Search for the Radiative Decay $B^0 \rightarrow \phi\gamma$

We perform a search for the exclusive radiative decay $B^0 \to \phi \gamma$, which is dominated by $t\bar{d}$ annihilation, in a sample of 124 million $B\bar{B}$ events collected with the BABAR detector at the PEP-II asymmetric-energy $e^+e^-$ storage ring at SLAC. No significant signal is seen. We set an upper limit on the branching fraction of $B(B^0 \to \phi \gamma) < 8.5 \times 10^{-7}$ at the 90% confidence level.

PACS numbers: 13.20.He

FIG. 1: One of the leading order Feynman diagrams contributing to the decay $B^0 \to \phi \gamma$ in the Standard Model.

Within the Standard Model (SM) the rare decay $B^0 \to \phi \gamma$ proceeds through a penguin annihilation process. No process of this kind has yet been observed. The largest short-distance contribution to the SM amplitude is illustrated in Fig. 1. The coupling of the top quark within the loop to the $d$ quark leads to a dependence of the amplitude on the Cabbibo-Kobayashi-Maskawa matrix element $V_{td}$ [1] that suppresses the decay rate. Helicity suppression and the smallness of the decay constants $f_B$ and $f_\phi$ also lead to a small decay rate [2]. A recent QCD factorization calculation predicts a SM branching fraction of $3.6 \times 10^{-12}$ [2]. However, contributions to the $B^0 \to \phi \gamma$ amplitude are possible from physics beyond the SM where new heavy particles enter the loop. For example, some models of supersymmetry that violate $R$-parity predict an enhancement of the $B^0 \to \phi \gamma$ branching fraction by up to four orders of magnitude [2]. A prior experiment has bounded the branching fraction to be $B(B^0 \to \phi \gamma) < 3.3 \times 10^{-6}$ at the 90% confidence level (C.L.) [3].

This analysis uses data collected with the BABAR detector at the PEP-II asymmetric-energy $e^+e^-$ storage ring at SLAC. The data sample consists of 124.1 ± 1.4 million $B\bar{B}$ events, corresponding to an integrated luminosity of 113 fb$^{-1}$ on the $\Upsilon(4S)$ resonance, which has a mass of 10.58 GeV/$c^2$.

The BABAR detector is described in detail in Ref. [4]. Charged particle trajectories are measured by the combination of a five-layer silicon vertex tracker (SVT) and a 40-layer drift chamber (DCH), which are embedded in the 1.5 T magnetic field of a solenoid. Photons are detected in a CsI(Tl) crystal electromagnetic calorimeter (EMC), with an energy resolution of $\sigma_E/E = 0.023(E/\text{GeV})^{-1/4} \pm 0.019$. Charged hadron identification is performed by the combination of energy-loss ($dE/dx$) information from the SVT and DCH and measurements from a ring-imaging Cherenkov detector (DIRC). The segmented return (IFR) of the magnet is instrumented with resistive plate chambers to identify muons.

Monte Carlo (MC) simulations of the response of the BABAR detector, based on GEANT4 [5], are used to optimize the selection criteria and determine the signal efficiencies. These simulations take into account the variations of detector conditions and beam induced backgrounds during the data-taking period.

The first stage of the analysis is to identify $\phi$ mesons and high-energy photons, which can be combined to form $B^0$ candidates [6]. The $\phi$ is reconstructed in the decay to $K^+K^-$, which corresponds to $(49.2^{+0.6}_{-0.7})\%$ of the total $\phi$ decay rate [7]. We require that the charged tracks used in reconstructing the $\phi$ candidates have associated hits in both the SVT and DCH and have a transverse momentum greater than 0.1 GeV/$c$. Tracks compatible with being kaons are identified via an algorithm that combines the candidate track’s measured $dE/dx$, Cherenkov angle and number of Cherenkov photons. Individual kaons produced by signal $\phi$ meson decays are identified with 80% efficiency by this algorithm, while the misidentification rate of pions as kaons is less than 2% over most of the relevant kaon momentum range. A $\phi$ candidate is composed of two identified kaons of opposite charge that are consistent with originating from a common vertex. We require $\phi$ candidates have a mass of $1.011 < m_{K^+K^-} < 1.029$ GeV/$c^2$ (a full width 4.2 times the natural width); this criterion is optimized by a procedure described below.

