Observation of a Significant Excess of $\pi^0\pi^0$ Events in $B$ Meson Decays

We present a study of the decay $B^0 \rightarrow \pi^0 \pi^0$ based on a sample of $124 \times 10^6$ $B\bar{B}$ pairs recorded by the BaBar detector at the PEP-II asymmetric-energy $B$ Factory at SLAC. We observe $46 \pm 13 \pm 3$ events, where the first error is statistical and the second is systematic, corresponding to a significance of 4.2 standard deviations including systematic uncertainties. We measure the branching fraction $\mathcal{B}(B^0 \rightarrow \pi^0 \pi^0) = (2.1 \pm 0.6 \pm 0.3) \times 10^{-6}$, averaged over $B^0$ and $\bar{B}^0$ decays.

DOI: 10.1103/PhysRevLett.91.241801

PACS numbers: 13.25.Hw, 11.30.Er, 12.15.Hh
The study of $B$ meson decays into charmless hadronic final states plays an important role in the understanding of $CP$ violation in the $B$ system. In the standard model, $CP$ violation arises from a single complex phase in the Cabibbo-Kobayashi-Maskawa (CKM) quark-mixing matrix $V$ [1]. Measurements of the time-dependent $CP$-violating asymmetry in the $B^0 \rightarrow \pi^+ \pi^-$ decay mode by the $BABAR$ and Belle Collaborations [2] provide information on the angle $\alpha = \arg \left(-V_{ud} V_{ub}^* / V_{td} V_{tb}^* \right)$ of the unitarity triangle. However, in contrast to the theoretically clean determination of the angle $\beta$ in $B^0$ decays to charmonium plus neutral-kaon final states [3,4], the extraction of $\alpha$ in $B^0 \rightarrow \pi^+ \pi^-$ is complicated by the interference of amplitudes with different weak phases. The difference between $\alpha_{\text{eff}}$, derived from the measured $B^0 \rightarrow \pi^+ \pi^-$ asymmetry, and $\alpha$ may be evaluated using isospin relations between the amplitudes for the decays $B^0(B^0) \rightarrow \pi^+ \pi^-$, $B^0(B^0) \rightarrow \pi^+ \pi^0$, and $B^- \rightarrow \pi^- \pi^0$ [5].

The primary contributions to the decay $B^0 \rightarrow \pi^0 \pi^0$ are expected to come from the so-called color-suppressed tree and gluonic penguin amplitudes [6]. The branching fraction for $B^0 \rightarrow \pi^0 \pi^0$ has been calculated in various QCD models [7]. All models use as inputs the values of the CKM angles, typically taken from unitarity-triangle fits. The predictions for $\mathcal{B}(B^0 \rightarrow \pi^0 \pi^0)$ are in the range $(0.3 - 1.1) \times 10^{-6}$. In particular, the prediction from the QCD factorization model of Beneke et al. has a central value of $0.3 \times 10^{-6}$. Alternatively, studies using phenomenological fits to experimental data for charmless $B$ decays [8] find values in the range $(1.6 - 2.5) \times 10^{-6}$ for the $B^0 \rightarrow \pi^0 \pi^0$ branching fraction.

In this Letter, we report the observation of a significant excess of $B^0 \rightarrow \pi^0 \pi^0$ decays based on $(124 \pm 1) \times 10^6$ $Y(4S) \rightarrow B\overline{B}$ pairs (on-resonance), collected with the $BABAR$ detector. We also use approximately $12 \text{fb}^{-1}$ of data recorded $40 \text{MeV}$ below the $B\overline{B}$ threshold (off-resonance).

$BABAR$ is a solenoidal detector optimized for the asymmetric-energy beams at PEP-II and is described in detail in Ref. [9]. Charged particle (track) momenta are measured with a 5-layer double-sided silicon vertex tracker and a 40-layer drift chamber inside a 1.5-T superconducting solenoidal magnet. Neutral cluster (photon) positions and energies are measured with an electromagnetic calorimeter (EMC) consisting of 6580 CsI(Tl) crystals. The photon energy resolution is $\sigma_E = \sqrt{2.3/E(\text{GeV})} \times 1.9\%$, and the angular resolution from the interaction point is $\sigma_\theta = 3.99/\sqrt{E(\text{GeV})}$. The photon energy scale is determined using symmetric $\pi^0 \rightarrow \gamma \gamma$ decays. Charged hadrons are identified with a detector of internally reflected Cherenkov light and ionization in the tracking detectors. The instrumented magnetic-flux return detects neutral hadrons and identifies muons. High efficiency for recording $B\overline{B}$ events in which one $B$ decays with low multiplicity is achieved with a two-level trigger with complementary tracking-based and calorimetry-based trigger decisions.

