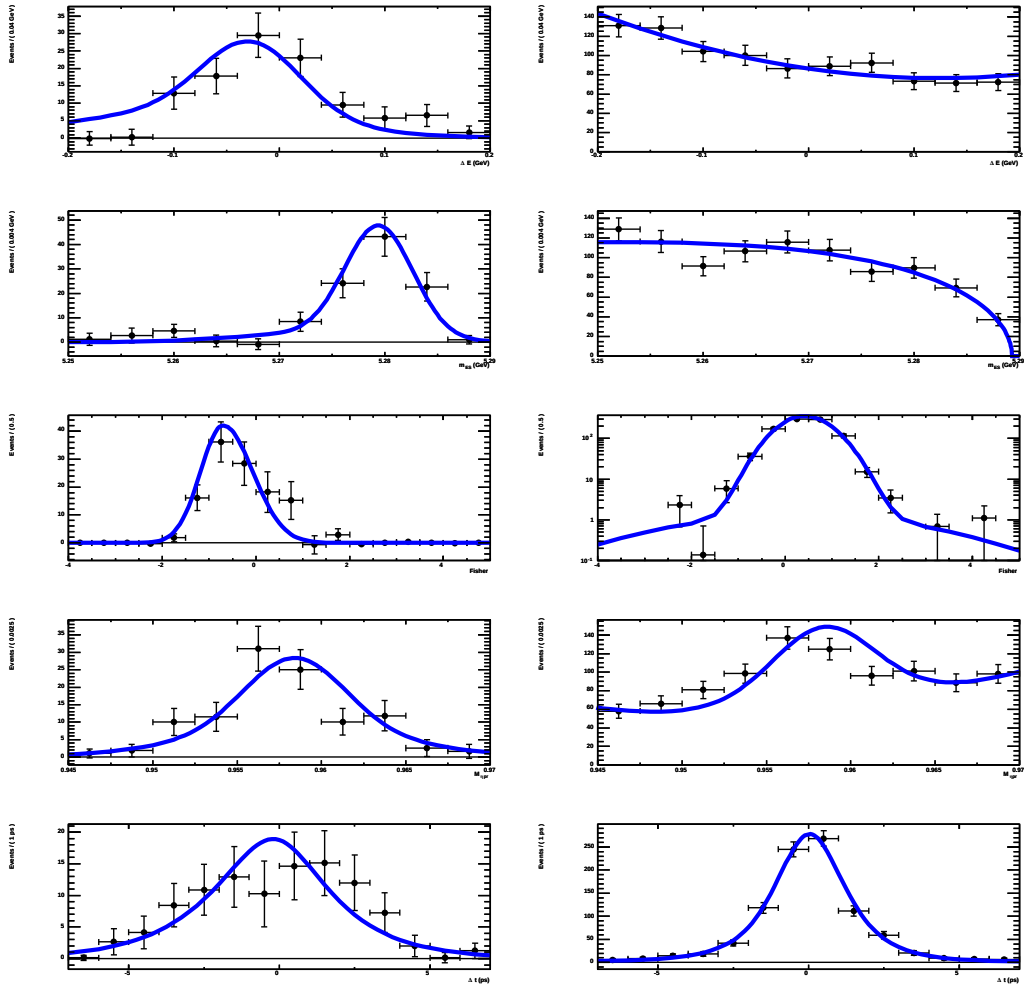


Figure E.5: *sPlots* for $\eta'_{\eta\pi\pi}K^0_{s00}$ (from top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η' mass, and Δt. Signal MC (left) and continuum background (right).

E.6 $\eta'_{\rho\gamma}K^0_{s00}$

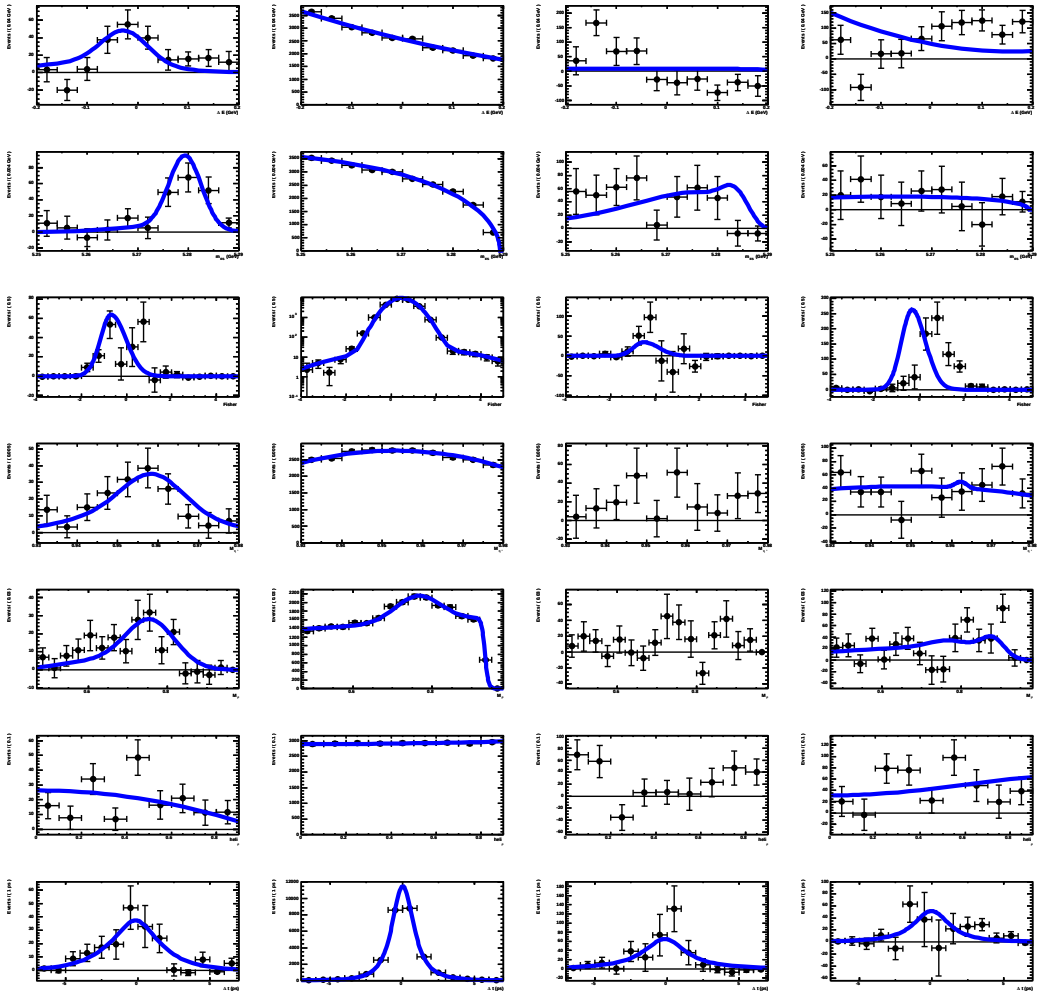


Figure E.6: $s\mathcal{P}lots$ for $\eta'_{\rho\gamma}K^0_{s00}$ (from top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η' mass, ρ mass, ρ helicity, and Δt. Signal MC (left), continuum background (left-center), charmless $B\overline{B}$ (right-center), and $b \to c$ (right).

E.7 $\eta'_{5\pi}K^0_S$

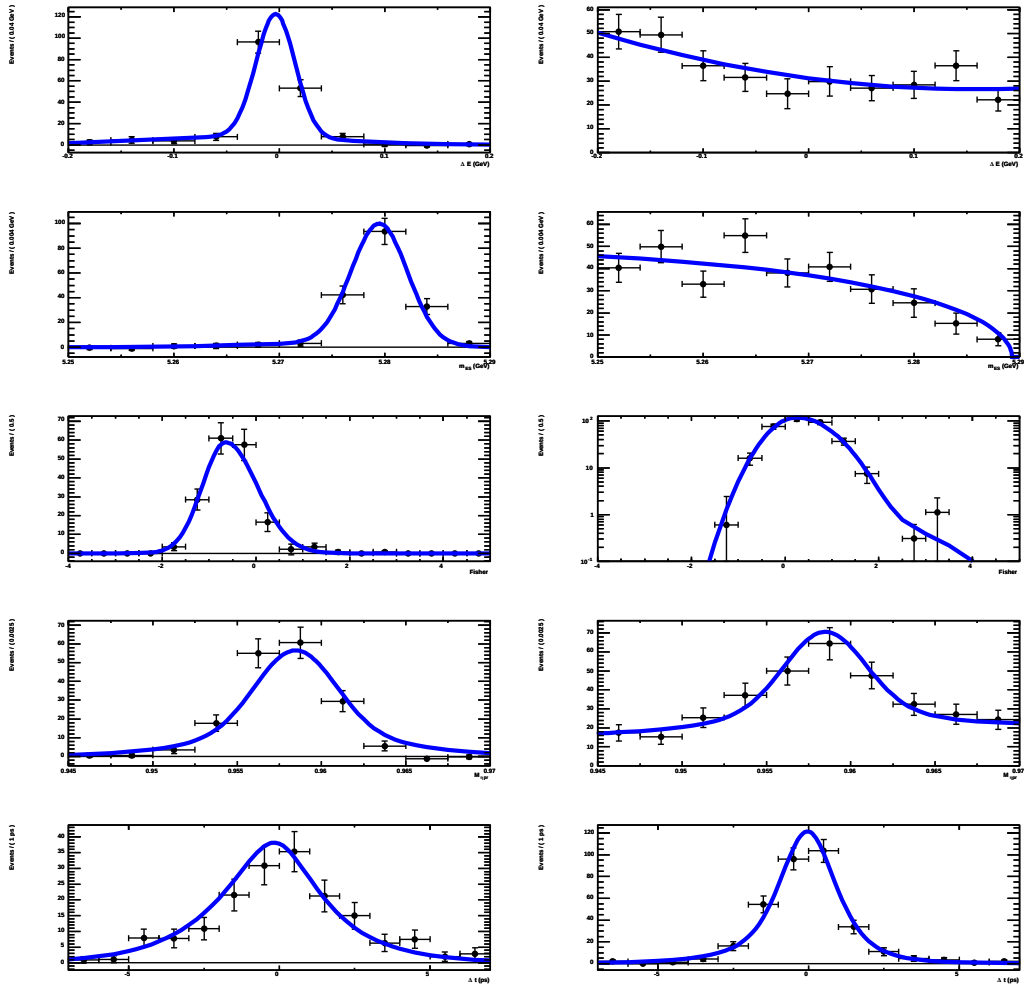


Figure E.7: *sPlots* for $\eta'_{5\pi}K^0_S$ (from top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η' mass, and Δt. Signal MC (left) and continuum background (right).

E.8 $\eta'_{\eta\pi\pi}K^0_L$

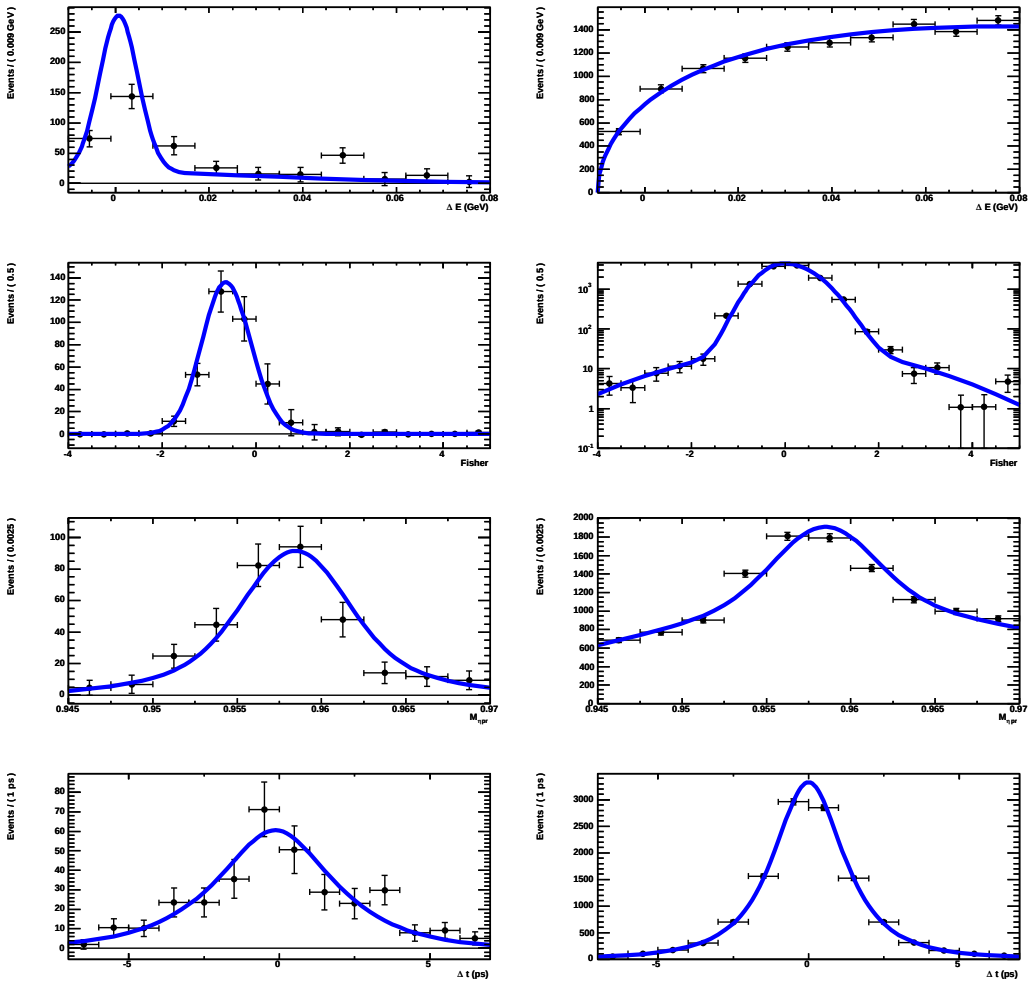


Figure E.8: $s\mathcal{P}lots$ for $\eta'_{\eta\pi\pi}K^0_L$ (from top to bottom) ΔE, $\mathcal{F}$, η' mass, and Δt. Signal MC (left) and continuum background (right).

E.9 $\eta'_{5\pi}K^0_L$

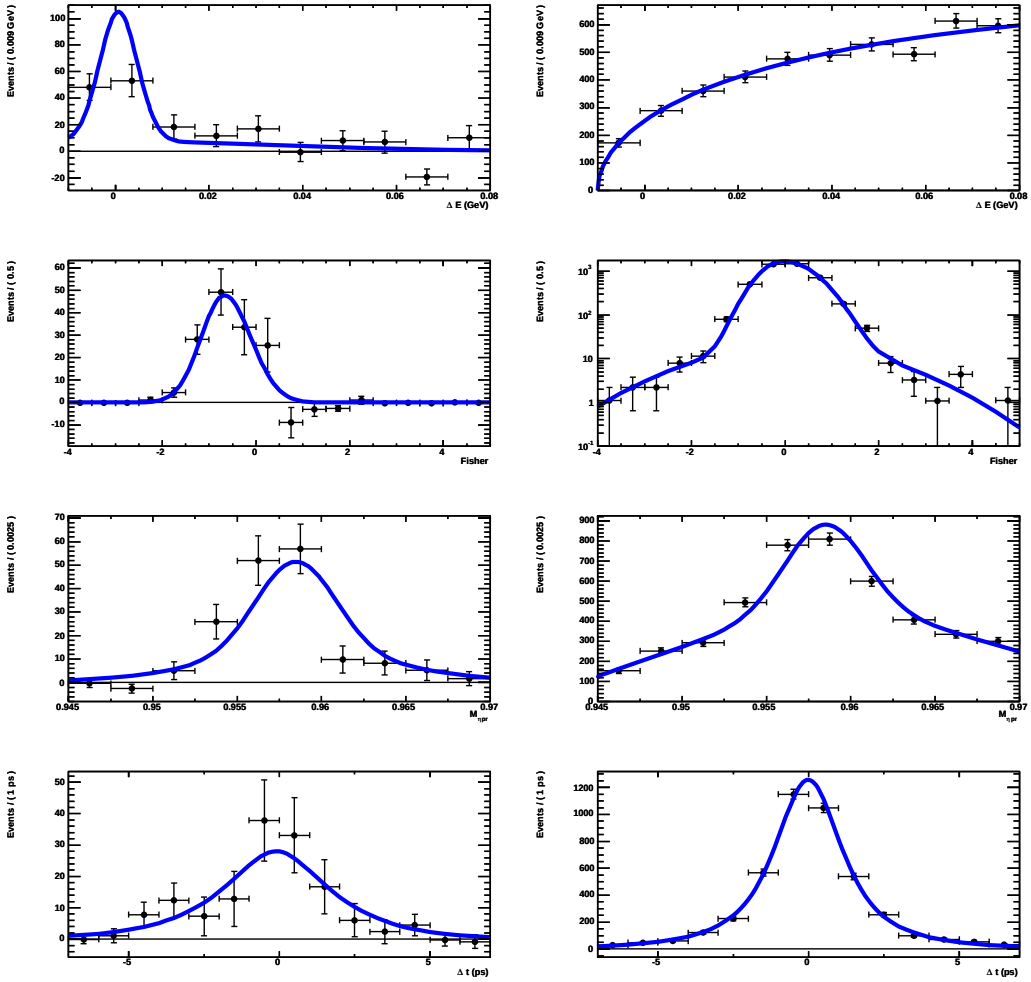


Figure E.9: *sPlots* for $\eta'_{5\pi}K^0_L$ (from top to bottom) ΔE, $\mathcal{F}$, η' mass, and Δt. Signal MC (left) and continuum background (right).

Appendix F

CP content of the $B\overline{B}$ backgrounds for $\eta' K^0$

When floating the eight $B\overline{B}$ *CP* parameters ($S_{\rm chrm}$, $C_{\rm chls}$, $S_{\rm chls}$, and $C_{\rm chls}$ in both $\eta'_{\rho\gamma}K^0_S$ and $\eta'_{\rho\gamma}K^0_{S00}$) in the fit to data, we find that the signal S and C change by 0.004 with respect to the nominal fit in which the $B\overline{B}$ *CP* parameters are fixed to zero. However, the errors on the background parameters are large, so we prefer to make conservative estimates of S and C in the charmless and charm backgrounds and fix them in the fit to obtain the systematic uncertainty.

F.1 Charmless $B\overline{B}$ backgrounds

We do not expect S in the charmless backgrounds (Table 5.3) to be 0, because 80 of the 280 expected charmless events come from the $K^0_S\pi^+\pi^-$ final state which contains $B^0 \to \rho^0 K^0_S$ ($-\eta_f S_f = 0.7$) and $B^0 \to f_0 K^0_S$ ($-\eta_f S_f = -0.7$). We also expect that there are 13 $B^0 \to \phi K^0_S$ ($-\eta_f S_f = 0.7$) events in the dataset. The time-dependence is simulated in these modes, so we can obtain the effective S and C for the charmless $B\overline{B}$ background by fitting the charmless $B\overline{B}$ MC sample. We find $S = 0.115 \pm 0.028$ and $C = 0.004 \pm 0.019$.

Based on the composition of the charmless $B\overline{B}$ background, $C = 0.0$ is very much expected. To understand whether $S \simeq 0.1$ is reasonable, we need to determine what fraction of the expected $K^0_S\pi^+\pi^-$ background events are $\rho^0 K^0_S$ and $f_0 K^0_S$. To that end, we show the distributions for the fit variables (ΔE, $m_{\rm ES}$, $\mathcal{F}$), the η' mass, and ρ

mass and helicity for all events in the mode $K^0_S\pi^+\pi^-$ that pass our selection in Fig. F.1.

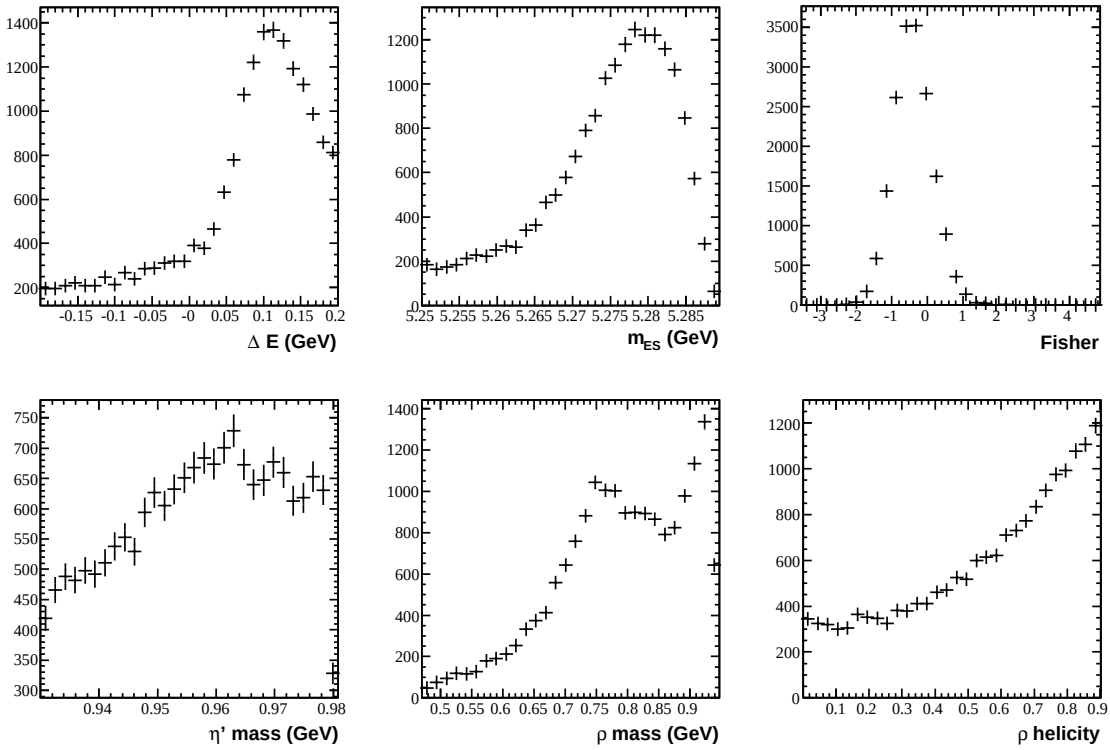


Figure F.1: Distributions of ΔE, $m_{\rm ES}$, $\mathcal{F}$, η' mass, ρ mass, and ρ helicity from $K^0_S\pi^+\pi^-$ Dalitz plot MC events that pass our selection for $\eta'_{\rho\gamma}K^0_S$.

From the ρ helicity plot in Fig. F.1, one can see that $\rho^0 K^0_S$, which has the shape $\cos^2\theta$ in helicity, makes up $\sim$50% of the $K^0_S\pi^+\pi^-$ events that pass our selection. The other 50% are $f_0K^0_S$, $K^*\pi$, and non-resonant events. The f_0 peak at high $\pi^+\pi^-$ invariant mass in Fig. F.1 suggests that $f_0K^0_S$ is $\sim$10% of the total $K^0_S\pi^+\pi^-$ events. Thus we expect the contribution of $K^0_S\pi^+\pi^-$ to the effective S of the charmless $B\overline{B}$ background to be $S_{\rm chls} = (80/280)(0.50-0.10)0.7 = 0.08$.

We also expect 13 ϕK^0_S events in the charmless background. These events contribute $(13/280)*0.7 = 0.03$, increasing $S_{\rm chls}$ to 0.11, in excellent agreement with the

value from fitting MC given above. To be conservative, we assume $S_{\rm chls} = 0.2$ and $C_{\rm chls} = 0.1$ when estimating the systematic. We also note that any effect of *CP* structure in the charmless $B\overline{B}$ background will appear in our determination of the fit bias with embedded toy studies. Since we correct for fit bias, we are actually correcting for this affect already.

F.2 Charm $B\overline{B}$ backgrounds

As can be seen in the $\eta'_{\rho\gamma}K^0_S$ PDFs (Fig. C.1), the charm background events are mostly combinatorial. The largest backgrounds ($D^-\rho^+$, $D^0\rho-$, $D^0\pi^-$, where the D decays via $K^0_S\pi$) have additional pions in the final state so we expect charm S and C to be near zero. Unfortunately, the time-dependence is not simulated in the MC, but even for the $D^-\pi^+$ mode that has the same final state as signal, the effective value of S is $\sim$ 0 because it is suppressed by the small ratio r between the interfering $b \to u$ and $b \to c$ amplitudes. We assume $S_{\rm chrm}$ and $C_{\rm chrm}$ of 0.1 when estimating the systematic error. We believe this is a conservative estimate.

F.3 Estimating the related systematic uncertainty

To estimate the systematic error we run two fits to the data. In the first, we simultaneously fix $S_{\rm chls}$ and $C_{\rm chls}$ (in both $\eta'_{\rho\gamma}K^0_S$ and $\eta'_{\rho\gamma}K^0_{S00}$) to 0.2 and 0.1, respectively. We keep the 4 charm *CP* parameters fixed to 0. In the second fit, we simultaneously fix all 4 charm parameters ($S_{\rm chrm}$ and $C_{\rm chrm}$ in $\eta'_{\rho\gamma}K^0_S$ and $\eta'_{\rho\gamma}K^0_{S00}$) to 0.1 while fixing all 4 charmless parameters to 0. We find the shifts in the signal S and C that are reported in Table F.1; we take the sum in quadrature from Table F.1 as the systematic.

Table F.1: Shifts in signal S and C from fits used to estimate the systematic uncertainty related to CP structure in the $B\overline{B}$ backgrounds.

Fit Configuration	ΔS	ΔC
$S_{\rm chls}$=0.2, $C_{\rm chls}$=0.1,$S_{\rm chrm}$=$C_{\rm chrm}$=0.0	0.007	0.003
$S_{\rm chls}$=$C_{\rm chls}$=0.0, $S_{\rm chrm}$=$C_{\rm chrm}$=0.1	0.004	0.003
Sum in Quadrature	0.008	0.004

Appendix G

Fixed charmless $B\overline{B}$ yields for $\eta' K^0$

In the nominal fit we fix the yields of the charmless $B\overline{B}$ fit components in modes with significant charmless $B\overline{B}$ backgrounds, $\eta'_{\rho\gamma}K^0_S$ and $\eta'_{\rho\gamma}K^0_{S00}$. Studies of the charmless backgrounds in Sec. 5.7 show that we know the charmless contribution to the data entering the fit to within 15% of itself. We perform blind, combined fits to the data in which we fix the charmless yields in $\eta'_{\rho\gamma}K^0_S$ and $\eta'_{\rho\gamma}K^0_{S00}$ to 80% and 120% of the expected yields. Compared to the nominal fit, the parameter S changes by $+0.001$ in the "Fix `nChls` Low" fit and -0.001 in the "Fix `nChls` High" fit; the C parameter does not change at all. Fit results are in Table G.1.

Table G.1: Blind results from combined fits to Run1-6 data in which we fix the charmless yields in $\eta'_{\rho\gamma}K^0_S$ and $\eta'_{\rho\gamma}K^0_{S00}$ at 80% (Fix `nChls` Low) and 120% (Fix `nChls` High) of the expected yields; and a combined fit in which we fix C at zero.

Fit vars	Nominal	Fix `nChls` Low	Fix `nChls` High
S	-0.481 ± 0.078	-0.480 ± 0.078	-0.482 ± 0.078
C	0.174 ± 0.058	0.174 ± 0.058	0.174 ± 0.058
Signal Yields			
$\eta'_{\eta\pi\pi}K^0_S$	468.8 ± 23.5	468.6 ± 23.5	468.5 ± 23.5
$\eta'_{\rho\gamma}K^0_S$	999.0 ± 39.6	1001.6 ± 39.7	996.2 ± 39.6
$\eta'_{\eta\pi\pi}K^0_{S00}$	104.3 ± 13.0	104.1 ± 12.9	104.1 ± 12.9
$\eta'_{\rho\gamma}K^0_{S00}$	201.9 ± 27.4	201.7 ± 27.3	201.3 ± 27.4
$\eta'_{5\pi}K^0_S$	170.5 ± 14.1	170.5 ± 14.1	170.5 ± 14.1
$\eta'_{\eta\pi\pi}K^0_L$	331.5 ± 31.4	335.5 ± 31.5	332.9 ± 31.5
$\eta'_{5\pi}K^0_L$	163.9 ± 21.8	160.1 ± 21.5	160.8 ± 21.6

Appendix H

ηK^* : PDFs, fit input correlations, and final free parameter values

We show here for each decay mode the signal and background PDFs used in ML fits, the correlation coefficients between the input variables used in the ML fits, and the initial and final values of the PDF parameters that are free in the fit. Correlations are less than 7% between all variables in the data (`onData`), which is predominantly $q\bar{q}$, and less than 10% in the signal and charmless $B\overline{B}$ MC; bias due to correlations in all components but $q\bar{q}$ is accounted for in embedded toy studies. PDFs for all ηK^* and charmless $B\overline{B}$ components are determined from MC; for continuum background PDFs we use on-peak sidebands.

H.1 $\eta_{\gamma\gamma} K^{*+}_{K^+\pi^0}(892)$

```
Correlation matrix for sigMC (15986 events):
                de       mes    fisher      mEta    mKstar
     mes   -0.0757
  fisher    0.0135   -0.0301
    mEta    0.0891    0.0089    0.0277
  mKstar    0.0592    0.0118   -0.0059   -0.0013

Correlation matrix for chlsMC (956 events):
                de       mes    fisher      mEta    mKstar
     mes    0.1125
  fisher   -0.1078   -0.0871
    mEta    0.0151    0.1520   -0.0237
  mKstar    0.0122    0.0018   -0.0049   -0.0780

Correlation matrix for onData (10039 events):
                de       mes    fisher      mEta    mKstar
     mes    0.0101
  fisher   -0.0530    0.0046
    mEta    0.0077    0.0081   -0.0228
  mKstar   -0.0015    0.0005   -0.0075   -0.0016
```

The $\eta_{\gamma\gamma} K^{*+}_{K^+\pi^0}(892)$ fit finds the following values for the parameters which were allowed to float:

Floating Parameter	InitialValue	FinalValue +/- Error	GblCorr.
Frac_nBkg_chgCat_Plus	5.1519e-01	5.1629e-01 +/- 5.06e-03	0.089659
Frac_nSig_chgCat_Plus	5.1519e-01	4.0272e-01 +/- 7.91e-02	0.093970
deBkg_P01	-1.6486e+00	-1.6247e+00 +/- 9.30e-02	0.168101
fisBkgC_asym	3.7529e-02	3.4156e-02 +/- 1.67e-02	0.268326
fisBkgC_mean	4.4135e-01	4.5228e-01 +/- 6.44e-03	0.390588
fisBkgC_rms	5.7645e-01	5.7690e-01 +/- 4.64e-03	0.358317
mEBkg_fracS	4.1857e-01	4.2310e-01 +/- 1.08e-02	0.082059
mEPolyBkg_P01	-2.9682e-01	-2.9442e-01 +/- 2.42e-02	0.128625
mKstarBkg_fracKst	1.3131e-01	1.2536e-01 +/- 1.09e-02	0.151688
mKstarPolyBkg_P01	-9.0137e-02	-9.4027e-02 +/- 1.89e-02	0.031628
mesBkg_c	-2.3678e+01	-2.0447e+01 +/- 2.61e+00	0.174634
nBkg	6.7000e+03	9.9109e+03 +/- 1.05e+02	0.296087
nChls	0.0000e+00	-2.3539e+00 +/- 2.97e+01	0.593074
nSig	8.8000e+01	9.8612e+01 +/- 1.62e+01	0.217194

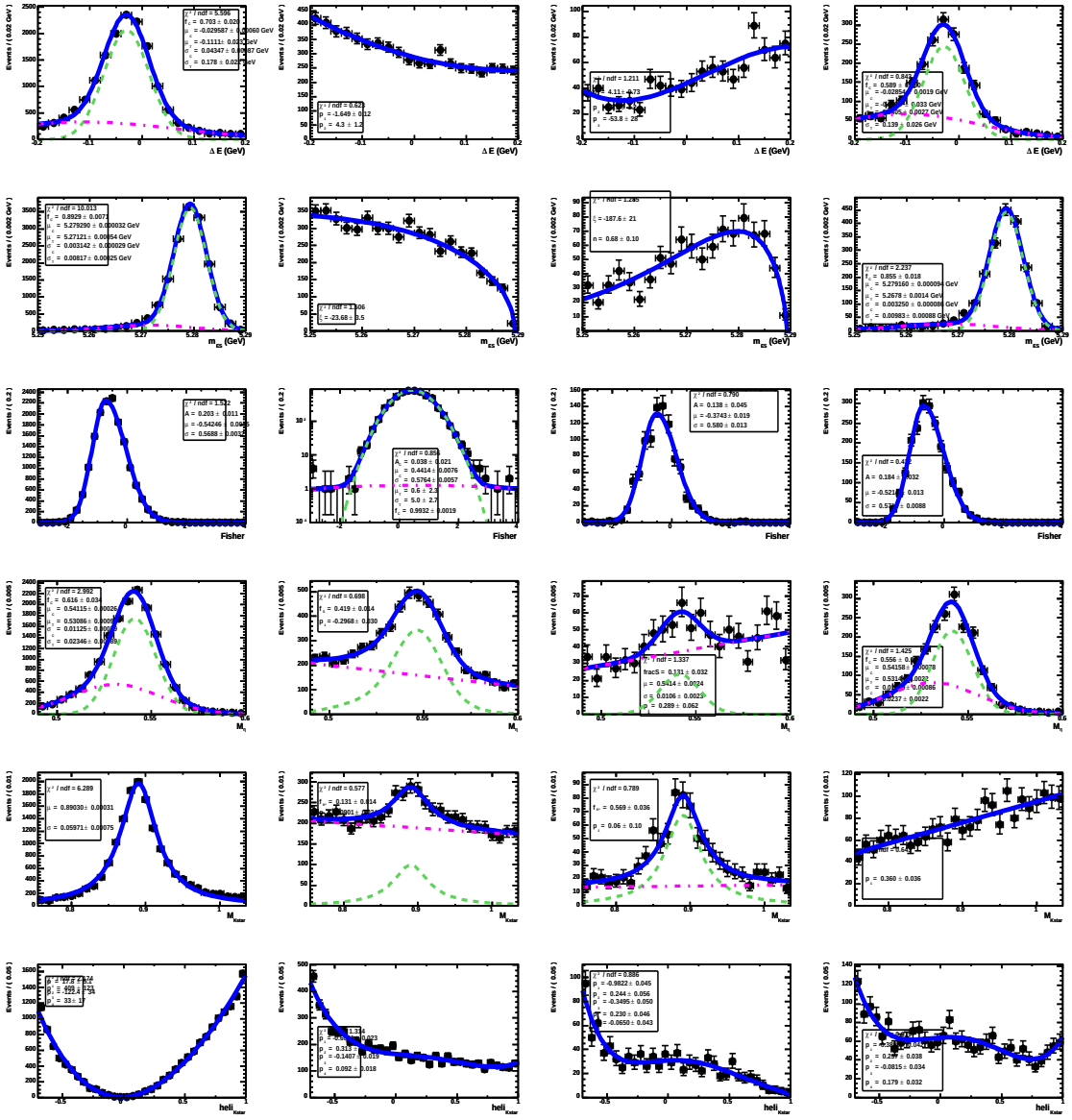


Figure H.1: PDFs for (top to bottom) ΔE, $m_{\rm ES}$, η mass, $\mathcal{F}$, K^* mass and $\mathcal{H}$, for (left to right) $B \to \eta K^*(892)$, on-peak sidebands, charmless $B\overline{B}$ background, and $B \to \eta K_0^*(S\text{-wave})$.

