Search for the decay $B^0 \to K_S^0 K_S^0 L_0$
SEARCH FOR THE DECAY $B^0 \rightarrow K^0_S K^0_L$

University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom

University of Iowa, Iowa City, Iowa 52242, USA

Iowa State University, Ames, Iowa 50011-3160, USA

Johns Hopkins University, Baltimore, Maryland 21218, USA

Universität Heidelberg, Physikalisches Institut, Philosophenweg 12, D-69120 Heidelberg, Germany

Imperial College London, London, SW7 2AZ, United Kingdom

University of Liverpool, Liverpool L69 7ZE, United Kingdom

Queen Mary, University of London, E1 4NS, United Kingdom

University of London, Royal Holloway and Bedford New College, Egham, Surrey TW20 0EX, United Kingdom

University of Louvain, Louvain, Belgium

University of Manchester, Manchester M13 9PL, United Kingdom

University of Maryland, College Park, Maryland 20742, USA

University of Massachusetts, Amherst, Massachusetts 01003, USA

Massachusetts Institute of Technology, Laboratory for Nuclear Science, Cambridge, Massachusetts 02139, USA

McGill University, Montréal, Québec, Canada H3A 2T8

Dipartimento di Fisica and INFN, Università di Milano, I-20133 Milano, Italy

University of Mississippi, University, Mississippi 38677, USA

Université de Montréal, Physique des Particules, Montréal, Québec, Canada H3C 3J7

Mount Holyoke College, South Hadley, Massachusetts 01075, USA

Dipartimento di Scienze Fisiche and INFN, Università di Napoli Federico II, I-80126 Napoli, Italy

NIKHEF, National Institute for Nuclear Physics and High Energy Physics, NL-1009 DB Amsterdam, The Netherlands

University of Notre Dame, Notre Dame, Indiana 46556, USA

Ohio State University, Columbus, Ohio 43210, USA

University of Oregon, Eugene, Oregon 97403, USA

Université de Padova, Dipartimento di Fisica e INFN, I-35131 Padova, Italy

Universités Paris VI et VII, Laboratoire de Physique Nucléaire et de Hautes Énergies, F-75252 Paris, France

University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

Dipartimento di Fisica and INFN, Università di Perugia, I-06100 Perugia, Italy

Dipartimento di Fisica, Scuola Normale Superiore and INFN, Università di Pisa, I-56127 Pisa, Italy

Prairie View A&M University, Prairie View, Texas 77446, USA

Princeton University, Princeton, New Jersey 08544, USA

Università di Roma La Sapienza, Dipartimento di Fisica and INFN, I-00185 Roma, Italy

University of Rochester, D-18051 Rochester, Germany

Rutherford Appleton Laboratory, Chilton, Didcot, Oxford, OX11 0QX, United Kingdom

DSM/Dapnia, CEA/Saclay, F-91191 Gif-sur-Yvette, France

University of South Carolina, Columbia, South Carolina 29208, USA

Stanford Linear Accelerator Center, Stanford, California 94309, USA

Stanford University, Stanford, California 94305-4060, USA

State University of New York, Albany, New York 12222, USA

University of Tennessee, Knoxville, Tennessee 37996, USA

University of Texas at Austin, Austin, Texas 78712, USA

University of Texas at Dallas, Richardson, Texas 75083, USA

Dipartimento di Fisica Sperimentale and INFN, Università di Torino-I10125 Torino, Italy

INFN, Università di Trieste, I-34127 Trieste, Italy

IFIC, Universitat de Valencia-CSIC, E-46071 Valencia, Spain

University of Victoria, Victoria, British Columbia, Canada V8W 3P6

Department of Physics, University of Warwick, Coventry CV4 7AL, United Kingdom

Also at Laboratoire de Physique Corpusculaire, Clermont-Ferrand, France

† Also with Università di Perugia, Dipartimento di Fisica, Perugia, Italy

‡ Also with Università della Basilicata, Potenza, Italy
We present the first search for the decay $B^0 \to K_S^0 K_S^0 K_L^0$ using a data sample of $232 \times 10^6 \, B \bar{B}$ pairs. We find no statistically significant evidence for the nonresonant component of this decay. Our central value for the branching fraction, assuming the true Dalitz distribution is uniform and excluding the $\phi$ resonance, is $B(B^0 \to K_S^0 K_S^0 K_L^0) = (2.4^{+2.7}_{-2.5} \pm 0.6) \times 10^{-6}$ where the errors are statistical and systematic, respectively.

We set a single-sided Bayesian upper limit of $B(B^0 \to K_S^0 K_S^0 K_L^0) < 7.4 \times 10^{-6}$ at 90% confidence level using a uniform prior probability for physical values. Assuming the worst-case true Dalitz distribution, where the signal is entirely in the region of lowest efficiency, the 90% confidence level upper limit is $B(B^0 \to K_S^0 K_S^0 K_L^0) < 16 \times 10^{-6}$.

