First observation of radiative $B^{0}\to\phi K^{0}\gamma$ decays and measurements of their time-dependent CP violation

H. Sahoo *et al.* (Belle Collaboration)

Phys. Rev. D **84**, 071101 — Published 6 October 2011

DOI: 10.1103/PhysRevD.84.071101

First Observation of Radiative $B^0 \to \phi K^0 \gamma$ Decays and Measurements of Their Time-Dependent CP Violation

H. Sahoo,[7] T. E. Browder,[7] I. Adachi,[8] D. M. Asner,[36] V. Aulchenko,[1,34] A. M. Bakich,[43] E. Barberio,[25] K. Belous,[13] V. Bhardwaj,[37] B. Bhuyan,[9] A. Bondar,[1,34] A. Bozek,[30] M. Bračko,[23,15] O. Brovchenko,[17] A. Chen,[28] P. Chen,[29] B. G. Cheon,[6] K. Cho,[18] Y. Choi,[42] J. Dalseno,[24,45] Z. Doležal,[2] Z. Drásal,[2] S. Eidelman,[1,34] D. Epifanov,[1,34] J. E. Fast,[36] V. Gaur,[44] N. Gabyshev,[1,34] J. Haba,[8] K. Hayasaka,[26] H. Hayashii,[27] Y. Horii,[47] Y. Hoshi,[46] W.-S. Hou,[29] Y. B. Hsiung,[29] H. J. Hyun,[20] T. Iijima,[26] K. Inami,[26] A. Ishikawa,[47] R. Itoh,[8] M. Iwabuchi,[54] Y. Iwasaki,[8] T. Iwashita,[27] N. J. Joshi,[44] T. Julius,[25] J. H. Kang,[54] T. Kawasaki,[32] C. Kiesling,[24] H. J. Kim,[20] H. O. Kim,[20] J. B. Kim,[19] J. H. Kim,[18] K. T. Kim,[19] M. J. Kim,[20] S. K. Kim,[41] Y. J. Kim,[18] K. Kinoshita,[3] B. R. Ko,[19] N. Kobayashi,[38,49] S. Koblitz,[24] P. Kodyš,[2] S. Korpar,[23,15] P. Križan,[22,15] T. Kuhr,[17] R. Kumar,[37] T. Kumita,[50] A. Kuzmin,[1,34] Y.-J. Kwon,[54] J. S. Lange,[4] M. J. Lee,[41] S.-H. Lee,[19] J. Li,[41] Y. Li,[52] J. Libby,[10] C.-L. Lim,[54] Z. Q. Liu,[11] D. Liventsev,[14] R. Louvot,[21] D. Matvienko,[1,34] S. McOnie,[43] K. Miyabayashi,[27] H. Miyata,[32] Y. Miyazaki,[26] R. Mizuk,[14] G. B. Mohanty,[44] T. Mori,[26] M. Nakao,[8] Z. Natkaniec,[30] C. Ng,[48] S. Nishida,[8] K. Nishimura,[7] O. Nitoh,[51] T. Nozaki,[8] T. Ohshima,[26] S. Okuno,[16] S. L. Olsen,[41,7] Y. Onuki,[47] P. Pakhlov,[14] G. Pakhlova,[14] C. W. Park,[42] H. K. Park,[20] T. Peng,[40] R. Pestotnik,[15] M. Petrič,[15] L. E. Piilonen,[52] M. Prim,[17] M. Röhrken,[17] S. Ryu,[41] K. Sakai,[8] Y. Sakai,[8] T. Sanuki,[47] O. Schneider,[21] C. Schwanda,[12] A. J. Schwartz,[3] K. Senyo,[26] O. Seon,[26] M. E. Sevior,[25] M. Shapkin,[13] V. Shebalin,[1,34] T.-A. Shibata,[38,49] J.-G. Shiu,[29] B. Shwartz,[1,34] F. Simon,[24,45] P. Smerkol,[15] Y.-S. Sohn,[54] A. Sokolov,[13] E. Solovieva,[14] S. Stanič,[33] M. Starič,[15] M. Sumihama,[38,5] K. Sumisawa,[8] T. Sumiyoshi,[50] S. Suzuki,[39] G. Tatishvili,[36] Y. Teramoto,[35] K. Trabelsi,[8] M. Uchida,[38,49] T. Uglov,[14] Y. Unno,[6] S. Uno,[8] Y. Ushiroda,[8] S. E. Vahsen,[7] G. Varner,[7] A. Vinokurova,[1,34] M. Watanabe,[32] Y. Watanabe,[16] K. M. Williams,[52] E. Won,[19] B. D. Yabsley,[43] Y. Yamashita,[31] C. Z. Yuan,[11] Y. Yusa,[52] C. C. Zhang,[11] Z. P. Zhang,[40] V. Zhilich,[1,34] P. Zhou,[53] V. Zhulanov,[1,34] and A. Zupanc[17]

