PHYSICAL REVIEW D **93,** 111101(R) (2016)

Search for the decay $B^0 \to \phi\gamma$

Z. King,[7] B. Pal,[7] A. J. Schwartz,[7] I. Adachi,[13,10] H. Aihara,[64] S. Al Said,[57,24] D. M. Asner,[48] H. Atmacan,[35] T. Aushev,[37] R. Ayad,[57] A. M. Bakich,[56] P. Behera,[17] V. Bhardwaj,[54] B. Bhuyan,[16] J. Biswal,[21] A. Bobrov,[4,46] A. Bozek,[43] T. E. Browder,[12] D. Červenkov,[5] V. Chekelian,[33] B. G. Cheon,[11] K. Chilikin,[29,36] R. Chistov,[29,36] K. Cho,[25] V. Chobanova,[33] Y. Choi,[55] D. Cinabro,[68] J. Dalseno,[33,59] N. Dash,[15] Z. Doležal,[5] D. Dutta,[58] S. Eidelman,[4,46] H. Farhat,[68] T. Ferber,[8] B. G. Fulsom,[48] V. Gaur,[58] N. Gabyshev,[4,46] A. Garmash,[4,46] R. Gillard,[68] R. Glattauer,[20] Y. M. Goh,[11] P. Goldenzweig,[23] B. Golob,[30,21] J. Haba,[13,10] T. Hara,[13,10] K. Hayasaka,[39] H. Hayashii,[40] T. Horiguchi,[62] W.-S. Hou,[42] C.-L. Hsu,[34] T. Iijima,[39,38] K. Inami,[38] G. Inguglia,[8] A. Ishikawa,[62] R. Itoh,[13,10] Y. Iwasaki,[13] W. W. Jacobs,[18] H. B. Jeon,[27] K. K. Joo,[6] T. Julius,[34] K. H. Kang,[27] E. Kato,[62] T. Kawasaki,[44] C. Kiesling,[33] D. Y. Kim,[53] H. J. Kim,[27] J. B. Kim,[26] K. T. Kim,[26] S. H. Kim,[11] Y. J. Kim,[25] K. Kinoshita,[7] P. Kodyš,[5] S. Korpar,[32,21] D. Kotchetkov,[12] P. Križan,[30,21] P. Krokovny,[4,46] T. Kuhr,[31] T. Kumita,[66] I. S. Lee,[11] C. H. Li,[34] H. Li,[18] L. Li,[50] Y. Li,[67] L. Li Gioi,[33] J. Libby,[17] T. Luo,[49] M. Masuda,[63] T. Matsuda,[71] D. Matvienko,[4,46] K. Miyabayashi,[40] H. Miyata,[44] R. Mizuk,[29,36,37] G. B. Mohanty,[58] A. Moll,[33,59] M. Nakao,[13,10] H. Nakazawa,[72] T. Nanut,[21] K. J. Nath,[16] K. Negishi,[62] S. Nishida,[13,10] S. Ogawa,[61] S. Okuno,[22] W. Ostrowicz,[43] C. W. Park,[55] S. Paul,[60] T. K. Pedlar,[73] L. Pesántez,[3] R. Pestotnik,[21] M. Petrič,[21] L. E. Piilonen,[67] C. Pulvermacher,[23] J. Rauch,[60] M. Ritter,[31] A. Rostomyan,[8] S. Ryu,[51] H. Sahoo,[12] Y. Sakai,[13,10] S. Sandilya,[58] L. Santelj,[13] T. Sanuki,[62] Y. Sato,[38] V. Savinov,[49] T. Schlüter,[31] O. Schneider,[28] G. Schnell,[1,14] C. Schwanda,[20] Y. Seino,[44] K. Senyo,[69] M. E. Sevior,[34] V. Shebalin,[4,46] C. P. Shen,[2] T.-A. Shibata,[65] J.-G. Shiu,[42] B. Shwartz,[4,46] F. Simon,[33,59] E. Solovieva,[29,37] S. Stanič,[45] M. Starič,[21] J. F. Strube,[48] J. Stypula,[43] M. Sumihama,[9] M. Takizawa,[52] N. Taniguchi,[13] Y. Teramoto,[47] K. Trabelsi,[13,10] M. Uchida,[65] T. Uglov,[29,37] Y. Unno,[11] S. Uno,[13,10] P. Urquijo,[34] Y. Usov,[4,46] P. Vanhoefer,[33] G. Varner,[12] K. E. Varvell,[56] V. Vorobyev,[4,46] C. H. Wang,[41] M.-Z. Wang,[42] P. Wang,[19] M. Watanabe,[44] Y. Watanabe,[22] S. Wehle,[8] K. M. Williams,[67] E. Won,[26] J. Yamaoka,[48] S. Yashchenko,[8] Y. Yook,[70] C. Z. Yuan,[19] Z. P. Zhang,[50] V. Zhilich,[4,46] and A. Zupanc[30,21]