DOI: 10.1103/PhysRevD.74.032005

PACS numbers: 13.25.Hw, 11.30.Er, 12.15.Hh

I. INTRODUCTION

Measurements from the BABAR and Belle experiments have confirmed the Cabibbo-Kobayashi-Maskawa quark-mixing matrix (CKM) mechanism [1] as the dominant source of $CP$ violation in flavor-changing weak interactions [2]. In the standard model, the mixing-induced, time-dependent $CP$ asymmetry in $B^0$ decays from $b \to s\bar{q}q$ penguin transitions should be the same as the precisely measured $CP$ asymmetry in charm-$K^0$ decays, namely, $\sin 2 \beta$, to within a few percent. New physics from higher mass scales may contribute to the loop in the penguin diagram, which could significantly alter the $CP$ asymmetry in penguin-dominated $B$ decays [3]. Initial $CP$ asymmetry measurements in $b \to s\bar{q}q$ penguin $B$ decays have suggested a possible violation of this test of the standard model [4–6].

The $b \to s\bar{q}q$ penguin decays fall into two categories. If the $\bar{q}q$ can be $\bar{u}u$, a CKM-suppressed tree-level $b \to u$ transition can contribute to the decay in addition to the dominant $b \to s\bar{q}q$ penguin. This introduces some uncertainty in the standard model prediction of the $CP$ asymmetry, since the $b \to u$ and the penguin amplitudes have different weak phases. On the other hand, decays such as $B^0 \to K_S^0 K_S^0 K_L^0$ [5] and $B^0 \to K_S^0 K_L^0 K_L^0$ are purely $b \to s\bar{s}s$ penguin transitions and they can only include $b \to u$ decay amplitudes through rescattering, thus the standard model uncertainty on the predicted $CP$ asymmetry for these decays is generally smaller [7].

It has been noted that three-body $B^0$ decays of the form $B^0 \to \Sigma^0 \Pi^0$, where $\Sigma$ and $\Pi$ are spin-0 $CP$ eigenstate neutral particles, are $CP$ eigenstates [8], thus the $CP$ asymmetry in $B^0 \to K_S^0 K_S^0 K_L^0$ is not diluted, as is generally the case in three-body $B^0$ decays. The $B^0 \to K_S^0 K_S^0 K_L^0$ would be a valuable addition to understanding the $b \to s\bar{q}q$ penguin $CP$ asymmetry anomaly, if the branching fraction is large enough. The resonant $\phi K_S^0$ contribution to this final state, neglecting interference effects, would not yield a sample of signal events large enough to make an interesting $CP$ asymmetry measurement. In the analysis described below, the expected and observed number of $K_S^0 K_S^0 K_L^0$ events with a $K_S^0 K_L^0$ invariant mass in the region of the $\phi$ resonance is about eight events. Prior to our search, the nonresonant component of the $K_S^0 K_S^0 K_L^0$ final state had not been experimentally investigated. This search was motivated by the possibility of discovering a large nonresonant contribution to the $K_S^0 K_S^0 K_L^0$ final state. Large nonresonant $S$-wave amplitudes have been found in the Dalitz amplitude analysis of $B^+ \to K^+ K^- K^+$ [9] and $B^0 \to K^+ K^- K_S^0$ [10]. The branching fraction was recently estimated to be [7] $B(B^0 \to K_S^0 K_S^0 K_L^0) = (5.23 \pm 2.52 \pm 6.86^{+0.05}_{-0.06}) \times 10^{-6}$ including the $\phi$ resonance. The difference between the upper and lower limits of this estimate is substantial. Another prediction, based on isospin and Bose symmetry [11], gives a somewhat small branching fraction of $B(B^0 \to K_S^0 K_S^0 K_L^0) = (1/3) B(B^0 \to K_S^0 K_S^0 K_S^0) = 2 \times 10^{-6}$. 

II. THE BABAR DETECTOR AND DATA SET

The results presented here are based on data collected with the BABAR detector [12] at the PEP-II asymmetric $e^+e^-$ collider [13] located at the Stanford Linear Accelerator Center. An integrated luminosity of 211 fb$^{-1}$, corresponding to $232 \times 10^6 \, B \bar{B}$ pairs, was recorded at the $Y(4S)$ resonance (center-of-mass energy $\sqrt{s} = 10.58$ GeV).

Charged particles from the $e^+e^-$ interactions are detected and their momenta are measured using five layers of double-sided silicon microstrip detectors and a 40-layer drift chamber, both operating in the 1.5 T magnetic field of a superconducting solenoid. Photons, electrons, and hadronic showers from $K_L^0$ interactions are identified with a CsI(Tl) electromagnetic calorimeter (EMC). Further charged particle identification is provided by the specific ionization ($dE/dx$) in the tracking devices and by an internally reflecting ring imaging Cherenkov detector covering the central region. The steel of the magnetic-flux return for the superconducting solenoid is instrumented with resistive plate chambers (instrumented flux return, or IFR), which are used to identify muons and hadronic showers from $K_L^0$ interactions.
III. ANALYSIS METHOD

A. Daughter candidate selection

We reconstruct $K^0_L$ candidates through the decay $K^0_S \rightarrow \pi^+ \pi^-$ only. We begin by forming all oppositely charged combinations of reconstructed tracks in the event. The invariant mass is required to be in the range of 490–506 MeV/$c^2$, assuming the tracks are pions. In addition, the $\chi^2$ probability of the $K^0_S$ vertex fit must be greater than 1% and the transverse component of the decay length is required to be greater than 8 mm. Finally, the angle between the $K^0_S$ flight direction and the $K^0_L$ momentum vector must be less than 0.2 radians.