(The Belle Collaboration)

[1]Budker Institute of Nuclear Physics, Novosibirsk
[2]Faculty of Mathematics and Physics, Charles University, Prague
[3]University of Cincinnati, Cincinnati, Ohio 45221
[4]Justus-Liebig-Universität Gießen, Gießen
[5]Gifu University, Gifu
[6]Hanyang University, Seoul
[7]University of Hawaii, Honolulu, Hawaii 96822
[8]High Energy Accelerator Research Organization (KEK), Tsukuba
[9]Indian Institute of Technology Guwahati, Guwahati
[10]Indian Institute of Technology Madras, Madras
[11]Institute of High Energy Physics, Chinese Academy of Sciences, Beijing
[12]Institute of High Energy Physics, Vienna
[13]Institute of High Energy Physics, Protvino
[14]Institute for Theoretical and Experimental Physics, Moscow
[15]J. Stefan Institute, Ljubljana
[16]Kanagawa University, Yokohama
[17]Institut für Experimentelle Kernphysik, Karlsruher Institut für Technologie, Karlsruhe
[18]Korea Institute of Science and Technology Information, Daejeon
[19]Korea University, Seoul
[20]Kyungpook National University, Taegu
[21]École Polytechnique Fédérale de Lausanne (EPFL), Lausanne
[22]Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana
[23]University of Maribor, Maribor
[24]Max-Planck-Institut für Physik, München
[25]University of Melbourne, School of Physics, Victoria 3010
[26]Nagoya University, Nagoya
[27]Nara Women's University, Nara
[28]National Central University, Chung-li
[29]Department of Physics, National Taiwan University, Taipei
[30]H. Niewodniczanski Institute of Nuclear Physics, Krakow
[31]Nippon Dental University, Niigata
[32]Niigata University, Niigata
[33]University of Nova Gorica, Nova Gorica
[34]Novosibirsk State University, Novosibirsk

[35]Osaka City University, Osaka
[36]Pacific Northwest National Laboratory, Richland, Washington 99352
[37]Panjab University, Chandigarh
[38]Research Center for Nuclear Physics, Osaka
[39]Saga University, Saga
[40]University of Science and Technology of China, Hefei
[41]Seoul National University, Seoul
[42]Sungkyunkwan University, Suwon
[43]School of Physics, University of Sydney, NSW 2006
[44]Tata Institute of Fundamental Research, Mumbai
[45]Excellence Cluster Universe, Technische Universität München, Garching
[46]Tohoku Gakuin University, Tagajo
[47]Tohoku University, Sendai
[48]Department of Physics, University of Tokyo, Tokyo
[49]Tokyo Institute of Technology, Tokyo
[50]Tokyo Metropolitan University, Tokyo
[51]Tokyo University of Agriculture and Technology, Tokyo
[52]CNP, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061
[53]Wayne State University, Detroit, Michigan 48202
[54]Yonsei University, Seoul