A photon is identified as a shower in the EMC that is not associated with a reconstructed track. We remove poorly reconstructed photons by rejecting showers in crystals with high noise rates or rejecting photons that are near the edge of the calorimeter’s acceptance. Furthermore, showers that contain energy depositions in
The shower profile is required to be compatible with a single photon to reject those that arise from \( \pi^0 \) decays where the two photons enter the calorimeter in close proximity to one another; this condition also rejects showers generated by neutral hadrons, predominantly \( \pi^0 \) and \( K^0_{\ell} \), that have a broader lateral profile than photon showers. To suppress photons from \( \pi^0 \) (\( \eta \)) decays, the photon candidate is combined in turn with all other photons in the event with a laboratory energy greater than 50 (250) MeV. If any of the resulting invariant masses are within 20 (40) MeV/\( c^2 \) of the \( \pi^0 \) (\( \eta \)) mass, the candidate is vetoed. To further remove photons from \( \pi^0 \) decays the shower is required to be isolated by at least 25 cm from any other shower in the event.

The photon and \( \phi \) meson candidates are combined to form \( B^0 \) meson candidates. We define \( \Delta E^* = E^*_B - E^*_\gamma \), where \( E^*_B = E^*_\phi + E^*_\gamma \) is the center-of-mass (CM) energy of the \( B^0 \) meson candidate and \( E^*_\gamma \) is the CM energy of each beam. The signal \( \Delta E^* \) distribution is peaked at zero with a resolution of approximately 50 MeV; there is a negative tail in \( \Delta E^* \) because of the asymmetric \( E^*_\gamma \) resolution that is the result of energy leakage from the EMC. We also define the beam-energy-substituted mass \( m_{ES} = \sqrt{E^*_\gamma^2 - P^2_B} \), where \( P_B \) is the CM three-momentum of the \( B^0 \) candidate. The signal \( m_{ES} \) distribution peaks at the mass of the \( B \) meson, \( m_B = 5.279 \) GeV/\( c^2 \) [7], and has a resolution of 3 MeV/\( c^2 \), which is dominated by the spread in \( E^*_\gamma \). The signal region in which we search for \( B^0 \to \phi \gamma \) events is defined as \(-0.2 < \Delta E^* < 0.1 \) GeV and \( 5.27 < m_{ES} < 5.29 \) GeV/\( c^2 \). We define three sideband regions for estimating backgrounds: (I) \( 5.10 < m_{ES} < 5.29 \) GeV/\( c^2 \) and \( 0.1 < \Delta E^* < 0.5 \) GeV, (II) \( 5.10 < m_{ES} < 5.29 \) GeV/\( c^2 \) and \(-0.5 < \Delta E^* < -0.2 \) GeV, and (III) \( 5.10 < m_{ES} < 5.27 \) GeV/\( c^2 \) and \(-0.2 < \Delta E^* < 0.1 \) GeV. At this stage in the analysis all candidates with \(-0.5 < \Delta E^* < 0.5 \) GeV and \( 5.10 < m_{ES} < 5.29 \) GeV/\( c^2 \) are retained.

The background comes predominantly from random combinations of real \( \phi \) mesons (Fig. 2) and high-energy photons produced in continuum \( u, d, s \) and \( c \) quark-antiquark events. The dominant sources of high-energy photons in continuum events are initial-state radiation (ISR) and \( \pi^0/\gamma \to \gamma \gamma \) decays where the second photon is undetected, or the measured two-photon mass lies outside the \( \pi^0/\gamma \) veto window. Since the continuum background does not peak in \( \Delta E^* \) or \( m_{ES} \), its magnitude can be evaluated from parts of the \( m_{ES} \) and \( \Delta E^* \) data distributions in the sideband regions. There are also potential backgrounds from charmless \( B \) decays that peak in \( m_{ES} \) and \( \Delta E^* \). The \( B\bar{B} \) background is estimated from simulation.