We reconstruct $K^0_L$ candidates by identifying clusters of energy in the EMC and hits in the IFR that are isolated from all charged tracks in the event. The $K^0_L$ candidates based on IFR clusters must have hits in at least two of the 16 to 19 layers in the detector. The $K^0_L$ candidates detected by a hadronic shower in the EMC must have a calorimeter energy of at least 200 MeV, where this energy is from interpreting the calorimeter signal as an electromagnetic shower. Clusters in the EMC that are consistent with photons from $\pi^0$ decay are vetoed by two methods. If the $K^0_L$ candidate EMC cluster forms an invariant mass in the range of 100–150 MeV/$c^2$ with another EMC cluster with a calorimeter energy of at least 100 MeV, it is rejected. If the $K^0_L$ candidate cluster has a calorimeter energy greater than 1.0 GeV and two local maxima (or two “bumps”) in the spatial distribution of the energy within the cluster, it may be from a high-energy $\pi^0$, where the electromagnetic showers from the two photons are merged into one cluster. If the two bumps within the cluster form an invariant mass greater than 110 MeV/$c^2$, the candidate is rejected.

Further background rejection for EMC $K^0_L$ candidates is achieved by using a neural network trained on signal, $q\bar{q}$ continuum (where $q = u, d, s, c$), and $B\bar{B}$ Monte Carlo samples to distinguish $K^0_L$ clusters from fake clusters. The neural network inputs are the EMC cluster energy and the following six shower-shape variables:

(i) The lateral moment LAT of the shower energy deposition [14] defined as

$$\text{LAT} = \sum_{i=3}^{n} E_i r_i^2 / (E_1 r_1^2 + E_2 r_2^2 + \sum_{i=3}^{n} E_i r_i^2)$$

where the $n$ crystals in the EMC cluster are ranked in order of deposited energy ($E_i$), $r_0 = 5$ cm is the average distance between crystal centers, and $r_i$ is the radial distance of crystal $i$ from the cluster center.

(ii) The second radial moment of the shower energy deposition, defined as $\sum E_i r_i^2$ where $r_i$ is the radial distance of crystal $i$ from the cluster center.

(iii) The energy sum of a $3 \times 3$ block of crystals, centered on the single crystal with the most energy, divided by the larger $5 \times 5$ block, also centered in the same way.

(iv) The energy of the most energetic crystal in the cluster divided by the energy sum of the $3 \times 3$

crystal block with the most energetic crystal in the center.

(v) The Zernike moments [15] $A_{2,0}$ and $A_{4,2}$ defined below.

The Zernike moment $A_{n,m}$ is defined as

$$A_{n,m} = \left| \sum_{i} E_i f_{n,m}(r_i/R_0)e^{im\phi_i} \right|,$$

with

$$f_{2,0}(x) = 2x^2 - 1 \quad \text{and} \quad f_{4,2}(x) = 4x^4 - 3x^2,$$

where $r_i$ and $\phi_i$ are the radial and angular separation of crystal $i$ with respect to the cluster center, $E_{\text{tot}}$ is the total cluster energy, and $R_0$ is a cutoff radius of 15 cm. Only $K^0_L$ candidates that pass an optimized cut on the neural network output are retained. This cut has a signal efficiency of 85% and rejects 70% of the EMC fake $K^0_L$ background.

B. $B^0$ candidate selection

We reconstruct $B^0$ candidates from selected $K^0_L$ clusters and pairs of $K^0_S$ candidates that do not share any tracks. We require the sum of the $K^0_S$ momentum magnitudes in the center-of-mass frame to be at least 2.1 GeV/$c$, which ensures consistency with the kinematics of a $B^0 \rightarrow KKK$ decay. Only the direction of the $K^0_S$ is reconstructed, from the vector defined by the primary vertex and the center of the neutral cluster. We compute the $K^0_L$ momentum by constraining the $K^0_S K^0_L K^0_L$ invariant mass to be the known mass of the $B^0$ [16]. We reject fake $K^0_L$ candidates by using the difference between the $K^0_L$ transverse momentum, computed from the $B^0$ mass constraint, and the transverse momentum along the $K^0_L$ direction that is missing from the event. The reconstructed event missing transverse momentum is calculated without using the $K^0_L$ cluster in the momentum sum and projected along the computed direction of the $K^0_L$. This missing transverse momentum difference (reconstructed minus calculated) is required to be greater than $-0.5$ GeV/$c$. This requirement and the previously mentioned requirement on the EMC $K^0_S$ neural network output were simultaneously optimized to give the greatest signal significance $S/\sqrt{S + B}$, where $S$ and $B$ are the expected number of signal and background events, assuming a signal branching fraction of $5 \times 10^{-6}$ [7].