We report the first observation of the radiative decay $B^0 \to \phi K^0\gamma$ using a data sample of 772×10^6 $B\overline{B}$ pairs collected at the $\Upsilon(4S)$ resonance with the Belle detector at the KEKB asymmetric-energy e^+e^- collider. We observe a signal of 37 ± 8 events with a significance of 5.4 standard deviations including systematic uncertainties. The measured branching fraction is $\mathcal{B}(B^0 \to \phi K^0\gamma) = (2.74 \pm 0.60 \pm 0.32) \times 10^{-6}$, where the uncertainties are statistical and systematic, respectively. We also report the first measurements of time-dependent CP violation parameters: $\mathcal{S}_{\phi K_S^0\gamma} = +0.74^{+0.72}_{-1.05}(\mathrm{stat})^{+0.10}_{-0.24}(\mathrm{syst})$ and $\mathcal{A}_{\phi K_S^0\gamma} = +0.35 \pm 0.58(\mathrm{stat})^{+0.23}_{-0.10}(\mathrm{syst})$. Furthermore, we measure $\mathcal{B}(B^+ \to \phi K^+\gamma) = (2.48 \pm 0.30 \pm 0.24) \times 10^{-6}$, $\mathcal{A}_{CP} = -0.03 \pm 0.11 \pm 0.08$ and find that the signal is concentrated in the $M_{\phi K}$ mass region near threshold.

PACS numbers: 14.40.Nd, 13.25.Hw, 11.30.Er

Rare radiative B meson decays play an important role in the search for physics beyond the standard model (SM). These are flavor changing neutral current decays, forbidden at tree level in the SM but allowed through electroweak loop processes. The loop can be mediated by non-SM particles (for example, charged Higgs or SUSY particles) and therefore is sensitive to new physics (NP). Here we report the first observation of a new $b \to s$ radiative penguin decay mode, $B^0 \to \phi K^0\gamma$, as well as measurements of its time-dependent CP asymmetry. This type of decay is sensitive to NP from right-handed currents [1] and will be useful for precise time-dependent measurements at future high-luminosity flavor facilities [2–4].

The emitted photons in $b \to s\gamma$ ($\overline{b} \to \overline{s}\gamma$) decays are predominantly left-handed (right-handed) in the SM, and hence the time-dependent CP asymmetry is suppressed by the quark mass ratio $(2m_s/m_b)$. The expected mixing-induced CP asymmetry parameter ($\mathcal{S}$) is $\mathcal{O}(3\%)$ and the direct CP asymmetry parameter ($\mathcal{A}$) is $\sim 0.6\%$ [1]. In several extensions of the SM, both photon helicities can contribute to the decay. Therefore, any significantly larger CP asymmetry would be clear evidence for NP. In contrast to $B^0 \to K^{*0}(\to K_S^0\pi^0)\gamma$ [5, 6], another related mode that is sensitive to NP, the time dependence of $B^0 \to \phi K_S^0\gamma$ can be measured from the $\phi \to K^+K^-$ decay and does not require a difficult measurement of the long lived K_S^0 decay inside the inner tracking volume or reconstruction of a low energy π^0. The $B \to \phi K\gamma$ mode can be used to search for a possible contribution from kaonic resonances decaying to ϕK. Furthermore, we can also probe the photon polarization using the angular distributions of the final state hadrons [7].

Results on $B \to \phi K\gamma$ decays have been reported by both Belle and BaBar collaborations based on 96×10^6 $B\overline{B}$ [8] and 228×10^6 $B\overline{B}$ [9] pairs, respectively. Only upper limits on $\mathcal{B}(B^0 \to \phi K^0\gamma)$ were given. Here we report the first observation of $B^0 \to \phi K^0\gamma$, the first measurements of time-dependent CP violation in this mode, as well as more precise measurements of $B^+ \to \phi K^+\gamma$ [10]. The data set used consists of 772×10^6 $B\overline{B}$ pairs collected at the $\Upsilon(4S)$ resonance with the Belle detector [11] at the KEKB asymmetric-energy e^+e^- (3.5 on 8.0 GeV) collider [12].

At KEKB, the $\Upsilon(4S)$ is produced with a Lorentz boost of $\beta\gamma = 0.425$ along the z axis, which is defined as opposite to the e^+ beam direction. In the decay chain $\Upsilon(4S) \to B^0\overline{B}^0 \to f_{\rm rec}f_{\rm tag}$, where one of the B mesons decays at time $t_{\rm rec}$ to the signal mode $f_{\rm rec}$ and the other decays at time $t_{\rm tag}$ to a final state $f_{\rm tag}$ that distinguishes between B^0 and $\overline{B}^0$, the decay rate has a time dependence given by

$$\mathcal{P}(\Delta t) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{4\tau_{B^0}}\Big\{1 + q\Big[\mathcal{S}\sin(\Delta m_d\Delta t) + \mathcal{A}\cos(\Delta m_d\Delta t)\Big]\Big\}. \quad (1)$$

