Measurements of the Branching Fraction and CP-Violation Asymmetries in $B^0 \to f_0(980)K_S^0$
dependent maximum likelihood fit, we measure the branching fraction $B_f^{0}$. The results are obtained from a data sample of $123 \times 10^6 \, Y(4S) \rightarrow \bar{B}\bar{B}$ decays. From a time-dependent maximum likelihood fit, we measure the branching fraction $B(B^0 \rightarrow f_0(980)(\pi^+\pi^-)K^0) = (6.0 \pm 0.9 \pm 0.6 \pm 1.2) \times 10^{-6}$, the mixing-induced CP violation parameter $S = -1.6_{-0.01}^{+0.56} \pm 0.09 \pm 0.04$, and the direct CP violation parameter $C = 0.27 \pm 0.36 \pm 0.10 \pm 0.07$, where the first errors are
In the standard model (SM), CP violation arises from a single phase in the three-generationCabibbo-Kobayashi-Maskawa (CKM) quark-mixing matrix [1]. Possible indications of physics beyond the SM may be observed in the time-dependent CP asymmetries of B decays dominated by penguin-type diagrams to states such as $\phi K^0$, $\eta'K^0$, $K^+K^-K^0$, and $f_0(980)K^0$ [2]. Neglecting CKM-suppressed amplitudes, these decays carry the same weak phase as the decay $B^0 \rightarrow J/\psi K^0$ [3]. As a consequence, their mixing-induced CP-violation parameter is expected to be $-\eta_1 \times \sin 2\beta = -\eta_1 \times 0.74 \pm 0.05$ [4] in the SM, where $\eta_1$ is the CP eigenvalue of the final state, which is +1 for $f_0(980)K^0$. There is no direct CP violation expected in these decays since they are dominated by a single amplitude in the SM. Because of the large virtual mass scales occurring in the penguin loops, additional diagrams with non-SM heavy particles in the loops and new CP-violating phases may contribute. Measurements of CP violation in these channels and their comparisons with the SM expectation are therefore sensitive probes for physics beyond the SM. The modes $\phi K^0$, $K^+K^-K^0$, and $\eta'K^0$ have been measured by both the BABAR [5] and Belle experiments. Interest in s-penguin modes has intensified since the Belle collaboration measured $\sin 2\beta = -0.96 \pm 0.50^{+0.09}_{-0.11}$ in the decay $B^0 \rightarrow \phi K^0_S$, while the BABAR collaboration (with a sample of approximately 114 million $BB$ pairs) measured $\sin 2\beta = 0.47 \pm 0.34$ (stat) $^{+0.08}_{-0.08}$ (syst) in the decay $B^0 \rightarrow \phi K^0$.

In this Letter we present the first measurement of the branching fraction and CP-violating asymmetries in the penguin-dominated decay $B^0 \rightarrow f_0 K^0_S$ [7] from a time-dependent maximum likelihood analysis. We also measure the mass and width of the $f_0$ resonance. We restrict the analysis to the region of the $\pi^+\pi^-K^0_S$ Dalitz plot that is dominated by the $f_0$ and we refer to this as the quasi-two-body (Q2B) approach. Effects due to the interference between the $f_0$ and the other resonances in the Dalitz plot are taken as systematic uncertainties.

The data we use in this analysis were accumulated with the BABAR detector [8] at the PEP-II asymmetric-energy $e^+e^-$ storage ring at SLAC. The data sample consists of an integrated luminosity of 111.2 fb$^{-1}$ collected at the Y(4S) resonance ("on-resonance") corresponding to $(122.6 \pm 0.7) \times 10^6$ $BB$ pairs, and 11.8 fb$^{-1}$ collected about 40 MeV below the Y(4S) ("off-resonance"). In Ref. [8] we describe the silicon vertex tracker and drift chamber used for track and vertex reconstruction, and the Čerenkov detector (DIRC), the electromagnetic calorimeter (EMC), and the instrumented flux return (IFR) used for particle identification.

If we denote by $\Delta t$ the difference between the proper times of the decay of the fully reconstructed $B^0 \rightarrow f_0 K^0_S$ ($B^0_{\text{rec}}$) and the decay of the other meson ($B^0_{\text{tag}}$), the time-dependent decay rate $f_{\text{tag}}$ is given by

$$f_{\text{tag}}(\Delta t) = \frac{e^{-|\Delta t|/\tau}}{4\tau} \left[ 1 + Q_{\text{tag}} S \sin(\Delta m_d \Delta t) - Q_{\text{tag}} C \cos(\Delta m_d \Delta t) \right],$$

where $Q_{\text{tag}} = 1(-1)$ when the tagging meson $B^0_{\text{tag}}$ is a $B^0(\bar{B}^0)$, $\tau$ is the mean $B^0$ lifetime, and $\Delta m_d$ is the $B^0\bar{B}^0$ oscillation frequency corresponding to the mass difference. The parameter $S$ is nonzero if there is mixing-induced CP violation while a nonzero value for $C$ would indicate direct CP violation.