H.2 $\eta_{\gamma\gamma} K^{*+}_{K^0_S\pi^+}(892)$

```
Correlation matrix for sigMC (29889 events):
               de       mes    fisher      mEta    mKstar
    mes   -0.0272
 fisher    0.0097   -0.0269
   mEta    0.1221    0.0183    0.0178
 mKstar    0.0161   -0.0242   -0.0032    0.0084

Correlation matrix for chlsMC (986 events):
               de       mes    fisher      mEta    mKstar
    mes    0.1331
 fisher   -0.0258   -0.0342
   mEta    0.0357    0.0819    0.0069
 mKstar   -0.0411    0.0181   -0.0022   -0.0179

Correlation matrix for onData (7663 events):
               de       mes    fisher      mEta    mKstar
    mes   -0.0087
 fisher   -0.0482   -0.0622
   mEta    0.0063    0.0034   -0.0303
 mKstar    0.0016    0.0225    0.0125    0.0107
```

The $\eta_{\gamma\gamma} K^{*+}_{K^0_S\pi^+}(892)$ fit finds the following values for the parameters which were allowed to float:

Floating Parameter	InitialValue	FinalValue +/- Error	GblCorr.
Frac_nBkg_chgCat_Plus	5.0581e-01	5.0560e-01 +/- 5.89e-03	0.110485
Frac_nSig_chgCat_Plus	5.0581e-01	5.1291e-01 +/- 6.27e-02	0.119026
deBkg_P01	-1.5420e+00	-1.5815e+00 +/- 1.05e-01	0.191763
fisBkgC_asym	1.1746e-01	1.2123e-01 +/- 2.14e-02	0.334922
fisBkgC_mean	4.4034e-01	4.3668e-01 +/- 8.31e-03	0.518201
fisBkgC_rms	5.7993e-01	5.7415e-01 +/- 5.87e-03	0.419020
mEBkg_fracS	3.8800e-01	4.1009e-01 +/- 1.27e-02	0.074737
mEPolyBkg_P01	-3.0383e-01	-3.0820e-01 +/- 2.76e-02	0.113960
mKstarBkg_fracKst	2.0776e-01	2.1031e-01 +/- 1.37e-02	0.164920
mKstarPolyBkg_P01	3.5003e-02	6.3354e-02 +/- 2.39e-02	0.061857
mesBkg_c	-1.7545e+01	-1.5683e+01 +/- 3.12e+00	0.279723
nBkg	5.1000e+03	7.4347e+03 +/- 9.67e+01	0.436840
nChls	1.8000e+02	4.1543e+01 +/- 4.50e+01	0.711513
nSig	9.9000e+01	1.4854e+02 +/- 1.94e+01	0.290298

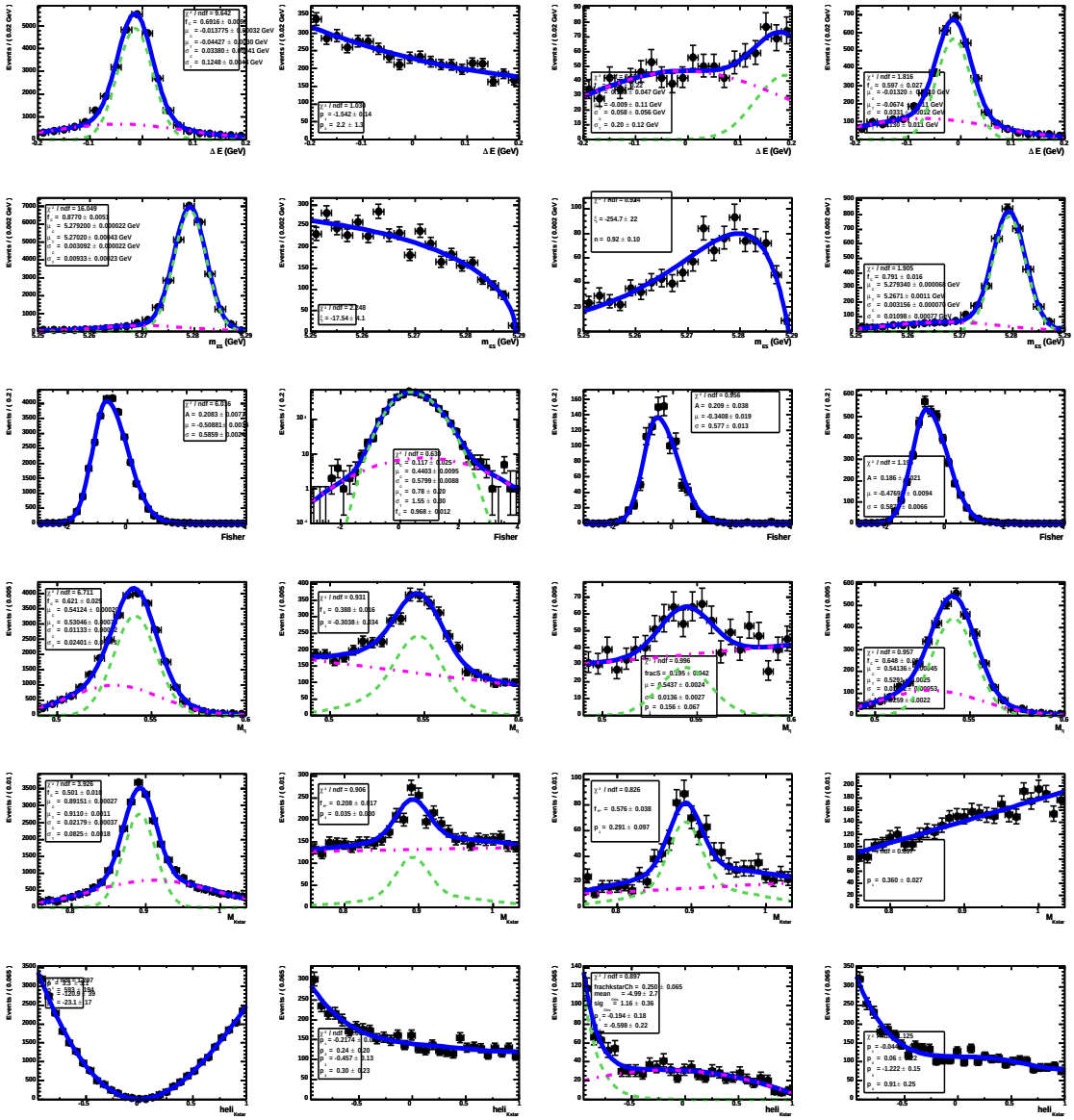


Figure H.2: PDFs for (top to bottom) ΔE, $m_{\rm ES}$, η mass, $\mathcal{F}$, K^* mass and $\mathcal{H}$, for (left to right) $B \to \eta K^*(892)$, on-peak sidebands, charmless $B\overline{B}$ background, and $B \to \eta K_0^*(S\text{-wave})$.

H.3 $\eta_{\gamma\gamma} K^{*0}_{K^+\pi^-}(892)$

```
Correlation matrix for sigMC (32451 events):
               de       mes    fisher      mEta    mKstar
    mes   -0.0084
 fisher    0.0031   -0.0208
   mEta    0.1353    0.0221    0.0115
 mKstar    0.0222   -0.0105   -0.0014   -0.0119

Correlation matrix for chlsMC (2434 events):
               de       mes    fisher      mEta    mKstar
    mes    0.1656
 fisher   -0.0335   -0.0070
   mEta    0.1000    0.1041   -0.0208
 mKstar   -0.0153    0.0178    0.0210   -0.0063

Correlation matrix for onData (18687 events):
               de       mes    fisher      mEta    mKstar
    mes    0.0089
 fisher   -0.0425   -0.0328
   mEta    0.0050   -0.0085   -0.0264
 mKstar    0.0097   -0.0004    0.0001   -0.0027
```

The $\eta_{\gamma\gamma} K^{*0}_{K^+\pi^-}(892)$ fit finds the following values for the parameters which were allowed to float:

```
  Floating Parameter  InitialValue    FinalValue +/-  Error     GblCorr.
--------------------  ------------  --------------------------  --------
Frac_nBkg_chgCat_Plus    5.0083e-01    5.0326e-01 +/-  3.76e-03  0.096089
Frac_nSig_chgCat_Plus    5.0083e-01    3.8170e-01 +/-  3.47e-02  0.115292
           deBkg_P01   -1.4486e+00   -1.4237e+00 +/-  6.63e-02  0.163647
        fisBkgC_asym    5.0466e-02    5.0452e-02 +/-  1.29e-02  0.255688
        fisBkgC_mean    4.1662e-01    4.2875e-01 +/-  4.89e-03  0.388517
         fisBkgC_rms    5.7719e-01    5.7632e-01 +/-  3.55e-03  0.328702
         mEBkg_fracS    3.9665e-01    3.9331e-01 +/-  8.14e-03  0.081597
       mEPolyBkg_P01   -2.9283e-01   -3.1704e-01 +/-  1.72e-02  0.116673
   mKstarBkg_fracKst    2.3100e-01    2.3024e-01 +/-  8.06e-03  0.133655
   mKstarPolyBkg_P01    6.6657e-02    6.1944e-02 +/-  1.53e-02  0.023917
            mesBkg_c   -2.2242e+01   -2.1961e+01 +/-  1.95e+00  0.200783
                nBkg    1.2000e+04    1.8156e+04 +/-  1.42e+02  0.299899
               nChls    9.4000e+01    5.1531e+01 +/-  4.26e+01  0.575145
                nSig    3.0000e+02    4.0698e+02 +/-  2.93e+01  0.226442
```

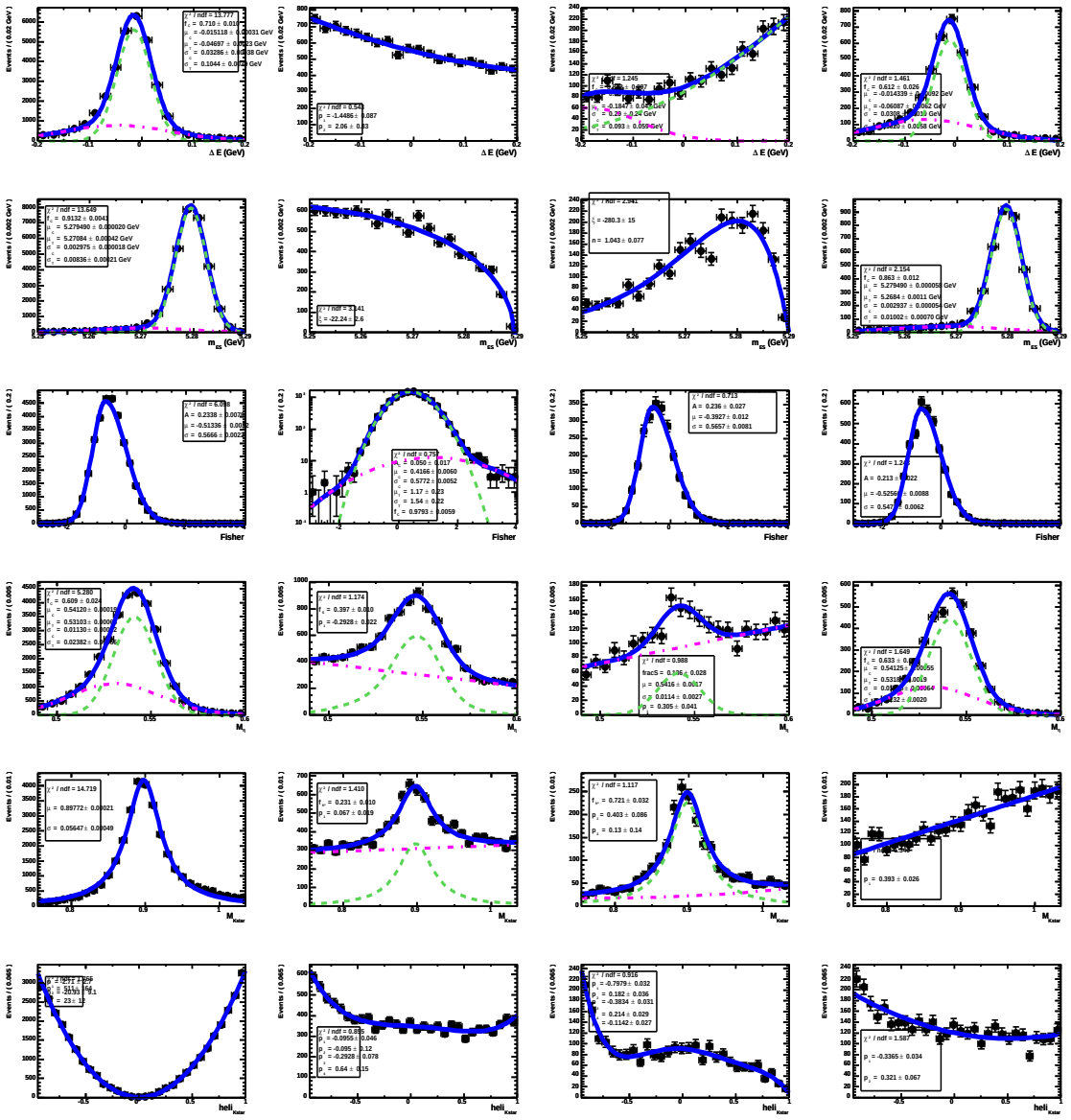


Figure H.3: PDFs for (top to bottom) ΔE, $m_{\rm ES}$, η mass, $\mathcal{F}$, K^* mass and $\mathcal{H}$, for (left to right) $B \to \eta K^*(892)$, on-peak sidebands, charmless $B\overline{B}$ background, and $B \to \eta K_0^*(S\text{-wave})$.

H.4 $\eta_{3\pi} K^{*+}_{K^+\pi^0}(892)$

```
Correlation matrix for sigMC (10580 events):
           fisher    mKstar    hKstar        de       mes
  mKstar   -0.0046
  hKstar   -0.0280    0.0454
      de    0.0009    0.1009   -0.1375
     mes   -0.0393    0.0112    0.1104   -0.0201

Correlation matrix for onData (5568 events):
           fisher    mKstar    hKstar        de       mes
  mKstar   -0.0174
  hKstar    0.0128    0.0286
      de   -0.0334    0.0045    0.0047
     mes   -0.0107   -0.0258    0.0147   -0.0069

Correlation matrix for chlsMC (219 events):
           fisher    mKstar    hKstar        de       mes
  mKstar   -0.0145
  hKstar    0.2134   -0.0203
      de   -0.0651    0.1392    0.0067
     mes   -0.0479   -0.0540   -0.0871    0.0196
```

The $\eta_{3\pi} K^{*+}_{K^+\pi^0}(892)$ fit finds the following values for the parameters which were allowed to float:

```
  Floating Parameter  InitialValue    FinalValue +/-  Error     GblCorr.
--------------------  ------------  --------------------------  --------
Frac_nBkg_chgCat_Plus    4.9371e-01    4.9350e-01 +/-  6.79e-03  0.077206
Frac_nSig_chgCat_Plus    4.9371e-01    5.2715e-01 +/-  9.94e-02  0.105003
           deBkg_P01   -1.5510e+00   -1.5674e+00 +/-  1.14e-01  0.012214
        fisBkgC_asym    5.6065e-02    7.9872e-02 +/-  2.25e-02  0.102783
        fisBkgC_mean    4.5647e-01    4.6183e-01 +/-  8.32e-03  0.151590
         fisBkgC_rms    5.8678e-01    5.8876e-01 +/-  6.16e-03  0.109988
         mEBkg_fracS    3.2567e-01    3.4588e-01 +/-  1.17e-02  0.105292
       mEPolyBkg_P01    3.4771e-01    3.4708e-01 +/-  2.90e-02  0.085742
   mKstarBkg_fracKst    1.1251e-01    9.7201e-02 +/-  1.52e-02  0.065145
   mKstarPolyBkg_P01   -8.0130e-02   -1.0015e-01 +/-  2.50e-02  0.031087
           mesBkg_c   -1.6884e+01   -1.4094e+01 +/-  3.47e+00  0.086227
               nBkg    3.0000e+03    5.4974e+03 +/-  7.47e+01  0.082140
               nSig    3.0000e+01    5.6243e+01 +/-  1.14e+01  0.197528
```

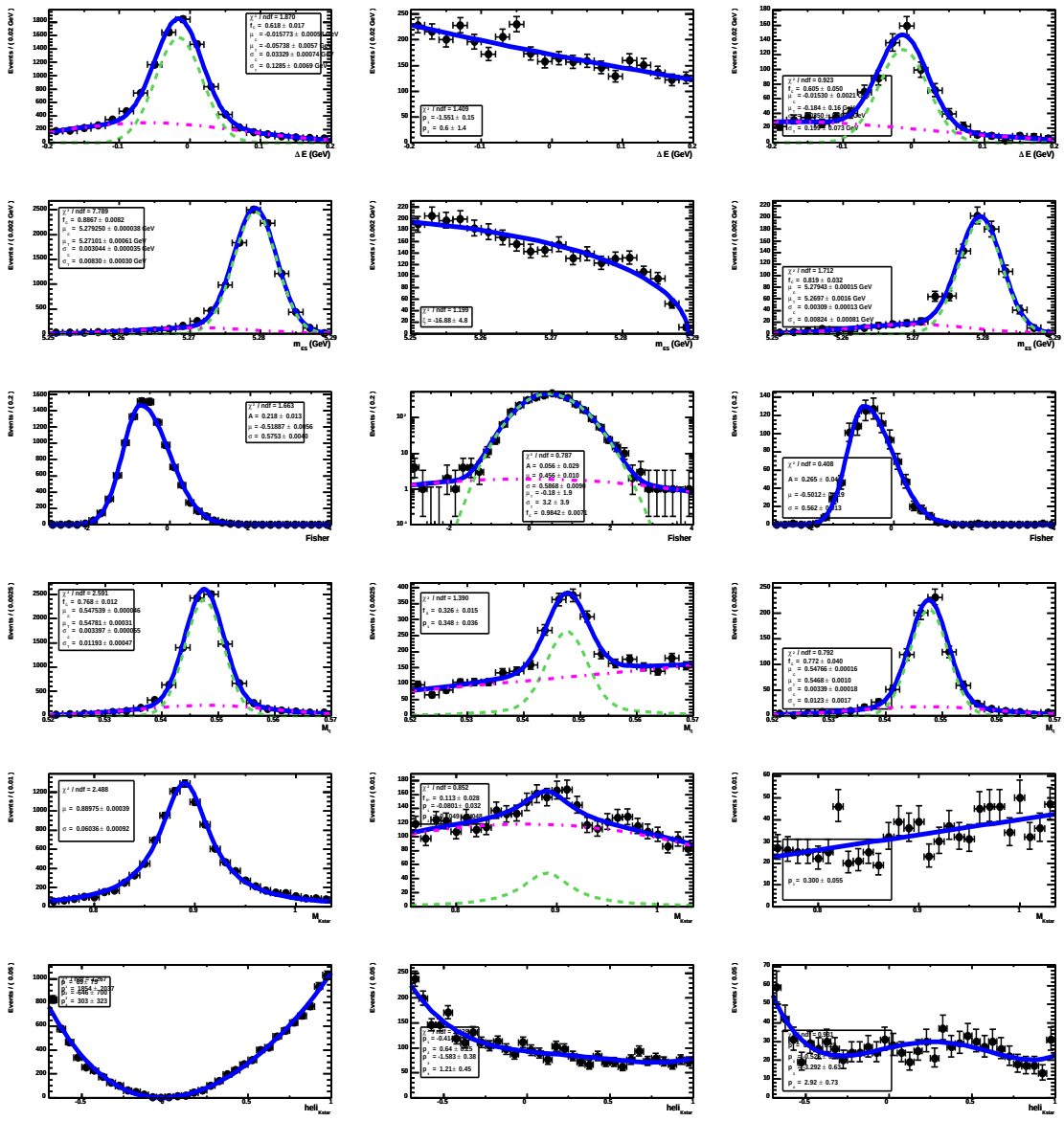


Figure H.4: PDFs for (top to bottom) ΔE, $m_{\rm ES}$, η mass, $\mathcal{F}$, K^* mass and $\mathcal{H}$, for (left to right) $B \to \eta K^*(892)$, on-peak sidebands, and $B \to \eta K_0^*(S\text{-wave})$.

H.5 $\eta_{3\pi}K^{*+}_{K^0_S\pi^+}(892)$

```
Correlation matrix for sigMC (20660 events):
           fisher    mKstar    hKstar        de       mes
  mKstar   -0.0002
  hKstar   -0.0220   -0.0297
      de   -0.0095    0.0242   -0.0231
     mes   -0.0428   -0.0158    0.1203    0.0203

Correlation matrix for onData (4083 events):
           fisher    mKstar    hKstar        de       mes
  mKstar    0.0200
  hKstar    0.0191   -0.0503
      de   -0.0389   -0.0115    0.0384
     mes   -0.0205    0.0052    0.0077    0.0159

Correlation matrix for chlsMC (567 events):
           fisher    mKstar    hKstar        de       mes
  mKstar   -0.0351
  hKstar    0.1467    0.0393
      de   -0.0507    0.0360   -0.0383
     mes    0.0069   -0.0293   -0.0389    0.0542
```

The $\eta_{3\pi}K^{*+}_{K^0_S\pi^+}(892)$ fit finds the following values for the parameters which were allowed to float:

```
  Floating Parameter  InitialValue    FinalValue +/-  Error     GblCorr.
--------------------  ------------  --------------------------  --------
Frac_nBkg_chgCat_Plus    4.9400e-01    4.9294e-01 +/-  7.94e-03  0.090778
Frac_nSig_chgCat_Plus    4.9400e-01    6.1493e-01 +/-  1.40e-01  0.222664
           deBkg_P01   -1.7294e+00   -1.7035e+00 +/-  1.43e-01  0.007984
        fisBkgC_asym    9.6339e-02    8.2974e-02 +/-  2.45e-02  0.091797
        fisBkgC_mean    4.3380e-01    4.3333e-01 +/-  9.79e-03  0.159523
         fisBkgC_rms    6.0239e-01    6.0287e-01 +/-  7.14e-03  0.125774
        mEBkg_fracS    3.5604e-01    3.5689e-01 +/-  1.39e-02  0.109695
       mEPolyBkg_P01    3.6019e-01    3.5628e-01 +/-  3.47e-02  0.087876
   mKstarBkg_fracKst    1.7857e-01    1.7051e-01 +/-  1.86e-02  0.073482
   mKstarPolyBkg_P01    8.4553e-02    9.6403e-02 +/-  3.11e-02  0.035117
           mesBkg_c   -7.2052e+00   -1.1266e+01 +/-  4.07e+00  0.105637
               nBkg    2.6000e+03    4.0321e+03 +/-  6.41e+01  0.097314
               nSig    2.9000e+01    3.5821e+01 +/-  1.02e+01  0.287435
```

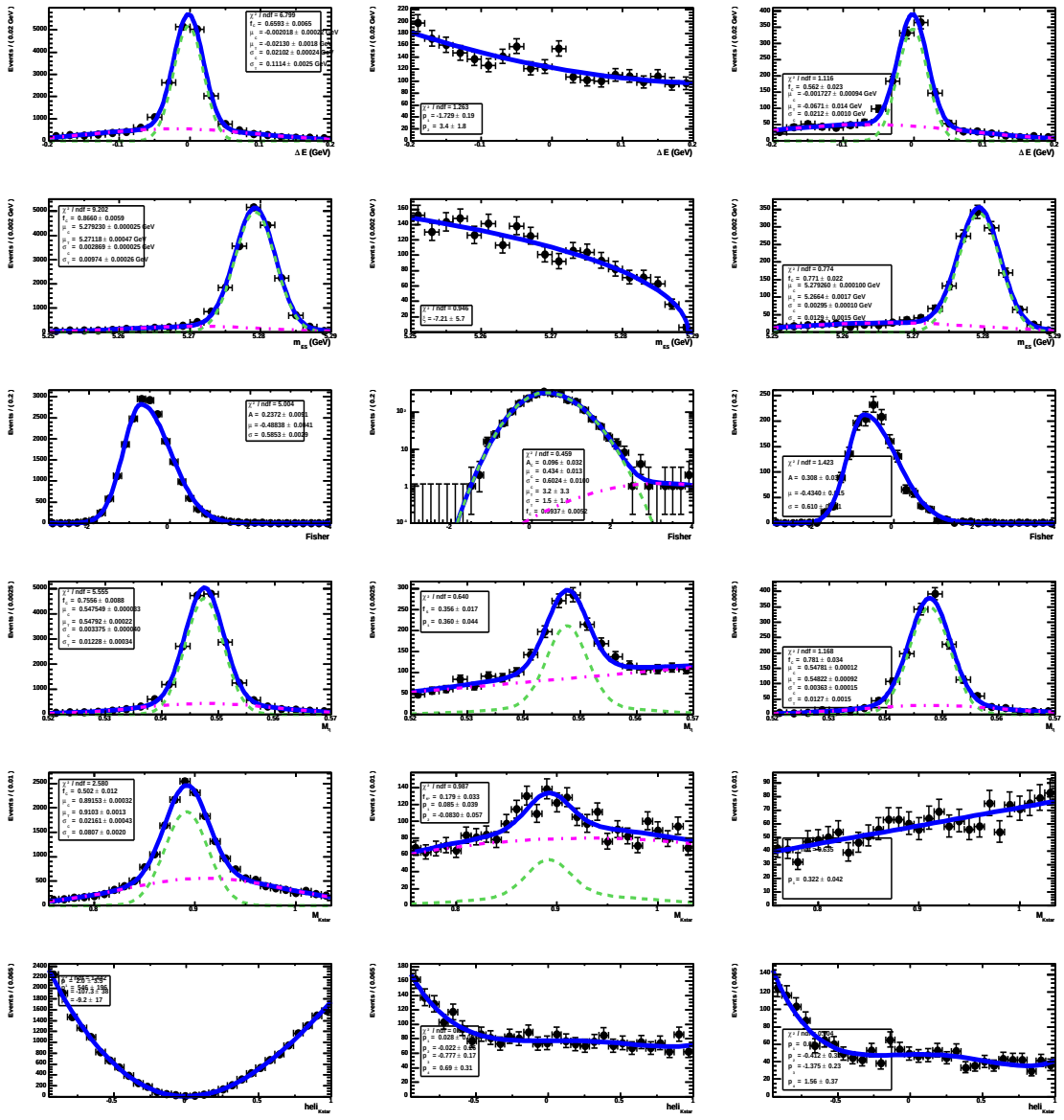


Figure H.5: PDFs for (top to bottom) ΔE, $m_{\rm ES}$, η mass, $\mathcal{F}$, K^* mass and $\mathcal{H}$, for (left to right) $B \to \eta K^*(892)$, on-peak sidebands, and $B \to \eta K_0^*(S\text{-wave})$.

H.6 $\eta_{3\pi}K^{*0}_{K^+\pi^-}(892)$

```
Correlation matrix for sigMC (21957 events):
               de       mes    fisher      mEta    mKstar
    mes    0.0442
 fisher   -0.0200   -0.0389
   mEta    0.0511   -0.0007    0.0215
 mKstar    0.0224   -0.0097   -0.0071   -0.0068

Correlation matrix for chlsMC (355 events):
               de       mes    fisher      mEta    mKstar
    mes    0.0497
 fisher   -0.0370   -0.1288
   mEta    0.0021    0.2148   -0.0078
 mKstar    0.0335    0.0015   -0.0636    0.0126

Correlation matrix for onData (9906 events):
               de       mes    fisher      mEta    mKstar
    mes    0.0038
 fisher   -0.0589   -0.0120
   mEta    0.0000   -0.0092   -0.0121
 mKstar   -0.0055    0.0072   -0.0067   -0.0101
```

The $\eta_{3\pi}K^{*0}_{K^+\pi^-}(892)$ fit finds the following values for the parameters which were allowed to float:

```
  Floating Parameter  InitialValue    FinalValue +/-  Error     GblCorr.
--------------------  ------------  --------------------------  --------
Frac_nBkg_chgCat_Plus    5.0061e-01    5.0101e-01 +/-  5.13e-03  0.076842
Frac_nSig_chgCat_Plus    5.0061e-01    4.3900e-01 +/-  7.15e-02  0.098181
           deBkg_P01   -1.7510e+00   -1.6836e+00 +/-  9.14e-02  0.027139
        fisBkgC_asym    7.7678e-02    1.0266e-01 +/-  1.78e-02  0.161790
        fisBkgC_mean    4.2280e-01    4.3642e-01 +/-  6.40e-03  0.276534
         fisBkgC_rms    5.7341e-01    5.7165e-01 +/-  4.69e-03  0.211917
         mEBkg_fracS    3.2978e-01    3.3632e-01 +/-  8.97e-03  0.122205
       mEPolyBkg_P01    3.9346e-01    3.8272e-01 +/-  2.14e-02  0.089722
   mKstarBkg_fracKst    2.1205e-01    2.1759e-01 +/-  1.10e-02  0.067208
   mKstarPolyBkg_P01    5.9866e-02    7.3501e-02 +/-  2.08e-02  0.032639
           mesBkg_c   -1.7734e+01   -1.5572e+01 +/-  2.63e+00  0.113806
               nBkg    6.0000e+03    9.6934e+03 +/-  1.01e+02  0.213548
              nChls    1.1000e+02    7.2239e+01 +/-  2.48e+01  0.380946
               nSig    8.6000e+01    1.1078e+02 +/-  1.62e+01  0.202646
```

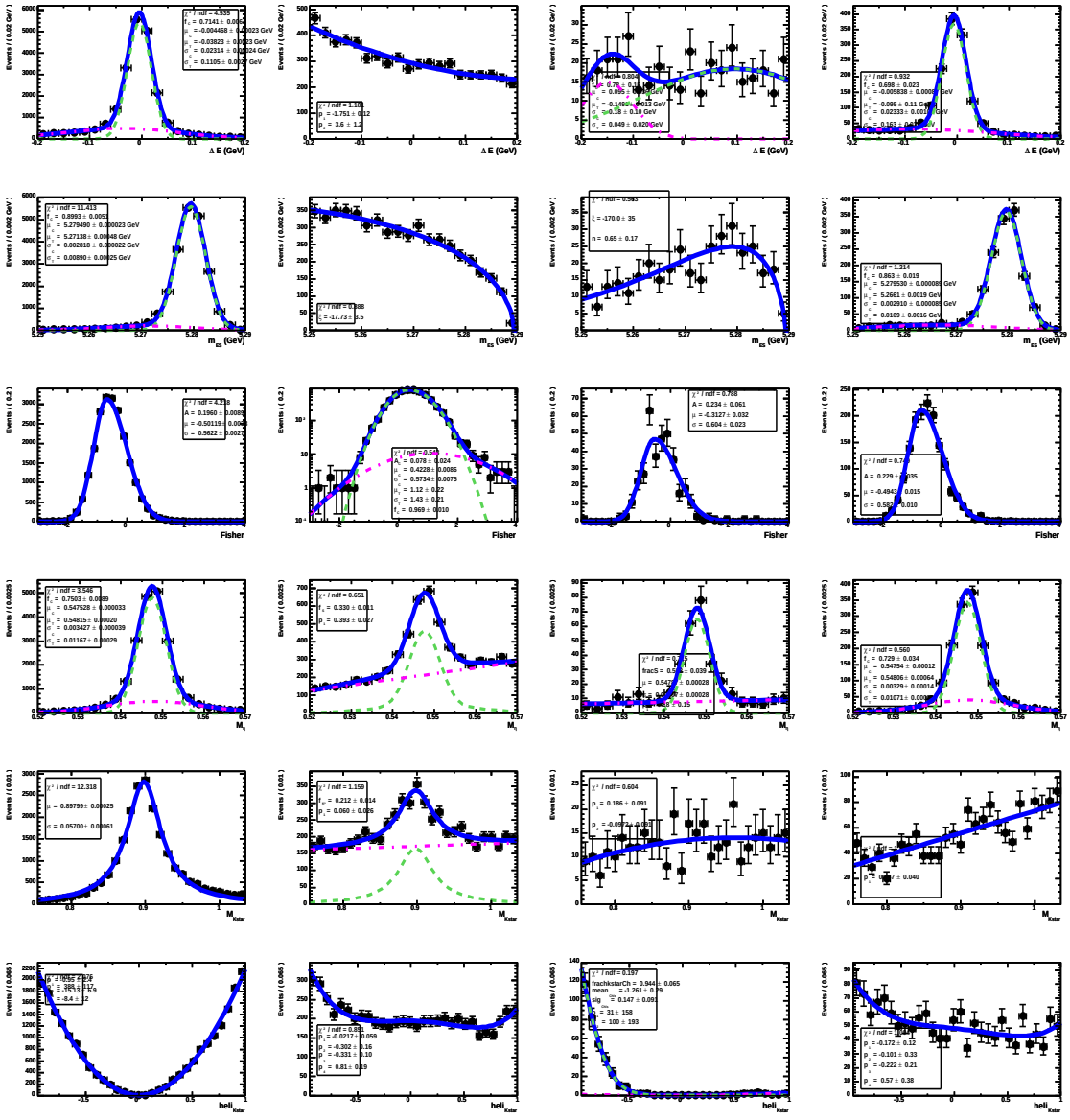


Figure H.6: PDFs for (top to bottom) ΔE, $m_{\rm ES}$, η mass, $\mathcal{F}$, K^* mass and $\mathcal{H}$, for (left to right) $B \to \eta K^*(892)$, on-peak sidebands, charmless $B\overline{B}$ background, and $B \to \eta K_0^*(S\text{-wave})$.