(Belle Collaboration)

[1]*University of the Basque Country UPV/EHU, 48080 Bilbao*
[2]*Beihang University, Beijing 100191*
[3]*University of Bonn, 53115 Bonn*
[4]*Budker Institute of Nuclear Physics SB RAS, Novosibirsk 630090*
[5]*Faculty of Mathematics and Physics, Charles University, 121 16 Prague*
[6]*Chonnam National University, Kwangju 660-701*
[7]*University of Cincinnati, Cincinnati, Ohio 45221*
[8]*Deutsches Elektronen–Synchrotron, 22607 Hamburg*
[9]*Gifu University, Gifu 501-1193*
[10]*SOKENDAI (The Graduate University for Advanced Studies), Hayama 240-0193*
[11]*Hanyang University, Seoul 133-791*
[12]*University of Hawaii, Honolulu, Hawaii 96822*
[13]*High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801*
[14]*IKERBASQUE, Basque Foundation for Science, 48013 Bilbao*
[15]*Indian Institute of Technology Bhubaneswar, Satya Nagar 751007*
[16]*Indian Institute of Technology Guwahati, Assam 781039*
[17]*Indian Institute of Technology Madras, Chennai 600036*
[18]*Indiana University, Bloomington, Indiana 47408*
[19]*Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049*
[20]*Institute of High Energy Physics, Vienna 1050*
[21]*J. Stefan Institute, 1000 Ljubljana*
[22]*Kanagawa University, Yokohama 221-8686*
[23]*Institut für Experimentelle Kernphysik, Karlsruher Institut für Technologie, 76131 Karlsruhe*
[24]*Department of Physics, Faculty of Science, King Abdulaziz University, Jeddah 21589*
[25]*Korea Institute of Science and Technology Information, Daejeon 305-806*
[26]*Korea University, Seoul 136-713*
[27]*Kyungpook National University, Daegu 702-701*
[28]*École Polytechnique Fédérale de Lausanne (EPFL), Lausanne 1015*
[29]*P.N. Lebedev Physical Institute of the Russian Academy of Sciences, Moscow 119991*
[30]*Faculty of Mathematics and Physics, University of Ljubljana, 1000 Ljubljana*
[31]*Ludwig Maximilians University, 80539 Munich*
[32]*University of Maribor, 2000 Maribor*
[33]*Max-Planck-Institut für Physik, 80805 München*

[34]School of Physics, University of Melbourne, Victoria 3010
[35]Middle East Technical University, 06531 Ankara
[36]Moscow Physical Engineering Institute, Moscow 115409
[37]Moscow Institute of Physics and Technology, Moscow Region 141700
[38]Graduate School of Science, Nagoya University, Nagoya 464-8602
[39]Kobayashi-Maskawa Institute, Nagoya University, Nagoya 464-8602
[40]Nara Women's University, Nara 630-8506
[41]National United University, Miao Li 36003
[42]Department of Physics, National Taiwan University, Taipei 10617
[43]H. Niewodniczanski Institute of Nuclear Physics, Krakow 31-342
[44]Niigata University, Niigata 950-2181
[45]University of Nova Gorica, 5000 Nova Gorica
[46]Novosibirsk State University, Novosibirsk 630090
[47]Osaka City University, Osaka 558-8585
[48]Pacific Northwest National Laboratory, Richland, Washington 99352
[49]University of Pittsburgh, Pittsburgh, Pennsylvania 15260
[50]University of Science and Technology of China, Hefei 230026
[51]Seoul National University, Seoul 151-742
[52]Showa Pharmaceutical University, Tokyo 194-8543
[53]Soongsil University, Seoul 156-743
[54]University of South Carolina, Columbia, South Carolina 29208
[55]Sungkyunkwan University, Suwon 440-746
[56]School of Physics, University of Sydney, New South Wales 2006
[57]Department of Physics, Faculty of Science, University of Tabuk, Tabuk 71451
[58]Tata Institute of Fundamental Research, Mumbai 400005
[59]Excellence Cluster Universe, Technische Universität München, 85748 Garching
[60]Department of Physics, Technische Universität München, 85748 Garching
[61]Toho University, Funabashi 274-8510
[62]Department of Physics, Tohoku University, Sendai 980-8578
[63]Earthquake Research Institute, University of Tokyo, Tokyo 113-0032
[64]Department of Physics, University of Tokyo, Tokyo 113-0033
[65]Tokyo Institute of Technology, Tokyo 152-8550
[66]Tokyo Metropolitan University, Tokyo 192-0397
[67]CNP, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061
[68]Wayne State University, Detroit, Michigan 48202
[69]Yamagata University, Yamagata 990-8560
[70]Yonsei University, Seoul 120-749
[71]University of Miyazaki, Miyazaki 889-2192
[72]National Central University, Chung-li 32054
[73]Luther College, Decorah, Iowa 52101