We reconstruct $B^0 \rightarrow f_0(\rightarrow \pi^+\pi^-)K^0_S$ candidates from combinations of two tracks and a $K^0_S$ decaying to $\pi^+\pi^-$. For the $\pi^+\pi^-$ pair from the $f_0$ candidate, we use information from the tracking system, EMC, and DIRC to remove tracks consistent with electron, kaon, or proton hypotheses. In addition, we require at least one track to have a signature in the IFR that is inconsistent with the muon hypothesis. The mass of the $f_0$ candidate must satisfy $0.86 < m(\pi^+\pi^-) < 1.10$ GeV/$c^2$. To reduce combinatorial background from low energy pions, we require $|\cos(\theta(\pi^+))| < 0.9$, where $\theta(\pi^+)$ is the angle between the positive pion in the $f_0$ rest frame and the $f_0$ flight direction in the laboratory frame. The $K^0_S$ candidate is required to have a mass within 10 MeV/$c^2$ of the nominal $K^0$ mass [4] and a decay vertex separated from the $B^0$ decay vertex by at least 5 standard deviations. In addition, the cosine of the angle between the $K^0_S$ flight direction and the vector between the $f_0$ and the $K^0_S$ vertices must be greater than 0.99.

Two kinematic variables are used to discriminate between signal-$B$ decays and combinatorial background. One variable is the difference $\Delta E$ between the measured center-of-mass (c.m.) energy of the $B$ candidate and $\sqrt{s}/2$, where $\sqrt{s}$ is the c.m. energy. The other variable is the beam-energy-substituted mass $m_{\text{ES}} = \sqrt{\left| s/2 + p_B \cdot p_B \right|^2 - p_B^2}$, where the $B$ momentum $p_B$ and the four-momentum of the initial state ($E_i$, $p_i$) are defined in the laboratory frame. We require $5.23 < m_{\text{ES}} < 5.29$ GeV/$c^2$ and $|\Delta E| < 0.1$ GeV.

Continuum $e^+e^- \rightarrow q\bar{q}$ ($q = u, d, s, c$) events are the dominant background. To enhance discrimination between signal and continuum, we use a neural network (NN) to combine four variables: the cosine of the angle between the $B^0_{\text{rec}}$ direction and the beam axis in the c.m., the cosine of the angle between the thrust axis of the $B^0_{\text{rec}}$ candidate and the beam axis, and the zeroth and second angular moments.
$L_{0.2}$ of the energy flow about the $B_0^0$ thrust axis. The moments are defined by $L_j = \sum p_i \times |\cos \theta_j|$, where $\theta_j$ is the angle with respect to the $B_0^0$ thrust axis of the track or neutral cluster $i$ and $p_i$ is its momentum. The sum excludes the $B_0^0$ candidate. The NN is trained with off-resonance data and simulated signal events. The final sample of signal candidates is selected with a cut on the NN output that retains $\sim 97\%$ ($52\%$) of the signal (continuum).

The signal efficiency determined from Monte Carlo (MC) simulation is $(37.2 \pm 3.1)\%$. MC simulation shows that $4.7\%$ of the selected signal events are misreconstructed, mostly due to combinatorial background from low-momentum tracks used to form the $f_0$ candidate. In total, 7556 on-resonance events pass all selection criteria.

We use MC-simulated events to study the background from other $B$ decays. The charmed decay modes are grouped into eight classes with similar kinematic and topological properties. The modes that decay to the $\pi^+ \pi^- K^0_S$ final state are of particular importance since they have signal-like $\Delta E$ and $m_{\text{ES}}$ distributions and their decay amplitudes interfere with the $f_0 K^0_S$ decay amplitude. Among these modes are $\rho^0 K^0_S$, $f_0(1370)K^0_S$, $f_2(1270)K^0_S$, $K^{*+} \pi^-$ (including other kaon resonances decaying to $K^0_S \pi^+$), and nonresonant $\pi^+ \pi^- K^0_S$ decays. The inclusive charmed $\pi^+ \pi^- K^0_S$ branching fraction $(23.4 \pm 3.3) \times 10^{-6}$ [9], together with the available exclusive measurements [9,10], are used to infer upper limits on the branching fractions of these decays. Along with selection efficiencies obtained from MC, these branching fractions are used to estimate the expected background. The charmed decays $B^0 \to D^- \pi^+ \to K^0_S \pi^+ \pi^-$ and $B^+ \to \overline{D}^0 \pi^+ \to K^0_S \pi^+ \pi^-$ contribute significantly to the selected data sample. Each of these modes is treated as a separate class. Two additional classes account for the remaining neutral and charged $b \to c$ decays. In the selected data sample, we expect $45 \pm 15$ charmed and $128 \pm 74$ $b \to c$ events.