H.7 $\eta_{\gamma\gamma}K^{*+}_{K^+\pi^0}$(*S*-wave), (*D*-wave)

```
Correlation matrix for onData (6566 events):
                de       mes    fisher      mEta    mKstar
     mes    0.0024
  fisher   -0.0483   -0.0280
    mEta    0.0121    0.0098   -0.0379
  mKstar   -0.0099    0.0161   -0.0660   -0.0234

Correlation matrix for kst1430MC (5284 events):
                de       mes    fisher      mEta    mKstar
     mes   -0.0019
  fisher   -0.0054   -0.0188
    mEta    0.1093    0.0247    0.0274
  mKstar    0.0416    0.0141    0.0035   -0.0057

Correlation matrix for k2MC (414 events):
                de       mes    fisher      mEta    mKstar
     mes   -0.1446
  fisher    0.0260   -0.0385
    mEta    0.0714    0.1414    0.0795
  mKstar    0.1068    0.0234   -0.0493    0.0230
```

The $\eta_{\gamma\gamma}K^{*+}_{K^+\pi^0}$(HMR) fit finds the following values for the parameters which were allowed to float:

```
  Floating Parameter  InitialValue    FinalValue +/-  Error     GblCorr.
--------------------  ------------  --------------------------  --------
Frac_nBkg_chgCat_Plus    5.0609e-01    5.0545e-01 +/-  6.31e-03  0.122587
Frac_nK2_chgCat_Plus    5.0609e-01    5.7813e-01 +/-  2.07e-01  0.468550
Frac_nKst1430_chgCat_Plus    5.0609e-01    5.2630e-01 +/-  1.03e-01  0.478204
           deBkg_P01   -1.4111e+00   -1.3360e+00 +/-  1.08e-01  0.124078
        fisBkgC_asym    6.9681e-02    5.0883e-02 +/-  2.28e-02  0.357117
        fisBkgC_mean    4.3366e-01    4.3666e-01 +/-  8.51e-03  0.446113
         fisBkgC_rms    5.6912e-01    5.6895e-01 +/-  6.28e-03  0.413920
         mEBkg_fracS    4.1040e-01    4.0470e-01 +/-  1.41e-02  0.125636
       mEPolyBkg_P01   -3.6315e-01   -3.6526e-01 +/-  2.94e-02  0.164886
       mKstarBkg_P01   -7.6638e-01   -7.6567e-01 +/-  6.36e-02  0.997996
       mKstarBkg_P02   -4.9623e-02   -4.9220e-02 +/-  5.87e-02  0.999174
       mKstarBkg_P03    2.6270e-02    2.6541e-02 +/-  3.97e-02  0.999306
       mKstarBkg_P04    5.5203e-02    5.5376e-02 +/-  1.79e-02  0.998618
            mesBkg_c   -2.2581e+01   -1.8419e+01 +/-  3.29e+00  0.237445
                nBkg    4.2000e+03    6.4422e+03 +/-  8.64e+01  0.357917
               nChls    6.8000e+01   -1.4426e+00 +/-  2.94e+01  0.641725
                 nK2    0.0000e+00    2.6338e+01 +/-  1.16e+01  0.479955
            nKst1430    1.1000e+02    9.1210e+01 +/-  1.95e+01  0.523698
```

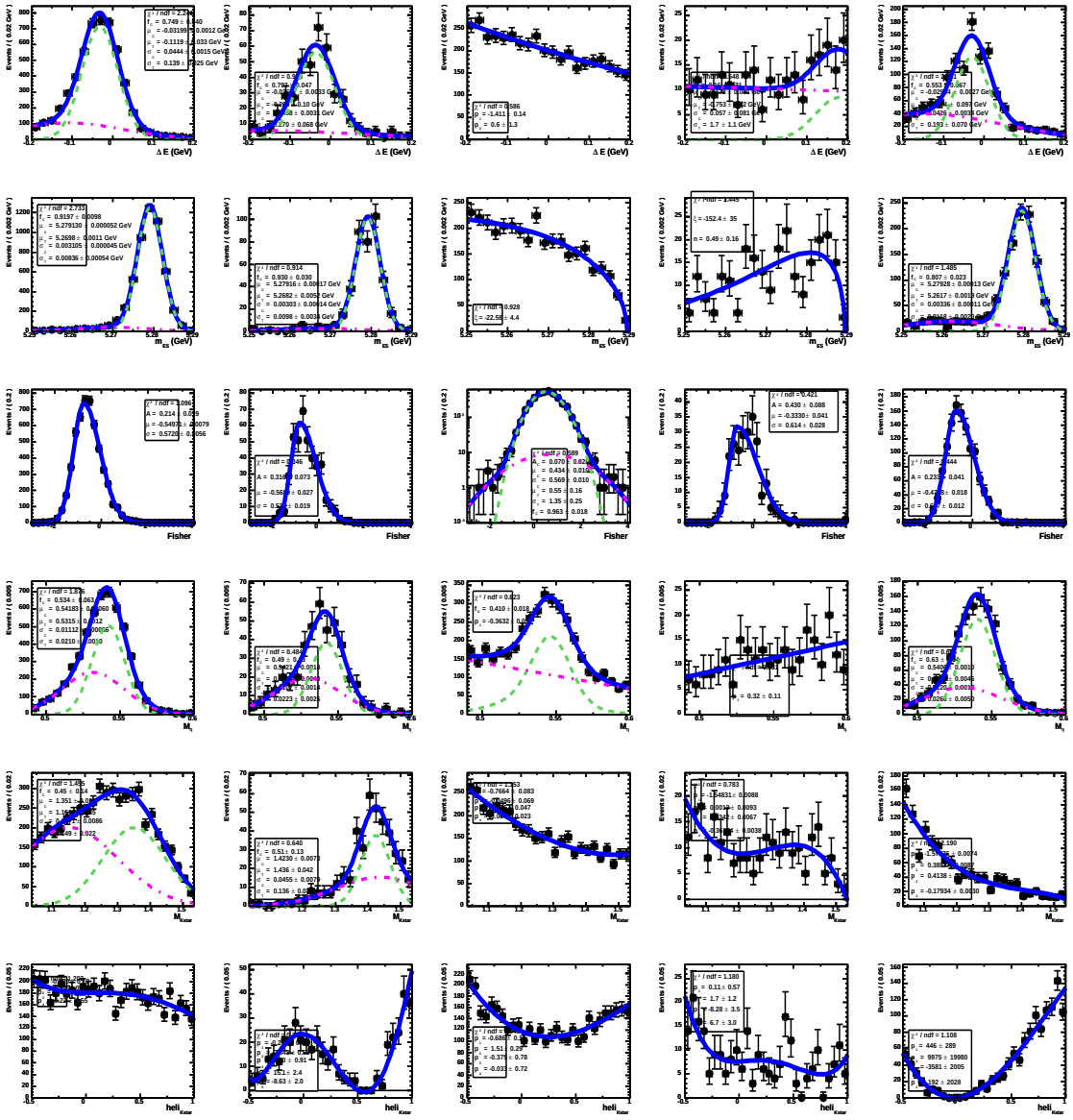


Figure H.7: PDFs for (top to bottom) ΔE, $m_{\rm ES}$, η mass, $\mathcal{F}$, K^* mass and $\mathcal{H}$, for (left to right) $B \to \eta K_0^*(S\text{-wave})$, $B \to \eta K_2^*(1430)$, on-peak sidebands, charmless $B\overline{B}$ background, and $B \to \eta K^*(892)$.

H.8 $\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}$(*S*-wave), (*D*-wave)

```
Correlation matrix for onData (5180 events):
               de       mes    fisher      mEta    mKstar
     mes   -0.0106
  fisher   -0.0344   -0.0197
    mEta   -0.0159   -0.0141    0.0021
  mKstar    0.0081    0.0319   -0.0433    0.0173

Correlation matrix for kst1430MC (7468 events):
               de       mes    fisher      mEta    mKstar
     mes    0.0142
  fisher   -0.0006   -0.0429
    mEta    0.1419    0.0125    0.0051
  mKstar    0.0294   -0.0034   -0.0201   -0.0052

Correlation matrix for k2MC (479 events):
               de       mes    fisher      mEta    mKstar
     mes    0.0335
  fisher    0.0049   -0.0062
    mEta    0.2249    0.0244    0.0011
  mKstar    0.0135   -0.0145   -0.0210    0.0345
```

The $\eta_{\gamma\gamma}K^{*+}_{K^0\pi^+}$(HMR) fit finds the following values for the parameters which were allowed to float:

```
  Floating Parameter  InitialValue    FinalValue +/-  Error     GblCorr.
--------------------  ------------  --------------------------  --------
Frac_nBkg_chgCat_Plus    5.1950e-01    5.2061e-01 +/-  7.09e-03  0.123644
Frac_nK2_chgCat_Plus    5.1950e-01    4.7501e-01 +/-  2.87e-01  0.496197
Frac_nKst1430_chgCat_Plus    5.1950e-01    4.3360e-01 +/-  1.45e-01  0.498190
           deBkg_P01   -1.4535e+00   -1.5325e+00 +/-  1.27e-01  0.158481
        fisBkgC_asym    5.8221e-02    5.0197e-02 +/-  2.56e-02  0.385364
        fisBkgC_mean    4.2342e-01    4.2774e-01 +/-  1.03e-02  0.564137
         fisBkgC_rms    5.7676e-01    5.8047e-01 +/-  7.27e-03  0.489288
        mEBkg_fracS    3.8761e-01    3.9403e-01 +/-  1.53e-02  0.175442
       mEPolyBkg_P01   -2.3527e-01   -2.6145e-01 +/-  3.27e-02  0.081217
       mKstarBkg_P01   -6.7058e-01   -6.8550e-01 +/-  2.98e-02  0.972059
       mKstarBkg_P02   -3.2838e-02   -3.3193e-02 +/-  2.02e-02  0.979770
       mKstarBkg_P03    1.4451e-02    1.8267e-02 +/-  1.32e-02  0.982489
       mKstarBkg_P04    3.6500e-02    3.9012e-02 +/-  7.87e-03  0.980342
            mesBkg_c   -7.6735e+00   -1.2218e+01 +/-  3.68e+00  0.278226
                nBkg    3.5000e+03    5.0971e+03 +/-  8.17e+01  0.477384
               nChls    1.2000e+01    5.2839e+00 +/-  4.06e+01  0.761133
                 nK2    0.0000e+00    1.7070e+01 +/-  1.05e+01  0.492058
            nKst1430    2.1000e+01    5.3515e+01 +/-  1.61e+01  0.519387
```

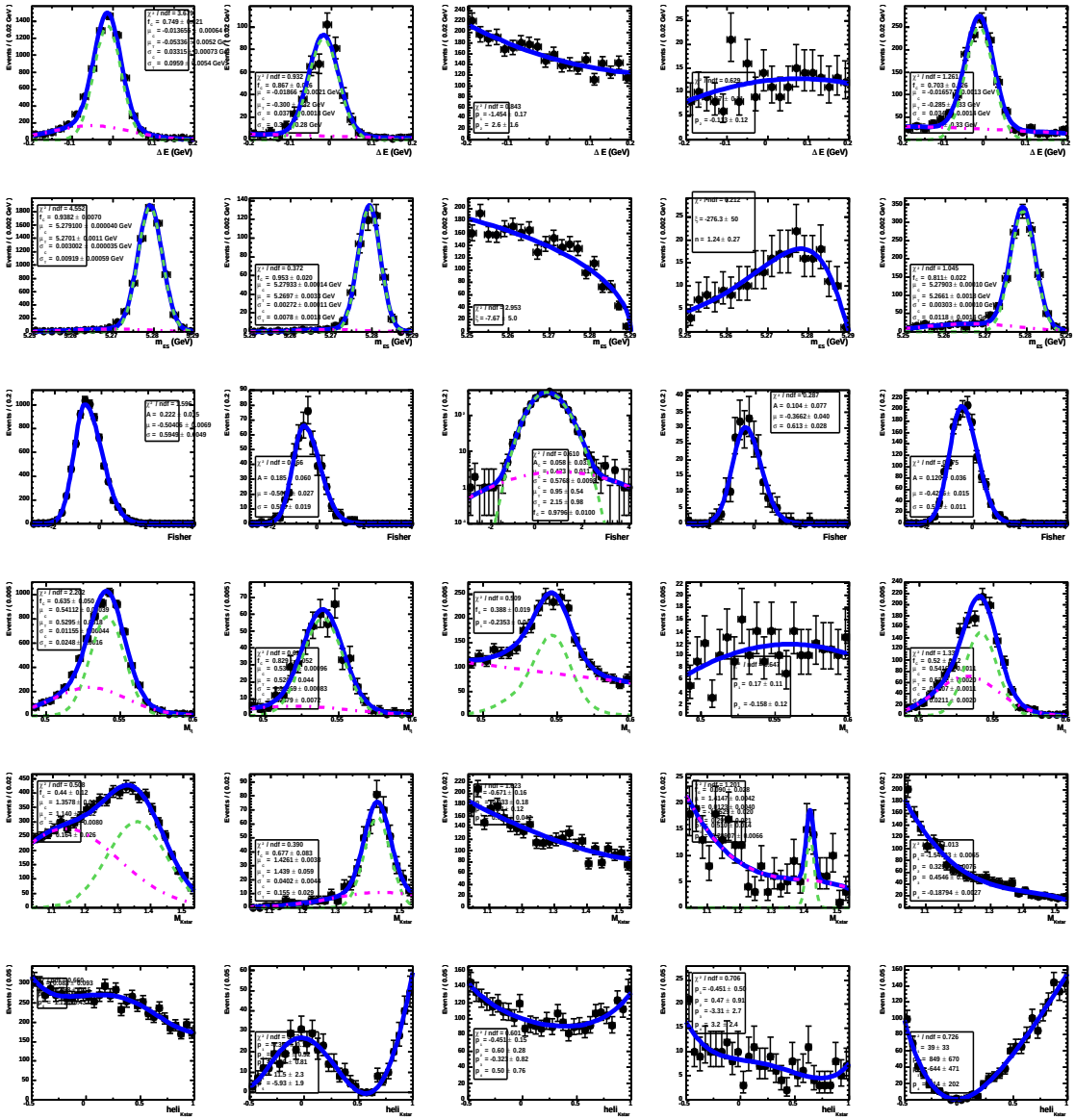


Figure H.8: PDFs for (top to bottom) ΔE, $m_{\rm ES}$, η mass, $\mathcal{F}$, K^* mass and $\mathcal{H}$, for (left to right) $B \to \eta K_0^*(S\text{-wave})$, $B \to \eta K_2^*(1430)$, on-peak sidebands, charmless $B\overline{B}$ background, and $B \to \eta K^*(892)$.

H.9 $\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}$ (S-wave), (D-wave)

```
Correlation matrix for onData (13117 events):
               de       mes    fisher      mEta    mKstar
     mes    0.0027
  fisher   -0.0332   -0.0390
    mEta    0.0108    0.0138   -0.0266
  mKstar    0.0110    0.0362   -0.0534    0.0001

Correlation matrix for kst1430MC (8891 events):
               de       mes    fisher      mEta    mKstar
     mes   -0.0001
  fisher    0.0107   -0.0246
    mEta    0.1224    0.0443    0.0227
  mKstar    0.0327   -0.0133   -0.0113    0.0148

Correlation matrix for k2MC (650 events):
               de       mes    fisher      mEta    mKstar
     mes    0.0048
  fisher    0.0972   -0.0138
    mEta    0.1958    0.1141    0.0876
  mKstar    0.0510   -0.0103    0.0078    0.0090
```

The $\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}$(HMR) fit finds the following values for the parameters which were allowed to float:

```
  Floating Parameter  InitialValue    FinalValue +/-  Error     GblCorr.
--------------------  ------------  --------------------------  --------
Frac_nBkg_chgCat_Plus    4.9783e-01    4.9810e-01 +/-  4.49e-03  0.112025
Frac_nK2_chgCat_Plus    4.9783e-01    6.0232e-01 +/-  1.16e-01  0.443028
Frac_nKst1430_chgCat_Plus    4.9783e-01    4.2828e-01 +/-  7.71e-02  0.440309
           deBkg_P01   -1.5023e+00   -1.5738e+00 +/-  7.96e-02  0.235431
        fisBkgC_asym    9.3404e-02    7.4207e-02 +/-  1.58e-02  0.327053
        fisBkgC_mean    4.2960e-01    4.3040e-01 +/-  6.67e-03  0.553763
         fisBkgC_rms    6.0394e-01    5.9681e-01 +/-  4.63e-03  0.448777
         mEBkg_fracS    3.7432e-01    3.7698e-01 +/-  9.66e-03  0.118630
       mEPolyBkg_P01   -3.6277e-01   -3.4865e-01 +/-  2.06e-02  0.240818
       mKstarBkg_P01   -9.8208e-01   -9.2384e-01 +/-  6.00e-02  0.997771
       mKstarBkg_P02    2.6394e-01    2.3339e-01 +/-  3.47e-02  0.997772
            mesBkg_c   -2.5260e+01   -2.3424e+01 +/-  2.37e+00  0.281241
                nBkg    8.8000e+03    1.2788e+04 +/-  1.30e+02  0.486200
               nChls    1.3000e+02    7.4315e+01 +/-  6.43e+01  0.742276
                 nK2    0.0000e+00    7.1872e+01 +/-  1.67e+01  0.448632
            nKst1430    1.4000e+02    1.6206e+02 +/-  2.56e+01  0.477690
```

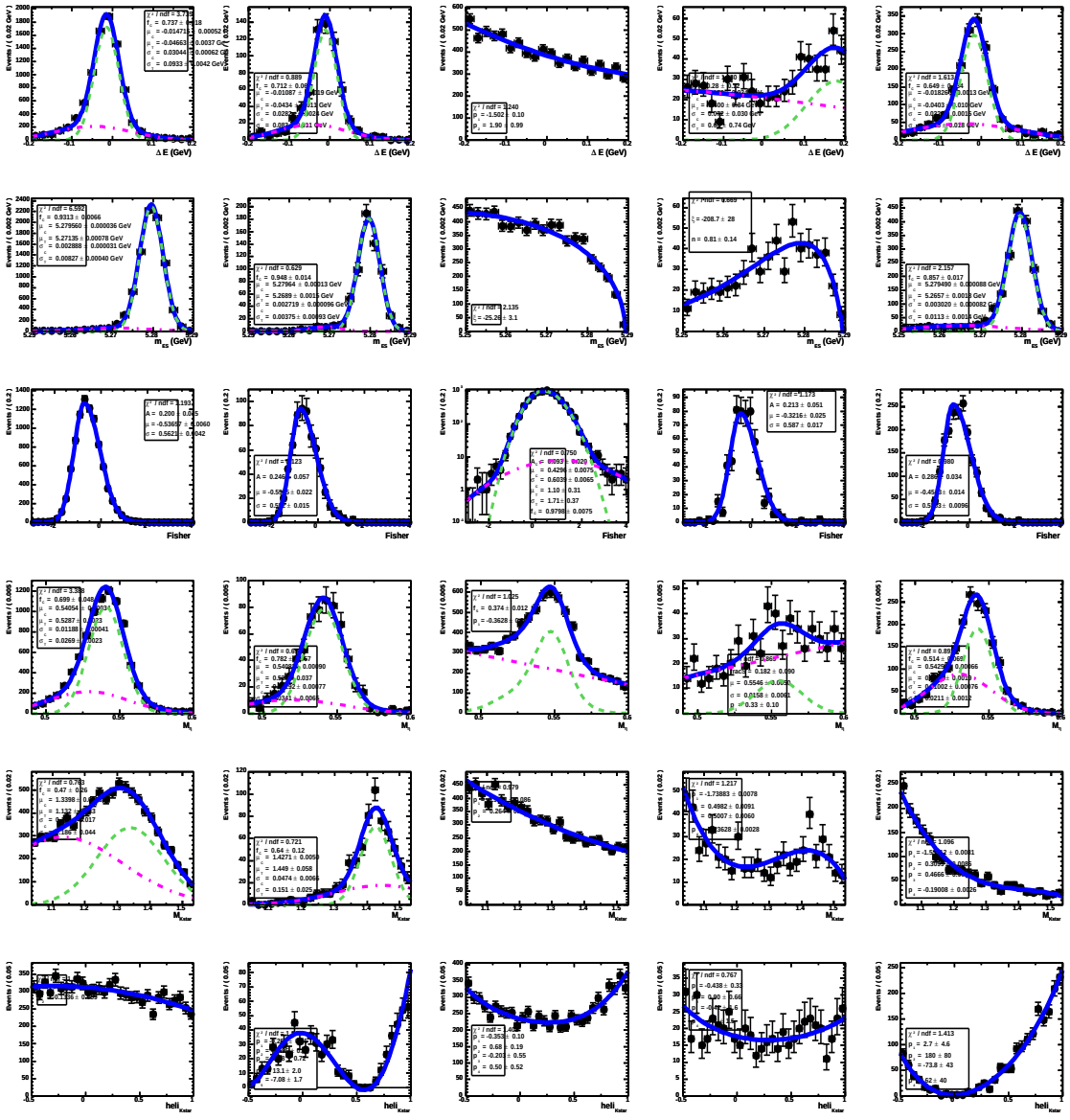


Figure H.9: PDFs for (top to bottom) ΔE, $m_{\rm ES}$, η mass, $\mathcal{F}$, K^* mass and $\mathcal{H}$, for (left to right) $B \to \eta K_0^*(S\text{-wave})$, $B \to \eta K_2^*(1430)$, on-peak sidebands, charmless $B\overline{B}$ background, and $B \to \eta K^*(892)$.

H.10 $\eta_{3\pi}K^{*+}_{K^+\pi^0}$(*S*-wave), (*D*-wave)

```
Correlation matrix for onData (3838 events):
           fisher    mKstar    hKstar        de       mes
  mKstar   -0.0546
  hKstar   -0.0075    0.0746
      de   -0.0587   -0.0129    0.0193
     mes   -0.0017    0.0241   -0.0039   -0.0090

Correlation matrix for kst1430MC (2343 events):
           fisher    mKstar    hKstar        de       mes
  mKstar    0.0195
  hKstar   -0.0004    0.0143
      de    0.0251    0.0511   -0.0787
     mes   -0.0799   -0.0167   -0.0550   -0.0191

Correlation matrix for k2MC (172 events):
           fisher    mKstar    hKstar        de       mes
  mKstar   -0.0047
  hKstar    0.1022   -0.0650
      de    0.0372   -0.0162   -0.0155
     mes   -0.0619   -0.0056   -0.1037   -0.0190
```

The $\eta_{3\pi}K^{*+}_{K^+\pi^0}$(HMR) fit finds the following values for the parameters which were allowed to float:

```
  Floating Parameter  InitialValue    FinalValue +/-  Error     GblCorr.
--------------------  ------------  --------------------------  --------
Frac_nBkg_chgCat_Plus    5.2788e-01    5.2659e-01 +/-  8.21e-03  0.102892
Frac_nK2_chgCat_Plus    5.2788e-01    9.1154e-01 +/-  2.37e-01  0.650693
Frac_nKst1430_chgCat_Plus    5.2788e-01    4.8550e-01 +/-  1.44e-01  0.478380
           deBkg_P01   -1.5190e+00   -1.5549e+00 +/-  1.48e-01  0.015989
        fisBkgC_asym    8.2502e-02    7.2824e-02 +/-  2.82e-02  0.130800
        fisBkgC_mean    4.5624e-01    4.6570e-01 +/-  1.04e-02  0.187196
         fisBkgC_rms    5.9907e-01    5.9936e-01 +/-  7.70e-03  0.148490
         mEBkg_fracS    3.1861e-01    3.2341e-01 +/-  1.50e-02  0.132475
       mEPolyBkg_P01    4.1386e-01    4.0574e-01 +/-  3.47e-02  0.103313
       mKstarBkg_P01   -1.4915e+00   -1.3552e+00 +/-  2.65e-02  0.997752
       mKstarBkg_P02    3.0184e-01    2.0284e-01 +/-  2.62e-02  0.999176
       mKstarBkg_P03    4.4120e-01    3.9326e-01 +/-  1.76e-02  0.999293
       mKstarBkg_P04   -1.7881e-01   -1.4059e-01 +/-  8.07e-03  0.998640
            mesBkg_c   -1.3145e+01   -1.8141e+01 +/-  4.20e+00  0.122390
                nBkg    2.5000e+03    3.7751e+03 +/-  6.21e+01  0.114519
                 nK2    0.0000e+00    1.6969e+01 +/-  8.63e+00  0.596081
            nKst1430    2.6000e+01    4.2517e+01 +/-  1.27e+01  0.510642
```

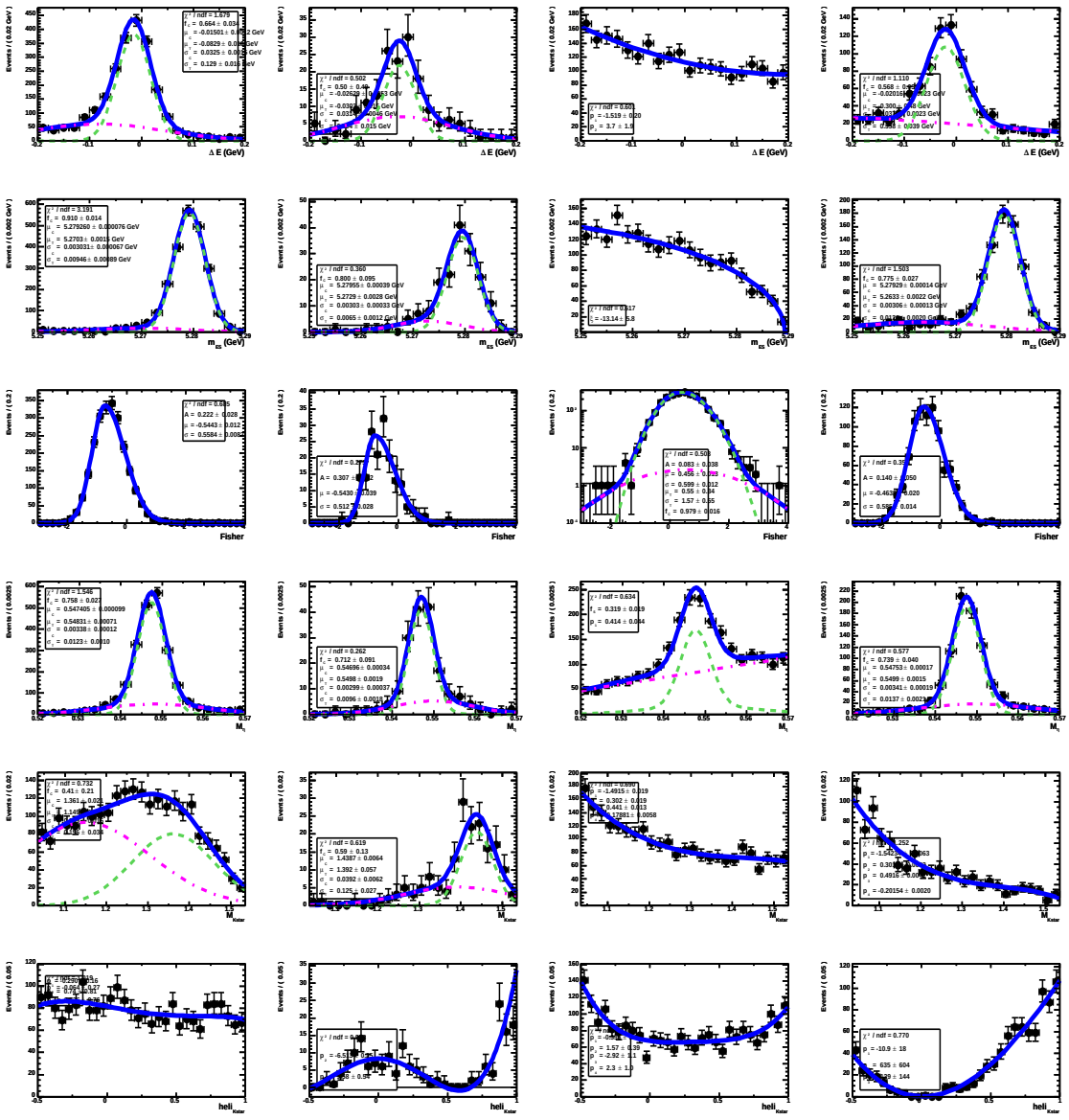


Figure H.10: PDFs for (top to bottom) ΔE, $m_{\rm ES}$, η mass, $\mathcal{F}$, K^* mass and $\mathcal{H}$, for (left to right) $B \to \eta K_0^*(S\text{-wave})$, $B \to \eta K_2^*(1430)$, on-peak sidebands, and $B \to \eta K^*(892)$.

H.11 $\eta_{3\pi}K^{*+}_{K^0_S\pi^+}$(*S*-wave), (*D*-wave)

```
Correlation matrix for onData (2847 events):
            fisher    mKstar    hKstar        de       mes
  mKstar   -0.0652
  hKstar    0.0172    0.1028
      de   -0.0132   -0.0214    0.0062
     mes    0.0030   -0.0048   -0.0166   -0.0110

Correlation matrix for kst1430MC (3219 events):
            fisher    mKstar    hKstar        de       mes
  mKstar   -0.0132
  hKstar    0.0349    0.0168
      de   -0.0113    0.0071   -0.0633
     mes   -0.0765   -0.0403   -0.0489    0.0788

Correlation matrix for k2MC (217 events):
            fisher    mKstar    hKstar        de       mes
  mKstar    0.0856
  hKstar    0.1507   -0.0150
      de    0.0208   -0.0633   -0.0958
     mes   -0.0952   -0.0866   -0.0379    0.2057
```

The $\eta_{3\pi}K^{*+}_{K^0\pi^+}$(HMR) fit finds the following values for the parameters which were allowed to float:

```
  Floating Parameter  InitialValue    FinalValue +/-  Error     GblCorr.
--------------------  ------------  --------------------------  --------
Frac_nBkg_chgCat_Plus    4.9245e-01    4.9356e-01 +/-  9.55e-03  0.097984
Frac_nK2_chgCat_Plus    4.9245e-01    9.9986e-01 +/-  8.16e-01  0.502754
Frac_nKst1430_chgCat_Plus    4.9245e-01    4.0771e-01 +/-  1.09e-01  0.344093
           deBkg_P01   -1.3479e+00   -1.3884e+00 +/-  1.65e-01  0.004273
        fisBkgC_asym    5.8920e-02    4.7882e-02 +/-  3.51e-02  0.171357
        fisBkgC_mean    3.6304e-01    3.7786e-01 +/-  1.26e-02  0.205493
         fisBkgC_rms    5.5731e-01    5.6467e-01 +/-  9.41e-03  0.121405
        mEBkg_fracS    3.2896e-01    3.3432e-01 +/-  1.70e-02  0.108539
       mEPolyBkg_P01    3.4222e-01    3.7272e-01 +/-  4.07e-02  0.077348
       mKstarBkg_P01   -7.0456e-01   -6.8866e-01 +/-  1.39e-01  0.997947
       mKstarBkg_P02   -5.7877e-02   -5.5996e-02 +/-  1.22e-01  0.999104
       mKstarBkg_P03    2.5496e-02    2.2869e-02 +/-  8.23e-02  0.999269
       mKstarBkg_P04    4.9363e-02    4.5642e-02 +/-  3.64e-02  0.998520
            mesBkg_c   -6.1087e+00   -6.7933e+00 +/-  4.92e+00  0.116577
                nBkg    2.0000e+03    2.7939e+03 +/-  5.34e+01  0.104730
                 nK2    0.0000e+00    1.9546e+00 +/-  4.08e+00  0.532044
           nKst1430    4.8000e+01    4.8779e+01 +/-  1.08e+01  0.343102
```

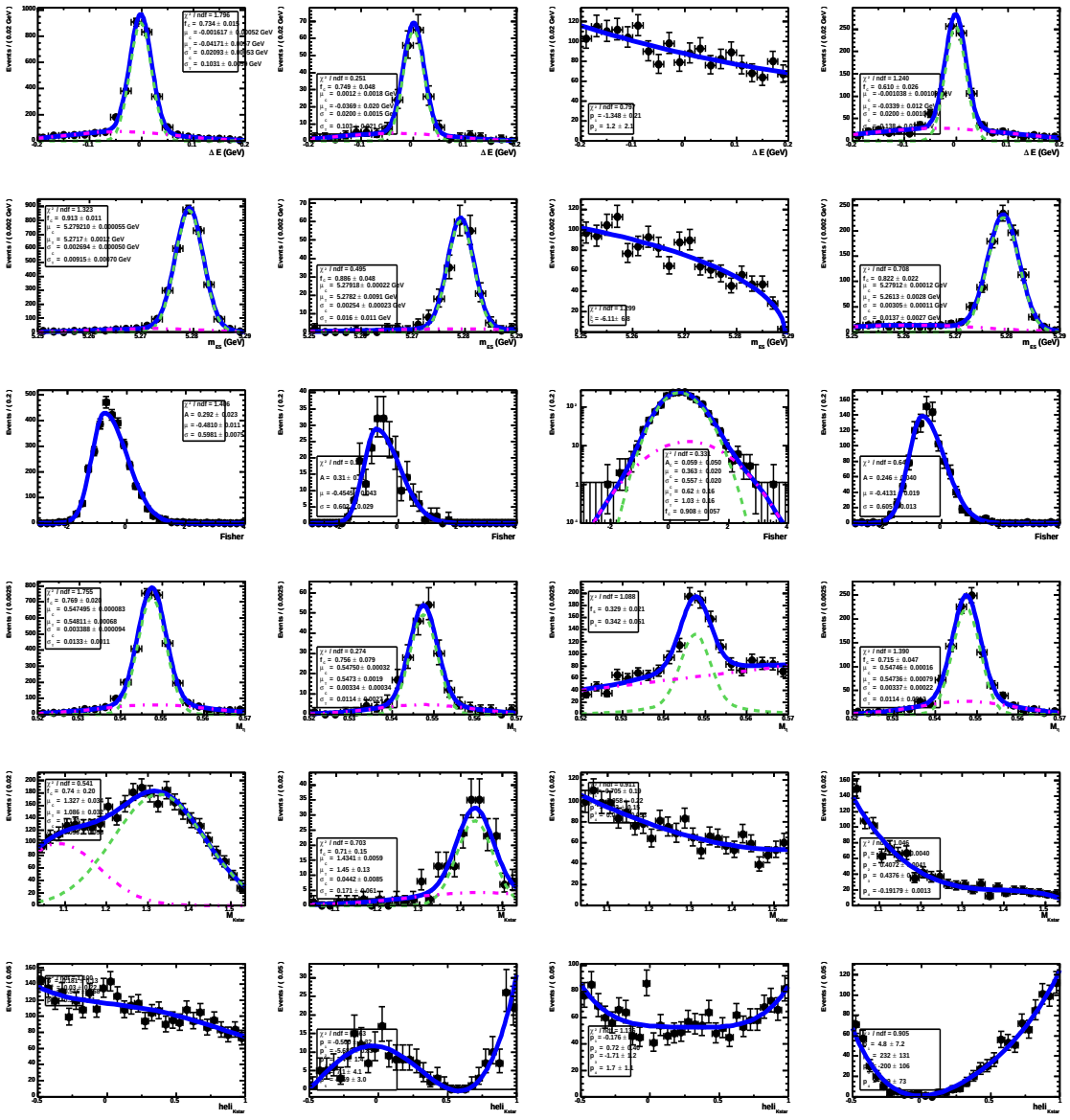


Figure H.11: PDFs for (top to bottom) ΔE, $m_{\rm ES}$, η mass, $\mathcal{F}$, K^* mass and $\mathcal{H}$, for (left to right) $B \to \eta K_0^*(S\text{-wave})$, $B \to \eta K_2^*(1430)$, on-peak sidebands, and $B \to \eta K^*(892)$.