(Received 21 March 2016; published 20 June 2016)

We have searched for the decay $B^0 \to \phi\gamma$ using the full Belle data set of 772×10^6 $B\bar{B}$ pairs collected at the $\Upsilon(4S)$ resonance with the Belle detector at the KEKB e^+e^- collider. No signal is observed, and we set an upper limit on the branching fraction of $\mathcal{B}(B^0 \to \phi\gamma) < 1.0 \times 10^{-7}$ at the 90% confidence level. This is the most stringent limit on this decay mode to date.

DOI: 10.1103/PhysRevD.93.111101

In the Standard Model (SM), the decay $B^0 \to \phi\gamma$ [1] proceeds through electroweak and gluonic $b \to d$ penguin annihilation processes as shown in Fig. 1. These amplitudes are proportional to the small Cabibbo-Kobayashi-Maskawa [2] matrix element V_{td} and thus are highly suppressed. The branching fraction has been estimated based on naive QCD factorization [3] and perturbative QCD [4] and found to be in the range 10^{-12} to 10^{-11}. However, the internal loop can also be mediated by non-SM particles such as a charged Higgs boson or supersymmetric squarks, and thus the decay is sensitive to new physics (NP). It is estimated that such NP could enhance the branching fraction to the level of 10^{-9} to 10^{-8} [3]. Experimentally, no evidence for this decay has been found, and the current upper limit on the branching fraction is 8.5×10^{-7} at the 90% confidence level (C.L.) [5]. Here, we present a search for this decay using the full Belle data set of 711 fb^{-1} recorded on the $\Upsilon(4S)$ resonance. This integrated luminosity corresponds to $(772 \pm 11) \times 10^6$ $B\bar{B}$ pairs, which is more than 6 times the amount of data used previously to search for this mode.

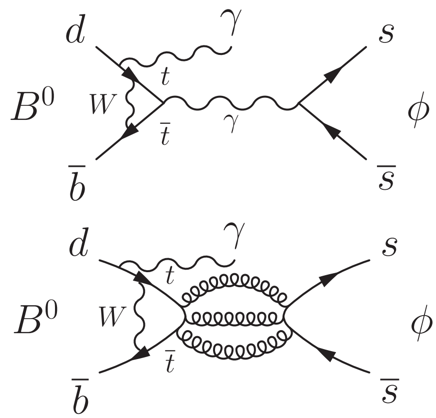

FIG. 1. Electroweak penguin (top) and gluonic penguin (bottom) contributions to $B^0 \to \phi\gamma$.

The Belle experiment ran at the KEKB asymmetric-energy e^+e^- collider located at the KEK laboratory [6]. The detector is a large-solid-angle magnetic spectrometer consisting of a silicon vertex detector (SVD), a 50-layer central drift chamber (CDC), an array of aerogel threshold Čerenkov counters (ACC), a barrel-like arrangement of time-of-flight scintillation counters (TOF), and an electromagnetic calorimeter (ECL) comprising CsI(Tl) crystals. These detector components are located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux-return located outside the coil (KLM) is instrumented to detect K_L^0 mesons and to identify muons. Two inner detector configurations were used: a 2.0 cm beampipe and a three-layer SVD were used for the first 140 fb^{-1} of data, while a 1.5 cm beampipe, a four-layer SVD, and a small-cell inner drift chamber were used for the remaining 571 fb^{-1} of data. The detector is described in detail elsewhere [7,8].