The time difference $\Delta t$ is obtained from the measured distance between the $z$ positions (along the beam direction) of the $B_{\text{rec}}$ and $B_{\text{tag}}$ decay vertices, and the boost $\beta \gamma = 0.56$ of the $e^+ e^-$ system [11,12]. To determine the flavor of the $B_{\text{tag}}$ we use the tagging algorithm of Ref. [12]. This produces four mutually exclusive tagging categories. We also retain untagged events in a fifth category to improve the efficiency of the signal selection.

We use an unbinned maximum likelihood fit to extract the $f_0 K^0_S$ event yield, the CP parameters defined in Eq. (1), and the $f_0$ resonance parameters. The likelihood function for the $N_k$ candidates tagged in category $k$ is

$$L_k = e^{-N_k} \prod_{i=1}^{N_k} \left[ N_{S\ell} \epsilon_{k,i} \mathcal{P}^S_{i,k} + N_{C\ell} \mathcal{P}^C_{i,k} + \sum N_{B\ell} \epsilon_{j,k} \mathcal{P}^B_{i,j,k} \right]^{n_k},$$

where $N_k$ is the sum of the signal, continuum, and the $n_k B$ background yields tagged in category $k$, $N_S$ is the number of $f_0 K^0_S$ signal events in the sample, $\epsilon_{k,i}$ is the fraction of signal events tagged in category $k$, $N_{C\ell}$ is the number of continuum-background events that are tagged in category $k$, and $N_{B\ell} \epsilon_{j,k}$ is the number of $B$-background events of class $j$ that are tagged in category $k$. The $B$-background event yields are fixed parameters, with the exception of the $D^- \pi^+$ yield. Since $B^0 \to D^- \pi^+$ events have a characteristic distribution in $\cos \theta(\pi^+)$, well separated from continuum and $f_0 K^0_S$ events, the $D^- \pi^+$ is free to vary in the fit along with the signal and continuum yields. The total likelihood $L$ is the product of the likelihoods for each tagging category.

The probability density functions (PDFs) $\mathcal{P}_k^S$, $\mathcal{P}_k^C$, and $\mathcal{P}_k^B$, for signal, continuum-background, and $B$-background class $j$, respectively, are the products of the PDFs of six discriminating variables. The signal PDF is thus given by $\mathcal{P}_k^S = \mathcal{P}^S(m_{\text{ES}}) \cdot \mathcal{P}^S(\Delta E) \cdot \mathcal{P}^S(\text{NN}) \cdot \mathcal{P}^S(|\cos \theta(\pi^+)|) \cdot \mathcal{P}^S(m(\pi^+ \pi^-)) \cdot \mathcal{P}^S(\Delta t)$, where $\mathcal{P}_k^S(\Delta t)$ contains the time-dependent CP parameters defined in Eq. (1), diluted by the effects of mistagging and the $\Delta t$ resolution.

The signal PDFs are decomposed into distinct distributions for correctly reconstructed and misreconstructed signal events. The fractions of misreconstructed signal events per tagging category are estimated by MC simulation. The $m_{\text{ES}}$, $\Delta E$, NN, $|\cos \theta(\pi^+)|$, and $m(\pi^+ \pi^-)$ PDFs for signal and $B$ background are taken from the simulation except for the means of the signal Gaussian PDFs for $m_{\text{ES}}$ and $\Delta E$, and the mass and width of the $f_0$, which are free to vary in the fit. We use a relativistic Breit-Wigner function to parametrize the $f_0$ resonance.