H.12 $\eta_{3\pi}K^{*0}_{K^+\pi^-}$ (S-wave), (D-wave)

```
Correlation matrix for onData (7427 events):
            fisher    mKstar    hKstar        de       mes
  mKstar   -0.0460
  hKstar   -0.0183    0.0425
      de   -0.0296    0.0102   -0.0035
     mes   -0.0052    0.0026    0.0198   -0.0018

Correlation matrix for kst1430MC (3597 events):
            fisher    mKstar    hKstar        de       mes
  mKstar   -0.0046
  hKstar    0.0387   -0.0168
      de    0.0307    0.0134    0.0300
     mes   -0.0555   -0.0187    0.0017    0.0447

Correlation matrix for k2MC (260 events):
            fisher    mKstar    hKstar        de       mes
  mKstar    0.0618
  hKstar   -0.0117   -0.0001
      de    0.1887    0.1137    0.1331
     mes   -0.0665    0.0227   -0.0869    0.0122
```

The $\eta_{3\pi}K^{*0}_{K^+\pi^-}$(HMR) fit finds the following values for the parameters which were allowed to float:

```
  Floating Parameter  InitialValue    FinalValue +/-  Error     GblCorr.
--------------------  ------------  --------------------------  --------
Frac_nBkg_chgCat_Plus    5.1245e-01    5.1246e-01 +/-  5.91e-03  0.104595
Frac_nK2_chgCat_Plus    5.1245e-01    3.8709e-01 +/-  1.57e-01  0.462437
Frac_nKst1430_chgCat_Plus    5.1245e-01    5.9204e-01 +/-  1.26e-01  0.499833
           deBkg_P01   -1.6797e+00   -1.6879e+00 +/-  9.94e-02  0.009713
        fisBkgC_asym    5.9680e-02    7.7352e-02 +/-  2.07e-02  0.098095
        fisBkgC_mean    4.0821e-01    4.1311e-01 +/-  7.45e-03  0.177131
         fisBkgC_rms    5.8767e-01    5.9043e-01 +/-  5.51e-03  0.123598
         mEBkg_fracS    3.2048e-01    3.2053e-01 +/-  1.03e-02  0.104640
       mEPolyBkg_P01    3.8843e-01    3.9734e-01 +/-  2.42e-02  0.076201
       mKstarBkg_P01   -9.7510e-01   -1.0251e+00 +/-  1.73e-02  0.962511
       mKstarBkg_P02    2.6693e-01    2.9419e-01 +/-  9.96e-03  0.962517
            mesBkg_c   -1.9015e+01   -1.4593e+01 +/-  3.03e+00  0.126093
                nBkg    4.9000e+03    7.3090e+03 +/-  8.64e+01  0.111581
                 nK2    0.0000e+00    4.1484e+01 +/-  1.34e+01  0.474847
           nKst1430    6.4000e+01    6.8055e+01 +/-  1.72e+01  0.488710
```

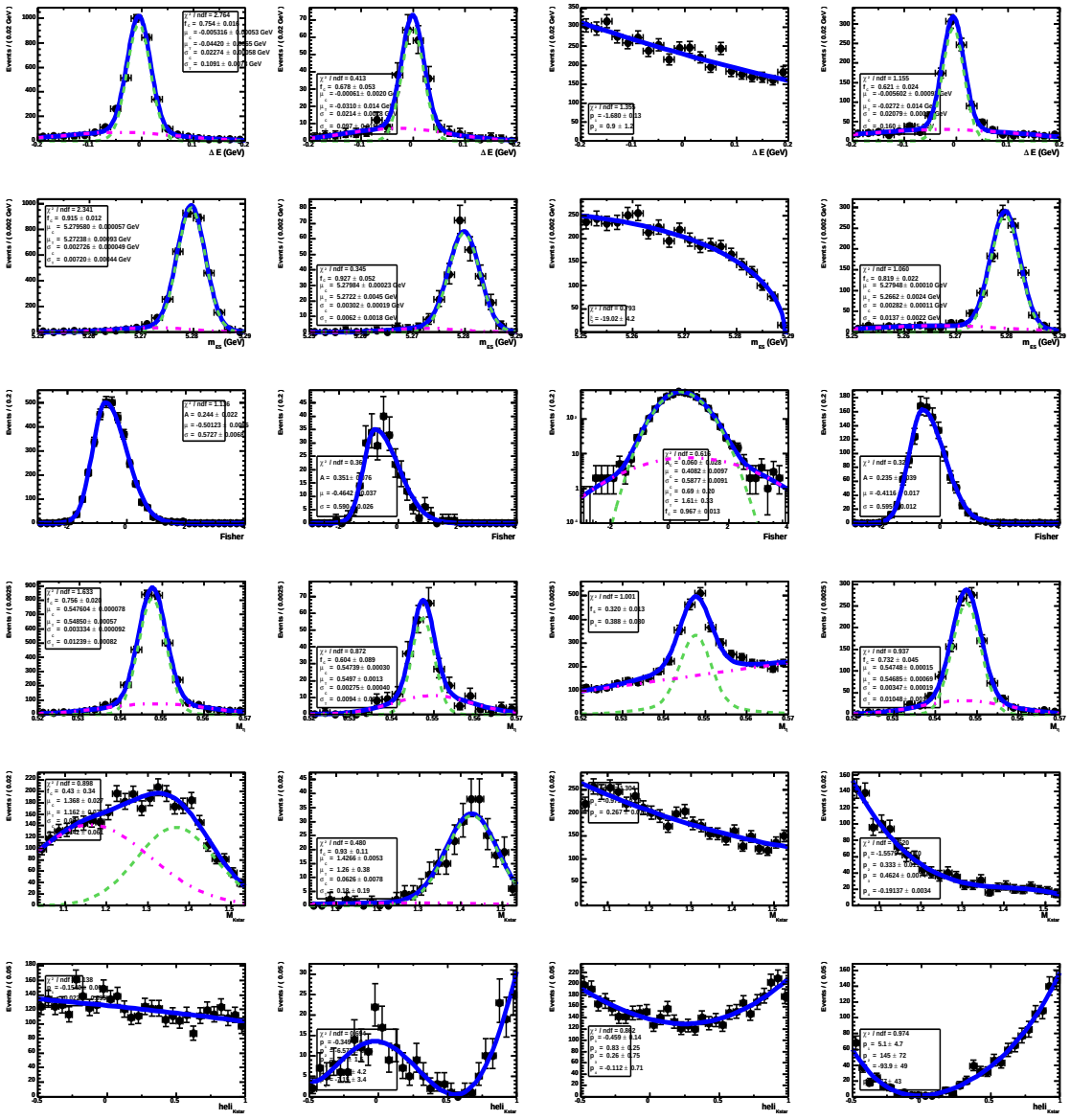


Figure H.12: PDFs for (top to bottom) ΔE, $m_{\rm ES}$, η mass, $\mathcal{F}$, K^* mass and $\mathcal{H}$, for (left to right) $B \to \eta K_0^*(S\text{-wave})$, $B \to \eta K_2^*(1430)$, on-peak sidebands, and $B \to \eta K^*(892)$.

Appendix I

Study of separate charmless $B\overline{B}$ fit components in ηK^*

In the LMR analysis, the dominant charmless $B\overline{B}$ background for $\eta K^*(892)$ (with $\eta \to \gamma\gamma$) is $K^*\gamma$ (see Tables 6.3, 6.4, and 6.5). The ΔE, resonance mass, and helicity PDFs for $K^*\gamma$ are qualitatively different from those of the total charmless component. We investigate whether the analysis benefits from the inclusion of a dedicated $K^*\gamma$ fit component.

For the $\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}(892)$ mode, which has the highest number of expected $B\overline{B}$ background events, we separate the charmless $B\overline{B}$ component into three separate components based on PDF shape and expected contribution: $K^{*0}_{K^+\pi^-}\gamma$, $K^*_X\gamma$ (where the K^*_X is a high-mass $K\pi$ resonance such as $K^*_1(1273)$), and the remaining modes. The weighted mix of modes and PDFs for distinct distributions for these charmless components are shown in Table I.1 and Fig. I.1.

For the $\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}(892)$ mode, which had the highest charmless yield of any mode in results from Run1-2 published in 2004, we separate the total component into two separate components: $K^{*0}_{K^0\pi^+}\gamma$ and all other modes. The weighted mix of modes and PDFs for distinct distributions for these separate components are shown in Table I.2 and Fig. I.2.

We perform the six-component fit for $\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}$ and the five-component fit $\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}$on the unblind Run1-2 data with various combinations of floated and fixed yields for the different charmless $B\overline{B}$ components. Comparing our results to the re-

sults from the nominal fit in Table I.3, we find the signal yield and errors to be very stable despite significant changes to the treatment of the background. We conclude that including multiple charmless components does not enhance the fit.

Table I.1: Separate background components for the $\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}(892)$ mode from exclusive charmless B decay. We show efficiency for the mode to pass selection cuts, the measured or estimated branching fraction, the appropriate product branching fraction given how the MC was produced, the estimated background normalized to 235.7 million $B\overline{B}$ events, the number of events we include in the relevant component file we use for making PDFs, and the total number of events in the file for each component. A * denotes that the branching fraction is estimated since no measurement exists.

Signal mode Bkg. channel	Mode Number	MC ϵ (%)	Est. $\mathcal{B}$ (10^{-6})	$\prod \mathcal{B}_i$ (%)	Norm. # $B\overline{B}$ Bkg.	# in PDF Bkg. file
$B\overline{B}$ Component 1:						
$B^0 \to K^{*0}_{K^+\pi^-}\gamma$	1710	0.91	40	0.667	57	4584
$B\overline{B}$ Component 2:						
$B^+ \to K_1^{*+}(1273)\gamma$	1453	0.1	43	1.000	10.5	229
$B^+ \to K^{*+}(1410)\gamma$	1970	0.13	15	1.000	4.7	104
$B^0 \to K_1^{*0}(1273)\gamma$	1452	0.11	16	1.000	4.1	90
$B^0 \to K^{*0}(1410)\gamma$	1971	0.1	15	1.000	3.4	74
$B^+ \to K_1^{*+}(1402)\gamma$	1454	0.16	8	1.000	3.1	67
$B^+ \to K_2^{*+}(1430)\gamma$	1765	0.06	15	1.000	2	45
$B^0 \to K_1^{*0}(1402)\gamma$	1455	0.1	8	1.000	1.9	42
$B^0 \to K_2^{*0}(1430)\gamma$	1766	0.06	12	1.000	1.6	35
$B^0 \to K^{*0}(1680)\gamma$	1973	0.05	2	1.000	0.2	5
$B^+ \to K^{*+}(1680)\gamma$	1972	0.05	2	1.000	0.2	4
					31.7	695
$B\overline{B}$ Component 3:						
$B^+ \to \eta_{\gamma\gamma}K^{*+}_{K^+\pi^0}$	1539	2.25	24	0.131	16.7	1668
$B^+ \to \eta_{\gamma\gamma}K^+$	1513	1.12	3	0.394	3.1	312
$B^0 \to K^{*0}_{K^+\pi^-}\pi^0$	1225	0.75	2	0.667	2.4	235
$B^0 \to \eta_{\gamma\gamma}\phi_{K^+K^-}$	2904	4.72	1*	0.193	2.1	214
$B^+ \to a_0^0(\eta_{\gamma\gamma}\pi^0)K^+$	6552	0.9	1*	0.394	0.8	83
$B^0 \to a_0^0(\eta_{\gamma\gamma}\pi^0)\ K^{*0}_{K^+\pi^-}$	6562	1.05	1*	0.263	0.7	65
					25.8	2577

Table I.2: Separate background components for the $\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}(892)$ mode from exclusive charmless B decay. See Table I.1 for explanation of column headers. A * denotes that the branching fraction is estimated since no measurement exists.

Signal mode Bkg. channel	Mode Number	MC ϵ (%)	Est. $\mathcal{B}$ (10^{-6})	$\prod \mathcal{B}_i$ (%)	Norm. # $B\overline{B}$ Bkg.	# in PDF Bkg. file
$B\overline{B}$ Component 1:						
$B^+ \to\ K^{*+}\gamma$	3135	0.21	40	1.000	19.6	4602
$B\overline{B}$ Component 2:						
$B^0 \to \eta_{\gamma\gamma}\ K^{*0}_{K_S\pi^0}$	6564	4.31	18	0.045	8.2	181
$B^+ \to\ K^{*+}_1(1273)\gamma$	1453	0.04	43	1.000	4.3	94
$B^0 \to\ K^{*0}(1410)\gamma$	1971	0.06	15	1.000	2.1	45
$B^0 \to\ K^{*0}_1(1273)\gamma$	1452	0.05	16	1.000	1.9	42
$B^0 \to\ K^{*0}_{K_S\pi^0}\gamma$	5197	0.16	40	0.115	1.7	37
$B^0 \to\ K^{*0}_1(1402)\gamma$	1455	0.06	8	1.000	1.2	25
$B^+ \to\ K^{*+}(1410)\gamma$	1970	0.03	15	1.000	1.1	24
$B^+ \to\ K^{*+}_2(1430)\gamma$	1765	0.03	15	1.000	1	22
$B^0 \to\ K^{*0}_2(1430)\gamma$	1766	0.03	12	1.000	0.9	20
$B^0 \to \eta_{\gamma\gamma}K_S$	2603	2.26	1*	0.136	0.7	15
$B^+ \to\ K^{*+}_1(1402)\gamma$	1454	0.04	8	1.000	0.7	15
$B^0 \to\ a^0_0(\eta_{\gamma\gamma}\pi^0)K_S$	6547	1.56	1*	0.136	0.5	11
$B^0 \to \eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}$	1541	0.04	19	0.262	0.5	10
$B^0 \to\ K^{*0}(1680)\gamma$	1973	0.02	2	1.000	0.1	2
$B^+ \to\ K^{*+}(1680)\gamma$	1972	0.02	2	1.000	0.1	2
					25	545

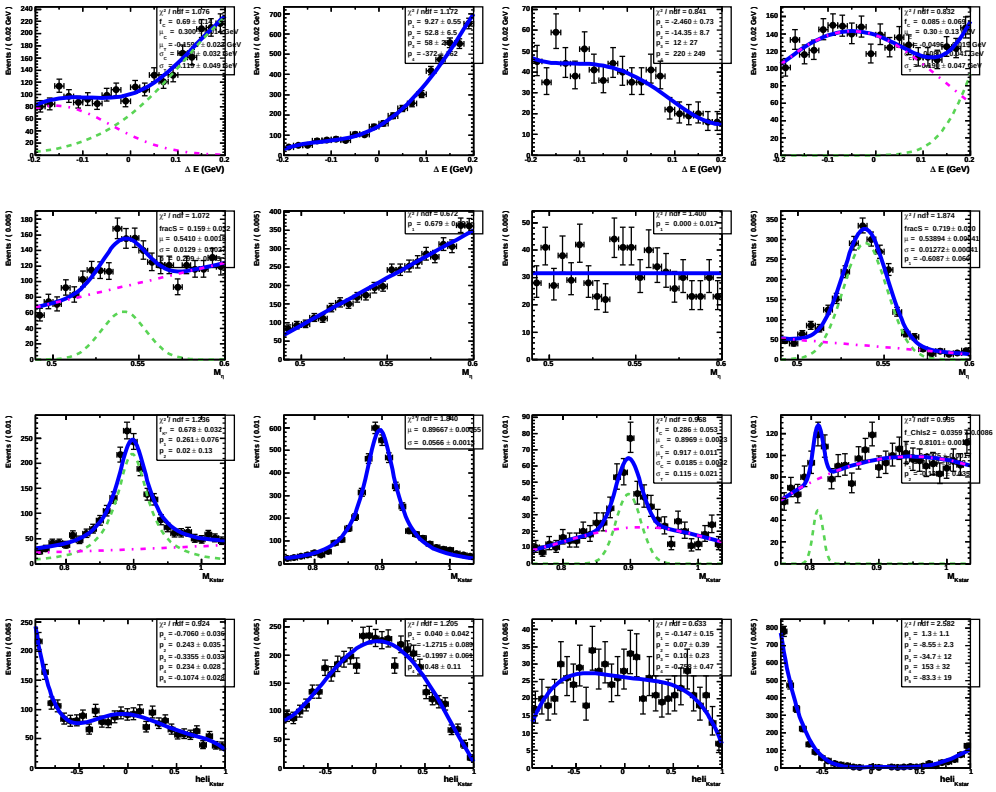


Figure I.1: For the mode $\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}(892)$, the ΔE, η mass, and helicity distributions for the original $B\overline{B}$ component (*left*); and the separate $K^{*0}_{K^+\pi^-}\gamma$ (*left-center*), high-mass $K\pi$ resonances (*right-center*), and the remaining-$B\overline{B}$ (*right*) components. The small peak near 800 MeV in the K^* mass for the remaining-$B\overline{B}$ component is from $\phi \to K^+K^-$.

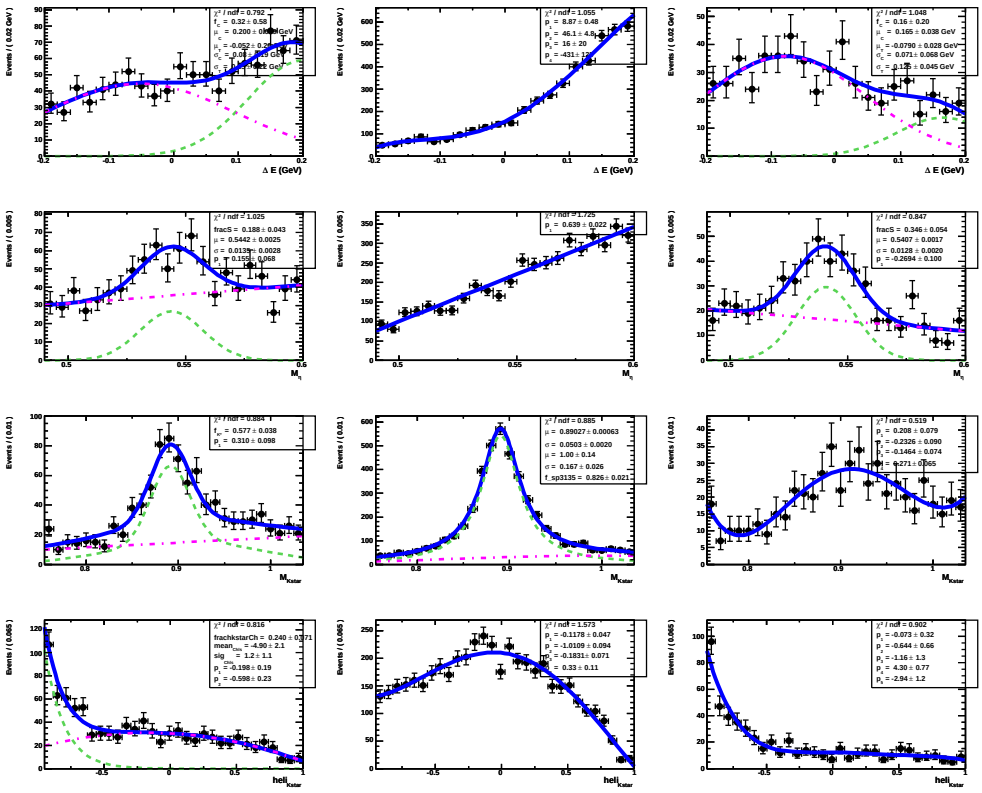


Figure I.2: For the mode $\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}(892)$, ΔE, η mass, K^* mass, and helicity distributions for the original $B\overline{B}$ component (*left*); and the separate $K^{*+}_{K^+\pi^0}\gamma$ (*center*) and remaining-$B\overline{B}$ (*right*) components.

Table I.3: For modes $\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}(892)$ and $\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}(892)$, results from the nominal fit to Run1-2 data (first column) and fits with multiple charmless $B\overline{B}$ components. Yield results with no errors mean that the yield was fixed; the $\eta K^*_0(S\text{-wave})$ yield is fixed in all fits as is done in the nominal fit. The choices of fixed yields in Test Fit 1 come from Tables I.1 and I.2. Results show that the ηK^* yields and errors are not sensitive to the treatment of charmless $B\overline{B}$ background.

Component	Nominal Fit	Test Fit 1 Fix Y	Test Fit 2 Float Y	Test Fit 3 No $K^*_X\gamma$
$\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}(892)$				
$Y_{\eta K^*}$	113.9 ± 15.6	109.2 ± 15.5	112.2 ± 15.6	111.2 ± 15.6
$Y_{S\text{-wave}}$	21.5	21.5	21.5	21.5
$Y_{K^*\gamma}$	–	57.0	25.8 ± 16.2	17.5 ± 14.6
$Y_{K^*_X\gamma}$	–	30.0	-69.0 ± 24.8	0.0
$Y_{B\overline{B}}$	–	25.8 ± 19.0	13.6 ± 18.6	20.2 ± 18.6
$Y_{\text{full}B\overline{B}}$	35.6 ± 28.4	–	–	–
$\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}(892)$				
Y_{sig}	37.9 ± 10.5	38.4 ± 10.5	37.2 ± 10.4	–
$Y_{S\text{-wave}}$	14.1	14.1	14.1	–
$Y_{K^*\gamma}$	–	19.6	15.2 ± 12.8	–
$Y_{B\overline{B}}$	–	68.5 ± 28.9	59.7 ± 29.1	–
$Y_{\text{full}B\overline{B}}$	68.2 ± 25.0	–	–	–

Appendix J

ηK^* : $s\mathcal{P}lot$ libraries

We show the signal, continuum, and charmless $B\overline{B}$ background $s\mathcal{P}lots$ for $B \to \eta K^*(892)$, $B \to \eta K_0^*(S\text{-wave})$, and $B \to \eta K_2^*(1430)$.

J.1 $\eta_{\gamma\gamma} K^{*+}_{K^+\pi^0}(892)$

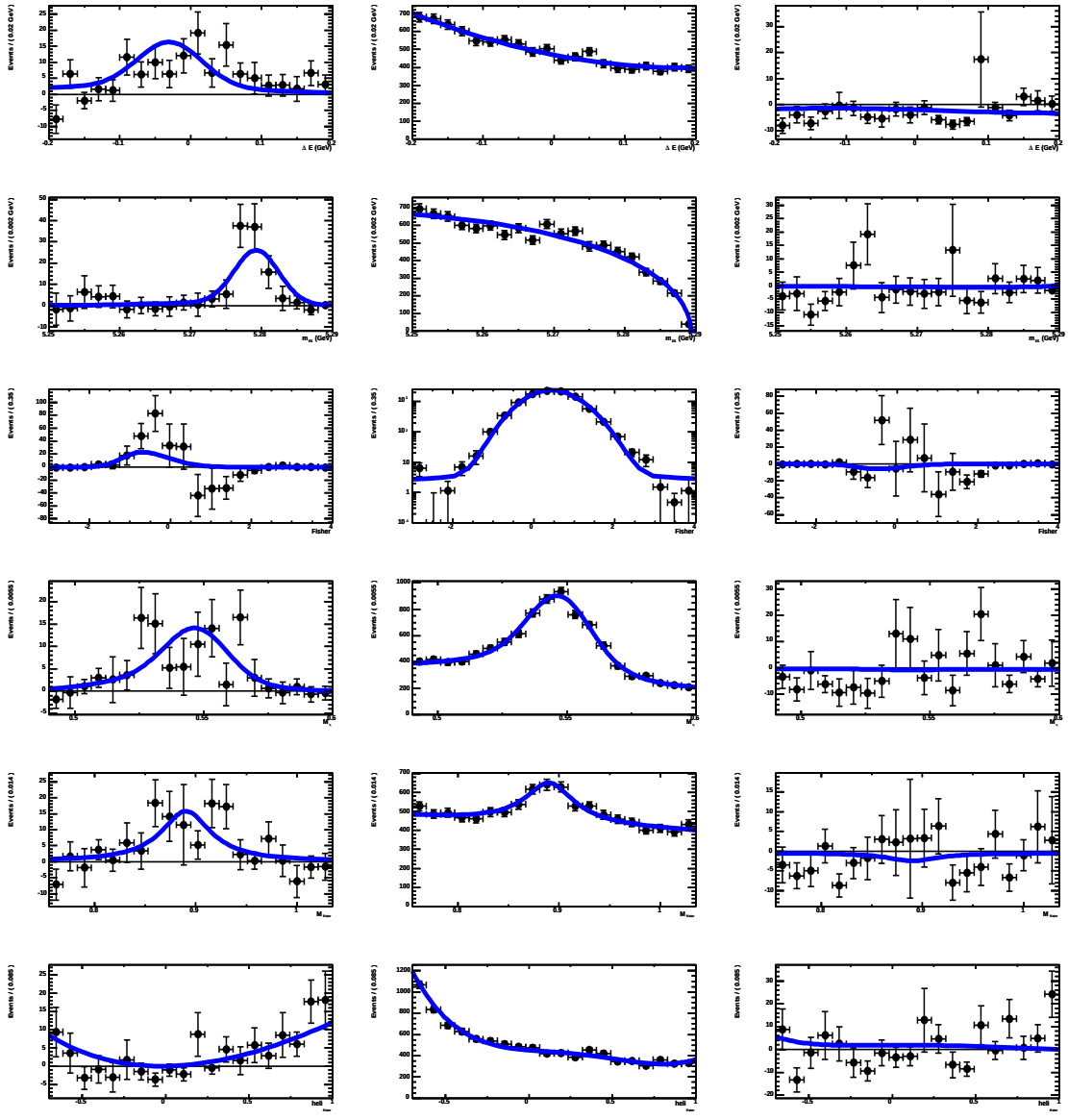


Figure J.1: *sPlots* for (top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η mass, K^* mass and $\mathcal{H}$ for (left to right) $B \to \eta K^*(892)$, continuum background, and charmless $B\overline{B}$ background.

J.2 $\eta_{\gamma\gamma} K^{*+}_{K^0_S \pi^+}(892)$

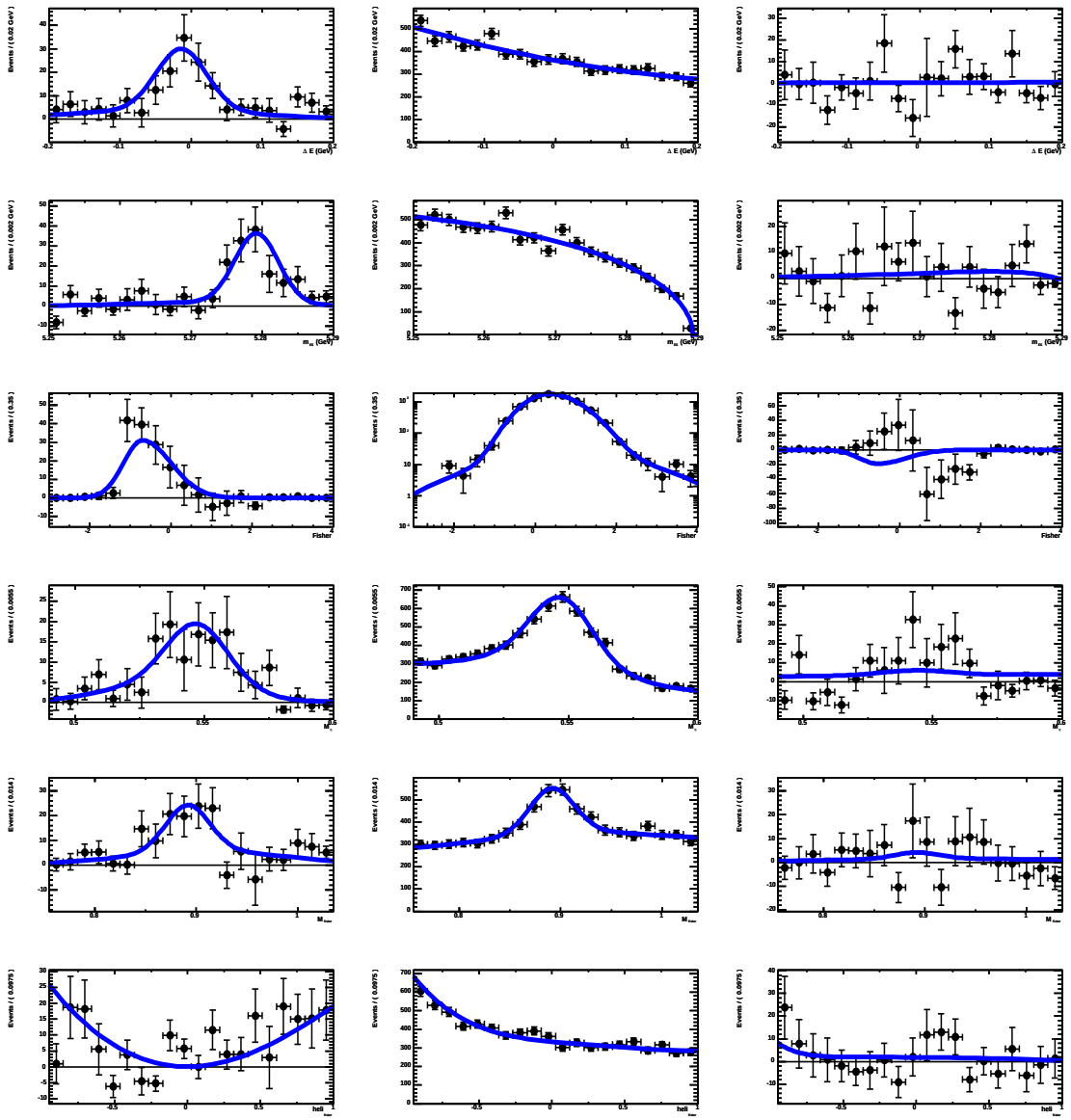


Figure J.2: $s\mathcal{P}lots$ for (top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η mass, K^* mass and $\mathcal{H}$ for (left to right) $B \to \eta K^*(892)$, continuum background, and charmless $B\overline{B}$ background.

J.3 $\eta_{\gamma\gamma} K^{*0}_{K^+\pi^-}(892)$

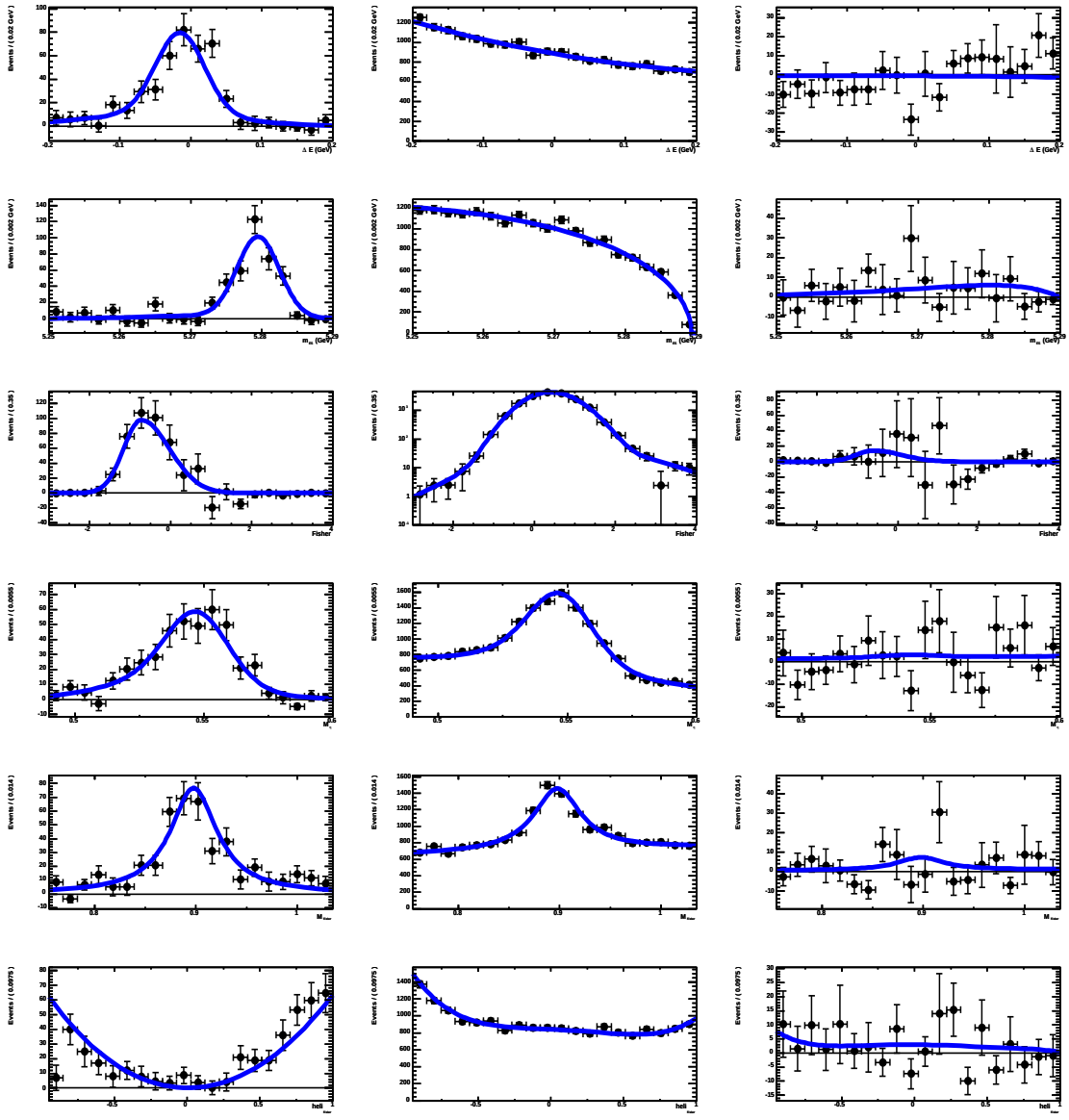


Figure J.3: *sPlots* for (top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η mass, K^* mass and $\mathcal{H}$ for (left to right) $B \to \eta K^*(892)$, continuum background, and charmless $B\overline{B}$ background.

J.4 $\eta_{3\pi} K^{*+}_{K^+\pi^0}(892)$

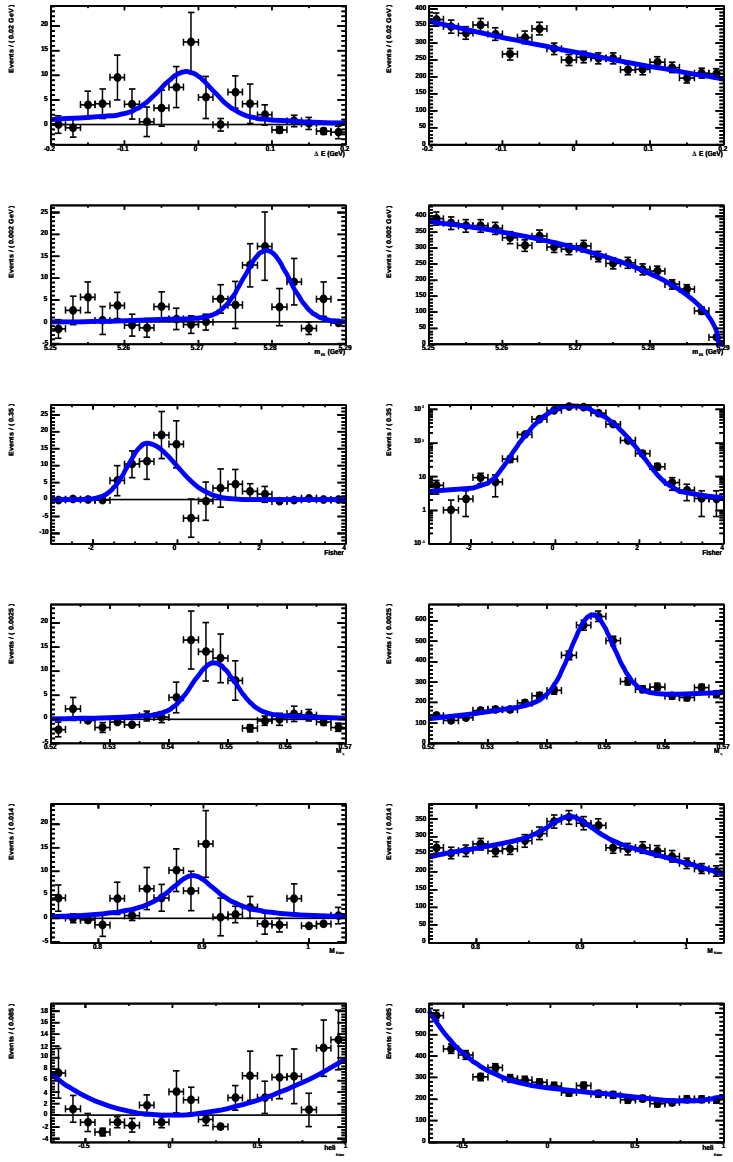


Figure J.4: *sPlots* for (top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η mass, K^* mass and $\mathcal{H}$ for (left to right) $B \to \eta K^*(892)$ and continuum background.

J.5 $\eta_{3\pi} K^{*+}_{K^0_S\pi^+}(892)$

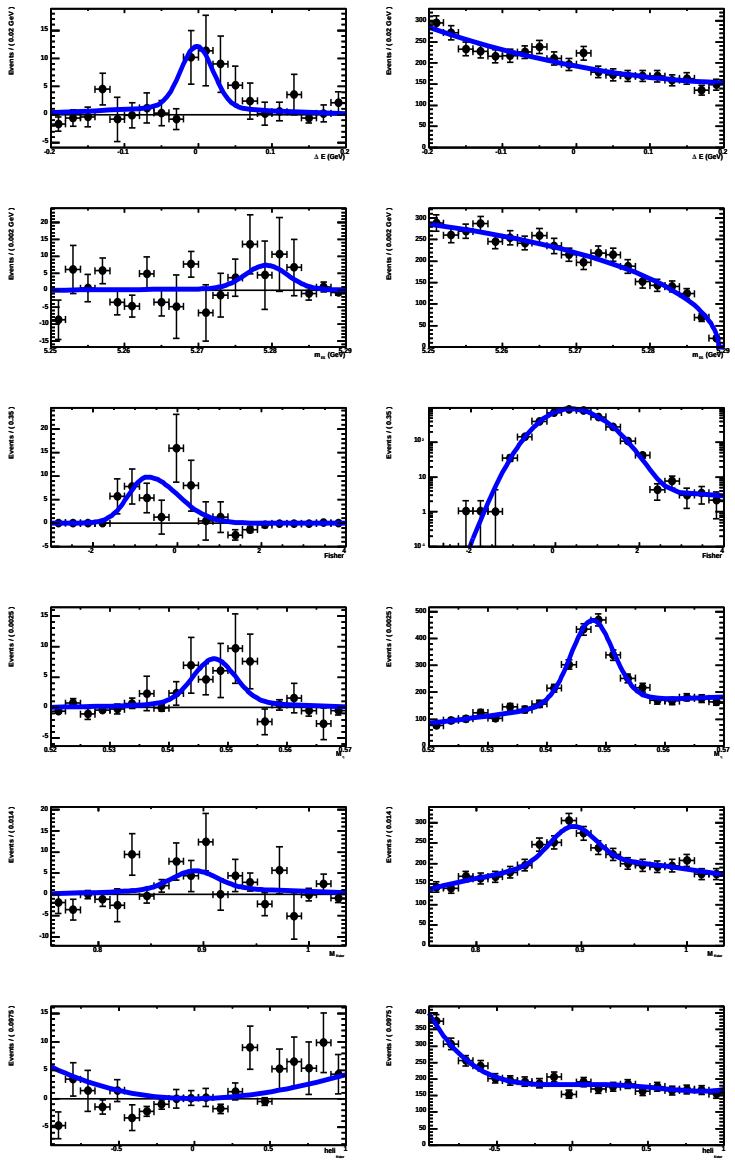


Figure J.5: $s\mathcal{P}lots$ for (top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η mass, K^* mass and $\mathcal{H}$ for (left to right) $B \to \eta K^*(892)$ and continuum background.