Candidate photons are required to have a momentum in the range $[2.0, 2.8]$ GeV$/c$ in the $\Upsilon(4S)$ center-of-mass (CM) frame. To reject neutral hadrons, the photon energy deposited in the 3×3 array of ECL crystals centered on the crystal with the highest energy must exceed 80% of the energy deposited in the corresponding 5×5 array of crystals. To reduce background from $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$ decays, we pair each photon candidate with all other photons in the event and, for each pairing, calculate π^0 and η likelihoods based on the invariant mass. We subsequently require these likelihoods to be less than 0.6, which preserves 97% of the signal while reducing the background by a factor of 2.

Candidate ϕ mesons are reconstructed via $\phi \to K^+K^-$ decays. Charged tracks are required to have a distance-of-closest-approach with respect to the interaction point of less than 3.0 cm along the z axis (antiparallel to the e^+ beam), and of less than 0.3 cm in the transverse plane. Kaons are identified using information from the CDC, TOF, and ACC detectors. This information is used to calculate relative likelihoods for hadron identification. A charged track with a likelihood ratio of $\mathcal{L}_K/(\mathcal{L}_\pi + \mathcal{L}_K) > 0.6$ is regarded as a kaon, where $\mathcal{L}_K(\mathcal{L}_\pi)$ is the relative likelihood of the track being a kaon (pion). The kaon identification efficiency is 85% and the probability for a pion to be misidentified as a kaon is 7%. Charged tracks that are consistent with the muon hypothesis based on information from the CDC and KLM are rejected, as are tracks consistent with the electron hypothesis based on information from the CDC and ECL. Oppositely charged kaon candidates are fit to a common vertex and required to have a vertex χ^2 less than 50. The K^+K^- invariant mass is required to be in the range $[1.000, 1.039]$ GeV$/c^2$, which corresponds to 4.5σ in resolution around the ϕ mass [9].

Candidate B mesons are identified using a modified beam-energy-constrained mass $M_{\mathrm{bc}} = \sqrt{E_{\mathrm{beam}}^2 - |\vec{p}_B c|^2}/c^2$, and the energy difference $\Delta E = E_B - E_{\mathrm{beam}}$, where E_{beam} is the beam energy and $\vec{p}_B$ and E_B are the momentum and energy, respectively, of the B^0 candidate. All quantities are evaluated in the CM frame. To improve the M_{bc} resolution, the momentum $\vec{p}_B$ is calculated as $\vec{p}_\phi + (\vec{p}_\gamma/|p_\gamma|)\sqrt{(E_{\mathrm{beam}} - E_\phi)^2}/c$, where $\vec{p}_\gamma$ is the photon momentum and $\vec{p}_\phi$ and E_ϕ are the momentum and energy, respectively, of the ϕ candidate. We require that events satisfy $M_{\mathrm{bc}} \in [5.25, 5.29]$ GeV$/c^2$ and $\Delta E \in [-0.30, 0.15]$ GeV; all events within this region are fitted. The signal yield is calculated in a smaller region $M_{\mathrm{bc}} \in [5.27, 5.29]$ GeV$/c^2$ and $\Delta E \in [-0.20, 0.10]$ GeV.

After applying the above selection criteria, less than 1% of events contain multiple B candidates. For these events we retain only the candidate that minimizes the difference $|M_{K^+K^-} - M_\phi|$. If there remains a choice of photons to be paired with the ϕ, we choose the one with the highest energy. According to Monte Carlo (MC) simulations, these criteria select the correct B candidate 96% of the time.

Charmless hadronic decays suffer from large backgrounds arising from continuum $e^+e^- \to q\bar{q}$ $(q = u, d, s, c)$ production. To suppress this background, we use a multivariate analyzer based on a neural network (NN) [10]. The NN uses the event topology and B-flavor-tagging information [11] to discriminate continuum events, which tend to be jet-like, from $B\bar{B}$ events, which tend to be spherical. The event shape variables include a set of 16 modified Fox-Wolfram moments [12]; the cosine of the angle between the z axis and the B flight direction; and the cosine of the angle between the B thrust axis [13] and the thrust axis of the non-B-associated tracks in the event. All of these quantities are evaluated in the CM frame.