The $\Delta t$-resolution function for signal and $B$-background events is a sum of three Gaussian distributions, with parameters determined by a fit to fully reconstructed $B^0$ decays [12]. The continuum $\Delta t$ distribution is parameterized as the sum of three Gaussian distributions with two distinct means and three distinct widths, which are scaled by the $\Delta t$ per-event error. For the $B$-background modes that are $CP$ eigenstates, the parameters $C$ and $S$ are fixed to 0 and $\pm \sin 2\beta$, respectively, depending on their $CP$ eigenvalues. For continuum, four tag asymmetries and the five yields $N_{C\ell}$ are free. The signal yield, $S$, $C$, and the $f_0$ mass and width are among the 41 parameters that are free to vary in the fit. The majority of the free parameters are used to describe the shape of the continuum background.

The contributions to the systematic error on the signal parameters are summarized in Table I. To estimate the errors due to the fit procedure, we perform fits on a large number of MC samples with the proportions of signal, continuum, and $B$-background events measured from data. Biases of a few percent observed in these fits are due to imperfections in the likelihood model such as neglected correlations between the discriminating variables.
of the signal and $B$-background PDFs and are assigned as a systematic uncertainty of the fit procedure. The error due to the fit procedure includes the statistical error on the bias added in quadrature with the observed bias. The expected event yields from the $B$-background modes are varied according to the uncertainties in the measured or estimated branching fractions. Since $B$-background modes may exhibit $CP$ violation, the corresponding parameters are varied within their physical ranges. We vary the parameters of the $\Delta t$ model and tagging fractions incoherently within their errors and assign as a systematic error the quadratic sum of the observed changes in our measured parameters. The uncertainties due to the simulated signal PDFs are obtained from a control sample of fully reconstructed $B^0 \to D^- (\to K^0_S \pi^-) \pi^+$ decays. The systematic errors due to interference between the doubly Cabibbo-suppressed (DCS) $b \to \bar{u}cd$ amplitude with the Cabibbo-favored $b \to \bar{c}ud$ amplitude for tag-side $B$ decays have been estimated from simulation by varying freely all relevant strong phases [13]. The errors associated with $\Delta m_d$ and $\tau$ are estimated by varying these parameters within the errors on the world average [4].

The systematic error introduced in the Q2B approximation by ignoring interference effects between the $f_0$ and the other resonances in the Dalitz plot (as listed earlier) is estimated from simulation by varying freely all relative strong phases and taking the largest observed change in each parameter as the error. While the systematic errors due to interference are comparable to the statistical error for the branching fraction and the $f_0$ mass and width, they are small compared to the statistical error for $S$ and $C$.

The maximum likelihood fit results in the $f_0K_S^0$ event yield $N_S = 93.6 \pm 13.6 \pm 6.3$, where the first error is statistical and the second systematic. The branching fraction corresponding to this yield is $B(f_0K_S^0) = (6.1 \pm 1.0 \pm 0.5) \times 10^{-6}$.

![Figure 1](image_url)

**Figure 1.** Distributions of (clockwise from bottom left) NN, $m(\pi^+\pi^-)$, $\Delta E$, $m_{ES}$, and $|\cos(\theta(\pi^+))|$ for samples enhanced in the $f_0K_S^0$ signal. For presentation purposes, the region $0.765 < |\cos(\theta(\pi^+))| < 0.81$ has been removed to suppress the contribution from $D^- \pi^+$ events. The bottom-right plot shows the distribution of the ratio of the signal likelihood to the total likelihood for all events entering the fit. In each plot, the solid curve represents a projection of the maximum likelihood fit result, the dashed curve represents the contribution from continuum events, and the dotted line indicates the combined contributions from continuum and $B$-background events.