The difference in energy $\Delta E$ between the reconstructed $B^0$ candidate and the beam energy in the center-of-mass frame is the main variable used to distinguish properly reconstructed signal events from combinatoric background. The missing transverse momentum difference distributions for the signal Monte Carlo sample and the background-dominated $\Delta E$ sideband ($\Delta E > 0.010$ GeV in data) are shown in Fig. 1.

We distinguish the nonresonant three-body $B^0$ decay from two-body $B^0$ decays to the same final state in this
analysis. There are four two-body $B^0$ decays through charmonium that can go to the final state $K_S^0 K_S^0 K_L^0$, namely, $J/\psi K_S^0$, $\chi_{c0} K_L^0$, $\chi_{c2} K_L^0$, and $\psi(2S) K_S^0$. These are particularly unwanted, since they are from a color-suppressed tree decay amplitude, not the $b \rightarrow s q \bar{q}$ penguin decay amplitude. The $\chi_c$ modes have unknown $B^0$ branching fractions. We veto $B^0$ candidates consistent with $B^0 \rightarrow \chi_{c0} K_L^0$ or $B^0 \rightarrow \chi_{c2} K_L^0$, where the $\chi_c$ decays to $K_S^0 K_S^0$, by removing candidates with a $K_S^0 K_S^0$ invariant mass in the range $3.400-3.429$ GeV/$c^2$ or $3.540-3.585$ GeV/$c^2$, respectively. The combined $B^0$ and daughter branching fractions for the $J/\psi K_S^0$ and $\psi(2S) K_S^0$ modes are 0.062 and 0.016 $\times$ 10$^{-6}$, respectively. We expect about two events from two-body $B^0$ decays through charmonium in our final sample.

We also remove $B^0$ candidates consistent with $B^0 \rightarrow \phi K_S^0$ with $\phi \rightarrow K_S^0 K_L^0$ by requiring the invariant mass of both $K_S^0 K_L^0$ combinations to be above 1.049 GeV/$c^2$, though as a cross-check, we measure the branching fraction in only the $\phi K_S^0$ region, where one $K_S^0 K_L^0$ combination has an invariant mass less than 1.049 GeV/$c^2$. These two-body $B^0$ decay vetoes are 96.6% efficient for the $B^0 \rightarrow K_S^0 K_S^0 K_L^0$ signal Monte Carlo sample generated with a uniform true Dalitz distribution.

C. Event selection

The main source of background is from random $K_S^0 K_L^0 K_L^0$ combinations from continuum $q\bar{q}$ events. We combine three variables in a neural network that is trained using Monte Carlo samples to distinguish signal from continuum events. The first input is the cosine of the angle of the $B^0$ momentum with respect to the beam axis in the center-of-mass frame. The other two inputs are topological variables that are commonly used to distinguish jetlike continuum events from the more isotropic particle distributions in $B\bar{B}$ events.

The first is the cosine of the angle $\theta_T$ between the thrust axis of the $B^0$ candidate in the center-of-mass frame and that of the rest of the charged tracks and neutral calorimeter clusters in the event. The distribution of $\cos \theta_T$ is sharply peaked near 1.0 for combinations drawn from jetlike $q\bar{q}$ pairs and nearly uniform for $B^0$ meson decays. The second is a linear combination of the zeroth and second angular moments $L_{0,2}$ of the energy flow about the $B^0$ thrust axis.

After all of the requirements stated thus far, if an event has more than one $B^0$ candidate, we choose the best one by selecting on the quality of the $K_S^0$ cluster. If there are one or more EMC $K_L^0$ candidates, the best $K_L^0$ candidate is the one with the highest cluster calorimeter energy. If there are no EMC $K_L^0$ candidates, the best $K_L^0$ candidate is the one with the highest number of IFR layers with hits in the $K_L^0$ cluster. If there is more than one $K_S^0$ pair that uses the same (best) $K_S^0$ cluster, the best $B^0$ candidate is the one with the lowest $K_S^0$ mass $\chi^2$, defined below:

$$\chi^2 = \left( \frac{\delta m_i}{\sigma} \right)^2 + \left( \frac{\delta m_2}{\sigma} \right)^2,$$

where $\delta m_i$ is the difference between the reconstructed invariant mass of the $K_S^0$ candidate $i$ and the known $K_S^0$ mass [16], and $\sigma$ is the invariant mass resolution (2.9 MeV/$c^2$).

Our final analysis sample contains 5892 events with $\Delta E$ in the range of $-0.010-0.080$ GeV. Signal events appear mostly in the range of $-0.010-0.010$ GeV. The

FIG. 1. Missing transverse momentum difference distributions for the signal Monte Carlo sample (solid) and the background-dominated $\Delta E$ sideband ($\Delta E > 0.010$ GeV) after all other selection criteria (see text) have been applied. The signal Monte Carlo sample distribution has been arbitrarily normalized to the number of entries in the sideband distribution.