J.6 $\eta_{3\pi} K^{*0}_{K^+\pi^-}(892)$

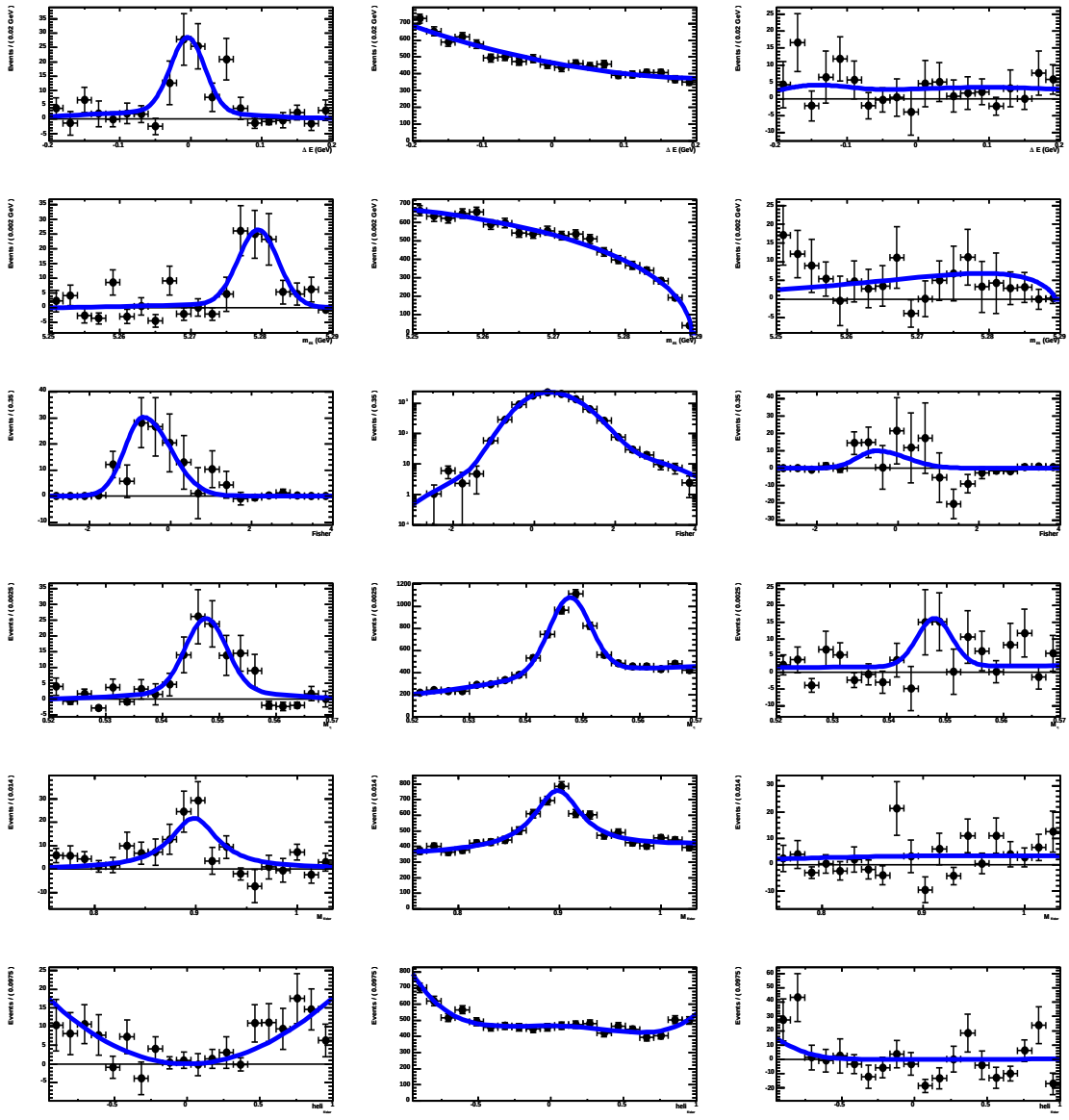


Figure J.6: *sPlots* for (top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η mass, K^* mass and $\mathcal{H}$ for (left to right) $B \to \eta K^*(892)$, continuum background, and charmless $B\overline{B}$ background.

J.7 $\eta_{\gamma\gamma}K^{*+}_{K^+\pi^0}$($S$-wave), ($D$-wave)

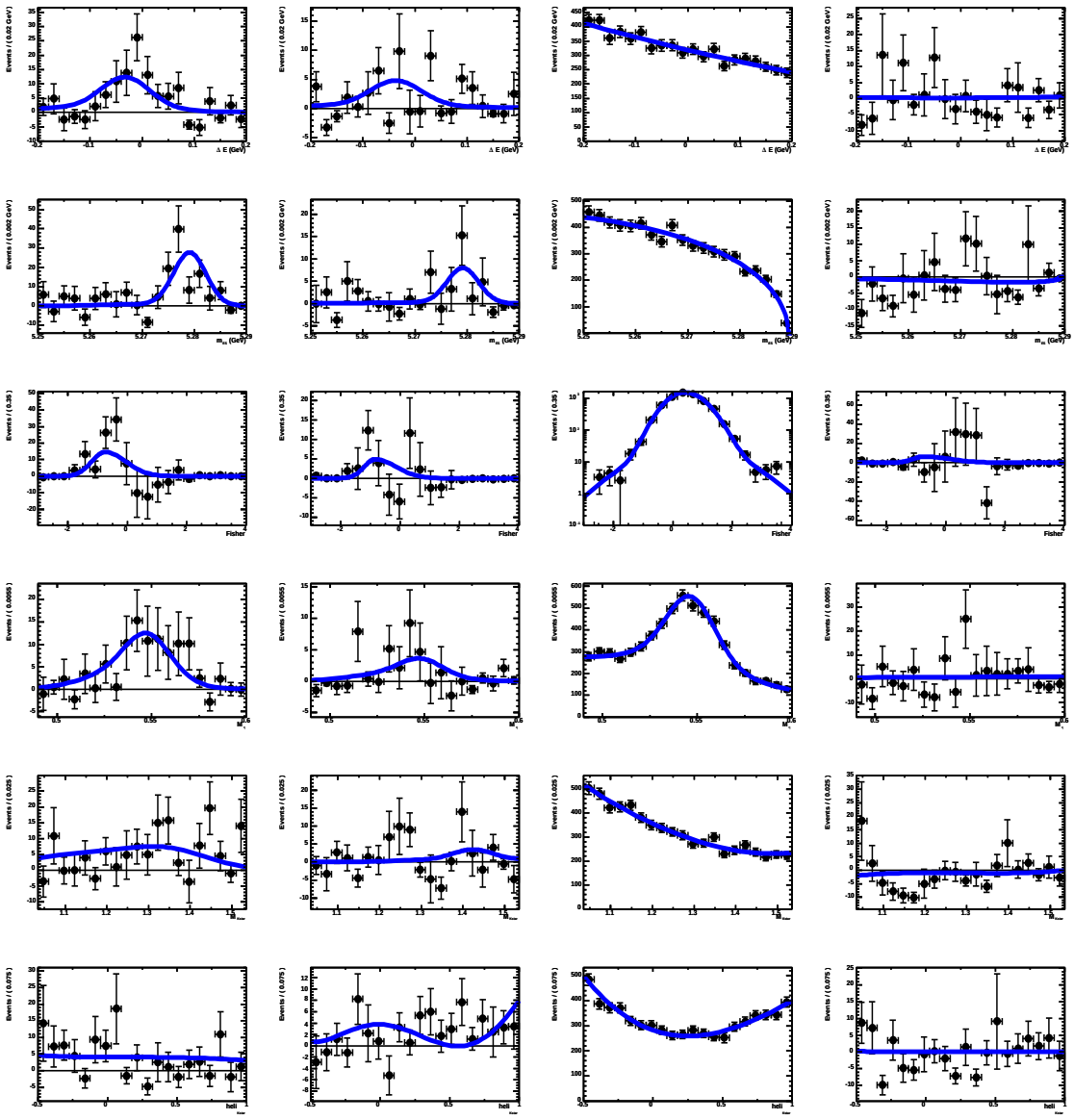


Figure J.7: *sPlots* for (top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η mass, K^* mass and $\mathcal{H}$ for (left to right) $B \to \eta K^*_0(S\text{-wave})$, $B \to \eta K^*_2(1430)$, continuum background, and charmless $B\overline{B}$ background.

J.8 $\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}$($S$-wave), ($D$-wave)

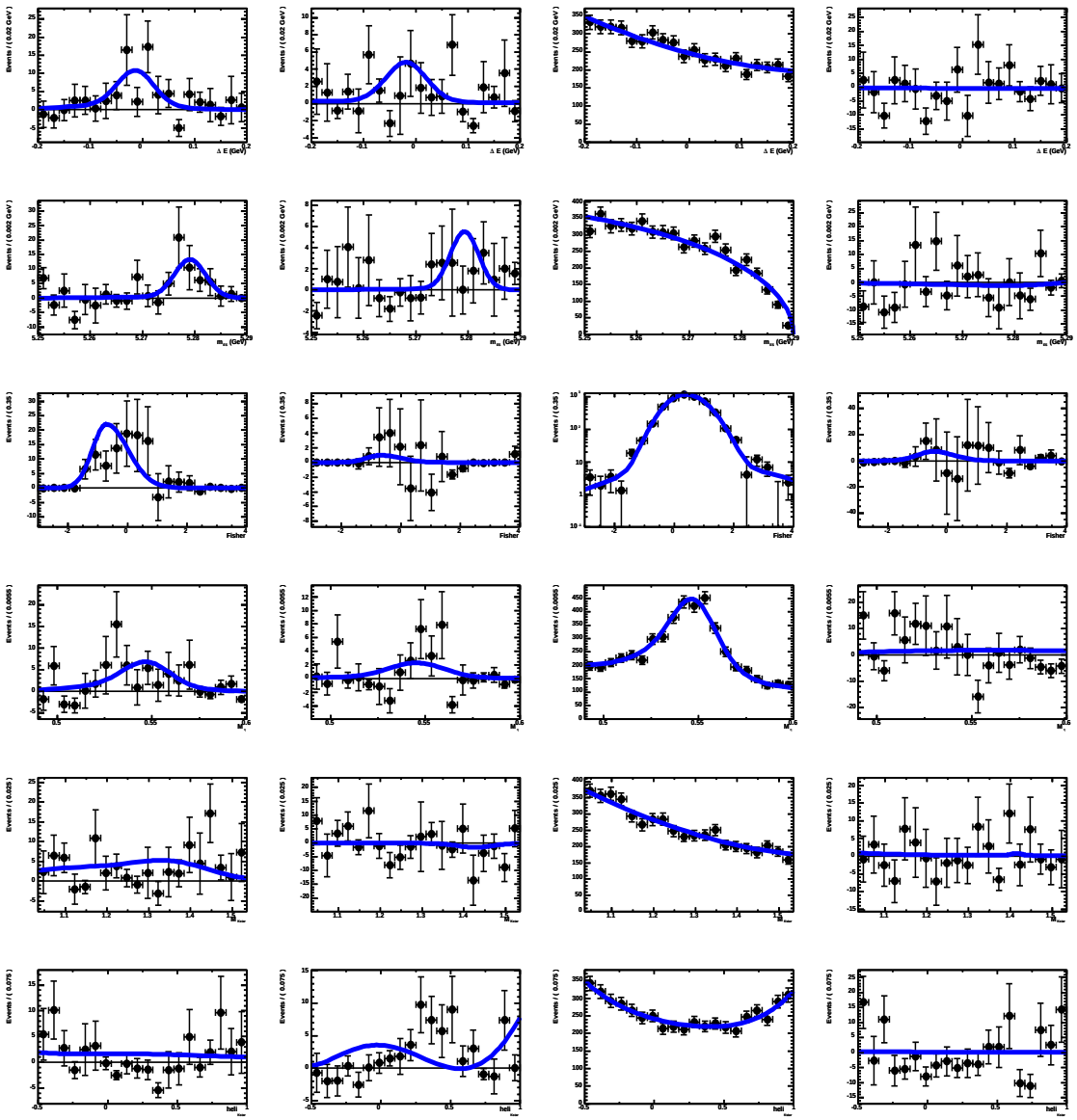


Figure J.8: $s\mathcal{P}lots$ for (top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η mass, K^* mass and $\mathcal{H}$ for (left to right) $B \to \eta K^*_0(S\text{-wave})$, $B \to \eta K^*_2(1430)$, continuum background, and charmless $B\overline{B}$ background.

J.9 $\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}$ (S-wave), (D-wave)

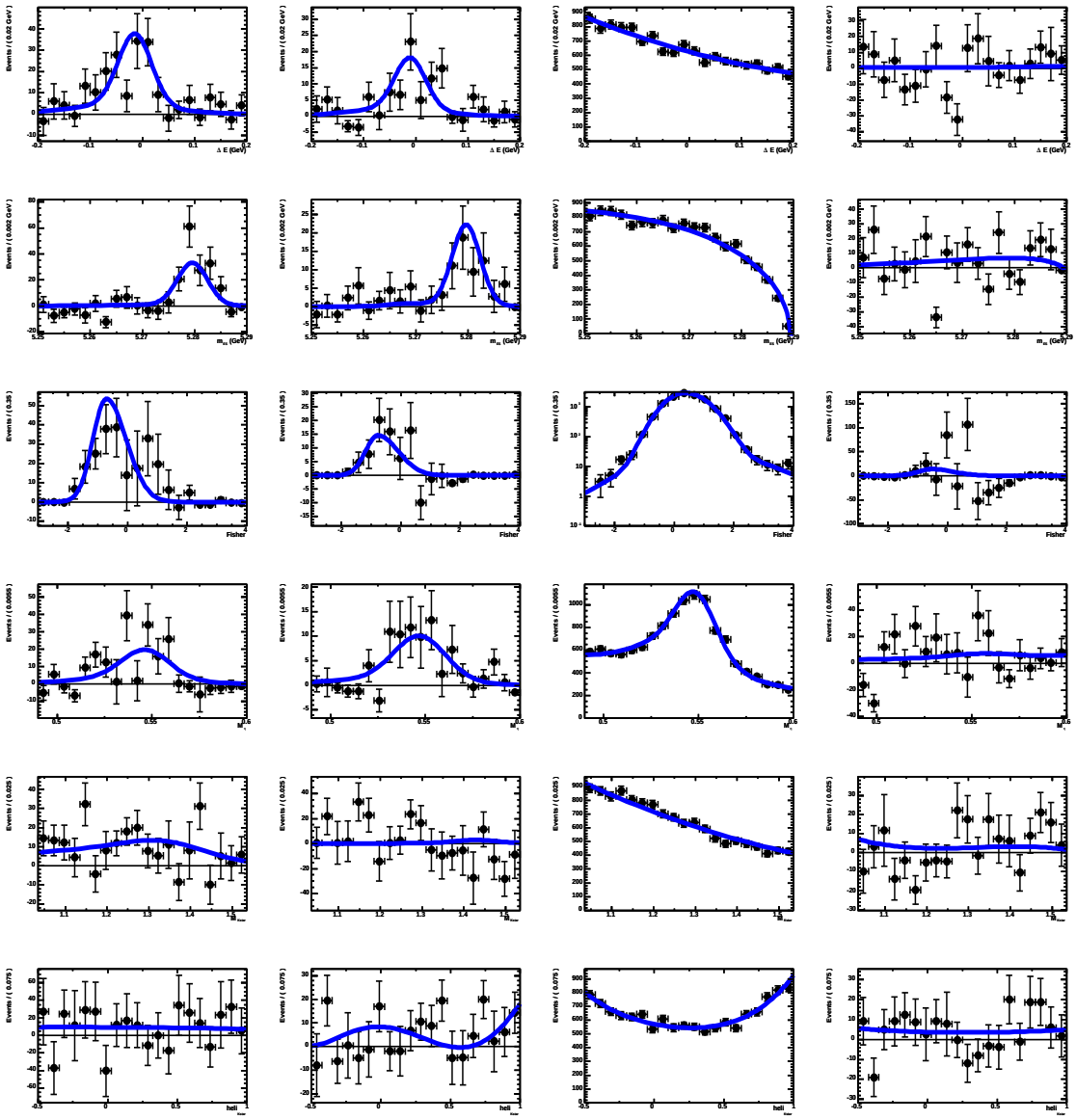


Figure J.9: *sPlots* for (top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η mass, K^* mass and $\mathcal{H}$ for (left to right) $B \to \eta K^*_0(S\text{-wave})$, $B \to \eta K^*_2(1430)$, continuum background, and charmless $B\overline{B}$ background.

J.10 $\eta_{3\pi} K^{*+}_{K^+\pi^0}$ (S-wave), (D-wave)

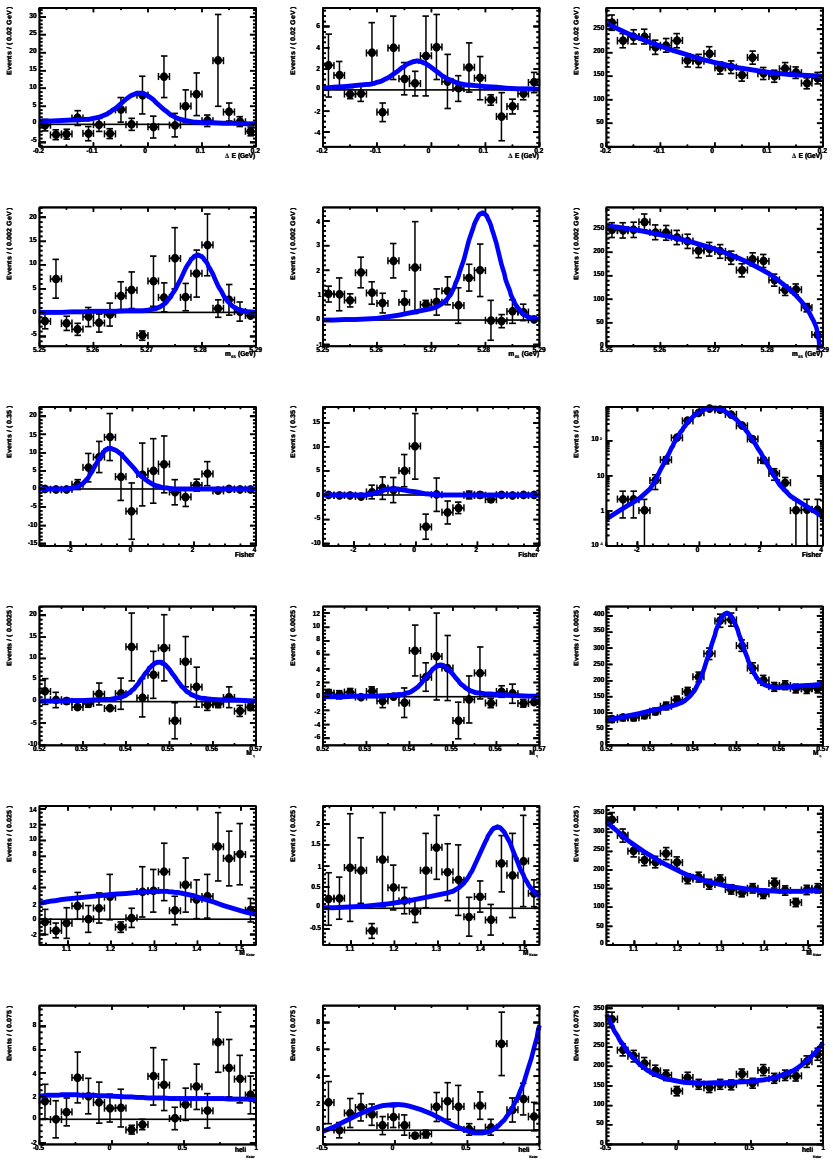


Figure J.10: *sPlots* for (top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η mass, K^* mass and $\mathcal{H}$ for (left to right) $B \to \eta K_0^*(S\text{-wave})$, $B \to \eta K_2^*(1430)$, and continuum background.

J.11 $\eta_{3\pi} K^{*+}_{K^0_S\pi^+}$ (S-wave), (D-wave)

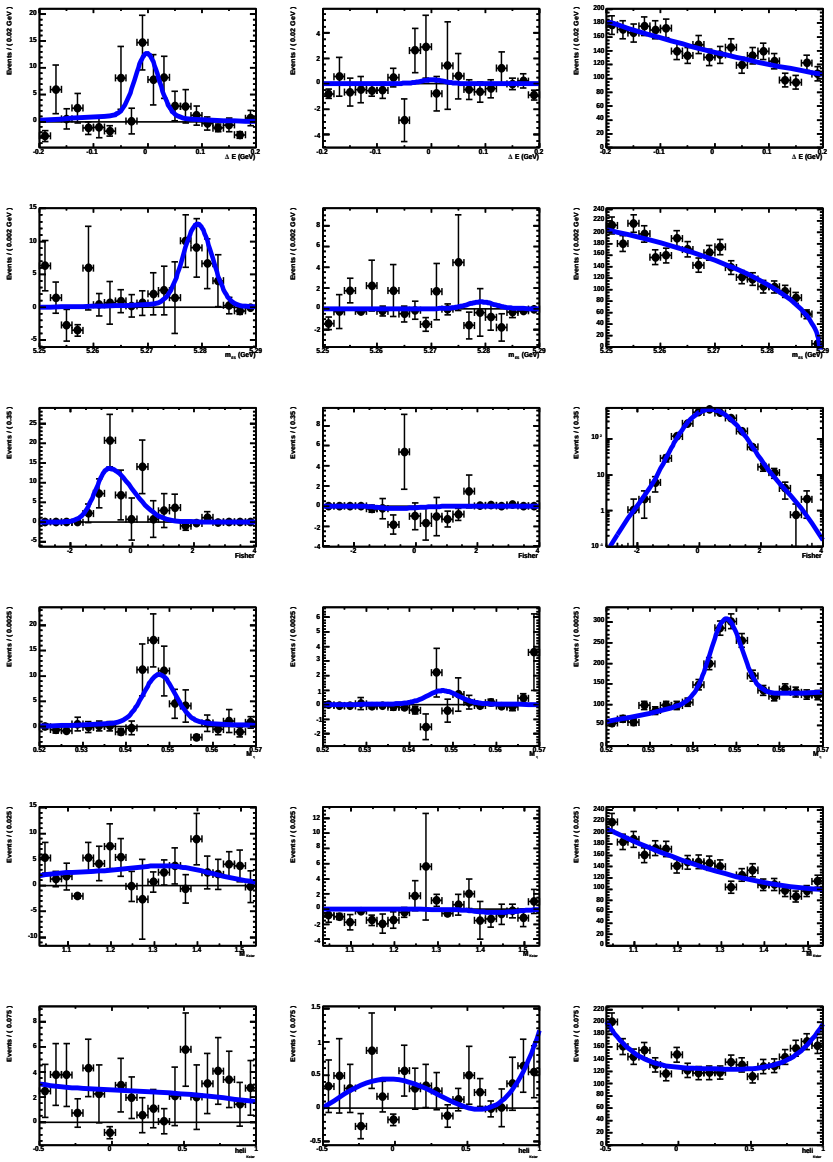


Figure J.11: $s\mathcal{P}lots$ for (top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η mass, K^* mass and $\mathcal{H}$ for (left to right) $B \to \eta K^*_0(S\text{-wave})$, $B \to \eta K^*_2(1430)$, and continuum background.

J.12 $\eta_{3\pi} K^{*0}_{K^+\pi^-}$ (S-wave), (D-wave)

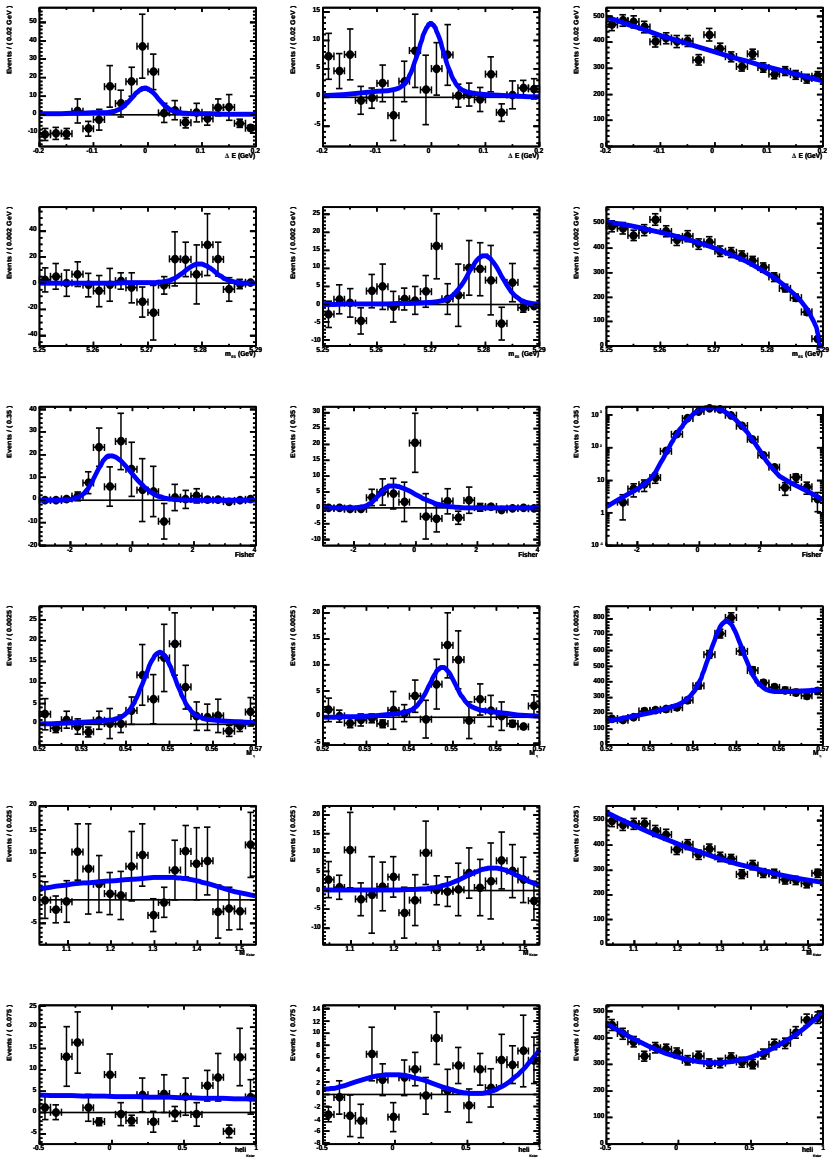


Figure J.12: $s\mathcal{P}lots$ for (top to bottom) ΔE, $m_{\rm ES}$, $\mathcal{F}$, η mass, K^* mass and $\mathcal{H}$ for (left to right) $B \to \eta K_0^*(S\text{-wave})$, $B \to \eta K_2^*(1430)$, and continuum background.

Appendix K

ηK^* : Signal-enhanced projection plots

Signal-enhanced projection plots for the $B \to \eta K^*(892)$, $B \to \eta K_0^*(S\text{-wave})$, and $B \to \eta K_2^*(1430)$ analyses.

K.1 $\eta_{\gamma\gamma} K^{*+}_{K^+\pi^0}(892)$

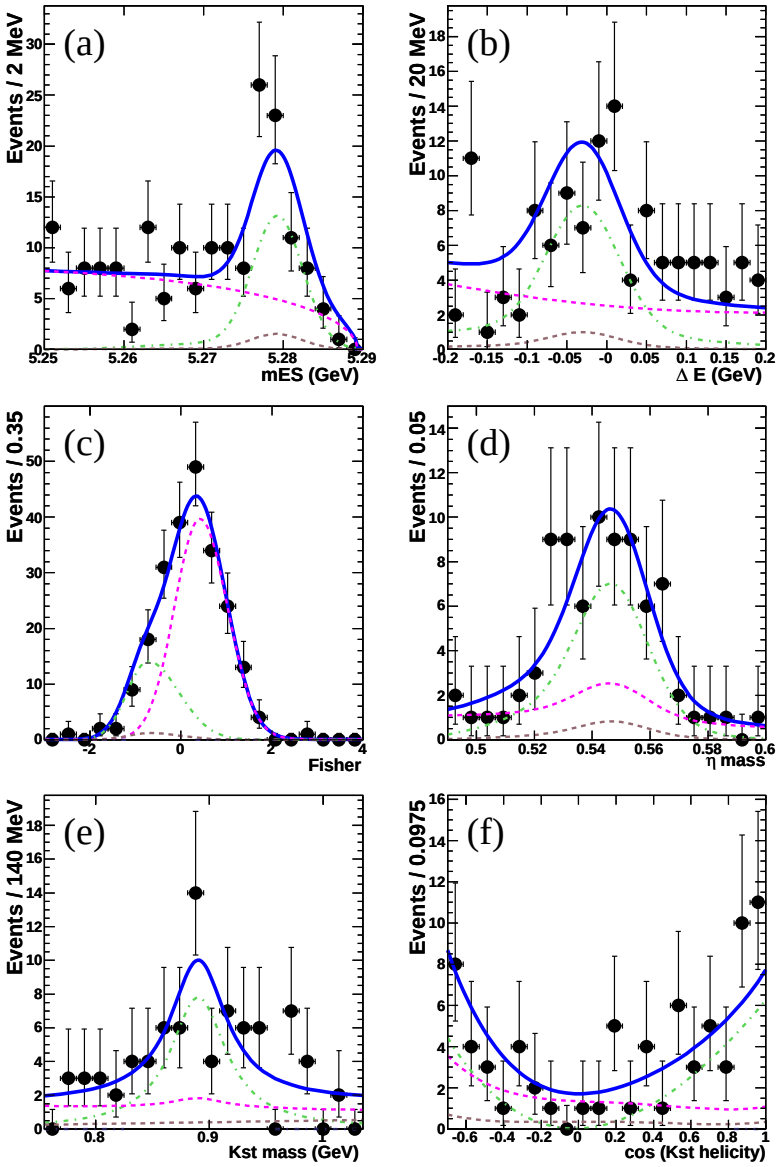


Figure K.1: Signal-enhanced projection plots for (a) $m_{\rm ES}$, (b) ΔE, (c) $\mathcal{F}$, (d) η mass, (e) K^* mass, and (f) $\mathcal{H}$.

K.2 $\eta_{\gamma\gamma} K^{*+}_{K^0_S\pi^+}(892)$

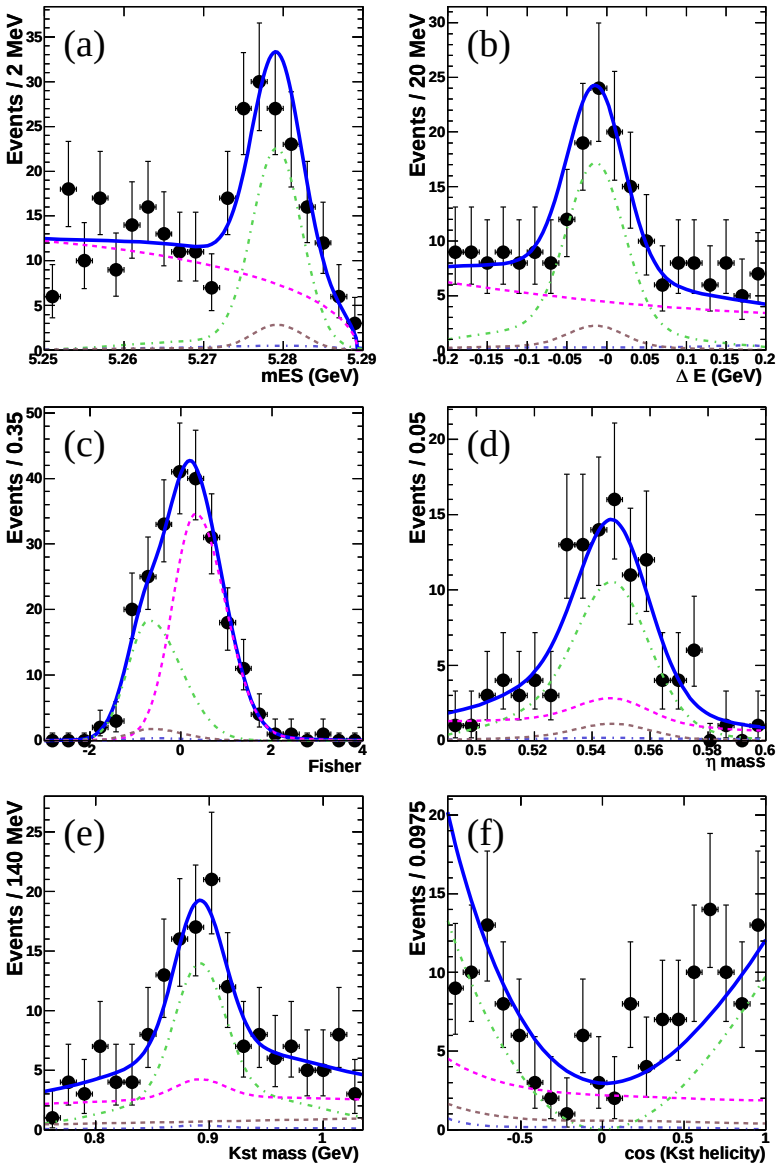


Figure K.2: Signal-enhanced projection plots for (a) $m_{\rm ES}$, (b) ΔE, (c) $\mathcal{F}$, (d) η mass, (e) K^* mass, and (f) $\mathcal{H}$.

K.3 $\eta_{\gamma\gamma} K^{*0}_{K^+\pi^-}(892)$

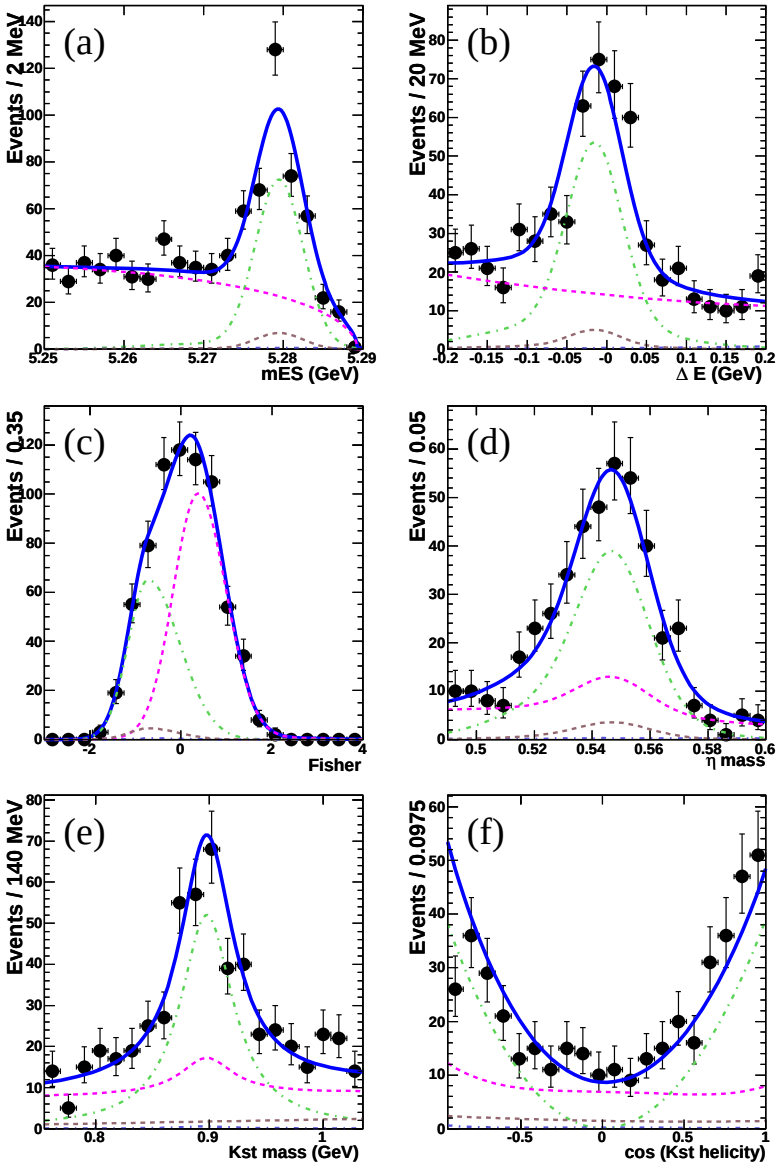


Figure K.3: Signal-enhanced projection plots for (a) $m_{\rm ES}$, (b) ΔE, (c) $\mathcal{F}$, (d) η mass, (e) K^* mass, and (f) $\mathcal{H}$.