The ratio of the second–to–zeroth Fox–Wolfram moments [8], for charged tracks in the event measured in the CM frame, is required to be less than 0.9 to reject some low multiplicity continuum final states. The combinatorial background within the signal region is reduced further by combining 24 input variables that distinguish between \( q\bar{q} \) continuum and \( B\bar{B} \) events into one discriminating variable via a neural network, as in Refs. [9, 10]. The neural network responds non-linearly to the input variables and exploits correlations among them [11]. To discriminate between jet-like continuum background and the more spherically-symmetric signal events, we include in the neural network the absolute value of the cosine of the angle between the high-energy photon and the CM thrust axis of the reconstructed particles in the r.o.e (rest of the event), \( |\cos \theta_T^\gamma| \), and the energy distribution of all
reconstructed particles in the r.o.e. binned into eighteen 10° intervals around the photon direction. The distribution of events in |cosθ_{B}^{|}| is uniform for signal events but is strongly peaked toward 1 for continuum events. The energy flow is collinear with the photon direction in continuum events but is more isotropic in signal events. To discriminate against ISR continuum events we incorporate in the neural network the ratio of the second–to–zeroth Fox-Wolfram moments computed in the photon recoil frame, in which ISR continuum events are more likely than signal events to have a back-to-back jet structure. To provide further discrimination from continuum background, we include in the selection two angular variables of the signal decay: |cosθ_{|B}^{|}|, where θ_{B}^{|} is the angle between the B meson candidate’s momentum and the beam axis in the CM frame, and |cosθ_{H}|, where θ_{H} is the angle between the flight directions of one of the daughters of the φ meson and the B meson candidate in the rest frame of the φ meson. The distribution of events of both of these variables is proportional to sin²θ, where θ is θ_{B}^{|} or θ_{H} for signal events. The |cosθ_{H}| distribution has this form because a vector state (T(4S)) is decaying to two pseudoscalars (BB), with the vector state having no helicity zero component. The |cosθ_{H}| distribution has this form because a pseudoscalar state (B₀) is decaying to two vector states (φγ), with the photon having no helicity zero component. Finally, to provide further sensitivity to B decays in the event, we admit two more variables. The first is the longitudinal separation between the decay vertex of the B meson candidate and the vertex of the other charged particles in the event, which tends to be non–zero for signal events due to the long lifetime of the B meson, whereas all particles in continuum events usually originate from a single vertex. The second is the net flavor of the rest of the event, \( N_{X} = N_{K_{S}^{0}} + \sum_{i} |N_{i}^{+} - N_{i}^{-}| \), where \( N_{K_{S}^{0}} \) is the number of reconstructed \( K_{S}^{0} \) and \( N_{i}^{\pm} \) are the numbers of reconstructed charged particles of type \( i = e^{\pm}, \mu^{\pm}, K^{\pm} \) or \( \pi_{slow}^{\pm} \) [12]. The discrimination power of this variable can be seen by noting how the final state particles (as seen by the detector) are produced. If the final state particle is produced through the decay of a B meson, then it most likely was generated through the weak interaction, which allows for flavor-changing currents. In contrast, if it is part of continuum u, d, and s quark-antiquark events, then there is no net flavor production.

The neural network is trained on samples of simulated signal and continuum events. A B → Dπ control sample was used to determine the systematic error in the selection efficiency of the neural network. Here, the “bachelor pion” is used in place of the signal photon. This sample is representative of signal events because the input variables are mainly derived from particles not associated with the B candidate. The exception is |cosθ_{H}|, which has a distribution for pions different from that for photons because of the different spins of these particles. Instead, this variable is drawn from the signal distribution of sin²θ. Fig. 3 shows the distribution of the neural network output for MC-simulated \( B^0 \to D^+π^- \) events compared to \( B^0 \to D^-π^+ \) events reconstructed in on-resonance data. Also shown is the distribution of the neural network output for MC simulated continuum background events compared to events in the on-resonance sideband. These distributions show reasonable agreement between MC and data samples.