Here τ_{B^0} is the neutral B lifetime, Δm_d is the mass difference between the two neutral B mass eigenstates, $\Delta t = t_{\rm rec} - t_{\rm tag}$, and the b-flavor charge q equals $+1$ (-1) when the tagging B meson is a B^0 ($\overline{B}^0$). Since the B^0 and $\overline{B}^0$ are approximately at rest in the $\Upsilon(4S)$ center-of-mass system (cms), Δt can be determined from Δz, the displacement in z between the two decay vertices: $\Delta t \simeq \Delta z/(\beta\gamma c)$.

Signal candidates are reconstructed in the $B^+ \to \phi K^+\gamma$ and $B^0 \to \phi K^0_S\gamma$ modes, with $\phi \to K^+K^-$ and $K^0_S \to \pi^+\pi^-$. Charged kaons are identified by requiring a likelihood ratio $\mathcal{L}_{K/\pi}$ [$= \mathcal{L}_K/(\mathcal{L}_K + \mathcal{L}_\pi)$] > 0.6, which is calculated using information from the aerogel Cherenkov, time-of-flight, and drift chamber detectors. This requirement has an efficiency of 90% for kaons and an 8% pion fake rate. A less restrictive selection $\mathcal{L}_{K/\pi} > 0.4$ is applied to the kaon candidates that are used to reconstruct the ϕ meson. The invariant mass of the ϕ candidates is required to satisfy $|M_{K^+K^-} - m_\phi| < 10$ MeV/c^2, where m_ϕ denotes the ϕ meson world-average mass [13]. The K^0_S selection criteria are the same as those described in Ref. [14]; the invariant mass of the pion pairs should be in the range $M_{\pi^+\pi^-} \in [482, 514]$ MeV/c^2. The high energy prompt photons must lie in the barrel region of the calorimeter (ECL), have a cms energy $E^{\rm cms}_\gamma \in [1.4, 3.4]$ GeV and a shower shape consistent with that of a photon. We also suppress the background photons from $\pi^0(\eta) \to \gamma\gamma$ using a likelihood $\mathcal{L}_{\pi^0}(\mathcal{L}_\eta) < 0.25$, as described in Ref. [15].

We combine a ϕ meson candidate, a charged or neutral kaon candidate, and a radiative photon to form a B meson. B candidates are identified using two kinematic variables: the energy difference $\Delta E \equiv E^{\rm cms}_B - E^{\rm cms}_{\rm beam}$ and the beam-energy-constrained mass $M_{\rm bc} \equiv \sqrt{(E^{\rm cms}_{\rm beam}/c^2)^2 - (p^{\rm cms}_B/c)^2}$, where $E^{\rm cms}_{\rm beam}$ is the beam energy in the cms, and $E^{\rm cms}_B$ and $p^{\rm cms}_B$ are the cms energy and momentum, respectively, of the reconstructed B candidate. In the $M_{\rm bc}$ calculation, the photon momentum is replaced by $(E^{\rm cms}_{\rm beam} - E^{\rm cms}_{\phi K})$ to improve its resolution. The candidates that satisfy the requirements $M_{\rm bc} > 5.2$ GeV/c^2 and $|\Delta E| < 0.3$ GeV are retained for further analysis. Using Monte Carlo (MC) simulations, we find nearly 12% (3%) of signal events in the charged (neutral) mode have more than one B candidate. In case of multiple candidates, we choose the best candidate based on a series of selection criteria, which depend on a χ^2 variable formed using the candidate's ϕ mass (and the K^0_S mass in the neutral mode) as well as the highest $E^{\rm cms}_\gamma$ and the highest $\mathcal{L}_{K/\pi}$ in the charged mode. For events with multiple candidates, this selection method chooses the correct B candidate for the charged (neutral) mode 57% (69%) of the time.