K.4 $\eta_{3\pi} K^{*+}_{K^+\pi^0}(892)$

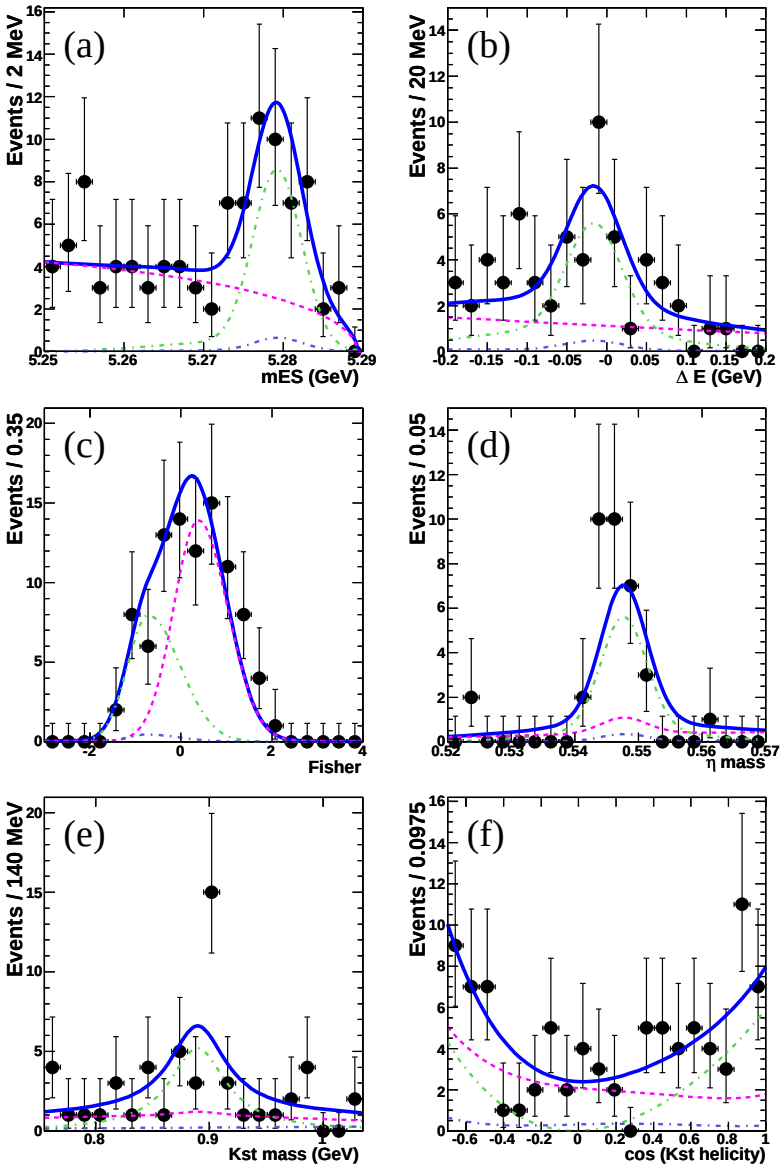


Figure K.4: Signal-enhanced projection plots for (a) $m_{\rm ES}$, (b) ΔE, (c) $\mathcal{F}$, (d) η mass, (e) K^* mass, and (f) $\mathcal{H}$.

K.5 $\eta_{3\pi} K^{*+}_{K^0_S\pi^+}(892)$

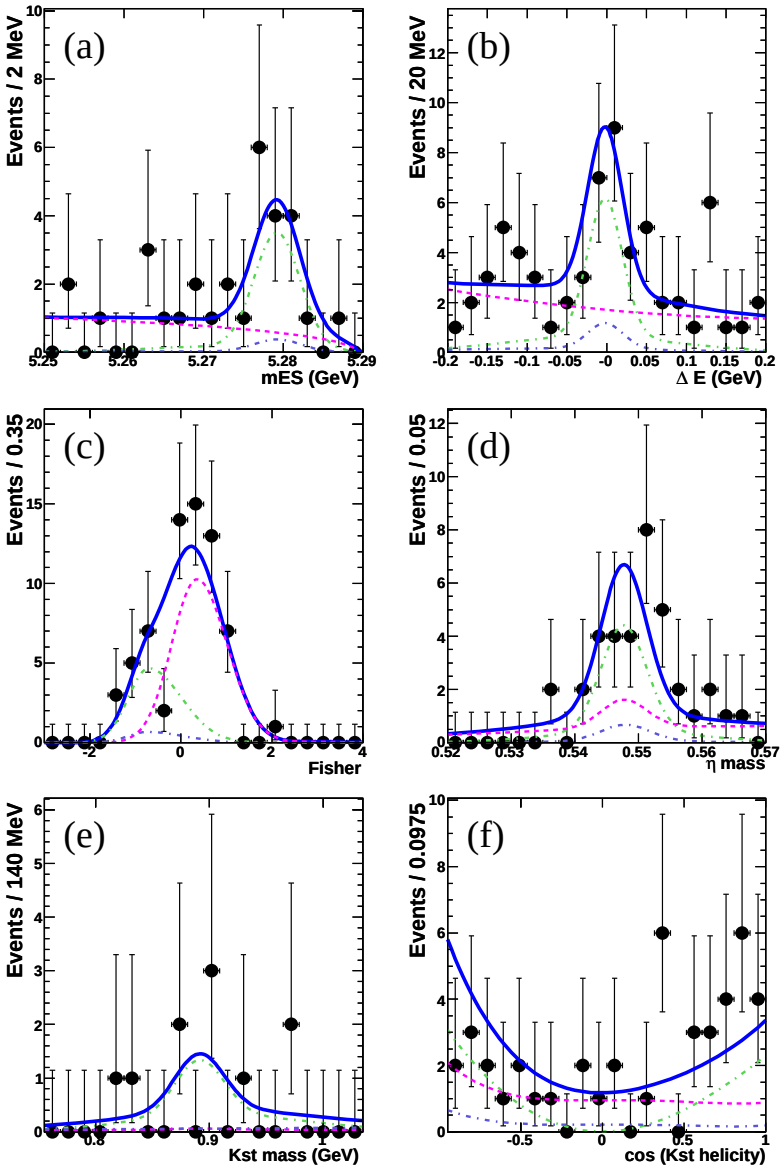


Figure K.5: Signal-enhanced projection plots for (a) $m_{\rm ES}$, (b) ΔE, (c) $\mathcal{F}$, (d) η mass, (e) K^* mass, and (f) $\mathcal{H}$.

K.6 $\eta_{3\pi} K^{*0}_{K^+\pi^-}(892)$

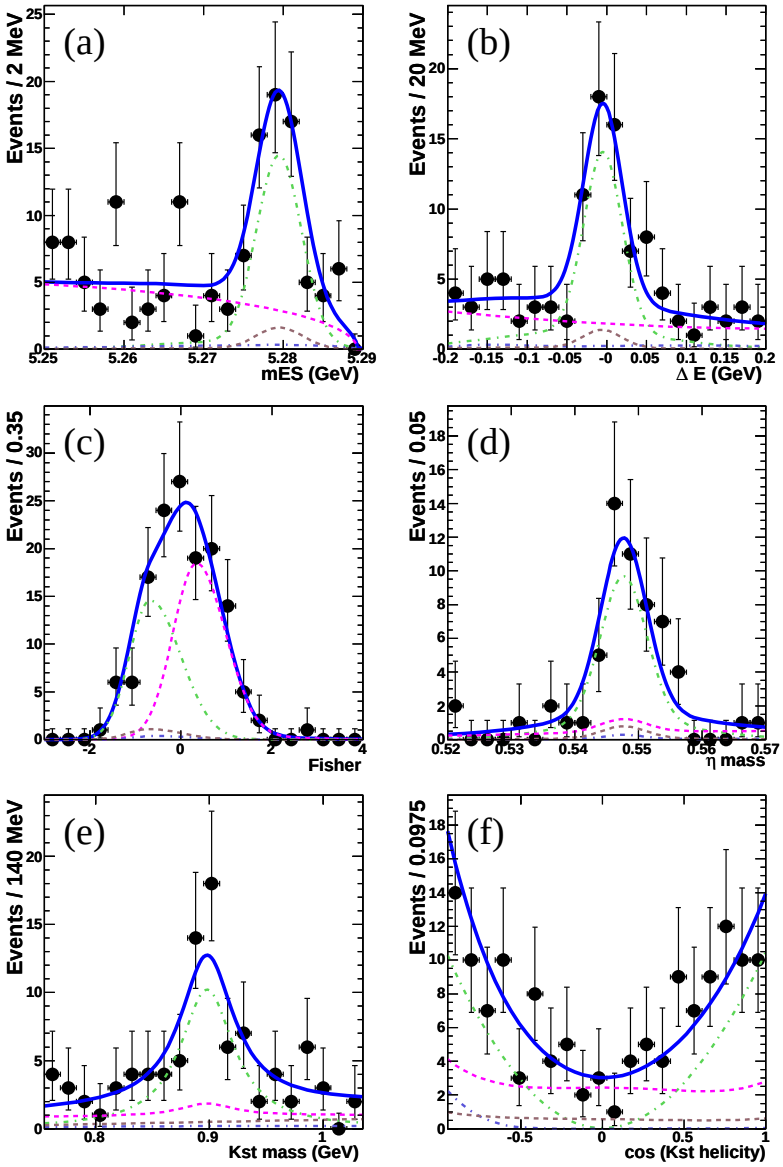


Figure K.6: Signal-enhanced projection plots for (a) $m_{\rm ES}$, (b) ΔE, (c) $\mathcal{F}$, (d) η mass, (e) K^* mass, and (f) $\mathcal{H}$.

K.7 $\eta_{\gamma\gamma} K^{*+}_{K^+\pi^0}$ (S-wave)

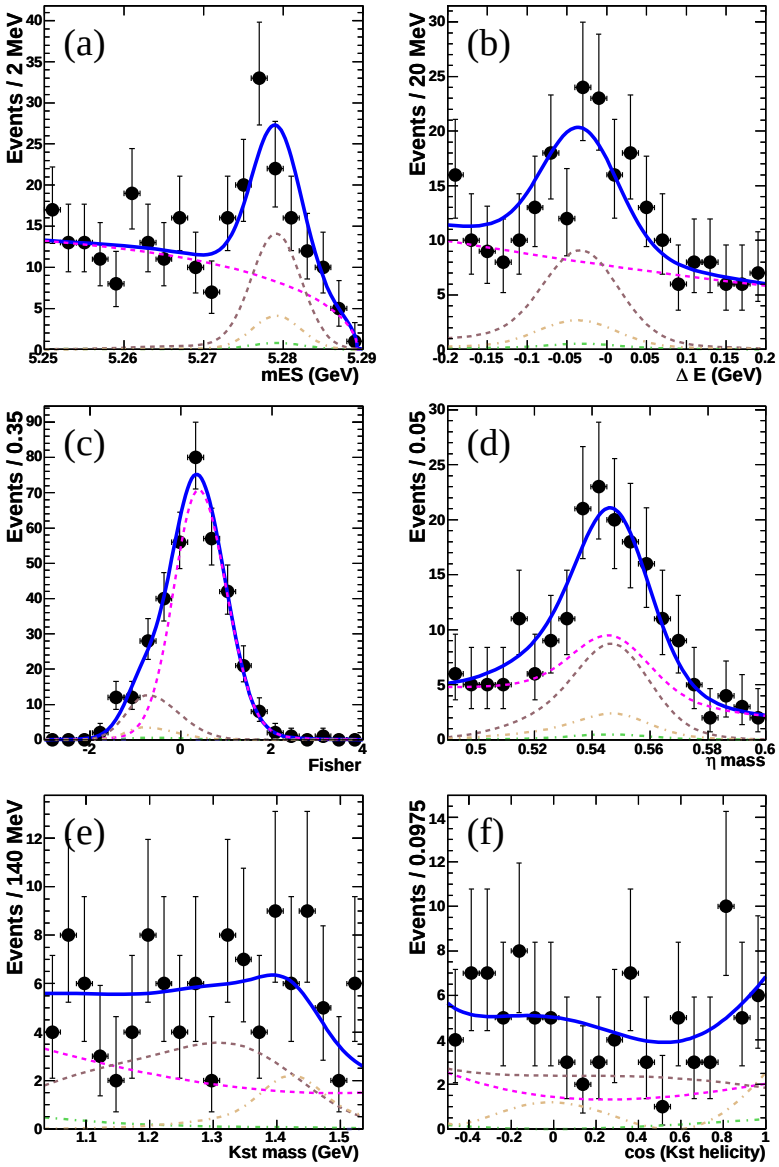


Figure K.7: Signal-enhanced projection plots for (a) $m_{\rm ES}$, (b) ΔE, (c) $\mathcal{F}$, (d) η mass, (e) K^* mass, and (f) $\mathcal{H}$.

K.8 $\eta_{\gamma\gamma} K^{*+}_{K^0_S\pi^+}$ (S-wave)

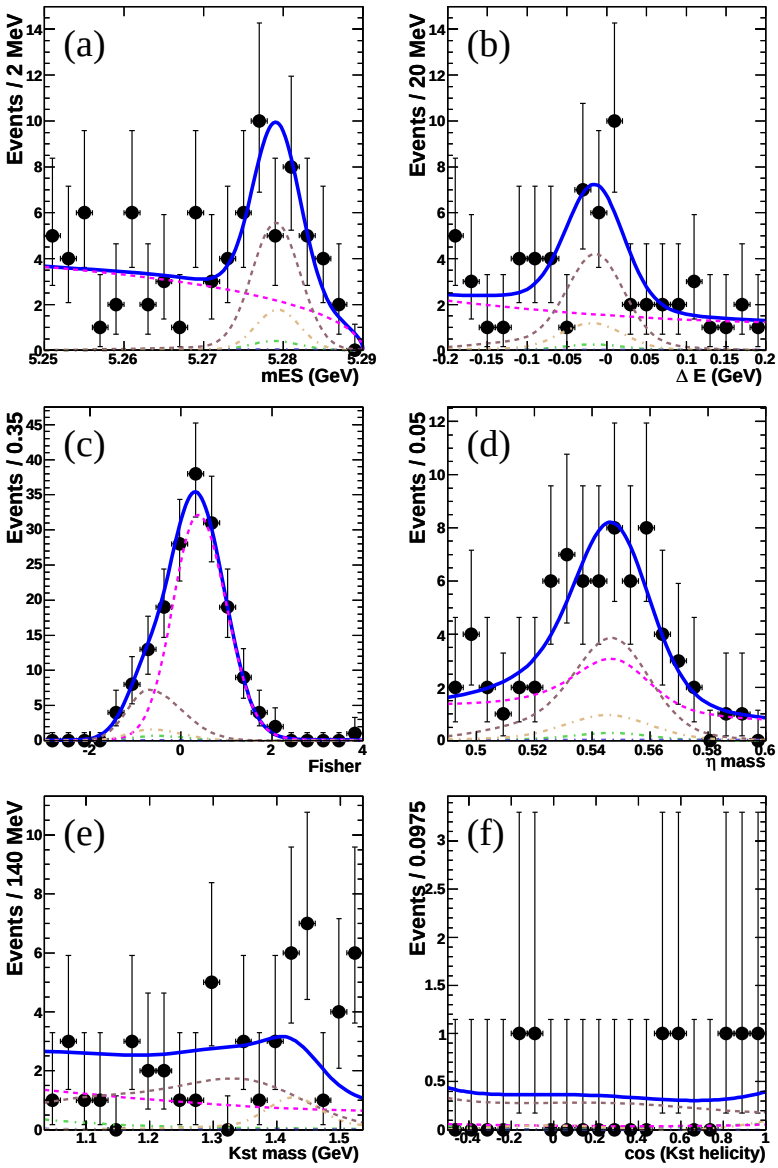


Figure K.8: Signal-enhanced projection plots for (a) $m_{\rm ES}$, (b) ΔE, (c) $\mathcal{F}$, (d) η mass, (e) K^* mass, and (f) $\mathcal{H}$.

K.9 $\eta_{\gamma\gamma} K^{*0}_{K^+\pi^-}$ (S-wave)

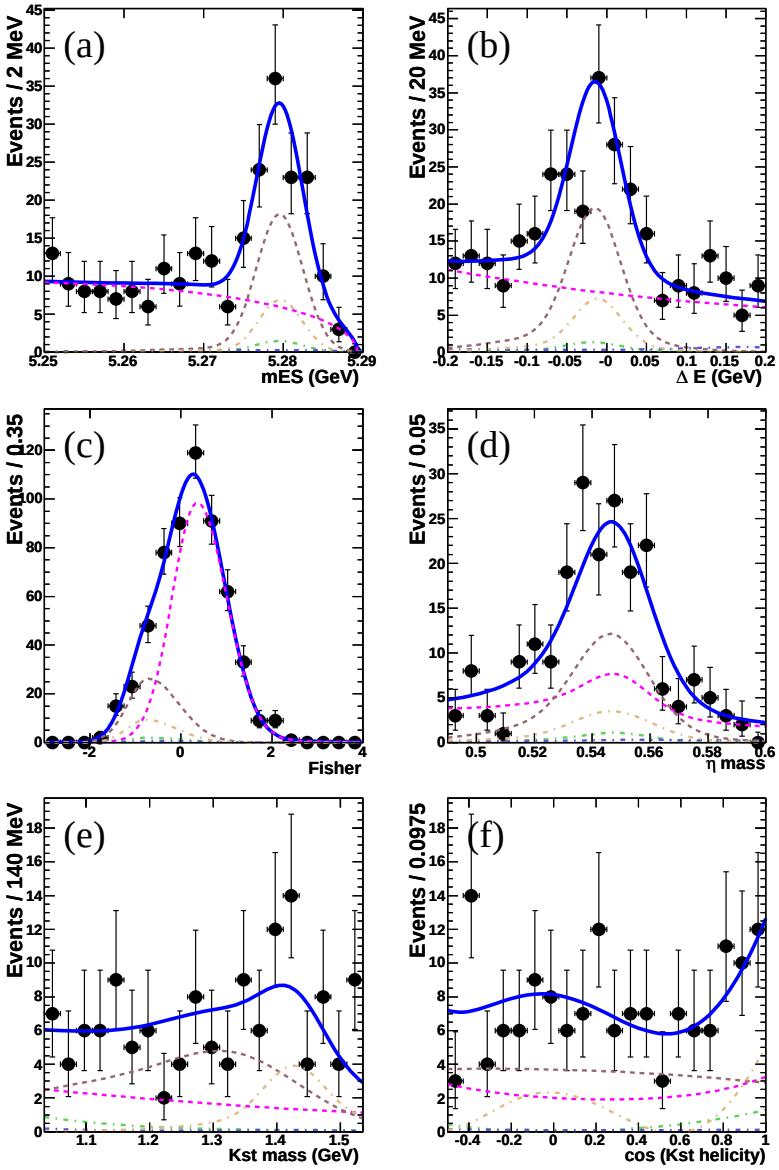


Figure K.9: Signal-enhanced projection plots for (a) $m_{\rm ES}$, (b) ΔE, (c) $\mathcal{F}$, (d) η mass, (e) K^* mass, and (f) $\mathcal{H}$.

K.10 $\eta_{3\pi} K^{*+}_{K^+\pi^0}$ (S-wave)

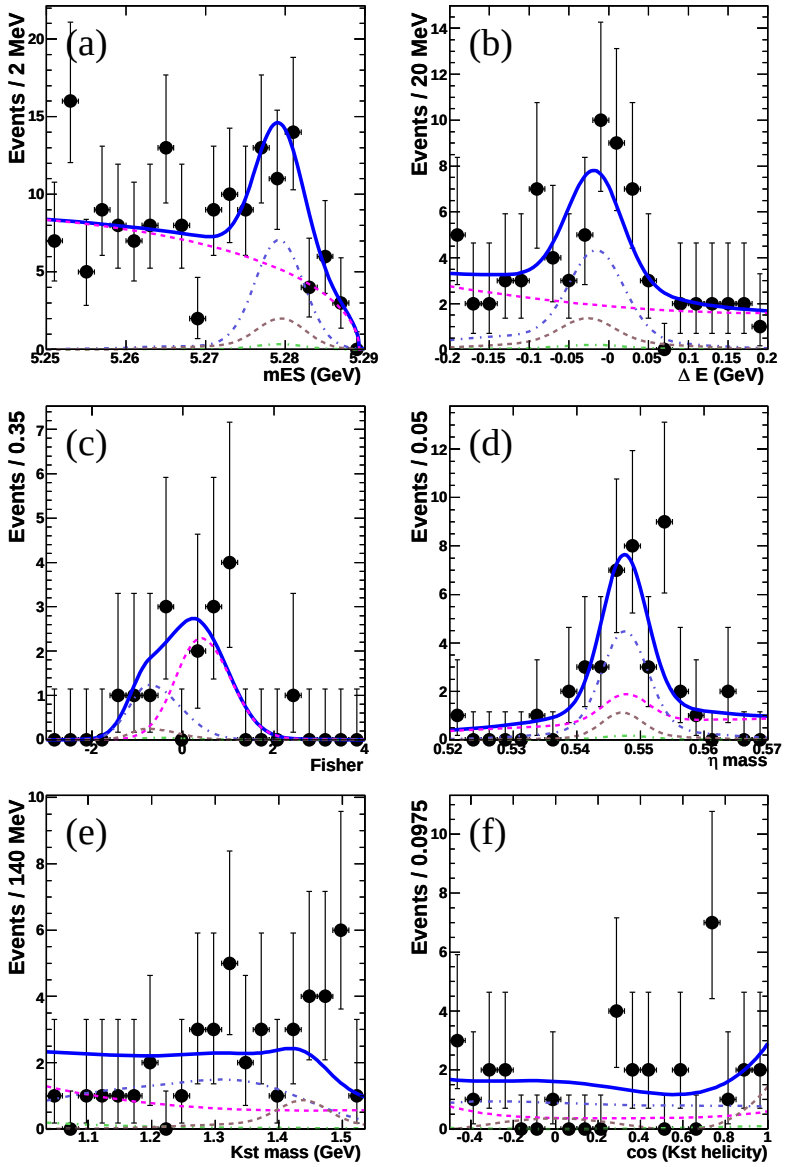


Figure K.10: Signal-enhanced projection plots for (a) $m_{\rm ES}$, (b) ΔE, (c) $\mathcal{F}$, (d) η mass, (e) K^* mass, and (f) $\mathcal{H}$.

K.11 $\eta_{3\pi} K^{*+}_{K^0_S\pi^+}$ (S-wave)

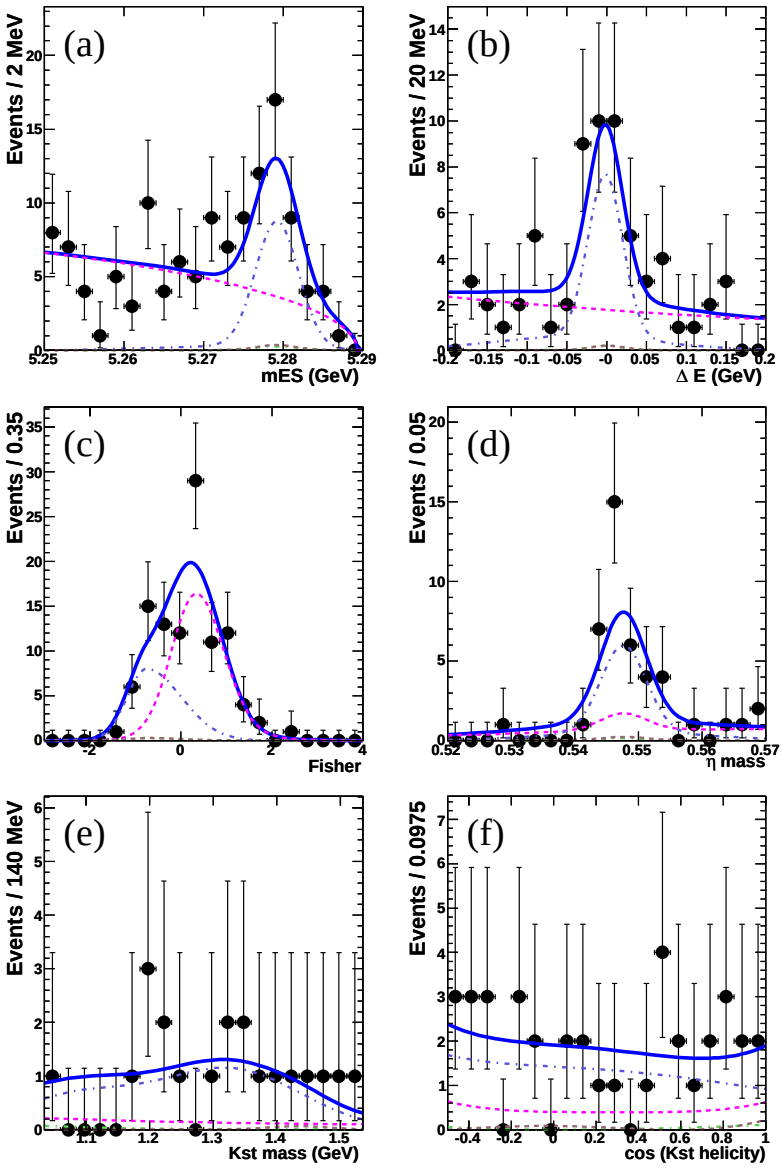


Figure K.11: Signal-enhanced projection plots for (a) $m_{\rm ES}$, (b) ΔE, (c) $\mathcal{F}$, (d) η mass, (e) K^* mass, and (f) $\mathcal{H}$.

K.12 $\eta_{3\pi} K^{*0}_{K^+\pi^-}$ (S-wave)

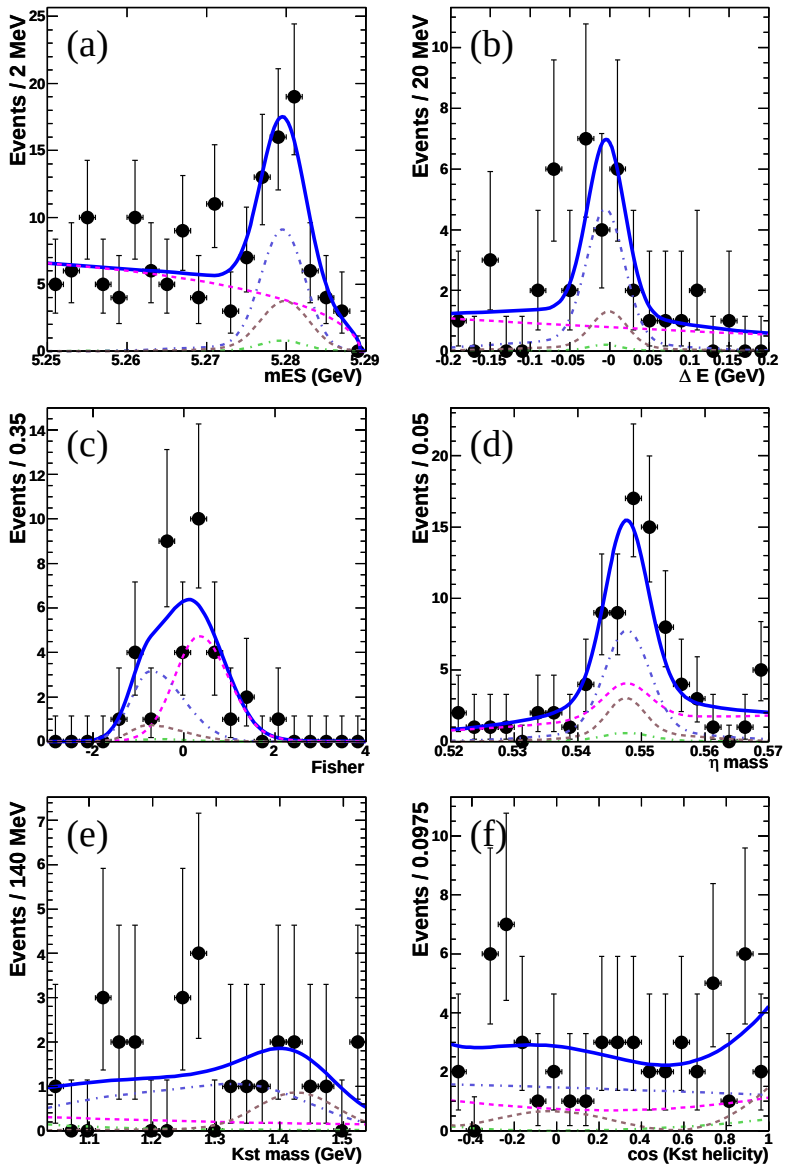


Figure K.12: Signal-enhanced projection plots for (a) $m_{\rm ES}$, (b) ΔE, (c) $\mathcal{F}$, (d) η mass, (e) K^* mass, and (f) $\mathcal{H}$.

K.13 $\eta_{\gamma\gamma} K_2^{*+}(K^+\pi^0)$

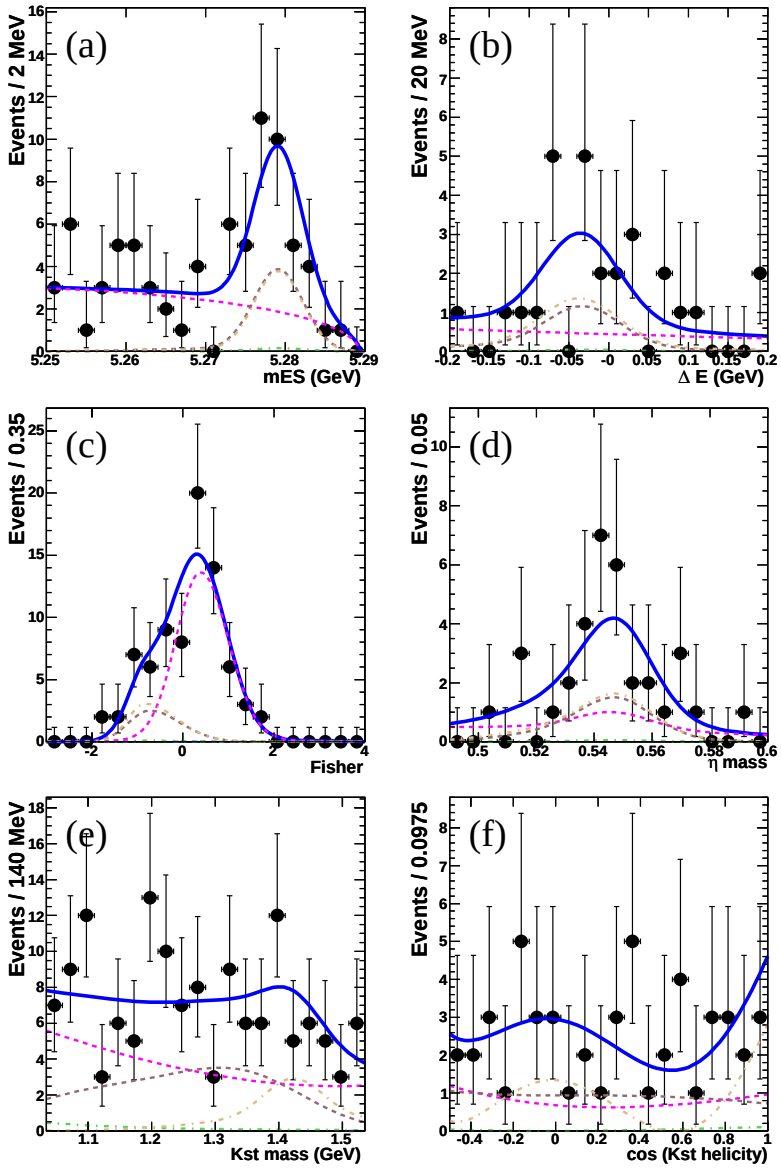


Figure K.13: Signal-enhanced projection plots for (a) $m_{\rm ES}$, (b) ΔE, (c) $\mathcal{F}$, (d) η mass, (e) K^* mass, and (f) $\mathcal{H}$.

K.14 $\eta_{\gamma\gamma} K_2^{*+}(K_S^0\pi^+)$

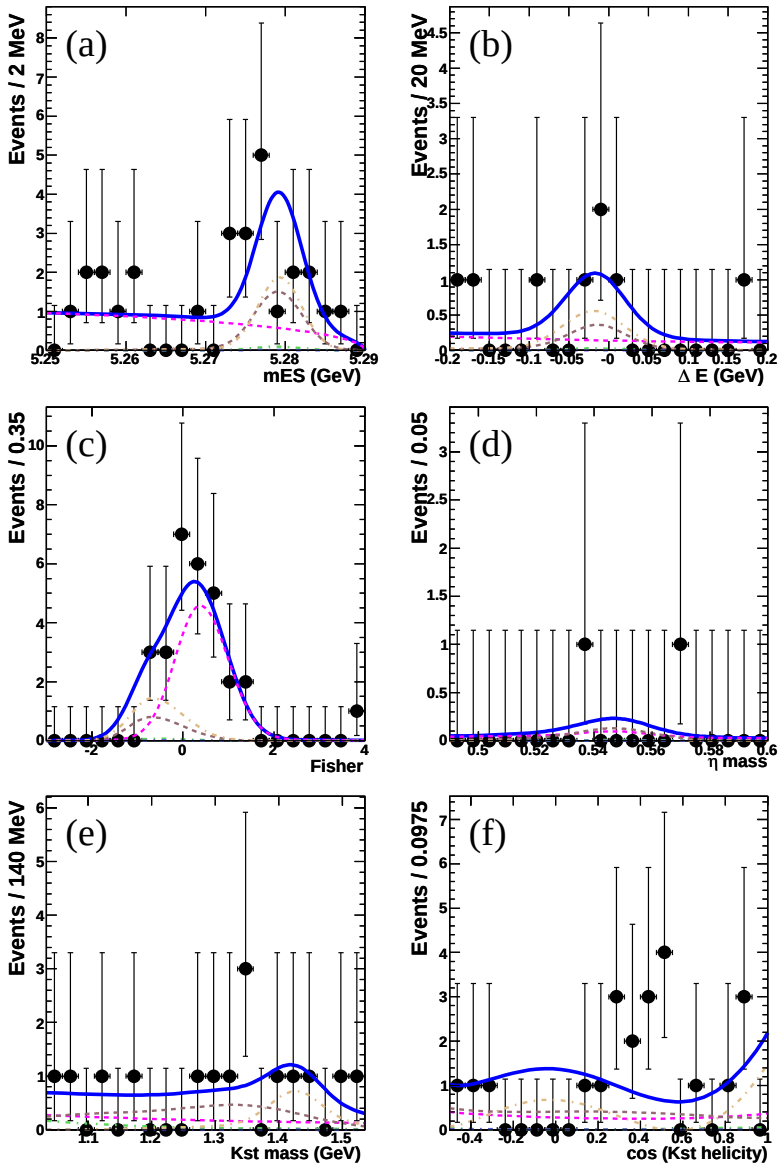


Figure K.14: Signal-enhanced projection plots for (a) $m_{\rm ES}$, (b) ΔE, (c) $\mathcal{F}$, (d) η mass, (e) K^* mass, and (f) $\mathcal{H}$.