The selection is optimized to achieve the best upper limit because of the very small SM expectation for \( B(B^0 \to φγ) \). The quantity minimized by the optimization is \( N/ε \), where \( N \) is the average 90% confidence level upper limit on the number of reconstructed signal events and \( ε \) is the signal efficiency. \( N \) is estimated from an ensemble of experiments with a given MC-simulated background and no true signal [13]. We optimize the selection on the neural network output and on the φ mass window simultaneously. The photon selection is identical to that optimized for the measurement of \( B \to K^{*}\gamma \) [9]. The continuum background within the signal region is estimated with likelihood fits to the \( ΔE^{*} \) and \( m_{ES} \) distributions of simulated continuum events for each set of selection criteria considered during the optimization. The probability density functions (PDFs) used to describe the \( m_{ES} \) and \( ΔE^{*} \) distributions are an ARGUS threshold function [14] and a first-order polynomial, respectively. The resulting functions are integrated over the signal region to estimate the background. The optimized selection criteria have a signal selection efficiency of (14.4 ± 0.1)% and a mean continuum background yield from MC simulation of 4.4 ± 0.5 events.

The selection leads to an expectation of 0.073 ± 0.004 \( B\bar{B} \) background events within the signal region, as determined from the MC simulation. The decays \( B^{0} \to φπ^{0} \) and \( B^{0} \to φη \) contribute most of the events; these charmless modes have not been observed, so the largest branching fractions predicted within the SM are used to compute the expectations [15]. The remaining contributions come from \( B^{0} \to φπ^{0}, B^{0} \to φK_{S}^{0} \) and \( B^{0} \to φK_{L}^{0} \) decays. The contributions of all other \( B\bar{B} \) events, including those with \( B^{0} \to K^{*0}\gamma \), were found to be negligible.

To make a more robust estimate of the continuum background we use the data themselves; this eliminates any uncertainty related to the simulation of the φγ continuum. To define functions that describe the background distribution we use a maximum-likelihood fit to the data within sideband regions I and II with the same form of the \( m_{ES} \) and \( ΔE^{*} \) PDFs as those used in the optimization procedure. The resulting functions are integrated over the sideband region III and the signal region to estimate the amount of continuum background within them. To validate the method the number of background expected in region III (mean of 79 ± 6 events) is compared to the number of events observed in this interval (71 events);
these two numbers are in good agreement. To assign the error on the background estimate, we take the difference of both PDFs evaluated with the central value of the fit parameters and the parameters varied by one standard deviation and then add the differences in quadrature. The mean continuum background estimate within the signal region is 6.0 ± 0.9 events.

In addition, we perform other consistency tests to ensure that both the background estimate and the method used to acquire it are reasonable. We use a second-order polynomial as the PDF for $\Delta E^*$; this leads to a mean continuum background estimate of 5.6 ± 0.8 events. The difference between 5.6 and 6.0 (the number obtained with the method defined previously) is taken as a systematic uncertainty on the background. To check the assumption that the PDFs factorize, we compute the ARGUS parameter and the slope of the first-order polynomial in different intervals of $\Delta E^*$ and $m_{ES}$; the parameters are found to be in reasonable agreement. We perform a fit to $m_{ES}$ and $\Delta E^*$ using regions I, II and III: this gives a mean expected background of 5.2 ± 0.8 events, which is in agreement with the expectation from using only regions I and III. The mean continuum background estimate within the signal region including the systematic uncertainty is 6.0 ± 1.0 events.

We observe 8 events in the signal region, which is consistent with the background estimation. The distribution of data events in $m_{ES}$ and $\Delta E^*$, together with the projections in these two variables, can be seen in Fig. 4.