FIG. 2. Neural network output (NN) distributions for the signal Monte Carlo sample (solid) and the background-dominated $\Delta E$ sideband ($\Delta E > 0.010$ GeV) after all other selection criteria (see text) have been applied. The signal Monte Carlo sample distribution has been arbitrarily normalized to the number of entries in the sideband distribution.
The ΔE > 0.010 GeV region is dominated by combinatoric background. The signal efficiency varies from 4% to 14% depending on the position in the Dalitz plot. The signal efficiency as a function of the Dalitz plot position and a histogram of the reconstruction efficiency for 322 uniformly distributed points in the Dalitz plot are shown in Figs. 3 and 4, respectively. For a uniform true Dalitz distribution, the average signal efficiency is 8.1%.

D. $J/\psi K^0_L$ control sample

We use a sample of $B^0 \rightarrow J/\psi K^0_L$ decays from the data to calibrate the reconstruction and selection efficiency and the ΔE resolution. The $J/\psi$ is reconstructed in the $e^+e^-$ and $\mu^+\mu^-$ channels. We apply the same $K^0_L$ selection, projected missing transverse momentum difference, and NN criteria as for our $K^0_SK^0_SK^0_L$ selection. The ratio of the $K^0_SK^0_SK^0_L$ and $J/\psi K^0_L$ Monte Carlo efficiencies for these criteria is $e_{K^0_SK^0_SK^0_L}/e_{J/\psi K^0_L} = 0.96$. We compare the number of fitted $J/\psi K^0_L$ events to the predicted number of events, based on the known branching fractions and the Monte Carlo efficiency. The $J/\psi K^0_L$ reconstruction efficiency for all selection criteria is 12.5%. We find $N_{\text{obs}} = 1420 \pm 56$ $J/\psi K^0_L$ events, consistent with the predicted yield of $N_0 = 1426 \pm 86$. We correct the $K^0_SK^0_SK^0_L$ efficiency by multiplying the MC efficiency by a correction factor of $0.96 \pm 0.08$, which is defined as $(N_{\text{obs}}/N_0) \times (e_{K^0_SK^0_SK^0_L}/e_{J/\psi K^0_L})$. The uncertainty on the correction factor includes the uncertainties on the relevant branching fractions for $J/\psi K^0_L$, the statistical error from the $J/\psi K^0_L$ fit, and our estimated uncertainty on relating the $J/\psi K^0_L$ and $K^0_SK^0_SK^0_L$ selection and reconstruction efficiencies.

The ΔE resolution is better for EMC candidates because the position of the hadronic shower is more precisely measured. We also use the $J/\psi K^0_L$ to measure the signal ΔE resolution separately for $K^0_L$ candidates reconstructed in the EMC and IFR. Figure 5 shows the fitted ΔE distributions for EMC and IFR $K^0_L$ candidates.

![FIG. 3](image1) The reconstruction efficiency for signal as a function of position in the symmetrized Dalitz plot.

![FIG. 4](image2) A histogram of the reconstruction efficiency for signal for 322 uniformly distributed points in the Dalitz plot.

![FIG. 5](image3) (color online). Fitted ΔE distributions of the $J/\psi K^0_L$ control sample for EMC (top) and IFR (bottom) $K^0_L$ candidates. The points with error bars are histograms of the data sample. The solid curve is the total PDF. The dotted-dashed curve is the signal PDF and the dashed curve is the background PDF.
E. Signal yield determination

We use an extended unbinned maximum likelihood fit to determine the number of signal events in our final sample. The likelihood for an event is the product of probability density functions (PDFs) for the two main discriminating variables in the analysis: $\Delta E$ and $NN$. Separate PDFs are used for EMC and IFR candidates due to the difference in $\Delta E$ resolution and the fact that some background channels mostly produce fake signal for EMC $K^0_L$ candidates only. The EMC and IFR $K^0_L$ samples are fitted simultaneously and the relative fraction of EMC signal events is taken from the $J/\psi K^0_L$ control sample. Separate PDFs also are used for signal, combinatoric background, and three classes of background from $B$ decays similar to our signal (“peaking backgrounds” described below in Section III F). The combinatoric background includes random $K^0_SK^0_SK^0_L$ combinations from $B\bar{B}$ events as well as from continuum events.

Figure 6 shows the $\Delta E$ and $NN$ PDFs for all five fit components. The numbers of signal and combinatoric background events are determined from the fit. The peak- ing background yields are fixed to values based on known and estimated branching fractions and then varied in the evaluation of systematic uncertainties.

The functional form of the signal PDF for $\Delta E$ is a triple Gaussian distribution. The mean and width of the core Gaussian distribution are determined separately for EMC and IFR $K^0_L$ candidates from the $J/\psi K^0_L$ control sample and held fixed in the $K^0_SK^0_SK^0_L$ fit. The remaining Gaussian parameters for the signal $\Delta E$ PDF are determined from the signal Monte Carlo sample and held fixed. The signal PDF for $NN$ is a fourth-order polynomial. The polynomial coefficients are determined from the signal Monte Carlo sample and held fixed in the fit. The combinatoric background $\Delta E$ PDF is an ARGUS function [17]. The $NN$ PDF for combinatoric background is the sum of a first-order polynomial and an ARGUS function. The PDF shape parameters for the combinatoric background component are free in the fit.