The NN technique requires a training procedure. For this training we use signal and continuum MC events. The MC samples are obtained using EVTGEN [14] for event generation and GEANT3 [15] for modeling the detector response. Final-state radiation is taken into account using PHOTOS [16]. The NN generates an output variable C_{NN}, which ranges from -1 for background-like events to $+1$ for signal-like events. We require $C_{\rm NN} > 0.3$, which rejects 89% of continuum background while retaining 85% of the signal. We then translate $C_{\rm NN}$ to an empirical variable $C'_{\rm NN}$, defined as

$$C'_{\rm NN} = \ln\left(\frac{C_{\rm NN} - C_{\rm min}}{C_{\rm max} - C_{\rm NN}}\right), \quad (1)$$

where $C_{\rm min} = 0.3$ and $C_{\rm max} = 1.0$. This translation is convenient, as the $C'_{\rm NN}$ distribution for both signal and background is well modeled by a sum of Gaussian functions.

After the above selections, 961 events remain. The remaining background consists of continuum events and rare charmless b-decay processes. The latter shows peaking structure in the $M_{\rm bc}$ distribution, with the dominant contribution coming from $B \to K_1(1270)\gamma$, $K_1(1270) \to K\pi\pi$ decays. From a large MC study we find a negligible background contribution from $b \to c$ processes.

We calculate signal yields using an unbinned extended maximum likelihood fit to the observables $M_{\rm bc}$, ΔE, $C'_{\rm NN}$, and $\cos\theta_\phi$. The helicity angle θ_ϕ is the angle between the K^+ momentum and the opposite of the B flight direction in the ϕ rest frame. Signal $B^0 \to \phi\gamma$ decays are distributed as $1 - \cos^2\theta_\phi$, whereas continuum events are distributed approximately flat in $\cos\theta_\phi$. Thus this variable provides additional discrimination between signal and background. The likelihood function $\mathcal{L}$ is defined as

$$e^{-\sum_j Y_j} \prod_i^N \left(\sum_j Y_j \mathcal{P}_j(M^i_{\rm bc}, \Delta E^i, C'^i_{\rm NN}, \cos\theta^i_\phi)\right), \quad (2)$$

where N is the number of candidate events (961), $\mathcal{P}_j(M^i_{\rm bc}, \Delta E^i, C'^i_{\rm NN}, \cos\theta^i_\phi)$ is the probability density function (PDF) of component j for event i, and j runs over all signal and background components. The parameter Y_j is the fitted yield of component j. These yields are the only free parameters in the fit.

All PDFs are obtained from MC simulation studies. Correlations among the fit variables are found to be small, except for a correlation between $M_{\rm bc}$ and ΔE for the charmless background. Thus, except for this background, we factorize the PDFs as

$$\begin{aligned}&\mathcal{P}_j(M_{\rm bc}, \Delta E, C'_{\rm NN}, \cos\theta_\phi)\\ &\quad = \mathcal{P}_j(M_{\rm bc}) \cdot \mathcal{P}_j(\Delta E) \cdot \mathcal{P}_j(C'_{\rm NN}) \cdot \mathcal{P}_j(\cos\theta_\phi). \quad (3)\end{aligned}$$

The $M_{\rm bc}$ and ΔE distributions for signal are modeled with Crystal Ball functions [17], while the $C'_{\rm NN}$ and $\cos\theta_\phi$ distributions are modeled with a bifurcated Gaussian and the function $1 - \cos^2\theta_\phi$, respectively. The peak positions and resolutions of the $M_{\rm bc}$, ΔE, and $C'_{\rm NN}$ PDFs are adjusted to account for small data-MC differences observed in a high-statistics control sample of $B^0 \to K^{*0}(\to K^+\pi^-)\gamma$ decays, which have a similar topology as $B^0 \to \phi\gamma$.