The fractional systematic uncertainties on $\epsilon$, the number of $B^0\bar{B}^0$ pairs in the data set ($N_{B^0\bar{B}^0}$), and $B(\phi \to K^+K^-)$ are summarized in Table I; their origins are briefly described below. The difference in tracking efficiency between data and simulation is evaluated with a sample of tracks that are well reconstructed in the SVT; these studies yield a relative shift of $(-0.8 \pm 1.3)%$ per charged track. The uncertainty related to the kaon identification is evaluated with simulation and data samples of the decay $D^{*+} \to D^0 (\to K^−\pi^+)\pi^+$ and its charge conjugate. The photon identification efficiency is studied in samples of $\pi^0$ decays from $\tau^+\tau^-$ events. The uncertainties and corrections due to the $\pi^0$ and $\eta$ vetoes and shower isolation criteria are evaluated by embedding simulated photons into both data and simulation samples of $B\bar{B}$ events. The difference between the neural network selection efficiency in the $B^0 \to D^-\pi^+$ data and simulation samples is $(-1.5 \pm 2.7)%$. There is a small uncertainty related to the statistics of the signal simulation sample used to calculate the efficiency. The corrected value of the efficiency after all corrections are applied is $(13.9 \pm 0.7)\%$.

The number of $B\bar{B}$ events used in the analysis is calculated from the change in the ratio of multihadron to $\mu^+\mu^-$ events between data taken at the $\Upsilon(4S)$ resonance and at $e^+e^-$ CM energy 40 MeV below the resonance; there is 1.1% fractional uncertainty in this measurement. We assume that half the number of $B\bar{B}$ events are $B^0\bar{B}^0$ events. The measured branching fraction $B(\phi \to K^+K^-)$ has a fractional uncertainty of 0.6% [7].

Using the signal efficiency, $B(\phi \to K^+K^-)$, $N_{B^0\bar{B}^0}$, the background estimation along with the associated uncer-
<table>
<thead>
<tr>
<th>Systematic effect</th>
<th>Correction</th>
<th>Uncertainty (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K^+ K^-$ tracking</td>
<td>0.984</td>
<td>2.6</td>
</tr>
<tr>
<td>$K^+ K^-$ identification</td>
<td>-</td>
<td>2.0</td>
</tr>
<tr>
<td>Shower separation</td>
<td>-</td>
<td>2.0</td>
</tr>
<tr>
<td>$\pi^0/\eta$ veto</td>
<td>-</td>
<td>1.0</td>
</tr>
<tr>
<td>Photon detection efficiency</td>
<td>0.997</td>
<td>2.5</td>
</tr>
<tr>
<td>Continuum suppression</td>
<td>0.985</td>
<td>2.7</td>
</tr>
<tr>
<td>Simulation statistics</td>
<td>-</td>
<td>0.1</td>
</tr>
<tr>
<td>Overall signal efficiency</td>
<td>0.966</td>
<td>5.4</td>
</tr>
</tbody>
</table>

TABLE I: The sources of systematic uncertainty on estimating the branching fraction. Any correction factor applied to the signal efficiency related to the systematic effect is also given.


tainties, we find by the procedure of Ref. [16] the upper limit:

$$B(B^0 \rightarrow \phi\gamma) < 8.5 \times 10^{-7},$$

at the 90% C.L. In conclusion, no evidence for the decay $B^0 \rightarrow \phi\gamma$ is observed. We set an upper limit that is 3.9 times lower than the previously published result.

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 (USA), 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.

---

1 Deceased

[6] Charge conjugate states are implied throughout this paper.
[9] B. Aubert et al., [BABAR Collaboration], hep-ex/0407003, submitted to PRL.
[11] We use the Stuttgart Neural Network Simulator (http://www-ra.informatik.uni-tuebingen.de/SNNS) to train the neural network with one hidden layer of seven nodes.
[12] A $\pi_{slow}^\pm$ is defined as a pion having a CM momentum less than 250 MeV/c, a cosine of the angle between the thrust axis of the event and the pion momentum less than 0.8 and a distance of closest approach to the primary vertex less than 0.5 cm. These criteria preferentially select pions from $D^*$ decays which are abundant in $B\bar{B}$ events.
[14] We use the distribution $x\sqrt{1-x^2} e^{-x(1-x^2)}$, where $x = m_{ES}/E_{\text{beam}}$. H. Albrecht et al., Z. Phys. C 48, 543 (1990).

* Also with Università della Basilicata, Potenza, Italy