F. Backgrounds from other $B$ decays

Backgrounds from $B$ decays to final states similar to $K^0_SK^0_SK^0_L$ can look similar to our signal in our discriminating variables $\Delta E$ and $NN$. We call events from these decays “peaking backgrounds.” The largest single source of peaking background is the decay $B^0 \rightarrow K^0_SK^0_SK^0_L$. One of the $K^0_S$ can decay to $\pi^0\pi^0$, where one or more of the photons from either $\pi^0$ fakes the $K^0_L$ cluster in the EMC. The world average $K^0_SK^0_SK^0_S$ branching fraction from the Heavy Flavor Averaging Group (HFAG) [5,6] is $B(B^0 \rightarrow K^0_SK^0_SK^0_S) = (6.2 \pm 0.9) \times 10^{-6}$. We expect $20 K^0_SK^0_SK^0_S$ events in our fit sample and vary this number by $\pm 5$ events in the evaluation of the systematic errors.

Decays of the type $KKK^*$ can produce a $K^0_SK^0_SK^0_L\pi$ final state, which can look similar to our signal if the momentum of the additional $\pi$ is low. The branching fractions for the relevant $KKK^*$ have not been measured. We assume they are each of the same order as our expected signal, based on comparing the relative branching ratios of $B \rightarrow K\pi$ decays with $B \rightarrow K^+\pi$ and $B \rightarrow K\rho$ decays. We estimate a total branching fraction of $30 \times 10^{-6}$ for this inclusive final state and vary this assumption by $\pm 15 \times 10^{-6}$ in the evaluation of systematic errors. This gives us an expected $70 \pm 35$ events from $KKK^*$ in the fit sample. In the evaluation of the 90% C. L. upper limit on the branching fraction, we set the $KKK^*$ event yield to zero. This is the
most conservative assumption, since the signal and $K^0 K^+ K^-$ event yields are anticorrelated in the fit.

Finally, the decay $B^0 \rightarrow K^0_SK^0_L\pi^0$ can also mimic our signal if the $\pi^0$ is misidentified as a $K^0_S$ in the EMC. The branching fraction for this mode has also not been measured. We estimate the branching fraction, relative to our signal branching fraction are listed in Table II. The additive contributions to the systematic error come from the uncertainties on the charmonium background subtraction (1.1 events). The multiplicative errors are subtracted from the fitted signal yield.

We find $2.1\pm0.3$ events from $K^0_SK^0_L\pi^0$ in the fit sample.

The $\Delta E$ and $NN$ PDFs for the three peaking background components above are determined from Monte Carlo samples. In general, the $\Delta E$ shape peaks near $\Delta E = 0$, though there is a substantial tail that extends to high $\Delta E$ values. We use the sum of an ARGUS function and a Landau function for the peaking background $\Delta E$ PDF. The $NN$ shape is quite similar to the signal shape for all three peaking backgrounds. The form of the $NN$ PDF is a polynomial. The $\Delta E$ and $NN$ PDFs for the three peaking background components are shown in Figs. 6(c)–6(e).

IV. RESULTS

Table I lists the results of the maximum likelihood fit. We find $23^{+23}_{-29}\pm6$ events for the $K^0_SK^0_LK^0_L$ signal yield, where the first errors are statistical and the second error is systematic. The maximum likelihood fit bias of $+0.3$ signal events was evaluated from an ensemble of data sets composed of fully simulated signal and $B$ background Monte Carlo events. The combinatoric background for these data sets was generated from the PDF parameters of the fit to the data. This average bias of 0.3 events and the expected 2.1 events from $B^0$ decays through charmonium are subtracted from the fitted signal yield.

The systematic errors on the fitted signal yield and the signal branching fraction are listed in Table II. The multiplicative contributions to the systematic error come from the uncertainty on our estimation of the $K^+ K^- K^0$ component normalization (5.2 events), varying the yields of the other fixed peaking background components and the fixed PDF parameters (2.6 events), our uncertainty on the fit bias correction (0.2 events), and the uncertainty on the charmonium background subtraction (1.1 events). The multiplicative contributions are from the uncertainties on the $K^0_S$ and $K^0_L$ reconstruction efficiencies (6.8% and 8.0%), the number of $BB$ events in the sample (1.1%), and the $K^0_S \rightarrow \pi^+ \pi^-$ branching fraction.

Figure 7 shows the fitted distributions of $\Delta E$ and $NN$, where a cut has been made on the fit variable not shown (e.g. there is an $NN$ cut applied for the $\Delta E$ plot) to enhance the signal. The signal efficiency of this cut is 57% for the $\Delta E$ distribution and 71% for the $NN$ distribution.