K.15 $\eta_{\gamma\gamma} K_2^{*0}(K^+\pi^-)$

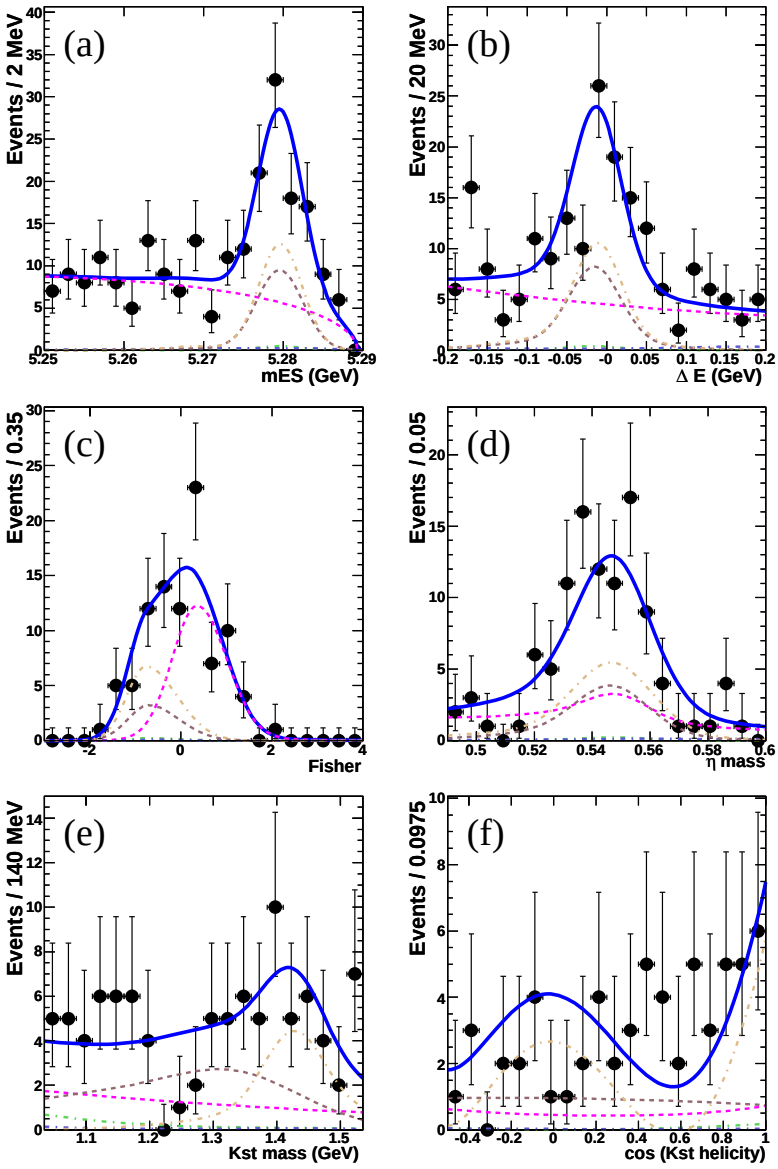


Figure K.15: Signal-enhanced projection plots for (a) $m_{\rm ES}$, (b) ΔE, (c) $\mathcal{F}$, (d) η mass, (e) K^* mass, and (f) $\mathcal{H}$.

K.16 $\eta_{3\pi} K_2^{*+}(K^+\pi^0)$

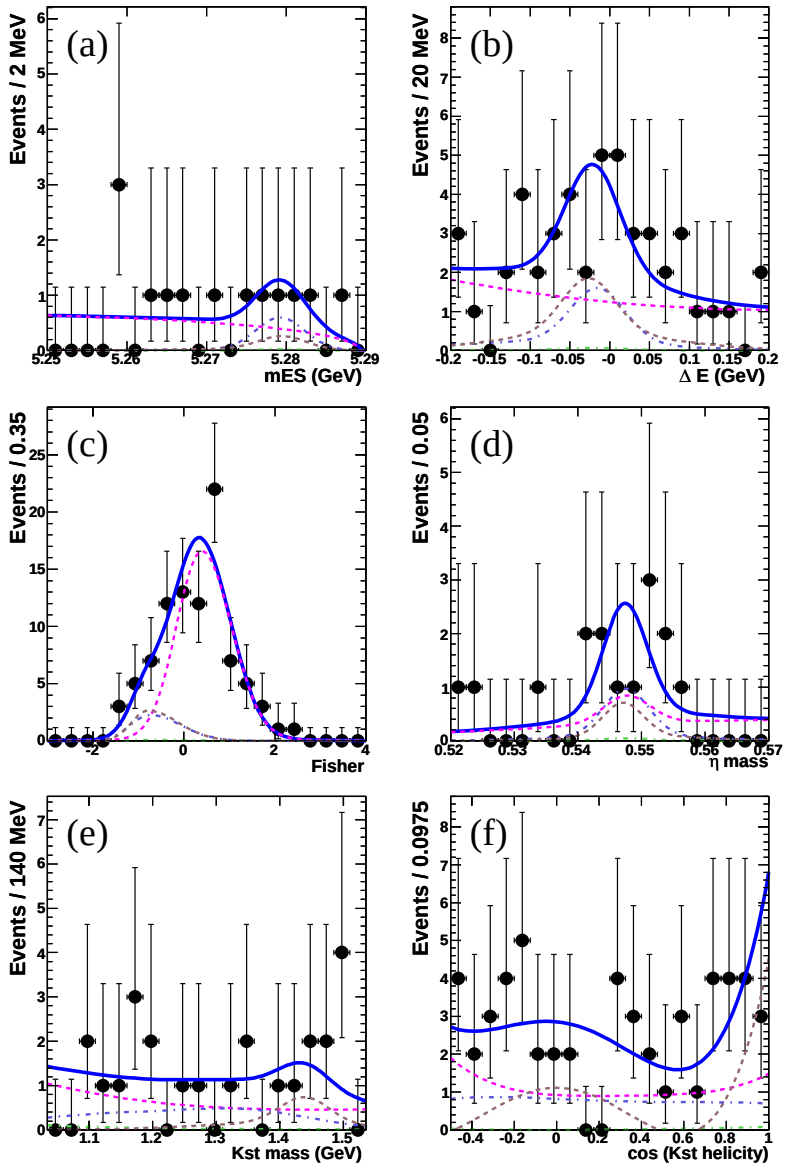


Figure K.16: Signal-enhanced projection plots for (a) $m_{\rm ES}$, (b) ΔE, (c) $\mathcal{F}$, (d) η mass, (e) K^* mass, and (f) $\mathcal{H}$.

K.17 $\eta_{3\pi} K_2^{*+}(K_S^0\pi^+)$

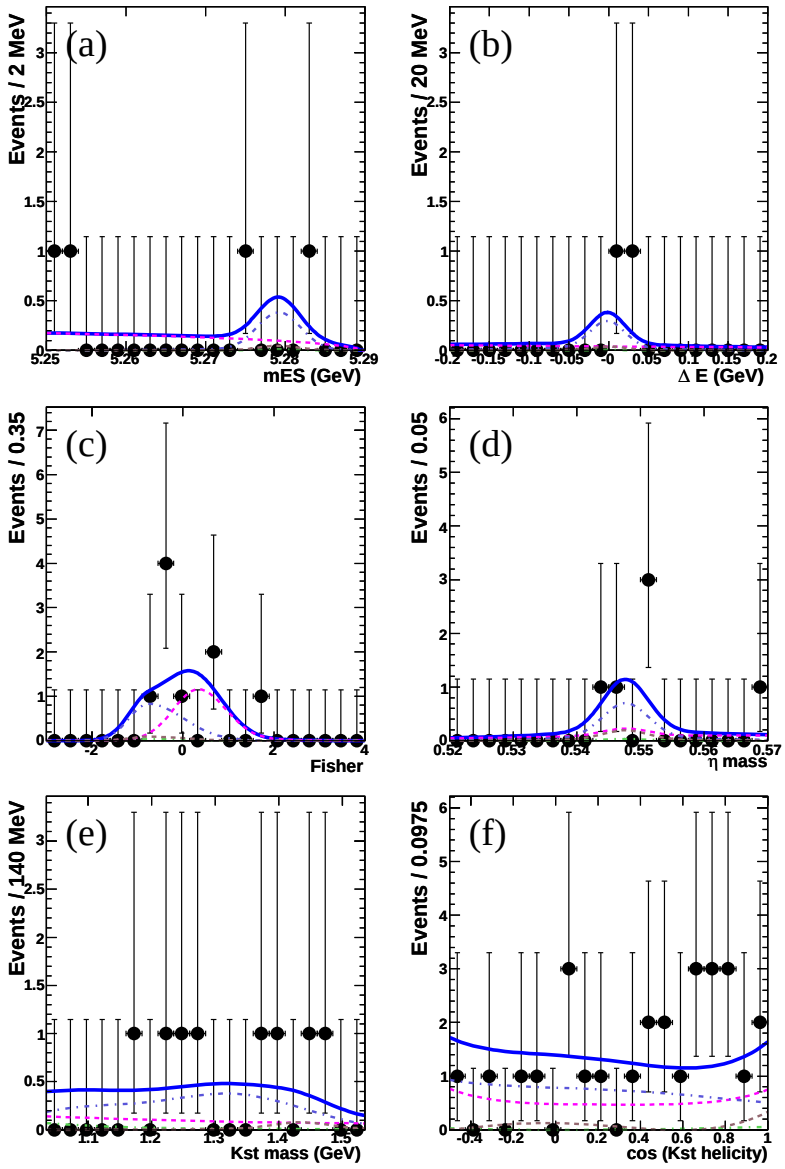


Figure K.17: Signal-enhanced projection plots for (a) $m_{\rm ES}$, (b) ΔE, (c) $\mathcal{F}$, (d) η mass, (e) K^* mass, and (f) $\mathcal{H}$.

K.18 $\eta_{3\pi} K_2^{*0}(K^+\pi^-)$

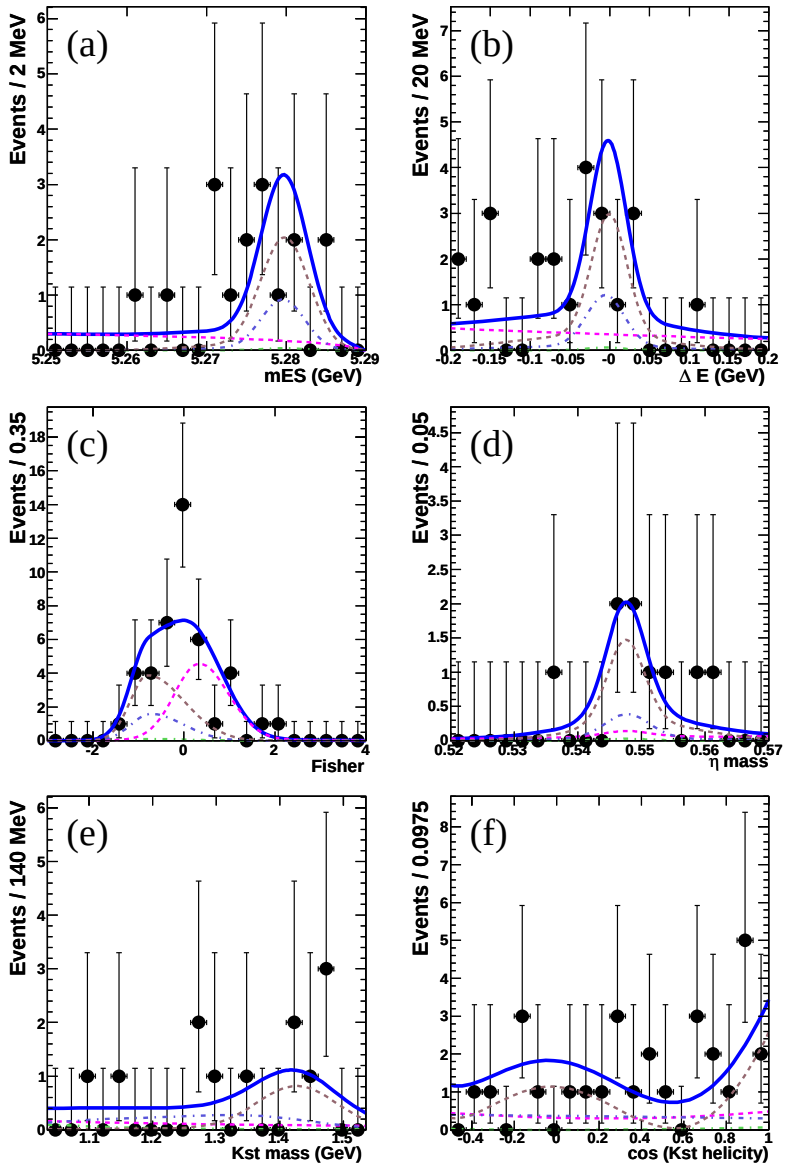


Figure K.18: Signal-enhanced projection plots for (a) $m_{\rm ES}$, (b) ΔE, (c) $\mathcal{F}$, (d) η mass, (e) K^* mass, and (f) $\mathcal{H}$.

Appendix L

Including partial wave interference in the fit

To check the reasonableness of the interference systematics quoted in our publication (Sec. 6.13), we measure the phase differences $\delta_0^{(\prime)}$ and branching fraction using an unfactorized, two-dimensional PDF in $K\pi$ mass and helicity modeled with Eqs. 6.16 and 6.17.

L.0.1 Mass-dependence of complex amplitudes

The mass-dependent amplitude for each partial wave can be written in terms of its phase δ_J, recall from Sec. 6.3 that $\delta_S = \delta_R + \delta_B$,

$$A_{BW_1}(m) = \frac{1}{\cot\delta_P - i} = \sin\delta_P e^{i\delta_P}, \tag{L.1}$$

$$A_{BW_2}(m) = \frac{1}{\cot\delta_D - i} = \sin\delta_D e^{i\delta_D}, \tag{L.2}$$

$$A_{LASS}(m) = \frac{1}{\cot(\delta_R+\delta_B) - i} \equiv \sin\delta_S e^{i\delta_S}. \tag{L.3}$$

The mass dependence of the P-wave, D-wave, and the resonant S-wave phases are described by a Breit-Wigner, and the non-resonant S-wave phase with an effective range,

$$\cot\delta_D = \frac{m^2_{K^*_2} - m^2}{m_{K^*_2}\Gamma_2(m)} \quad \cot\delta_P = \frac{m^2_{K^*} - m^2}{m_{K^*}\Gamma_1(m)}$$
$$\cot\delta_R = \frac{m^2_{K^*_0} - m^2}{m_{K^*_0}\Gamma_0(m)} \quad \cot\delta_B = \frac{1}{aq} + \frac{1}{2}rq,$$

where q is the center-of-mass momentum, and $\Gamma_J(m)$ is the mass-dependent width of a J-spin Breit-Wigner, which can be written in terms of the mass and width of the

resonance (m_{K^*}, Γ_{K^*}), the interaction radius r, and q evaluated at the resonance mass q_0,

$$\Gamma_0(m) = \Gamma_{K_0^*} \frac{m_{K_0^*}}{m} \frac{q}{q_0}, \tag{L.4}$$

$$\Gamma_1(m) = \Gamma_{K^*} \frac{m_{K^*}}{m} \frac{1 + r^2 q_0^2}{1 + r^2 q^2} [\frac{q}{q_0}]^3, \tag{L.5}$$

$$\Gamma_2(m) = \Gamma_{K_2^*} \frac{m_{K_2^*}}{m} \frac{9 + 3r^2 q_0^2 + r^4 q_0^4}{9 + 3r^2 q^2 + r^4 q^4} [\frac{q}{q_0}]^5. \tag{L.6}$$

L.0.2 Acceptance functions

Since the PDFs are normalized, the acceptance is only important to the degree that it varies over $m_{K\pi}$ or $\mathcal{H}$. Using MC, we observe that the acceptance does vary with $K\pi$ mass and helicity for all modes and partial waves. The divergence of the observed helicity distributions from the expected distributions is primarily due to low reconstruction efficiency for slow pions, which are near $\mathcal{H} = -1$. Because the LASS shape is very wide (300 MeV), the observed S-wave mass distribution differs from the expected distribution because events outside the $m_{K\pi}$ window are misreconstructed inside the mass window. For the more narrow P-wave and D-wave, there are very few events outside the mass window, and so this misreconstruction is minimal.

The acceptance effects are different for each sub-mode, so there are 48 different helicity and mass acceptance functions, $\mathrm{accH}_i^J(\mathcal{H})$ and $\mathrm{accM}_i^J(m)$, one for each of six sub-modes i of each of two interfering partial waves J in each of two mass ranges. We use a simple one-dimensional fit to obtain the reconstructed helicity and mass distributions from MC, modeling the helicity distributions with a polynomial and a Fermi-Dirac roll-off function and the mass distributions with a Breit-Wigner plus Gaussian.

L.0.3 Normalization of PDFs

The first two terms of each PDF (Eqs. 6.16 and 6.17) are normalized by hand to unity. The normalization of the third term depends on the values of $\delta_0^{(\prime)}$ and f. We obtain 24 normalization factors c_i^J ($J = P,\ D,\ S_{\rm l},\ S_{\rm h};\ \ i = 1...6$ sub-modes; $S_{l,h}$ denotes the S-wave in the LMR, HMR) by normalizing the mass acceptance function, theoretical mass distribution, Dalitz two body phase space factor (m/q) [70], $m^2 \to m$ Dalitz Jacobian $(2m)$, and simple integration over $m^2_{\eta\pi}$ dimension of the Dalitz plot ($\mathrm{dal}(m) = (m^2_{\eta\pi})_{\rm max} - (m^2_{\eta\pi})_{\rm min}$) for each mode, partial wave, and mass range combination, such that

$$\int_{0.755}^{1.035} dm\ c_i^P 2m\ \mathrm{dal}(m)\ \mathrm{accM}_i^P(m) |\sin\delta_P|^2 \frac{m^2}{q^2} = 1, \quad \text{(L.7)}$$

$$\int_{0.755}^{1.035} dm\ c_i^{S_{\rm l}} 2m\ \mathrm{dal}(m)\ \mathrm{accM}_i^{S_{\rm l}}(m) |\sin\delta_S|^2 \frac{m^2}{q^2} = 1, \quad \text{(L.8)}$$

$$\int_{1.035}^{1.535} dm\ c_i^D 2m\ \mathrm{dal}(m)\ \mathrm{accM}_i^D(m) |\sin\delta_D|^2 \frac{m^2}{q^2} = 1, \quad \text{(L.9)}$$

$$\int_{1.035}^{1.535} dm\ c_i^{S_{\rm h}} 2m\ \mathrm{dal}(m)\ \mathrm{accM}_i^{S_{\rm h}}(m) |\sin\delta_S|^2 \frac{m^2}{q^2} = 1. \quad \text{(L.10)}$$

For the helicity parts of the amplitudes, we obtain another 24 normalization factors, h_i^J ($J = P,\ D,\ S_{\rm l},\ S_{\rm h};\ \ i = 1...6$ sub-modes), by separately normalizing each acceptance function and corresponding theoretical helicity distribution over the appropriate helicity range. In Eqs. L.11 and L.12, *lowcut* refers to the LMR lower helicity limit of $\mathcal{H} = -0.7$ for modes in which the K^* final state includes a π^0 and $\mathcal{H} = -0.95$ for the other four modes.

$$\int_{\rm lowcut}^{1.0} d\mathcal{H}\ h_i^P\ \mathrm{accH}_i^P(\mathcal{H})\ \mathcal{H}^2 = 1, \quad \text{(L.11)}$$

$$\int_{\rm lowcut}^{1.0} d\mathcal{H}\ h_i^{S_{\rm l}}\ \mathrm{accH}_i^{S_{\rm l}}(\mathcal{H}) = 1, \quad \text{(L.12)}$$

$$\int_{-0.5}^{1.0} d\mathcal{H}\ h_i^D\ \mathrm{accH}_i^D(\mathcal{H})\ (3\mathcal{H}^2 - 1)^2 = 1, \quad \text{(L.13)}$$

$$\int_{-0.5}^{1.0} d\mathcal{H}\ h_i^{S_{\rm h}}\ \mathrm{accH}_i^{S_{\rm h}}(\mathcal{H}) = 1. \quad \text{(L.14)}$$

L.0.4 Constant phase difference

As mentioned, there is an overall, constant phase difference between the interfering amplitudes which appears only in the interference terms of the PDFs in Eqs. 6.16 and 6.17. We write the interference terms for S-wave and J-waves with a phase shift δ_0 (δ_0') for S/P (S/D) interference as

$$
\begin{aligned}
&2Re[A_{BW_J} A^*_{LASS} e^{-i\delta_0^{(\prime)}}] \\
&= 2\mathrm{Re}[\sin\delta_J e^{i\delta_J}\ \sin\delta_S e^{-i\delta_S}\ e^{-i\delta_0^{(\prime)}}], \\
&= 2\ \sin\delta_J\ \sin\delta_S\ \cos(\delta_J - \delta_S - \delta_0^{(\prime)}), \\
&= 2(\sin\delta_J \cos(\delta_J - \delta_0^{(\prime)}) \frac{\cot\delta_S}{1+\cot^2\delta_S} + \sin\delta_J \sin(\delta_J - \delta_0^{(\prime)}) \frac{1}{1+\cot^2\delta_S}), \\
&= \frac{2\sin\delta_J}{1+\cot^2\delta_S}[(\cot\delta_S\cos\delta_J + \sin\delta_J)\cos\delta_0^{(\prime)} + (\cot\delta_S\sin\delta_J - \cos\delta_J)\sin\delta_0^{(\prime)}], \\
&= 2\sin^2\delta_J \sin^2\delta_S[(1+\cot\delta_S\cot\delta_J)\cos\delta_0^{(\prime)} + (\cot\delta_S - \cot\delta_J)\sin\delta_0^{(\prime)}]. \qquad \text{(L.15)}
\end{aligned}
$$

In practice, the above expression (Eq. L.15) is readily implemented with the mass dependence of the phases $\delta_P,\ \delta_D,\ \delta_R,\ \delta_B$ in Sec. L.0.1. The phase shifts $\delta_0^{(\prime)}$ are floated in the fits.

L.0.5 Mode specific PDFs

Because there is a unique acceptance function for each sub-mode, the two PDFs in Eqs. 6.16 and 6.17 become twelve; these equations can be written, for $i = 1...6$ sub-modes

$$
\begin{aligned}
&\mathrm{PDF}_i^{LMR}(m,\mathcal{H}) \\
&= f_P\ c_i^P h_i^P\ \mathrm{accM}_i^P(m)\mathrm{accH}_i^P(\mathcal{H})\ \sin^2\delta_P\ \mathcal{H}^2 \\
&\quad + (1-f_P)\ c_i^{S_1} h_i^{S_1}\ \mathrm{accM}_i^{S_1}(m)\mathrm{accH}_i^{S_1}(\mathcal{H})\ \sin^2\delta_S \\
&\quad + \sqrt{f_P(1-f_P)c_i^P c_i^{S_1} h_i^P h_i^{S_1}\ \mathrm{accH}_i^P \mathrm{accH}_i^{S_1} \mathrm{accM}_i^P \mathrm{accM}_i^{S_1}} \\
&\qquad \times 2\mathrm{Re}[A_{BW_1} A^*_{LASS} e^{-i\delta_0}]\ \mathcal{H},
\end{aligned}
$$

$$
\mathrm{PDF}_i^{HMR}(m,\mathcal{H})
$$

$$
\begin{aligned}
= \quad & f_D \; c_i^D h_i^D \; \mathrm{accM}_i^D(m)\mathrm{accH}_i^D(\mathcal{H}) \; \sin^2\delta_D \; (3\mathcal{H}^2-1)^2 \\
& + (1-f_D) \; c_i^{S_\mathrm{h}} h_i^{S_\mathrm{h}} \; \mathrm{accM}_i^{S_\mathrm{h}}(m)\mathrm{accH}_i^{S_\mathrm{h}}(\mathcal{H}) \; \sin^2\delta_S \\
& + \sqrt{f_D(1-f_D)c_i^D c_i^{S_\mathrm{h}} h_i^D h_i^{S_\mathrm{h}} \; \mathrm{accH}_i^D \mathrm{accH}_i^{S_\mathrm{h}} \mathrm{accM}_i^D \mathrm{accM}_i^{S_\mathrm{h}}} \\
& \qquad \times 2\mathrm{Re}[A_{BW_2} A^*_{LASS} e^{-i\delta'_0}] \; (3\mathcal{H}^2-1).
\end{aligned}
$$

The interference terms ($2\mathrm{Re}[AB^*e^{-i\delta}]$) are written explicitly in Eq. L.15. The only floating parameters are f_P, f_D, δ_0, δ'_0, and the total $\eta K\pi$ yield, which scales each entire PDF.

L.0.6 Results: comparison with published Run1-5 results

In Table L.1, we compare results of the nominal LMR analysis of Run1-5 data with S-wave yield fixed, the same analysis with S-wave yield floated, and the analysis in which we employ the two-dimensional $K\pi$ mass/helicity PDF. The P-wave yield results in the third column of Table L.1 differ slightly from those in Table 6.13 because the PDFs were improved slightly between analyses.

In Table L.2, we compare results of the nominal HMR analysis of Run1-5 data and the same analysis with a Breit-Wigner D-wave $m_{K\pi}$ PDF instead of a double Gaussian, and the analysis in which we employ the two-dimensional $K\pi$ mass/helicity PDF.

In Table L.3 we show results for $\delta_0^{(\prime)}$, f_D, and f_P. We report the weighted averages of the $\delta_0^{(\prime)}$ for the interference between S-wave/P-wave and S-wave/D-wave in each of the neutral and charged modes.

L.0.6.1 Results: simultaneous fit in HMR

In addition to finding the weighted average of sub-mode results, we also perform a simultaneous fit in which we constrain the phase differences for appropriate HMR sub-modes to be the equal. In Table L.4 we show results for δ'_0, D-wave fraction (f_D), and total ηK^* yield ($Y_{\eta K^*}$) from the simultaneous fits for δ'_0.

Table L.1: Comparison of results from three LMR analyses of Run1-5 data: the nominal analysis, which includes 1D $K\pi$ mass and helicity PDFs with a fixed S-wave yield, an analysis that incorporates the same 1D PDFs with a floated S-wave yield, and an analysis that includes a 2D $K\pi$ mass/helicity PDF with a floated partial wave fraction and phase difference. (The yields are calculated from these fractions and phase differences.)

	Nominal			Float $Y_{K_0^*}$			2D Pdf		
	Y_{tot}	$Y_{K_0^*}$	$Y_{K_1^*}$	Y_{tot}	$Y_{K_0^*}$	$Y_{K_1^*}$	Y_{tot}	$Y_{K_0^*}$	$Y_{K_1^*}$
$\eta_{\gamma\gamma}K^{*+}_{K^+\pi^0}$	133.8 ± 16.3	34.5	99.3 ± 16.3	110.4 ± 24.8	-1.9 ± 17.2	112.3 ± 17.8	117.3 ± 17.3	13.7 ± 11.1	103.6 ± 19.0
$\eta_{3\pi}K^{*+}_{K^+\pi^0}$	71.1 ± 11.4	14.9	56.2 ± 11.4	76.8 ± 16.6	8.6 ± 11.4	58.2 ± 12.0	64.2 ± 12.3	12.4 ± 13.3	51.8 ± 17.0
$\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}$	188.3 ± 19.3	39.8	148.5 ± 19.3	176.2 ± 29.6	19.1 ± 20.5	157.1 ± 21.4	181.4 ± 20.6	46.3 ± 14.9	135.1 ± 20.8
$\eta_{3\pi}K^{*+}_{K^0_S\pi^+}$	52.5 ± 10.2	16.5	36.0 ± 10.2	41.9 ± 17.1	-0.2 ± 12.6	42.1 ± 11.5	47.8 ± 11.4	19.0 ± 9.5	28.8 ± 10.9
$\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}$	480.9 ± 29.4	70.4	410.5 ± 29.4	486.9 ± 44.4	82.0 ± 30.3	404.9 ± 32.5	469.4 ± 30.7	66.2 ± 20.6	403.1 ± 33.4
$\eta_{3\pi}K^{*0}_{K^+\pi^-}$	140.0 ± 16.3	27.4	112.6 ± 16.3	149.5 ± 25.2	45.3 ± 17.9	104.2 ± 17.8	137.9 ± 17.5	36.1 ± 16.9	101.8 ± 21.1

Table L.2: Comparison of results from three HMR analyses of Run1-5 data: the nominal analysis, which includes 1D $K\pi$ mass and helicity PDFs, the nominal analysis with Breit-Wigner D-wave $K\pi$ mass PDF instead of a double Gaussian, and an analysis that includes a 2D $K\pi$ mass/helicity PDF with a floated partial wave fraction and phase difference. Results for ηK^{*0} are from the "right" solution, results for ηK^{*+} are from "left" solution (Sec. L.0.7).

	Nominal			BW $m_{K\pi}$ PDF			2D PDF		
	Y_{tot}	$Y_{K_0^*}$	$Y_{K_2^*}$	Y_{tot}	$Y_{K_0^*}$	$Y_{K_2^*}$	Y_{tot}	$Y_{K_0^*}$	$Y_{K_2^*}$
$\eta_{\gamma\gamma} K^{*+}_{K^+\pi^0}$	118.5 ± 22.8	92.9 ± 19.6	25.6 ± 11.6	117.8 ± 22.9	89.8 ± 19.5	28.0 ± 12.0	118.1 ± 17.8	96.7 ± 20.5	21.4 ± 14.2
$\eta_{3\pi} K^{*+}_{K^+\pi^0}$	58.8 ± 14.2	39.0 ± 11.7	19.8 ± 8.1	59.6 ± 15.6	34.4 ± 12.2	25.2 ± 9.7	60.2 ± 11.9	35.9 ± 10.7	24.3 ± 9.2
$\eta_{\gamma\gamma} K^{*+}_{K^0_S\pi^+}$	67.7 ± 18.4	55.3 ± 15.7	12.4 ± 9.6	70.2 ± 19.6	43.9 ± 16.0	26.3 ± 11.4	74.1 ± 14.9	49.0 ± 14.4	25.1 ± 11.3
$\eta_{3\pi} K^{*+}_{K^0_S\pi^+}$	50.8 ± 11.9	48.9 ± 10.9	1.9 ± 4.7	50.5 ± 13.2	47.5 ± 11.5	3.0 ± 6.4	52.9 ± 10.6	50.7 ± 10.4	2.3 ± 2.3
$\eta_{\gamma\gamma} K^{*0}_{K^+\pi^-}$	234.5 ± 30.3	162.6 ± 25.4	71.9 ± 16.6	234.3 ± 31.6	148.6 ± 25.8	85.7 ± 18.2	231.9 ± 24.1	147.2 ± 23.3	84.8 ± 19.7
$\eta_{3\pi} K^{*0}_{K^+\pi^-}$	108.8 ± 21.6	69.0 ± 17.1	39.8 ± 13.2	110.2 ± 21.6	65.2 ± 16.9	45.0 ± 13.4	110.1 ± 16.3	61.0 ± 15.3	49.1 ± 14.2

Table L.3: Results for phase difference $\delta_0^{(\prime)}$, partial wave fraction f_J (J is P-wave in the LMR and D-wave in the HMR), and total ηK^* yield in the analysis that includes a 2D $K\pi$ mass/helicity PDF of Run1-5 data. We also show the weighted averages of $\delta_0^{(\prime)}$ with the χ^2 per degree of freedom for each weighted average. In the HMR where there are multiple minima in the $-2\ln(\mathcal{L}/\mathcal{L}_0)$ distribution as a function of δ_0', see Section L.0.7, we show both results.

	$\delta_0^{(\prime)}$	f_J	$Y_{\eta K^*}$	$\delta_0^{(\prime)}$	f_J	$Y_{\eta K^*}$
LMR						
$\eta_{\gamma\gamma}K^{*+}_{K^+\pi^0}$	-0.27 ± 0.38	0.88 ± 0.09	117.3 ± 17.3			
$\eta_{3\pi}K^{*+}_{K^+\pi^0}$	0.25 ± 0.51	0.81 ± 0.20	64.2 ± 12.3			
$\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}$	0.05 ± 0.19	0.74 ± 0.08	181.4 ± 20.6			
$\eta_{3\pi}K^{*+}_{K^0_S\pi^+}$	0.87 ± 0.33	0.60 ± 0.17	47.8 ± 11.4			
Combined δ_0	0.18 ± 0.14	$\chi^2/\text{dof} = 2.1$				
$\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}$	-0.24 ± 0.15	0.86 ± 0.04	469.4 ± 30.7			
$\eta_{3\pi}K^{*0}_{K^+\pi^-}$	-0.43 ± 0.25	0.74 ± 0.12	137.9 ± 17.5			
Combined δ_0	-0.29 ± 0.13	$\chi^2/\text{dof} = 0.4$				
HMR	Right δ_0' Solution			Left δ_0' Solution		
$\eta_{\gamma\gamma}K^{*+}_{K^+\pi^0}$	0.88 ± 0.21	0.26 ± 0.12	112.9 ± 17.4	-2.16 ± 0.23	0.18 ± 0.12	118.1 ± 17.8
$\eta_{3\pi}K^{*+}_{K^+\pi^0}$	0.17 ± 0.59	0.66 ± 0.13	50.2 ± 11.0	-1.60 ± 0.24	0.40 ± 0.12	60.2 ± 11.9
$\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}$	0.87 ± 0.33	0.34 ± 0.26	62.6 ± 14.9	-1.90 ± 0.25	0.34 ± 0.13	74.1 ± 14.9
$\eta_{3\pi}K^{*+}_{K^0_S\pi^+}$	-0.55 ± 0.95	0.04 ± 0.04	52.9 ± 10.6	-0.55 ± 0.95	0.04 ± 0.04	52.9 ± 10.6
Combined δ_0'	0.78 ± 0.17	$\chi^2/\text{dof} = 1.1$		-1.87 ± 0.14	$\chi^2/\text{dof} = 1.6$	
$\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}$	0.63 ± 0.13	0.37 ± 0.08	231.9 ± 24.1	-2.02 ± 0.13	0.34 ± 0.08	239.5 ± 24.7
$\eta_{3\pi}K^{*0}_{K^+\pi^-}$	0.48 ± 0.18	0.45 ± 0.11	110.1 ± 16.3	-1.85 ± 0.18	0.40 ± 0.12	110.1 ± 16.5
Combined δ_0'	0.58 ± 0.11	$\chi^2/\text{dof} = 0.5$		-1.96 ± 0.11	$\chi^2/\text{dof} = 0.6$	

Table L.4: Results for phase difference δ_0, partial wave fraction f_D, and total ηK^* yield from separate-fit and simultaneous-fit HMR analyses with 2D $K\pi$ mass/helicity PDF of Run1-5 data. We also show the weighted averages of δ_0 from the separate fits with the χ^2 per degree of freedom for each weighted average. We show results associated with the smaller $-2\ln(\mathcal{L}/\mathcal{L}_0)$ in scans of δ_0 (the "right" solution).

	δ_0	f_D	$Y_{\eta K^*}$	f_D	$Y_{\eta K^*}$
HMR		Separate Fit		Combined Fit	
$\eta_{\gamma\gamma} K^{*+}_{K^+\pi^0}$	-2.16 ± 0.23	0.18 ± 0.12	118.1 ± 17.8	0.10 ± 0.13	119.5 ± 17.2
$\eta_{3\pi} K^{*+}_{K^+\pi^0}$	-1.60 ± 0.24	0.40 ± 0.12	60.2 ± 11.9	0.41 ± 0.13	61.3 ± 11.7
$\eta_{\gamma\gamma} K^{*+}_{K^0_S\pi^+}$	-1.90 ± 0.25	0.34 ± 0.13	74.1 ± 14.9	0.33 ± 0.13	75.8 ± 14.8
$\eta_{3\pi} K^{*+}_{K^0_S\pi^+}$	-0.55 ± 0.95	0.04 ± 0.04	52.9 ± 10.6	0.04 ± 0.07	51.1 ± 10.3
Combined δ_0	-1.87 ± 0.14	$\chi^2/\text{dof} = 1.6$		-1.82 ± 0.19	
$\eta_{\gamma\gamma} K^{*0}_{K^+\pi^-}$	0.63 ± 0.13	0.37 ± 0.08	231.9 ± 24.1	0.37 ± 0.08	230.84 ± 23.7
$\eta_{3\pi} K^{*0}_{K^+\pi^-}$	0.48 ± 0.18	0.45 ± 0.11	110.1 ± 16.3	0.45 ± 0.11	106.15 ± 15.8
Combined δ_0	0.58 ± 0.11	$\chi^2/\text{dof} = 0.5$		0.58 ± 0.11	

In Fig. L.1 we plot $\chi^2 = -2\ln(\mathcal{L}/\mathcal{L}_0)$ as a function of δ_0' from the two simultaneous fits. Notice that the ηK^{*+} modes prefer the left solution while the two $B \to \eta K^{*0}$ sub-modes slightly prefer the right solution.