The dominant background comes from $e^+e^- \to q\overline{q}$ ($q = u, d, s, c$) continuum events. We use two event-shape variables (a Fisher discriminant formed from modified Fox-Wolfram moments [16] and the cosine of the angle between the B flight direction and the beam axis, $\cos\theta_B$, in the cms frame) to distinguish spherically symmetric $B\overline{B}$ events from the jet-like continuum background. From these variables we form a likelihood ratio, denoted by $\mathcal{R}_{s/b}$. We require $\mathcal{R}_{s/b} > 0.65$, which removes 91% of the continuum while retaining 76% of the signal. In addition to the continuum, various $B\overline{B}$ background sources are also studied. In the $B^0 \to \phi K^0_S\gamma$ mode, backgrounds from some $b \to c$ decays such as $D^0\pi^0$, $D^0\eta$ and $D^-\rho^+$, peak in the $M_{\rm bc}$ distribution. We remove the dominant peaking backgrounds by applying a veto to ϕK^0_S combinations consistent within detector resolution ($\pm 4\,\sigma$) with the nominal D mass [13]. Some of the charmless backgrounds, where the B meson decays to $\phi K^*(892)$, $\phi K\pi^0$ and $\phi K\eta$ also peak in $M_{\rm bc}$ but shift towards lower ΔE. Another significant background is non-resonant (NR) $B \to K^+K^-K\gamma$, which peaks in the ΔE-$M_{\rm bc}$ signal region; it is estimated using the ϕ mass sideband, $M_{K^+K^-} \in [1.05, 1.30]$ GeV/c^2, in data.

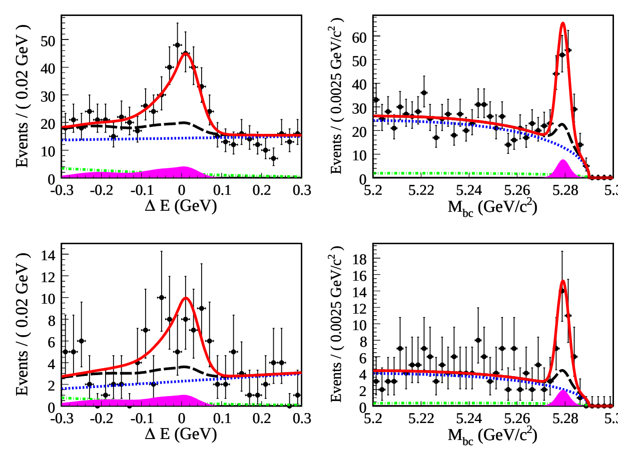

FIG. 1: ΔE and $M_{\rm bc}$ projections for $B^+ \to \phi K^+\gamma$ (upper) and $B^0 \to \phi K^0_S\gamma$ (lower). The ΔE projections include the requirement $M_{\rm bc} \in [5.27, 5.29]$ GeV/c^2 while the $M_{\rm bc}$ projections require $\Delta E \in [-0.08, 0.05]$ GeV. The points with error bars are the data. The curves show the total fit function (solid red), total background function (long-dashed black), continuum component (dotted blue), the $b \to c$ component (dashed-dotted green) and the non-resonant component as well as other charmless backgrounds (filled magenta histogram).

The signal yield is obtained from an extended unbinned maximum-likelihood (UML) fit to the two-dimensional ΔE-$M_{\rm bc}$ distribution. We model the shape for the signal component using the product of a Crystal Ball line shape [17] for ΔE and a Gaussian for $M_{\rm bc}$. The continuum background is represented by the product of a first-order polynomial for ΔE and an ARGUS [18] function for $M_{\rm bc}$. The $b \to c$ background is described by the product of a second-order polynomial for ΔE and the sum of an ARGUS and a Gaussian function for $M_{\rm bc}$. For the small charmless backgrounds (except for the NR component), we use the sum of two Gaussians for ΔE and a Gaussian for $M_{\rm bc}$. The probability density function (PDF) is the product of these two functional forms [19]. In the final fit the continuum parameters are allowed to vary while all other background parameters are fixed to the values from MC simulations. The shapes of the $b \to c$ and NR peaking background components are fixed to that of the signal. In the charged mode, the NR background yield, $(12.5 \pm 6.7)\%$ of the signal, is fixed from the ϕ mass sideband. Since the neutral mode is limited by statistics, we assume isospin symmetry and use the same NR fraction. The signal shapes are adjusted for small differences between MC simulations and data using a $B^0 \to K^*(892)^0(\to K^+\pi^-)\gamma$ control sample, with $M_{K^+\pi^-} \in [820, 970]$ MeV$/c^2$. The fit yields a signal of 144 ± 17 $B^+ \to \phi K^+\gamma$ and 37 ± 8 $B^0 \to \phi K^0_S\gamma$ events. The projections of the fit results onto ΔE and $M_{\rm bc}$ are shown in Fig. 1. The signal significance is defined as $\sqrt{-2\ \ln(\mathcal{L}_0/\mathcal{L}_{\rm max})}$, where $\mathcal{L}_{\rm max}$ is the maximum likelihood for the nominal fit and $\mathcal{L}_0$ is the corresponding value with the signal yield fixed to zero. The additive sources of systematic uncertainty (described below) are included in the significance by varying each by its error and taking the lowest significance. The signal in the charged mode has a significance of $9.6\,\sigma$, whereas that for the neutral mode is $5.4\,\sigma$.