<table>
<thead>
<tr>
<th>Mode</th>
<th>$K^0_S K^0_S K^0_L$</th>
<th>$\phi K^0_S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Events to fit</td>
<td>5892</td>
<td>210</td>
</tr>
<tr>
<td>Signal yield</td>
<td>$23^{+23}_{-29}\pm6$</td>
<td>8.3$^{+2.3}_{-2.5}$</td>
</tr>
<tr>
<td>Combinatoric background</td>
<td>$7777 \pm 79$</td>
<td>202 $\pm 15$</td>
</tr>
<tr>
<td>$K^0_S K^0_S K^0_L$ background</td>
<td>70 (fixed)</td>
<td>NA</td>
</tr>
<tr>
<td>$K^0_S K^0_L K^0_L$ background</td>
<td>20 (fixed)</td>
<td>NA</td>
</tr>
<tr>
<td>$K^0_S K^0_S \pi^0$ background</td>
<td>2.3 (fixed)</td>
<td>NA</td>
</tr>
<tr>
<td>Estimated fit bias (evt)</td>
<td>$+0.3$</td>
<td>0.0</td>
</tr>
<tr>
<td>$c\bar{c}K_{S/L}$</td>
<td>2.1</td>
<td>NA</td>
</tr>
<tr>
<td>MC error (%)</td>
<td>8.7</td>
<td>6.1</td>
</tr>
<tr>
<td>$K^0_S$ corr. (%)</td>
<td>96.2</td>
<td></td>
</tr>
<tr>
<td>$K^0_L$ corr. (%)</td>
<td>96.1</td>
<td></td>
</tr>
<tr>
<td>Corr. error (%)</td>
<td>8.1</td>
<td>5.6</td>
</tr>
<tr>
<td>$\prod B_i$ (%)</td>
<td>47.5</td>
<td>16.2</td>
</tr>
<tr>
<td>$c\bar{c}K_{S/L}$ error (%)</td>
<td>3.8</td>
<td>0.9</td>
</tr>
<tr>
<td>$\beta (\times 10^{-6})$</td>
<td>$2.4^{+2.2}_{-2.3}$</td>
<td>$4.0^{+2.6}_{-2.2}$</td>
</tr>
<tr>
<td>Stat. signif. ($\sigma$)</td>
<td>1.0</td>
<td>2.2</td>
</tr>
<tr>
<td>Signif. w/ syst. ($\sigma$)</td>
<td>0.9</td>
<td></td>
</tr>
</tbody>
</table>

Table II. Estimates of systematic errors. Multiplicative systematic errors are in percent while additive systematic errors are in events. The fit yield systematic error is due to fixed fitting parameters, mainly from $K^0_S K^0_L K^0_L$ $\Delta E$ PDF parameters, and the uncertainties of fixed peaking $B$ background yields. The fit bias error is one half of the bias. The $c\bar{c}K_{S/L}$ error is the uncertainty of charmonium background subtraction, estimated as one half of the subtraction.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>$K^0_S K^0_S K^0_L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additive errors (events)</td>
<td>$K^0_S K^0_L$ normalization</td>
</tr>
<tr>
<td>$K^0_S K^0_L$ efficiency</td>
<td>6.8</td>
</tr>
<tr>
<td>$K^0_L$ efficiency</td>
<td>8.0</td>
</tr>
<tr>
<td>Number $BB$</td>
<td>$B(B^0 \rightarrow K^0_S \rightarrow \pi^+ \pi^-)$</td>
</tr>
<tr>
<td>Total multiplicative (%)</td>
<td>10.6</td>
</tr>
<tr>
<td>Total errors [$B(10^{-6})$]</td>
<td>0.6</td>
</tr>
</tbody>
</table>
function of the $K^0_{L3}$ is unknown. We use the average signal efficiency assuming $0.0030$ uncertainties of peaking backgrounds. The dominant systematic error is from the dependence of the signal efficiency of this cut is $0.0003$ for the $K^0_{L3}$ peaking backgrounds. The dotted-dashed curves are signal PDFs.

The true distribution of $K^0_{S3}K^0_{S3}K^0_{L}$ events in the Dalitz plot is unknown. We use the average signal efficiency assuming a uniform true Dalitz distribution (8.1%) to compute the $K^0_{S3}K^0_{S3}K^0_{L}$ branching fraction. We found no significant dependence of the signal $\Delta E$ resolution or the signal $NN$ shape on the Dalitz plot variables. With these assumptions, we find a $K^0_{S3}K^0_{S3}K^0_{L}$ branching fraction, excluding the $\phi$ resonance, of $\mathcal{B}(B^0 \rightarrow K^0_{S3}K^0_{S3}K^0_{L}) = (2.4^{+2.7}_{-2.5} \pm 0.6) \times 10^{-6}$, where the first error is statistical and the second error is systematic. The dominant systematic error is from the uncertainties of peaking $B$ backgrounds (23% relative).