Candidate $\pi^0$ mesons are reconstructed as pairs of photons, spatially separated in the EMC, with an invariant mass within $3\sigma$ of the $\pi^0$ mass. The mass resolution $\sigma$ is approximately $8 \text{MeV}/c^2$ for high-momentum $\pi^0$ mesons. Photon candidates are required to be consistent with the expected lateral shower shape, not to be matched to a track, and to have a minimum energy of 30 MeV. To reduce the background from false $\pi^0$ candidates, the angle $\theta_\gamma$ between the photon momentum vector in the $\pi^0$ rest frame and the $\pi^0$ momentum vector in the laboratory frame is required to satisfy $|\cos \theta_\gamma| < 0.95$.

$B$ meson candidates are reconstructed by combining two $\pi^0$ candidates. Two kinematic variables, used to isolate the $B^0 \rightarrow \pi^0 \pi^0$ signal, take advantage of the kinematic constraints of $B$ mesons produced at the $Y(4S)$. The first is the beam-energy-substituted mass $m_{\text{ES}} = \sqrt{(s/2 + p_1 \cdot p_B)^2/E_1^2 - p_B^2}$, where $\sqrt{s}$ is the total $e^+ e^-$ center-of-mass (c.m.) energy, $(E_1, p_1)$ is the four-momentum of the initial $e^+ e^-$ system and $p_B$ is the $B$ candidate momentum, both measured in the laboratory frame. The second variable is $\Delta E = E_B - \sqrt{s}/2$, where $E_B$ is the $B$ candidate energy in the c.m. frame. The $\Delta E$ resolution for signal is approximately $80 \text{MeV}$.

The primary source of background is $e^+ e^- \rightarrow q\overline{q}$ ($q = u, d, s, c$) events where a $\pi^0$ from each quark jet randomly combine to mimic a $B$ decay. The jetlike $q\overline{q}$ background is suppressed by requiring that the angle $\theta_S$ between the sphericity [10] axes of the $B$ candidate and of the remaining tracks and photons in the event, in the c.m. frame, satisfies $|\cos \theta_S| < 0.7$. The other source of background is $B^z \rightarrow \rho^\pm \pi^0 \rho^- \pi^+ \pi^0$ decays in which the charged pion is emitted nearly at rest in the $B$ rest frame so that the remaining two $\pi^0$ mesons are kinematically consistent with a $B^0 \rightarrow \pi^0 \pi^0$ decay. Energy resolution smearing causes some $B^z \rightarrow \rho^\pm \pi^\pm$ events to have $\Delta E$ above the kinematic limit of $m_B - m_{\pi^\pm}$. From simulation, other $B$ decays contribute no more than one background event.

The number of signal $B^0 \rightarrow \pi^0 \pi^0$ candidates is determined in an extended, unbinned maximum-likelihood fit. The variables used in the fit are $m_{\text{ES}}, \Delta E,$ and a Fisher-discriminant $F$. The $F$ discriminant is a linear combination of three variables, optimized to separate signal from $q\overline{q}$ background. The first two variables are sums: $L_1 = \sum p_i$ and $L_2 = \sum p_i \cos^2 \theta_i$, where $p_i$ is the momentum and $\theta_i$ is the angle with respect to the thrust axis of the $B$ candidate, both in the c.m. frame, for all tracks and neutral clusters not used to reconstruct the $B$ meson. The third variable in $F$ is the output of a neural network designed to separate $B$ events from $q\overline{q}$ background, whose inputs are information from the remaining tracks and photons in the event. The inputs include information about high-momentum leptons, low-momentum leptons, charged kaons, and slow pions (from $D^{*+} \rightarrow D^0 \pi_{\text{slow}}$) in the event; these are the same inputs used in the $B$-tagging algorithm of Ref. [3]. All neural-network training and
Fisher-discriminant optimization is performed using simulated events.