For the charmless background, the $C'_{\rm NN}$ component is modeled with a Gaussian function. The peak position and resolution are adjusted from data-MC differences observed for the charmless background in the $B^0 \to K^{*0}(\to K^+\pi^-)\gamma$ control sample. The $M_{\rm bc}$ and ΔE components are modeled by a joint two-dimensional nonparametric function based on kernel estimation [18], to account for their correlation. The $\cos\theta_\phi$ distribution is modeled by a one-dimensional nonparametric function. For continuum background, the $M_{\rm bc}$ shape is modeled by an ARGUS function [19], and the $C'_{\rm NN}$ shape is modeled by the sum of two Gaussians having a common mean. The peak positions and resolutions are adjusted from data-MC differences observed for the continuum background of the control sample. The ΔE and $\cos\theta_\phi$ distributions are modeled by Chebyshev polynomials of the first and second order, respectively. All shape parameters of these PDFs are fixed to the corresponding MC values. To test the stability of the fitting procedure, we perform numerous fits on large ensembles of MC events; in all cases the input value is recovered within the statistical error.

The projections of the fit are shown in Fig. 2. The resulting branching fraction is calculated as

$$\mathcal{B}(B^0 \to \phi\gamma) = \frac{Y_{\rm sig}}{N_{B\bar{B}} \cdot \varepsilon \cdot \mathcal{B}(\phi \to K^+K^-)}, \quad (4)$$

where $Y_{\rm sig} = 3.4^{+4.6}_{-3.8}$ is the signal yield in the signal region; $\varepsilon = 0.296 \pm 0.001$ is the signal efficiency in this region as calculated from MC simulation; $N_{B\bar{B}} = (772 \pm 11) \times 10^6$ is the number of $B\bar{B}$ events; and $\mathcal{B}(\phi \to K^+K^-) = (48.9 \pm 0.5)\%$ is the branching fraction for $\phi \to K^+K^-$ [9]. The efficiency ε is corrected by a factor 1.024 ± 0.010 to account for a small difference in particle identification efficiencies between data and simulations. This correction is estimated from a sample of $D^{*+} \to D^0(\to K^-\pi^+)\pi^+$ decays [20]. In Eq. (4) we assume equal production of $B^0\bar{B}^0$ and B^+B^- pairs at the $\Upsilon(4S)$ resonance.

We observe no statistically significant signal and set an upper limit on the number of signal events by integrating the area under the likelihood function $\mathcal{L}(Y_{\rm sig})$. The value of $Y_{\rm sig}$ that corresponds to 90% of the total area from zero to infinity is taken as the 90% C.L. upper limit [21]. This value is converted to an upper limit on the branching fraction $\mathcal{B}$ using Eq. (4); the result is

$$\mathcal{B}(B^0 \to \phi\gamma) < 1.0 \times 10^{-7}. \quad (5)$$

We include systematic uncertainties (discussed below) in the upper limit by convolving the likelihood function with a Gaussian function whose width is set equal to the total

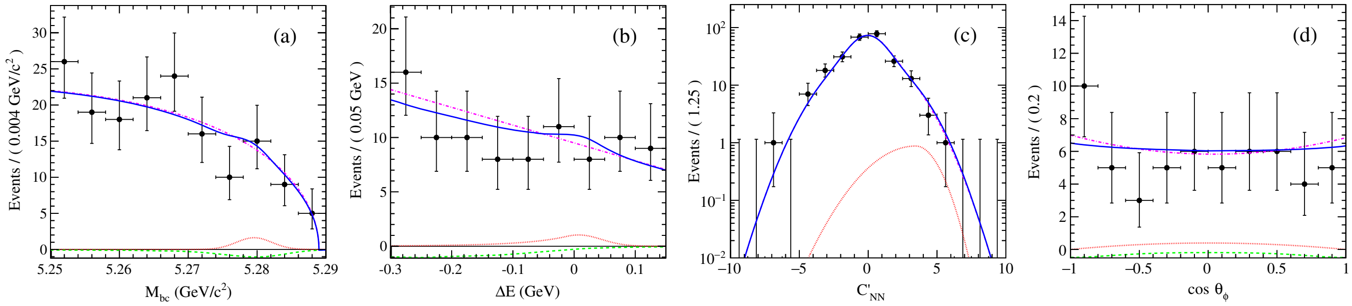

FIG. 2. Projections of the four-dimensional fit: (a) $M_{\rm bc}$ in the ΔE signal region; (b) ΔE in the $M_{\rm bc}$ signal region; (c) $C'_{\rm NN}$ in the $M_{\rm bc}$ and ΔE signal regions; and (d) $\cos\theta_\phi$ in the $M_{\rm bc}$ and ΔE signal regions. Plots (a), (b), and (d) also require $C'_{\rm NN} > 1$. The points with error bars show the data; the dotted (red) curves represent the signal; the dashed-dotted (magenta) curves represent continuum events; the dashed (green) curves represent the charmless background; and the solid (blue) curves represent the total.