Figure 8 shows a scan of the negative log likelihood as a function of the $K^0_{S3}K^0_{S3}K^0_{L}$ branching fraction, where the minimum negative log likelihood has been subtracted. In order to remove any dependence on our estimate of the total $KKK^*$ branching fraction, we conservatively fix the $KKK^*$ yield to zero in the scan of the log likelihood. We compute a one-sided Bayesian 90% confidence level upper limit on the branching fraction assuming a uniform prior probability for positive (physical) branching fraction values. Systematic errors are included by convolving the fit likelihood with a Gaussian distribution with a width corresponding to the total systematic error, excluding the uncertainty on the $KKK^*$ yield, since it is fixed to zero in the likelihood scan. The result for the nonresonant three-body branching fraction is $\mathcal{B}(B^0 \rightarrow K^0_{S3}K^0_{S3}K^0_{L}) < 7.4 \times 10^{-6}$ at 90% C.L. Assuming the worst-case true Dalitz distribution, where the signal is entirely in the region of lowest efficiency (4%), the 90% C.L. upper limit on the branching fraction is $16 \times 10^{-6}$.

As a cross-check on the analysis, we have performed the fit only in the $\phi K^0_{S3}$ region of the Dalitz plot where the invariant mass of one of the $K^0_{S3}$ pairs is less than 1.049 GeV/$c^2$. The results are given in the second column of Table I. We find $8.3^{+2.8}_{-3.4}$ signal events, which corresponds to a branching fraction of $\mathcal{B}(B^0 \rightarrow \phi K^0_{S3}) = (4.0^{+2.6}_{-3.2}) \times 10^{-6}$, where the errors are statistical only. This is consistent with the world average value of $0.5 \cdot \mathcal{B}(B^0 \rightarrow \phi K^0) = (4.3^{+0.7}_{-0.6}) \times 10^{-6}$ [16].

We checked our estimation of the $KKK^*$ peaking background yield by allowing it to float in the fit. This fit gave a $KKK^*$ yield of $-54 \pm 170$ events, which is consistent with our estimation of $70 \pm 35$ events.

V. SUMMARY

We have searched for the decay $B^0 \rightarrow K^0_{S3}K^0_{S3}K^0_{L}$ using $232 \times 10^6 BB$ pairs recorded by the BABAR experiment. We find no significant evidence for this decay. The central value for the branching fraction, assuming a uniform true Dalitz distribution for the signal and excluding the $\phi$ resonance, is $\mathcal{B}(B^0 \rightarrow K^0_{S3}K^0_{S3}K^0_{L}) = (2.4^{+2.7}_{-2.5} \pm 0.6) \times 10^{-6}$, where the first error is statistical and the second error is systematic. This corresponds to a Bayesian 90% C.L. upper limit of $\mathcal{B}(B^0 \rightarrow K^0_{S3}K^0_{S3}K^0_{L}) < 7.4 \times 10^{-6}$. 
Assuming the worst-case true Dalitz distribution, where the signal is entirely in the region of lowest efficiency, the upper limit on the branching fraction is $16 \times 10^{-6}$. Our results show that the $K_{S}^{0}K_{S}^{0}K_{L}^{0}$ channel will be of limited use in understanding the $b \to s\bar{q}q$ penguin $CP$ anomaly, due to the low efficiency times branching fraction, which limits the yield of signal events.

**ACKNOWLEDGMENTS**

We are grateful for the extraordinary contributions of our PEP-II colleagues in achieving the excellent luminosity and machine conditions that have made this work possible. The success of this project also relies critically on the expertise and dedication of the computing organizations that support BABAR. The collaborating institutions wish to thank SLAC for its support and the kind hospitality extended by them. This work is supported by the U.S. Department of Energy and National Science Foundation, the Natural Sciences and Engineering Research Council (Canada), Institute of High Energy Physics (China), the Commissariat à l’Energie Atomique and Institut National de Physique Nucléaire et de Physique des Particules (France), the Bundesministerium für Bildung und Forschung and Deutsche Forschungsgemeinschaft (Germany), the Istituto Nazionale di Fisica Nucleare (Italy), the Foundation for Fundamental Research on Matter (The Netherlands), the Research Council of Norway, the Ministry of Science and Technology of the Russian Federation, and the Particle Physics and Astronomy Research Council (United Kingdom). Individuals have received support from the Marie-Curie IEF program (European Union) and the A.P. Sloan Foundation.

[6] For the most recent summary of all $b \to s\bar{q}q$ penguin $CP$ asymmetry measurements, see the Heavy Flavor Averaging Group summary at http://www.slac.stanford.edu/xorg/hfag/ and hep-ex/0603003.
[10] A preliminary Dalitz analysis of the decay $B^0 \to K^+K^-K^0_S$ is given in B. Aubert et al. (BABAR Collaboration), hep-ex/0507094.
[17] $f(x) = x\sqrt{1 - x^2} \exp[-\xi(1 - x^2)]$ with $x = (C + \Delta E_{\text{min}} - \Delta E)/C$, $C = 5.29$ GeV or $x = NN/NN_{\text{max}}$. See H. Albrecht et al. (ARGUS Collaboration), Phys. Lett. B 241, 278 (1990).