The data are divided into two samples: a signal sample with candidates satisfying $m_{ES} > 5.2 \text{ GeV}/c^2$ and $|\Delta E| < 0.2 \text{ GeV}$, and a sideband sample with candidates from off-resonance data with $m_{ES} > 5.2 \text{ GeV}/c^2$ and $0.2 < |\Delta E| < 0.4 \text{ GeV}$ (and well outside the triangular region in $m_{ES}$ and $\Delta E$ populated by $B^\pm \to \rho^\mp \pi^0$ decays) and candidates from off-resonance data with $m_{ES} > 5.2 \text{ GeV}/c^2$ and $|\Delta E| < 0.4 \text{ GeV}$. The sideband sample contains only $q\bar{q}$ background candidates and is used in the fit to improve the statistical precision of the $F$ distribution for $q\bar{q}$ events. There are 4470 events in the signal sample and 3253 events in the sideband sample. The $m_{ES}$, $\Delta E$, and $F$ distributions are shown in Fig. 1 for all data used in the fit. The reconstruction efficiency for $B^0 \to \pi^0 \pi^0$ is $(17.7 \pm 2.7)\%$, and for $B^\pm \to \rho^\mp \pi^0$ is $(0.8 \pm 0.1)\%$, derived from simulation. The errors are due to a systematic uncertainty in the efficiency for high-momentum $\pi^0$ mesons to pass the selection criteria; statistical uncertainties are negligible.

For candidates in the signal sample the probabilities $\mathcal{P}_i(\tilde{x}_j; \tilde{\alpha}_i)$ used in the maximum-likelihood fit are the product of probability density functions (PDFs) for the variables $\tilde{x}_j = \{m_{ES}, \Delta E, F\}$, given the set of parameters $\tilde{\alpha}_i$. The likelihood function is given by a product over all $j = 1$ to $N$ candidates and a sum over the $i = \{B^0 \to \pi^0 \pi^0, B^\pm \to \rho^\mp \pi^0, q\bar{q}\}$ hypotheses:

$$\mathcal{L} = \exp \left(- \sum_{j=1}^{N} \sum_{i=1}^{3} n_j \mathcal{P}_i(\tilde{x}_j; \tilde{\alpha}_i) \right).$$

The coefficients $n_j$ are the numbers of $B^0 \to \pi^0 \pi^0$ signal, $B^\pm \to \rho^\mp \pi^0$ background, and $q\bar{q}$ background events in the sample. The number of $B^\pm \to \rho^\mp \pi^0$ events is fixed in the fit to the expected value based on the measured branching fraction $B(B^\pm \to \rho^\mp \pi^0) = (11.0 \pm 2.7) \times 10^{-6}$ [11]. For candidates in the sideband sample, the likelihood function includes only the PDF for the $F$ variable, and only the component for $q\bar{q}$ background. A simultaneous fit to both signal sample and sideband sample data is performed. Monte Carlo simulations are used to verify that the likelihood fit is unbiased.

The PDFs are determined from data and simulation. The $m_{ES}$ and $\Delta E$ variables are correlated for both $B^0 \to \pi^0 \pi^0$ and $B^\pm \to \rho^\mp \pi^0$, so a two-dimensional PDF derived from a smoothed, simulated distribution is used. The $m_{ES}$ distribution for $q\bar{q}$ events is modeled as a threshold function [12] whose shape parameter is determined from data with $|\cos \theta_q| > 0.9$. The $\Delta E$ distribution for $q\bar{q}$ events is modeled as a quadratic polynomial with parameters determined from data with $m_{ES} < 5.26 \text{ GeV}/c^2$. 

FIG. 1. The distributions of (a) $m_{ES}$, (b) $\Delta E$, and (c) Fisher discriminant $F$ for candidates in the signal data sample, and (d) the $F$ distribution for candidates in the sideband data sample. The solid lines show the PDF for signal plus background. For $m_{ES}$ and $\Delta E$, the dashed lines show the PDF for the $q\bar{q}$ background. The abscissa in (c) and (d) is the $F$ bin number, where the bins have been chosen so that each bin contains approximately 10% of the distribution for the $B^0 \to D^{(c)}n\pi$ data sample.
Before fitting the data. The fraction of signal events included in the plots is 24%, 42%, and 74% for $m_{ES}$, $\Delta E$, and $F$, respectively. The PDF projections are shown as a dashed line for $q\bar{q}$ background, a dotted line for $B^+ \rightarrow \rho^+ \pi^0$, and a dash-dotted line for $B^0 \rightarrow \pi^0 \pi^0$ signal. The solid line shows the sum of all PDF projections. The PDF projections are scaled by the expected fraction of events passing the probability-ratio requirement. The ratio $\sim 2 \ln (L/L_{\text{max}})$ is shown in (d) where the dashed line corresponds to statistical errors only and the solid line corresponds to statistical and systematic errors, as applied for the calculation of significance.