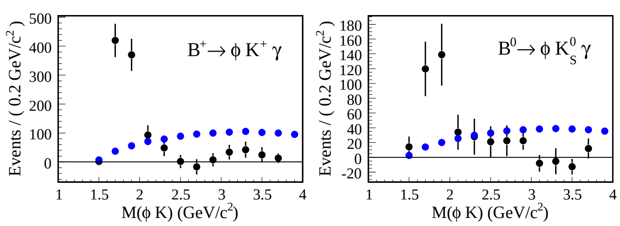

FIG. 2: Background-subtracted and efficiency-corrected ϕK mass distributions for the charged (left) and neutral (right) modes. The points with error bars represent the data. The yield in each bin is obtained by the fitting procedure described in the text. A three-body phase-space model from MC simulation is shown by the filled circles (blue points) and normalized to the total data signal yield.

To measure the $M_{\phi K}$ distribution, we repeat the fit in bins of ϕK mass and the resulting signal yields are corrected for the detection efficiency. Nearly 72% of the signal events are concentrated in the low-mass region, $M_{\phi K} \in [1.5, 2.0]$ GeV$/c^2$, as shown in Fig. 2. The MC efficiencies are reweighted according to this $M_{\phi K}$ dependence. These spectra are consistent with the expectations from the pQCD model for non-resonant $B \to \phi K\gamma$ decays [20]. With the present statistics no clear evidence is found for the existence of a kaonic resonance decaying to ϕK.

From the signal yield ($N_{\rm sig}$), we calculate the branching fraction ($\mathcal{B}$) as $N_{\rm sig}/(\epsilon \times N_{B\overline{B}} \times \mathcal{B}_{\rm sec})$, where ϵ is the weighted efficiency $[(15.3 \pm 0.1({\rm stat}))\%$ for the charged mode and $(10.0 \pm 0.1({\rm stat}))\%$ for the neutral mode], $N_{B\overline{B}}$ is the number of $B\overline{B}$ pairs in the data sample, and $\mathcal{B}_{\rm sec}$ is the product of daughter branching fractions [13]. We obtain $\mathcal{B}(B^+ \to \phi K^+\gamma) = (2.48 \pm 0.30 \pm 0.24) \times 10^{-6}$ and $\mathcal{B}(B^0 \to \phi K^0\gamma) = (2.74 \pm 0.60 \pm 0.32) \times 10^{-6}$, where the uncertainties are statistical and systematic, respectively.

We evaluate the systematic uncertainties on the signal yield by fitting the data with each fixed parameter varied by its $\pm 1\,\sigma$ error, and then taking the quadratic sum of all differences from the nominal value. The largest contribution of 8.0% arises from the NR yield. The other sources of systematic error are from charged track efficiency ($\sim$ 1.1% per track), photon detection efficiency (2.4%), particle identification (1.4%), number of produced $B\overline{B}$ pairs (1.4%), ϕ and K^0_S branching fractions (1.2%), K^0_S reconstruction (4.6%), and the requirement on $\mathcal{R}_{s/b}$ (0.3%). The statistical uncertainty on the MC efficiency after reweighting is 1.0% (1.2%) in the charged (neutral) mode. Furthermore, we assign a systematic error of 0.2% (2.7%) for possible fit bias, which is obtained from ensemble tests with MC pseudo-experiments. The total systematic uncertainty on the branching fraction is 9.5% (11.7%).