---

**TABLE I.** Summary of systematic uncertainties. The uncertainties on $m_{f_0}$ and $f_0$ are in units of MeV/c$^2$.

<table>
<thead>
<tr>
<th>Error source</th>
<th>$S$</th>
<th>$C$</th>
<th>$B \times 10^{-6}$</th>
<th>$m_{f_0}$</th>
<th>$f_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fitting procedure</td>
<td>0.06</td>
<td>0.07</td>
<td>0.26</td>
<td>0.5</td>
<td>1.0</td>
</tr>
<tr>
<td>$B$ background</td>
<td>0.05</td>
<td>0.05</td>
<td>0.30</td>
<td>0.2</td>
<td>3.0</td>
</tr>
<tr>
<td>$\Delta t$ model</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
<td>0.0</td>
<td>0.1</td>
</tr>
<tr>
<td>Tagging fraction</td>
<td>0.04</td>
<td>0.02</td>
<td>0.01</td>
<td>0.0</td>
<td>0.2</td>
</tr>
<tr>
<td>Signal model</td>
<td>0.02</td>
<td>0.01</td>
<td>0.03</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>DCS decays</td>
<td>0.01</td>
<td>0.04</td>
<td>0.00</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>$\Delta m_d$ and $\tau$</td>
<td>0.01</td>
<td>0.00</td>
<td>0.01</td>
<td>0.0</td>
<td>0.1</td>
</tr>
<tr>
<td>Tracking and PID</td>
<td>0.00</td>
<td>0.00</td>
<td>0.49</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Subtotal</td>
<td>0.09</td>
<td>0.10</td>
<td>0.63</td>
<td>0.5</td>
<td>3.2</td>
</tr>
<tr>
<td>Q2B approximation</td>
<td>0.04</td>
<td>0.07</td>
<td>1.21</td>
<td>4.0</td>
<td>8.5</td>
</tr>
</tbody>
</table>
The time-dependent distributions and asymmetry in quasi-two-body approximation. The systematic error in the data and the model. Also presented in Fig. 1, showing good agreement between likelihood to total likelihood for all events entering the convergence. The distribution of the ratio of signal-like variables for both data and the likelihood model after continuum likelihood ratios of the other discriminating variables, we obtain

\[ \mathcal{B}(B^0 \rightarrow f_0(980)K^0) \times \mathcal{B}(f_0(980) \rightarrow \pi^+ \pi^-) \]

\[ = (6.0 \pm 0.9 \pm 0.6 \pm 1.2) \times 10^{-6}, \]

where the first error is statistical, the second systematic, and the third accounts for the model dependence in the quasi-two-body approximation. The systematic error includes an uncertainty of 8.2% from differences between data and MC in tracking, particle identification (PID), and \( K_S^0 \) detection efficiencies. Figure 1 shows distributions of \( \Delta E, m_{ES}, |\cos(\theta^+)|, m(\pi^+ \pi^-), \) and NN that are enhanced in signal content by cuts on the signal-to-continuum likelihood ratios of the other discriminating variables for both data and the likelihood model after fit convergence. The distribution of the ratio of signal likelihood to total likelihood for all events entering the fit is also presented in Fig. 1, showing good agreement between the data and the model.

For the CP-violation parameters, we obtain

\[ S = -1.62^{+0.56}_{-0.51} \pm 0.09 \pm 0.04, \]

\[ C = 0.27 \pm 0.36 \pm 0.10 \pm 0.07. \]

The time-dependent distributions and asymmetry \( A_{B^0} = (N_{B^0} - N_{\bar{B}^0})/(N_{B^0} + N_{\bar{B}^0}) \) in the tagged events are represented in Fig. 2. The model-dependent mass and width of the \( f_0 \) are found to be

\[ m_{f_0} = (980.6 \pm 4.1 \pm 0.5 \pm 4.0) \text{ MeV}/c^2, \]

\[ \Gamma_{f_0} = (43^{+12}_{-9} \pm 3 \pm 9) \text{ MeV}/c^2. \]

These results are in agreement with previous mass and width measurements [4,14].

In summary, we have presented the first observation of \( B^0 \rightarrow f_0(980)K^0 \) and measurements of the branching fraction, resonance parameters, and CP-violating asymmetries in \( B^0 \rightarrow f_0(980)(\rightarrow \pi^+ \pi^-)K^0 \) decays. As determined from a large number of simulated experiments, our results for \( S \) and \( C \) are consistent with the standard model within 1.7 and 0.8 standard deviations, respectively. The result for \( S \) is 1.2 standard deviations from the physical limit, and the hypothesis of no mixing-induced CP violation is excluded within 2.7 standard deviations.

We are grateful for the excellent luminosity and machine conditions provided by our PEP-II colleagues, and for the substantial dedicated effort from the computing organizations that support BABAR. The collaborating institutions wish to thank SLAC for its support and kind hospitality. This work is supported by DOE and NSF (U.S.A.), NSERC (Canada), IHEP (China), CEA and CNRS-IN2P3 (France), BMBF and DFG (Germany), INFN (Italy), FOM (The Netherlands), NFR (Norway), MIST (Russia), and PPARC (United Kingdom). Individuals have received support from CONACyT (Mexico), A.P. Sloan Foundation, Research Corporation, and Alexander von Humboldt Foundation.

FIG. 2. The signal-enhanced time distributions tagged as \( B^0 \) (top) and \( \bar{B}^0 \) (middle), and the asymmetry, \( A_{B^0} \) (bottom). The solid curve is a projection of the fit result. The dashed line is the distribution for continuum background and the dotted line is the total \( B \)- and continuum-background contribution.