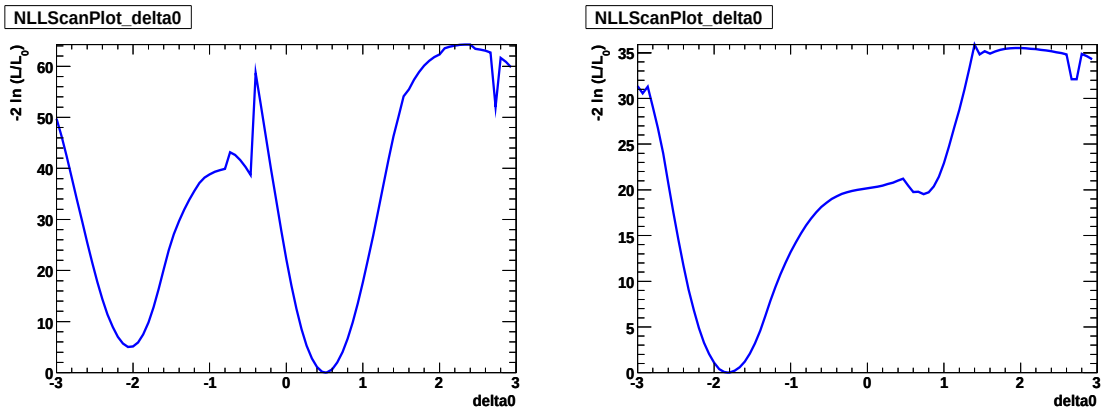


Figure L.1: Scans of χ^2 as functions of δ_0' for the simultaneous fits in the HMR: ηK^{*0} (*left*) and ηK^{*+} (*right*).

L.0.7 Multiple minima in HMR and δ_0' ambiguity

In all sub-modes of the HMR fit, except $\eta_{3\pi} K^{*+}_{K^0_S\pi^+}$, there are multiple minima in $-\ln\mathcal{L}$ as a function of δ_0'; one is near $\delta_0' \simeq -2$ and the other is near $\delta_0' \simeq 0.65$. Scans of $\chi^2 = -2\ln(\mathcal{L}/\mathcal{L}_{\rm max})$ as a function of δ_0' are shown in Fig. L.2. In Table L.2, results for ηK^{*0} are from fits that have chosen the right minima and results for ηK^{*+} are from fits that have settled in the left minima; these yield results depend very little on the value of δ_0'. In Table L.3 we show results for both minima.

These multiple minima are similar to those seen in the study of S-wave and D-wave interference in the *BABAR* ϕK^{*0} analysis [85]. In Fig. L.2 we reproduce the $\chi^2 = -2\ln(\mathcal{L}/\mathcal{L}_{\rm max})$ plot from the ϕK^{*0} analysis. The minima are at 3.54 and ~ 1.25 over a range of $0 \to 2\pi$ corresponding to minima of -2.74 and 1.25 over a range of $-\pi \to \pi$.

In Fig. L.3, we show plots of the helicity projection of 2D $K\pi$ mass/helicity

PDF for difference values of δ_0' from 10,000 signal events generated from PDFs of the $\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}$ mode. One can see that the shape is very similar for values $\delta_0' = -2.02$ and $\delta_0' = 0.63$, which were the values of δ_0' at the multiple minima of $-\ln\mathcal{L}$ for the $\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}$ mode.

L.0.8 Branching fraction extraction

In cases where there is no interference, such as when integrating over the entire helicity range, we use two signal yields to calculate two branching fractions. These signal yields come solely from squared amplitude terms. We write the each branching fraction in terms of the relevant signal yield Y, the bias, the number of $B\overline{B}$ pairs (nBB), the efficiency (ϵ), and the product of daughter branching fractions (prodBR).

$$\mathcal{B} \;=\; \frac{Y - \text{bias}}{\text{nBB}\cdot\epsilon\cdot\text{prodBR}} \tag{L.16}$$

In the present case, Y is the the total yield from two squared amplitudes and interference. We must remove the contribution of interference to the yield and obtain the branching fractions that we would have found from the squared amplitudes were there no interference; i.e., we need to determine two effective yields that we can use in Eq. L.16 to determine two partial wave branching fractions. These effective yields will be functions of the fit parameters f_J, $\delta_0^{(\prime)}$, and Y. (As seen in the PDFs in Sec. L.0.5, the variable f_J is the fraction of the total yield Y that is from the P-wave (D-wave) in the LMR (HMR) fit; the variable $\delta_0^{(\prime)}$ is the phase difference between S and P (D) partial waves in the LMR (HMR) fit.)

The PDFs in Sec. L.0.5 were constructed so that the first two squared amplitude terms would each integrate to unity. The remaining interference term integrates to some number determined by the values of f_J and $\delta_0^{(\prime)}$. In terms of the interference integral over the Dalitz plot (DP)

$$\text{intIntgrl}(\delta_0^{(\prime)}, f_J) \;=\; \iint_{\text{DP}} \text{interferenceTerm}(m_{K\pi}, \mathcal{H}; \delta_0^{(\prime)}, f_J)\; \mathrm{d}m_{K\pi}\; \mathrm{d}\mathcal{H}, \tag{L.17}$$

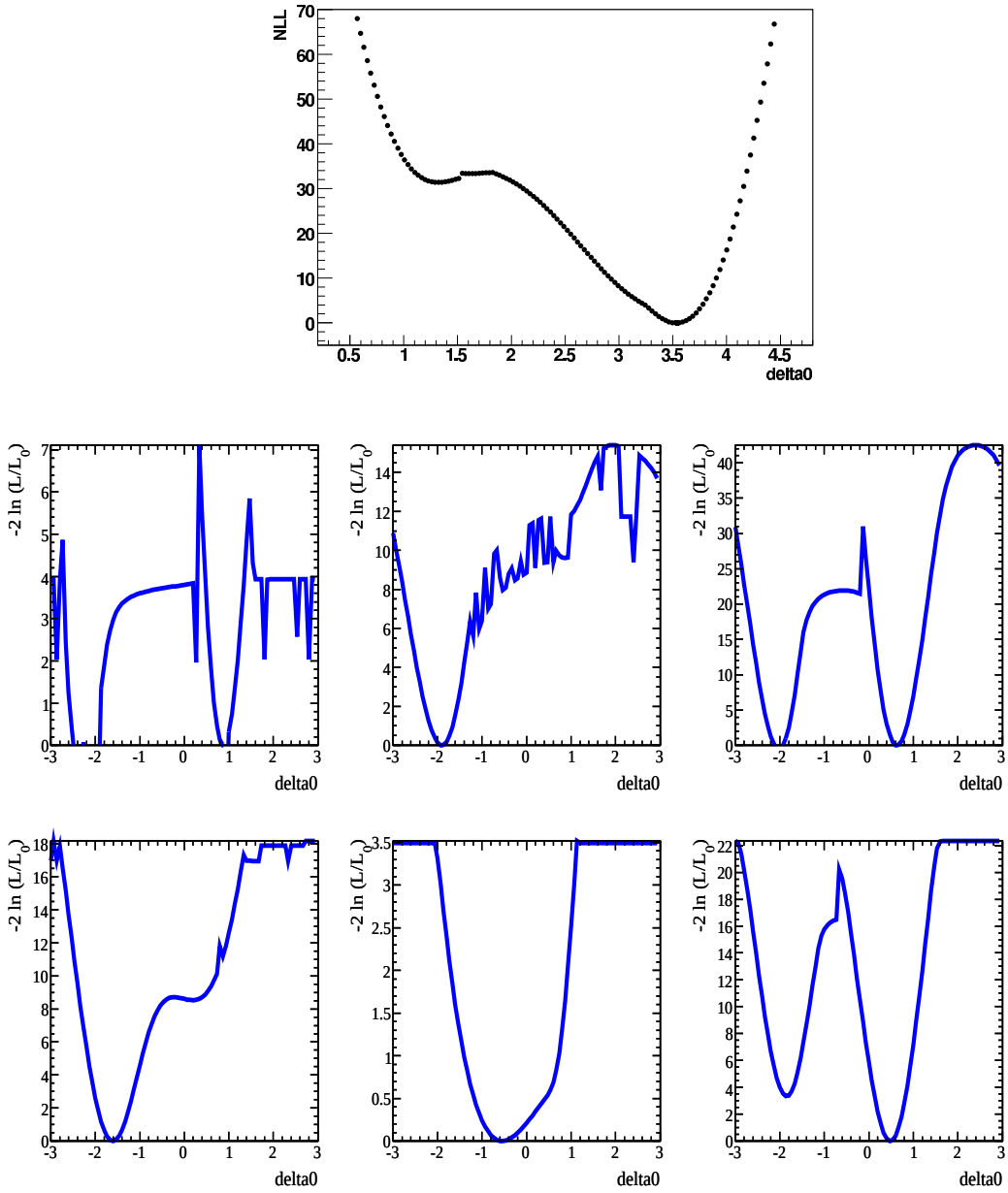


Figure L.2: Distribution of $\chi^2 = -2\ln(\mathcal{L}/\mathcal{L}_{\rm max})$ for δ_0 in the $B^0 \to \phi K_2^{*0}(1430)$ fit with S-wave/D-wave interference included (*top*). Scans of χ^2 as functions of δ_0' for the HMR sub-modes: $\eta_{\gamma\gamma}K^{*+}_{K^+\pi^0}$, $\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}$, $\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}$(*middle, left-to-right*) and $\eta_{3\pi}K^{*+}_{K^+\pi^0}$, $\eta_{3\pi}K^{*+}_{K^0_S\pi^+}$, $\eta_{3\pi}K^{*0}_{K^+\pi^-}$ (*bottom, left-to-right*).

we can write the fraction of the total yield that is contributed by interference

$$\text{intFrac}(\delta_0^{(\prime)}, f_J) \quad = \quad \frac{\text{intIntgrl}(\delta_0^{(\prime)}, \text{f}_\text{J})}{1 + \text{intIntgrl}(\delta_0^{(\prime)}, \text{f}_\text{J})}. \tag{L.18}$$

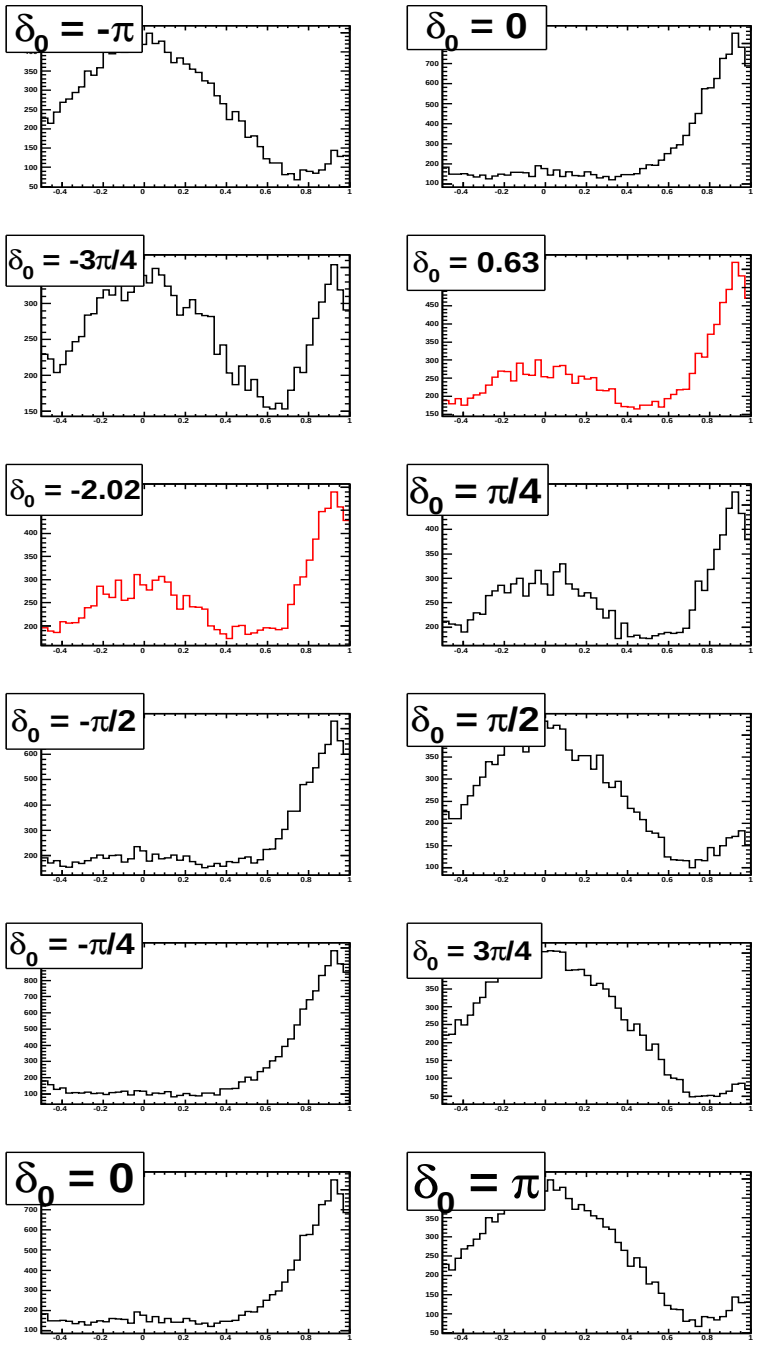


Figure L.3: Plots of helicity projection of 2D $K\pi$ mass/helicity PDF for different values of δ_0' from 10,000 signal events generated from PDFs of the $\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}$ mode. One can see that the shape is very similar for multiple fit values $\delta_0' = -2.02$ and $\delta_0' = 0.63$.

In terms of the fit parameters Y, $\delta_0^{(\prime)}$, and f_J, we can write the interference-free effective

yield as:

$$Y_J = Y \cdot f_J \cdot (1 - \text{intFrac}(\delta_0^{(\prime)}, f_J)) \quad \text{(L.19)}$$

$$= \frac{Y \cdot f_J}{1 + \text{intIntgrl}(\delta_0^{(\prime)}, f_J)}, \quad \text{(L.20)}$$

and the branching fraction is

$$\mathcal{B}(Y, \delta_0^{(\prime)}, f_J) = \frac{1}{\text{nBB} \cdot \epsilon \cdot \text{prodBR}} \left[\frac{Y \cdot f}{1 + \text{intIntgrl}(\delta_0^{(\prime)}, f_J)} - \text{bias} \right]. \quad \text{(L.21)}$$

We calculate the branching fractions and related uncertainties using Eq L.21. To correctly treat the errors, we use the variances (σ_Y^2, $\sigma^2_{\delta_0^{(\prime)}}$, and σ_f^2), covariances ($\rho_{\delta_0^{(\prime)},f}$, $\rho_{Y,\delta_0^{(\prime)}}$, and $\rho_{Y,f}$), and partial derivatives ($\frac{\partial}{\partial f}\mathcal{B}$, $\frac{\partial}{\partial \delta_0^{(\prime)}}\mathcal{B}$, and $\frac{\partial}{\partial Y}\mathcal{B}$) related to the fit parameters as reported by the fit. Table L.5 lists the numbers that go into the branching fraction calculation, and Table L.6 lists the interference fraction, the published branching fractions, the branching fractions calculated with numbers from a 2D fit assuming no interference, and branching fractions calculated with numbers from a 2D fit treating interference correctly.

L.0.9 Interpreting results

The most important conclusion to draw from the results of the 2D fit is that the interference-related systematic uncertainties reported in the publication (Sec. 6.13) are reasonable. We quoted systematic uncertainty related to interference to be 4.7% for $\eta K^*(892)$ modes in which $K^* \to K^+\pi^0$, 0.9% for other $\eta K^*(892)$ modes, and 10.5% for $\eta K_0^*(S\text{-wave})$ and $\eta K_2^*(1430)$ modes. As seen in the first column of Table L.6, these estimates are very good for all modes except $\eta_{\gamma\gamma} K^{*+}_{K^0_S\pi^+}$ and $\eta_{3\pi} K^{*+}_{K^0_S\pi^+}$ in the LMR, for which the interference fraction is much larger than expected. This discrepancy can be related to several sources. Most important, the $\eta K_0^*(S\text{-wave})$ yields are floating in the 2D fit and are fixed in the nominal fit. Second, we have not used the simultaneous fit for the LMR, so that determination of δ_0 is not optimal.

Table L.5: For each sub-mode we list the final values and errors for floated parameters Y, $\delta_0^{(\prime)}$, f, and the covariances between these parameters as they go into the branching fraction calculation.

Mode	Y (evts)	$\delta_0^{(\prime)}$ (rad)	f	$\rho_{\delta_0,f}$	$\rho_{Y,f}$	$\rho_{Y,\delta_0^{(\prime)}}$
HMR ηK_2^*						
$\eta_{\gamma\gamma}K^{*+}_{K^+\pi^0}$	119.5 ± 17.2	-1.82 ± 0.19	0.10 ± 0.13	0.04	0.67	0.10
$\eta_{3\pi}K^{*+}_{K^+\pi^0}$	61.3 ± 11.7	-1.82 ± 0.19	0.41 ± 0.13	0.03	-0.07	-0.20
$\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}$	75.8 ± 14.8	-1.82 ± 0.19	0.33 ± 0.13	0.02	-0.18	-0.22
$\eta_{3\pi}K^{*+}_{K^0_S\pi^+}$	51.1 ± 10.3	-1.82 ± 0.19	0.04 ± 0.07	0.06	0.11	0.01
$\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}$	230.8 ± 23.7	0.58 ± 0.11	0.37 ± 0.08	0.01	0.07	-0.20
$\eta_{3\pi}K^{*0}_{K^+\pi^-}$	106.2 ± 15.8	0.58 ± 0.11	0.45 ± 0.11	-0.02	0.01	-0.20
HMR ηK_0^*						
$\eta_{\gamma\gamma}K^{*+}_{K^+\pi^0}$	119.5 ± 17.2	-1.82 ± 0.19	0.10 ± 0.13	0.04	0.67	0.10
$\eta_{3\pi}K^{*+}_{K^+\pi^0}$	61.3 ± 11.7	-1.82 ± 0.19	0.41 ± 0.13	0.03	-0.07	-0.20
$\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}$	75.8 ± 14.8	-1.82 ± 0.19	0.33 ± 0.13	0.02	-0.18	-0.22
$\eta_{3\pi}K^{*+}_{K^0_S\pi^+}$	51.1 ± 10.3	-1.82 ± 0.19	0.04 ± 0.07	0.06	0.11	0.01
$\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}$	230.8 ± 23.7	0.58 ± 0.11	0.37 ± 0.08	0.01	0.07	-0.20
$\eta_{3\pi}K^{*0}_{K^+\pi^-}$	106.2 ± 15.8	0.58 ± 0.11	0.45 ± 0.11	-0.02	0.01	-0.20
LMR ηK^*						
$\eta_{\gamma\gamma}K^{*+}_{K^+\pi^0}$	117.3 ± 17.3	-0.27 ± 0.38	0.88 ± 0.09	-0.05	-0.07	-0.14
$\eta_{3\pi}K^{*+}_{K^+\pi^0}$	64.2 ± 12.3	0.25 ± 0.51	0.81 ± 0.20	0.03	-0.18	-0.21
$\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}$	181.4 ± 20.6	0.05 ± 0.19	0.74 ± 0.08	-0.03	-0.001	-0.07
$\eta_{3\pi}K^{*+}_{K^0_S\pi^+}$	47.8 ± 11.4	0.87 ± 0.33	0.60 ± 0.17	-0.14	-0.11	-0.15
$\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}$	469.4 ± 30.7	-0.24 ± 0.15	0.86 ± 0.04	-0.04	-0.04	-0.11
$\eta_{3\pi}K^{*0}_{K^+\pi^-}$	137.9 ± 17.5	-0.43 ± 0.25	0.74 ± 0.12	-0.12	0.08	-0.16

Table L.6: We list the fraction of the total event yield contributed by interference, the published branching fraction, the branching fraction calculated with the results from our current 2D fit ($Y,\ \delta_0^{(\prime)},\ f$) having naively ignored interference, and the branching fraction calculated having correctly removed interference.

Mode	intFrac (%)	$\mathcal{B}(10^{-6})$(pub)	$\mathcal{B}(10^{-6})$(noInt)	$\mathcal{B}(10^{-6})$
HMR ηK_2^*				
$\eta_{\gamma\gamma}K^{*+}_{K^+\pi^0}$	$-\ 7.7 \pm 6.1$	9.1 ± 4.0	4.4 ± 5.5	4.7 ± 6.3
$\eta_{3\pi}K^{*+}_{K^+\pi^0}$	-10.7 ± 5.6	17.8 ± 7.2	22.4 ± 7.3	24.7 ± 8.2
$\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}$	-12.5 ± 6.3	6.4 ± 4.7	12.7 ± 4.6	14.2 ± 5.3
$\eta_{3\pi}K^{*+}_{K^0_S\pi^+}$	$-\ 4.6 \pm 4.6$	0.9 ± 5.1	1.0 ± 3.9	1.1 ± 4.2
$\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}$	-10.7 ± 3.7	8.4 ± 1.9	10.4 ± 2.3	11.5 ± 2.6
$\eta_{3\pi}K^{*0}_{K^+\pi^-}$	-11.1 ± 3.8	12.5 ± 4.1	15.2 ± 3.9	16.9 ± 4.4
HMR ηK_0^*				
$\eta_{\gamma\gamma}K^{*+}_{K^+\pi^0}$	$-\ 7.7 \pm 6.1$	19.2 ± 4.5	22.5 ± 5.3	24.4 ± 5.2
$\eta_{3\pi}K^{*+}_{K^+\pi^0}$	-10.7 ± 5.6	18.0 ± 6.3	16.4 ± 5.1	18.5 ± 5.6
$\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}$	-12.5 ± 6.3	13.3 ± 4.2	12.1 ± 3.2	13.8 ± 3.5
$\eta_{3\pi}K^{*+}_{K^0_S\pi^+}$	$-\ 4.6 \pm 4.6$	28.1 ± 6.7	28.2 ± 6.5	29.6 ± 6.5
$\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}$	-10.7 ± 3.7	10.8 ± 1.9	9.5 ± 1.6	10.7 ± 1.7
$\eta_{3\pi}K^{*0}_{K^+\pi^-}$	-11.1 ± 3.8	11.4 ± 3.2	9.4 ± 2.5	10.6 ± 2.8
LMR ηK^*				
$\eta_{\gamma\gamma}K^{*+}_{K^+\pi^0}$	2.2 ± 1.5	18.0 ± 3.2	18.9 ± 3.4	18.5 ± 11.7
$\eta_{3\pi}K^{*+}_{K^+\pi^0}$	5.3 ± 2.4	25.4 ± 5.5	23.4 ± 7.0	22.0 ± 50.0
$\eta_{\gamma\gamma}K^{*+}_{K^0_S\pi^+}$	$-\ 7.8 \pm 1.7$	20.5 ± 2.9	18.4 ± 3.0	20.0 ± 10.1
$\eta_{3\pi}K^{*+}_{K^0_S\pi^+}$	-11.1 ± 3.9	11.9 ± 3.9	9.2 ± 3.8	10.4 ± 16.9
$\eta_{\gamma\gamma}K^{*0}_{K^+\pi^-}$	$-\ 0.8 \pm 0.2$	18.2 ± 1.4	18.1 ± 1.4	$18.2 \pm\ 2.0$
$\eta_{3\pi}K^{*0}_{K^+\pi^-}$	$-\ 0.6 \pm 0.5$	10.9 ± 2.0	10.8 ± 2.3	$10.9 \pm\ 5.5$

One can also see that the uncertainties on the branching fractions in the last column of Table L.6 are much larger than the published branching fraction uncertainties for some LMR modes. Again, this is mostly because we did not use a combined fit in the LMR, and we propagated the statistical uncertainties on the fit parameters other than the yield (f_J, $\delta_0^{(\prime)}$) in the calculation of the branching fraction, using the correlations between the parameters. One might argue that these uncertainties should be regarded as a separate systematic uncertainty, but that is not how they were treated in this study.

L.0.10 Future work

If this method is to be used in future analyses one needs to understand the issue of multiple $-2\ln\mathcal{L}$ minima in δ_0' and how statistical errors on Y, f_J, and $\delta_0^{(\prime)}$ should be propagated during the branching fraction calculation.

L.0.11 Comparison with other *BABAR* interference measurements

Five other *BABAR* analyses have studied phase differences between $K\pi$ partial waves. Some of these analyses have reported measurements of the unknown phase differences between S/P and S/D partial waves. We compare the results of these analyses and our results in Table L.7.

L.0.12 Forward-backward asymmetry in LMR

In the case of S-wave and P-wave interference, where the interference contribution as a function of $K\pi$ helicity is proportional to $K\pi$ helicity, one can observe the forward-backward asymmetry (A_{FB}) in the signal. We show signal A_{FB} distributions from three *BABAR* analyses in Fig. L.4. In the ψK^* and DK^* analyses, this asymmetry is defined in terms of the number of forward events with $\mathcal{H} > 0$ (N_F) and backward events with $\mathcal{H} < 0$ (N_B) as

$$A_{FB} = \frac{N_F - N_B}{N_F + N_B} \tag{L.22}$$

Table L.7: Comparison of measurements of phase differences between S/P and S/D partial waves in four *BABAR* analyses. For each result, we list the analysis, the *BABAR* analysis document (BAD) number, and an attempt to convert the quoted result for a sensible comparison with our result.

Analysis	BAD	S/P Result	Converted δ_0	S/D Result	Converted δ_0'
ψK^*	673	from A_{FB} plot	$\sim \pi$	–	–
$D^0 K^{*-}$	697	-1.57 ± 0.65	$\sim -3\pi/4$	–	–
$\rho^+ K^{*0}$	1435	4.33 ± 0.41	-1.95 ± 0.41	–	–
$K^+\pi^-\pi^+$	859	2.92 ± 0.11	~ 0	–	–
ϕK^{*0}	1428	2.78 ± 0.17	2.78 ± 0.17	3.54 ± 0.12	-2.74 ± 0.12
ηK^{*0}	1341	–	-0.29 ± 0.13	–	-1.96 ± 0.11
ηK^{*+}	1341	–	0.18 ± 0.14	–	-1.82 ± 0.19

However, it is actually more useful to plot the unnormalized forward-backward difference ($N_F - N_B$) as function of $K\pi$ mass, since the normalization washes out the shape of the forward-backward difference when $\delta_0 = \pi/4$ or $-3\pi/4$. When δ_0 is near these values, the mass projection of the interference term is essentially a Breit-Wigner which, when normalized with a Breit-Wigner, is flat. This explains the flat distribution of A_{FB} from the DK^* analysis in Fig. L.4.

In Fig. L.5, we show the *sPlot* for the unnormalized forward-backward difference as a function of $K\pi$ mass for ηK^{*+} (*l*eft) and ηK^{*0} (*r*ight) from the LMR analysis. We accumulate the signal sWeights from a helicity *sPlot* in bins of $K\pi$ mass in two histograms, one for forward events and one for backward events. We obtain the $N_F - N_B$ histogram by simply subtracting the backward histogram from the forward histogram, bin-by-bin. We also plot the theoretically expected forward-backward difference in $\eta_{\gamma\gamma} K^{*0}$ for values of the phase difference δ_0 from three analyses (ϕK^*, ρK^*, and this ηK^* analysis) to show how the $N_F - N_B$ distribution varies for different values of the parameter.

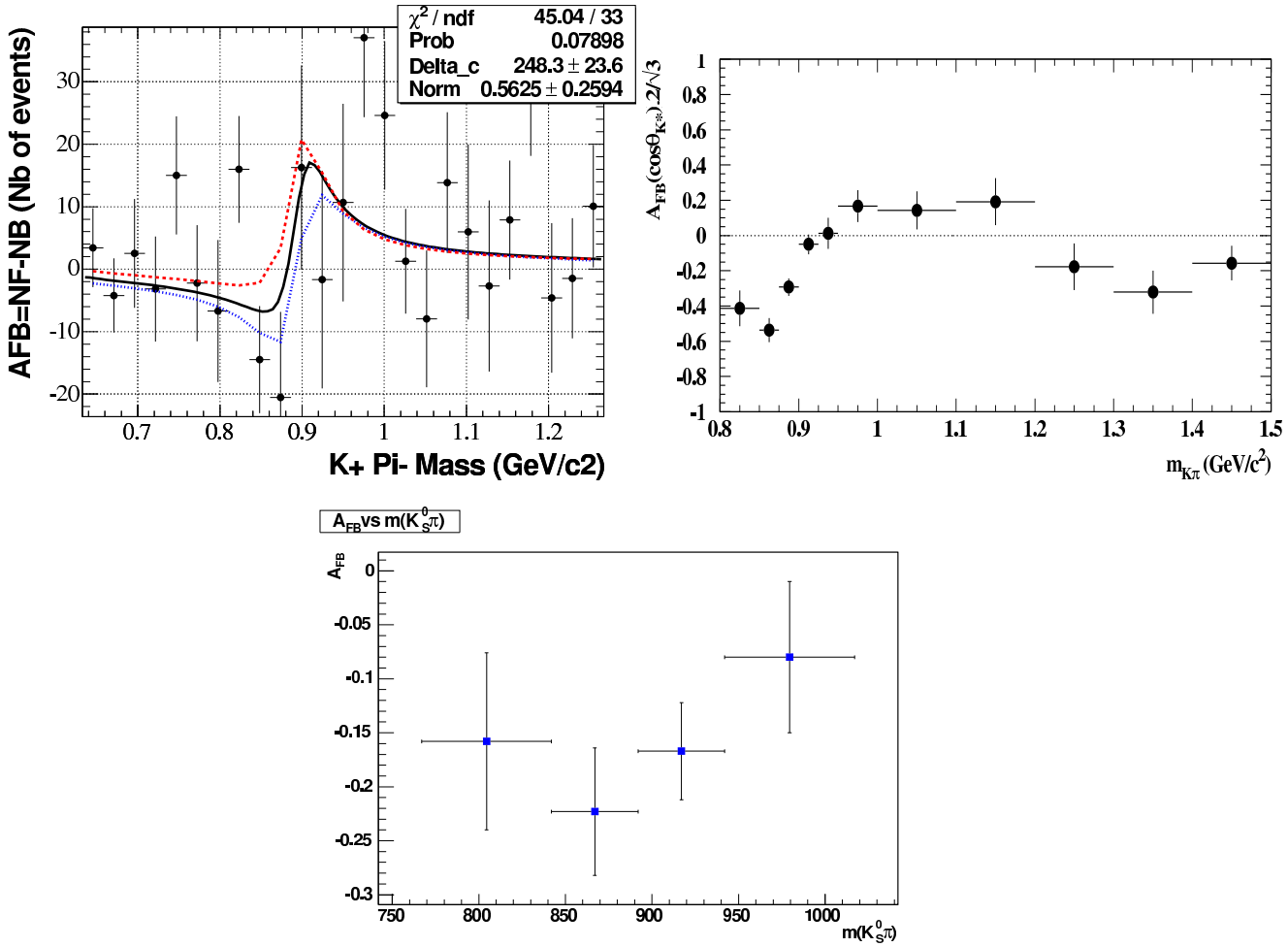


Figure L.4: Forward-backward asymmetries as functions of $K\pi$ mass from ρK^* BAD-1435 (*upper-left*), ψK^* BAD-673 (*upper-right*), and DK^* BAD-697 (*bottom*).

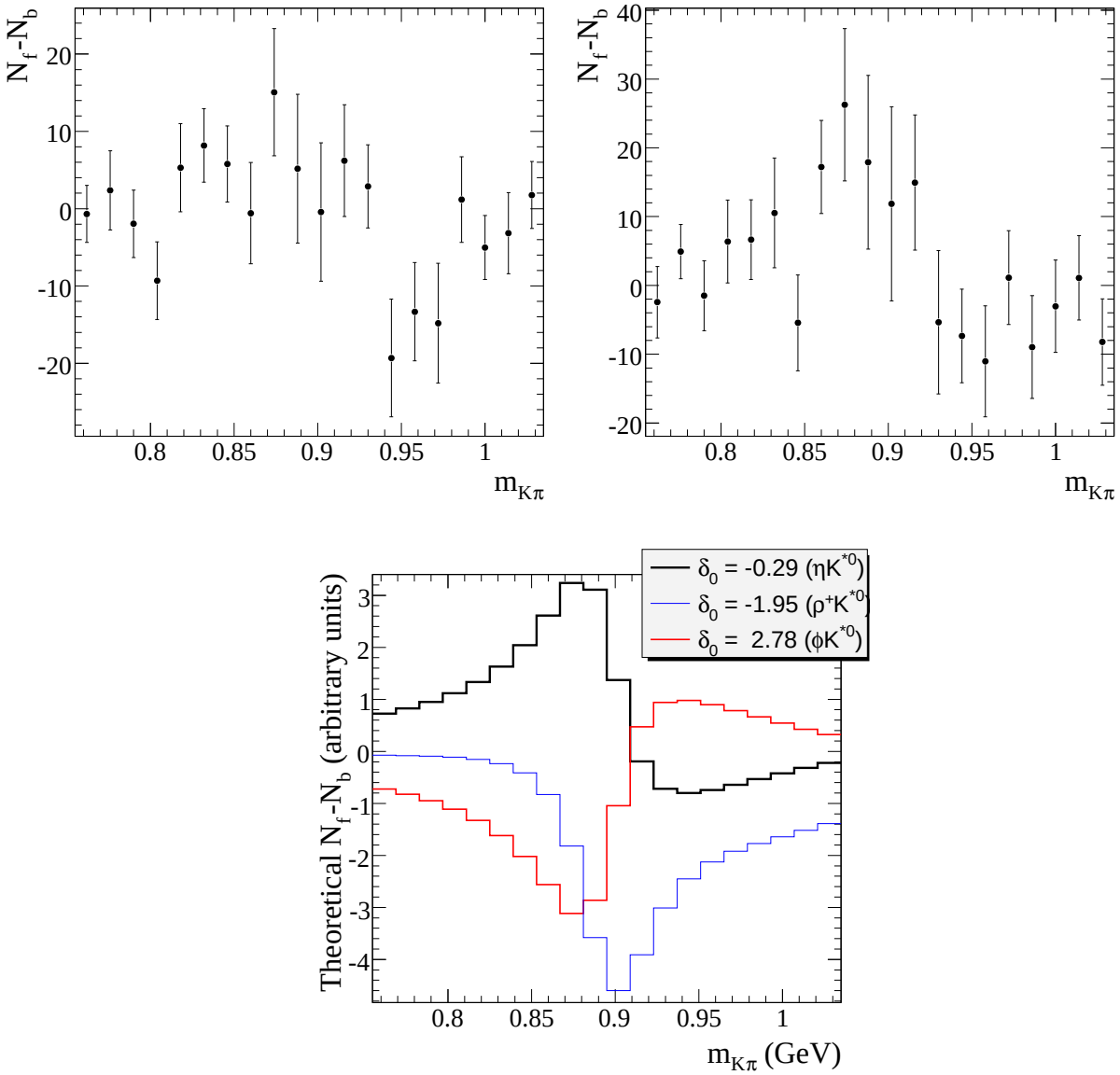


Figure L.5: *sPlot* of forward-backward difference as function of $K\pi$ mass for ηK^{*+} (*top-left*) and ηK^{*0} (*top-right*) from the LMR analysis indicating S/P-wave interference with a phase difference of ~ 0. Plot of theoretically expected forward-backward difference in $\eta_{\gamma\gamma}K^{*0}$ for values of the phase difference δ_0 from three *BABAR* analyses (ϕK^*, ρK^*, and this ηK^* analysis) (*bottom*). Notice the agreement between observed and expected $N_F - N_B$ for this analysis.