systematic uncertainty. We perform this convolution before calculating the upper limit on $Y_{\rm sig}$.

The systematic uncertainties on the branching fraction are listed in Table I. The largest uncertainty is due to the fixed parameters in the PDFs. We evaluate this by varying each parameter individually according to its statistical uncertainty. The resulting changes in $Y_{\rm sig}$ are added in quadrature to obtain the systematic uncertainty. We evaluate, in a similar manner, the uncertainty due to errors in the calibration factors. The sum in quadrature of these two uncertainties is listed in Table I as the uncertainty due to PDF parametrization.

To test for potential bias in our fitting procedure, we fit a large ensemble of MC events. By comparing the mean of the yields obtained with the input value, a potential bias of -0.08 events is found. We attribute this to neglecting small correlations between the fitted variables and take this bias as a systematic uncertainty. The uncertainty due to the $C_{\rm NN}$ selection is determined by applying different $C_{\rm NN}$ criteria to the $B^0 \to K^{*0}\gamma$ control sample; the change in the branching fraction is taken as the systematic uncertainty. The uncertainty due to the background sample used to train the NN is taken to be the change in the control sample branching fraction when two different training samples are used: one from a sideband region in data, and the other from the same sideband region in MC simulation.

TABLE I. Systematic uncertainties on $\mathcal{B}(B^0 \to \phi\gamma)$ in units of number of events. We convert fractional errors to number of events for easy comparison. Uncertainties listed in the lower section are external to our analysis.

Source	Uncertainty (events)
PDF parameterization	$^{+1.21}_{-1.14}$
Fit bias	$^{+0.00}_{-0.08}$
$C_{\rm NN}$ selection efficiency	0.03
$C_{\rm NN}$ background sample	0.02
Tracking efficiency	0.02
PID efficiency	0.05
Photon reconstruction	0.08
MC statistics	0.01
$\mathcal{B}(\phi \to K^+K^-)$	0.03
Number of $B\bar{B}$ events	0.05
Total	$^{+1.22}_{-1.15}$

The systematic uncertainty due to charged track reconstruction is determined from a study of partially reconstructed $D^{*+} \to D^0(\to K_S^0\pi^+\pi^-)\pi^+$ decays and found to be 0.35% per track. An uncertainty due to particle identification of 0.8% per kaon is obtained from a study of $D^{*+} \to D^0(\to K^-\pi^+)\pi^+$ decays. The uncertainty on ε due to MC statistics is 0.2%, and the uncertainty on the number of $B\bar{B}$ pairs is 1.4%. The total systematic uncertainty is obtained by summing all individual contributions in quadrature; the result corresponds to ± 1.2 events.

In summary, we have searched for the decay $B^0 \to \phi\gamma$ using the full Belle data set. We find no evidence for this decay and set an upper limit on the branching fraction of $\mathcal{B}(B^0 \to \phi\gamma) < 1.0 \times 10^{-7}$ at 90% C.L. This limit is almost an order of magnitude lower than the previous most stringent result [5].