The PDF for the $F$ variable is modeled as a parametric step function (PSF) for $B^0 \rightarrow \pi^0 \pi^0$, $B^+ \rightarrow \rho^+ \pi^0$, and $q\bar{q}$ events. A PSF is a binned distribution (as in a histogram), whose parameters are the heights of each bin. Since the parent distribution of $F$ is not known, any functional form (such as a multiple Gaussian) assumed for the PDF will suffer from a systematic uncertainty due to the choice of function. By binning the data, the PSF substantially reduces this uncertainty. The PSF is normalized to one, so that the number of free parameters is the number of bins minus one. For both $B^0 \rightarrow \pi^0 \pi^0$ and $B^+ \rightarrow \rho^+ \pi^0$, the $F$ PSF parameters are taken from a sample of $3.2 \times 10^4$ fully reconstructed $B^0 \rightarrow D^{(*)} n \pi (n = 1, 2, 3)$ events in data. The $F$ PSF has ten bins, with bin limits chosen so that each bin contains approximately 10% of the $B^0 \rightarrow D^{(*)} n \pi$ events. Simulation is used to verify that the same distribution can be used for both $B^0 \rightarrow \pi^0 \pi^0$ and $B^+ \rightarrow \rho^+ \pi^0$. For $q\bar{q}$ background, the $F$ PSF parameters are free parameters in the fit; these parameters are determined from data in both the signal and sideband samples.

All event-selection requirements, PDF parameters, and maximum-likelihood fit conditions were determined before fitting the data.

The result of the fit is $n_{B^0 \rightarrow \pi^0 \pi^0} = 46 \pm 13$ events, corresponding to a branching fraction of $B(B^0 \rightarrow \pi^0 \pi^0) = (2.1 \pm 0.6) \times 10^{-6}$. $B^0$ and $\bar{B}^0$ decays are not separated, so the branching fraction is measured for the average of $B^0$ and $\bar{B}^0$. The $m_{ES}$, $\Delta E$, and $F$ distributions for events that pass a requirement on the signal probability ratio are shown in Fig. 2. This requirement is optimized to maximize the ratio $S/(S+B)$, where $S$ is the number of signal events and $B$ is the number of background events in the plot. The significance of the event yield is evaluated from the square root of the change in $-2 \ln L$ between the nominal fit and a separate fit in which the signal yield is fixed to zero, and is found to be 4.7$\sigma$ with statistical errors only.

The number of signal events is stable when the $q\bar{q}$ $m_{ES}$ and $\Delta E$ PDF parameters, or $n_{B^0 \rightarrow \rho^+ \pi^0}$, are allowed to vary in the fit. A validation of the maximum-likelihood fit is made by performing a simpler event-counting analysis, based on the number of events satisfying tighter $m_{ES}$, $\Delta E$, and $F$ requirements. The event-counting analysis finds $13 \pm 6$ events with an efficiency of 31% relative to the maximum-likelihood fit. This agrees well with the fitted result, and has a statistical significance of 2.7$\sigma$.

This result is consistent with our previous limit for this decay [13] based on $88 \times 10^6 BB$ pairs. The data
TABLE I. A summary of systematic uncertainties listed as the change in the fitted event yield, $\Delta n_{B^0\to\pi^0\pi^0}$, for different parameter variations.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\Delta n_{B^0\to\pi^0\pi^0}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q\bar{q}$ mES Shape parameter</td>
<td>$+2.0$</td>
</tr>
<tr>
<td>$q\bar{q}$ $\Delta E$ quadratic polynomial</td>
<td>$+0.9$</td>
</tr>
<tr>
<td>$n_{B^0\to\pi^0\pi^0}$</td>
<td>$+0.9$</td>
</tr>
<tr>
<td>$B^0 \to \pi^0\pi^0\Delta E$ mean</td>
<td>$+0.6$</td>
</tr>
</tbody>
</table>

*Also with Università di Perugia, Perugia, Italy.
†Also with Università della Basilicata, Potenza, Italy.
‡Also with IFIC, Instituto de Física Corpuscular, CSIC-Universidad de Valencia, Valencia, Spain.
§Deceased.

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 the A.P. Sloan Foundation, Research Corporation, and Alexander von Humboldt Foundation.

---

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 the A.P. Sloan Foundation, Research Corporation, and Alexander von Humboldt Foundation.

---

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 the A.P. Sloan Foundation, Research Corporation, and Alexander von Humboldt Foundation.