We thank the KEKB group for the excellent operation of the accelerator; the KEK cryogenics group for the efficient operation of the solenoid; and the KEK computer group, the National Institute of Informatics, and the PNNL/EMSL computing group for valuable computing and SINET4 network support. We acknowledge support from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan, the Japan Society for the Promotion of Science (JSPS), and the Tau-Lepton Physics Research Center of Nagoya University; the Australian Research Council; Austrian Science Fund under Grant No. P 22742-N16 and P 26794-N20; the National Natural Science Foundation of China under Contracts No. 10575109, No. 10775142, No. 10875115, No. 11175187, No. 11475187 and No. 11575017; the Chinese Academy of Science Center for Excellence in Particle Physics; the Ministry of Education, Youth and Sports of the Czech Republic under Contract No. LG14034; the Carl Zeiss Foundation, the Deutsche Forschungsgemeinschaft, the Excellence Cluster Universe, and the Volkswagen-Stiftung; the Department of Science and Technology of India; the Istituto Nazionale di Fisica Nucleare of Italy; the WCU program of the Ministry of Education, National Research Foundation (NRF) of Korea Grants No. 2011-0029457, No. 2012-0008143, No. 2012R1A1A2008330, No. 2013R1A1A3007772, No. 2014R1A2A2A01005286, No. 2014R1A2A2A01002734, No. 2015R1A2A2A01003280, No. 2015H1A2A1033649; the Basic Research Lab program under NRF Grant No. KRF-2011-0020333, Center for Korean J-PARC Users, No. NRF-2013K1A3A7A06056592; the Brain Korea 21-Plus program and Radiation Science Research Institute; the Polish Ministry of Science and Higher Education and the National Science Center; the Ministry of Education and Science of the Russian Federation and the Russian Foundation for Basic Research; the Slovenian Research Agency; Ikerbasque, Basque Foundation for Science and the Euskal Herriko Unibertsitatea (UPV/EHU) under program UFI 11/55 (Spain); the Swiss National Science Foundation; the Ministry of Education and the Ministry of Science and Technology of Taiwan; and the U.S. Department of Energy and the National Science Foundation. This work is supported by a Grant-in-Aid from MEXT for Science Research in a Priority Area ("New Development of Flavor Physics") and from JSPS for Creative Scientific Research ("Evolution of Tau-lepton Physics").

[1] Throughout this paper, charge-conjugate decay modes are implicitly included unless stated otherwise.

[2] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963); M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).

[3] X.-Q. Li, G.-R. Lu, R.-M. Wang, and Y. Yang, Eur. Phys. J. C **36**, 97 (2004); J. Hua, C. Kim, and Y. Li, Eur. Phys. J. C **69**, 139 (2010).

[4] C.-D. Lu, Y.-L. Shen, and W. Wang, Chin. Phys. Lett. **23**, 2684 (2006).

[5] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **72**, 091103 (2005).

[6] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods Phys. Res., Sect. A **499**, 1 (2003), and other papers included in this volume; T. Abe *et al.*, Prog. Theor. Exp. Phys. **2013** (2013) 03A001 and references therein.

[7] A. Abashian *et al.* (Belle Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 117 (2002); also see the detector section in J. Brodzicka *et al.*, Prog. Theor. Exp. Phys. **2012**, 04D001 (2012).

[8] Z. Natkaniec *et al.* (Belle SVD2 Group), Nucl. Instrum. Methods Phys. Res., Sect. A **560**, 1 (2006).

[9] K. A. Olive *et al.* (Particle Data Group), Chin. Phys. C **38**, 090001 (2014).

[10] M. Feindt and U. Kerzel, Nucl. Instrum. Methods Phys. Res., Sect. A **559**, 190 (2006).

[11] H. Kakuno *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **533**, 516 (2004).

[12] G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978); The modified moments used in this paper are described in S. H. Lee *et al.* (Belle Collaboration), Phys. Rev. Lett. **91**, 261801 (2003).

[13] S. Brandt, C. Peyrou, R. Sosnowski, and A. Wroblewski, Phys. Lett. **12**, 57 (1964).

[14] D. J. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 152 (2001).

[15] R. Brun *et al.*, GEANT 3.21, CERN Report No. DD/EE/84-1, 1984.

[16] P. Golonka and Z. Wăs, Eur. Phys. J. C **45**, 97 (2006).

[17] T. Skwarnicki, Ph.D. thesis, Institute of Nuclear Physics, Krakow, 1986 (Report No. DESY-F31-86-02).

[18] K. S. Cranmer, Comput. Phys. Commun. **136**, 198 (2001).

[19] H. Albrecht *et al.* (ARGUS Collaboration), Phys. Lett. B **241**, 278 (1990).

[20] E. Nakano, Nucl. Instrum. Methods Phys. Res., Sect. A **494**, 402 (2002).

[21] As our upper limit is obtained using a Bayesian method, it corresponds to a 90% "credibility level." However, we use the (frequentist) term "confidence level" following common convention.