For the CP asymmetry fit, we select events in the signal region defined as $M_{\rm bc} \in [5.27, 5.29]$ GeV$/c^2$ and $\Delta E \in [-0.2, 0.1]$ GeV. Different selection criteria on $\mathcal{R}_{s/b}$ are used depending upon the flavor-tagging information. In addition, ECL endcap region photons are included in the analysis. We use a flavor tagging algorithm [21] to obtain the b-flavor charge q and a tagging quality factor $r \in [0, 1]$. The value $r = 0$ signifies no flavor discrimination while $r = 1$ implies unambiguous flavor assignment. The data are divided into seven r intervals. The vertex position for the $f_{\rm rec}$ decay is reconstructed using the two kaon tracks from the ϕ meson and that of the $f_{\rm tag}$ decay is from well-reconstructed tracks that are not assigned to $f_{\rm rec}$ [22]. The typical vertex reconstruction efficiency (z resolution) is 96% (115 μm) for $f_{\rm rec}$ and 94% (104 μm) for $f_{\rm tag}$. After all selection criteria are applied, we obtain 75 (436) events in the signal region for the CP fit with a purity of 45% (37%) in the neutral (charged) mode.

We determine $\mathcal{S}$ and $\mathcal{A}$ by performing an UML fit to the observed Δt distribution by maximizing the likelihood function $\mathcal{L}(\mathcal{S}, \mathcal{A}) = \prod_i P_i(\mathcal{S}, \mathcal{A}; \Delta t_i)$, where the product is over all events in the signal region. The likelihood P_i for each event is given by

$$P_i = (1 - f_{\rm ol}) \int \left[\sum_j f_j \mathcal{P}_j(\Delta t') R_j(\Delta t_i - \Delta t') \right] d(\Delta t') + f_{\rm ol} P_{\rm ol}(\Delta t_i), \quad (2)$$

where j runs over signal and all background components. $\mathcal{P}_j(\Delta t)$ is the corresponding PDF and $R_j(\Delta t)$ is the Δt resolution function. The fraction of each component (f_j) depends on the r region and is calculated for each event as a function of ΔE and $M_{\rm bc}$. The signal PDF is given by a modified form of Eq. (1) by fixing τ_{B^0} and Δm_d to their world-average values [13] and incorporating the effect of incorrect flavor assignment. The distribution is then convolved with a resolution function to take into account the finite vertex resolution. Since the NR component is expected to have the same NP as the signal $B \to \phi K \gamma$, we treat this as signal for the time-dependent fit [23]. For the other $B\overline{B}$ components, we use the same functional forms as signal with an effective lifetime taken from MC and CP parameters fixed to zero. For the continuum background, we use the functional form described in Ref. [22]; the parameters are determined from a fit to the Δt distribution of events in the data sideband $M_{\rm bc} < 5.26$ GeV/c^2 and $\Delta E \in [0.1, 0.3]$ GeV. The term $P_{\rm ol}(\Delta t)$ is a broad Gaussian function that represents an outlier component with a small fraction $f_{\rm ol}$. The PDFs and resolution functions are described in detail elsewhere [22].

We perform various consistency checks of the CP fitting technique. A lifetime fit to the $B^0 \to K^{*0}(\to K^+\pi^-)\gamma$, $B^+ \to \phi K^+ \gamma$ and $B^0 \to \phi K^0_S \gamma$ data sample yields 1.56 ± 0.03 ps, 1.70 ± 0.20 ps and 2.09 ± 0.45 ps, respectively. These are all consistent with the world-average values of the B lifetimes. The results of the CP asymmetry fit to the $B^0 \to K^{*0}(\to K^+\pi^-)\gamma$ ($\mathcal{S} = +0.02 \pm 0.06$, $\mathcal{A} = -0.06 \pm 0.04$) and $B^+ \to \phi K^+ \gamma$ ($\mathcal{S} = +0.25 \pm 0.33$, $\mathcal{A} = +0.18 \pm 0.26$) are consistent with zero. A fit to the sideband events in the $B^0 \to \phi K^0_S \gamma$ data sample gives an asymmetry consistent with zero ($\mathcal{S} = -1.77 \pm 1.30$, $\mathcal{A} = -0.04 \pm 0.14$).

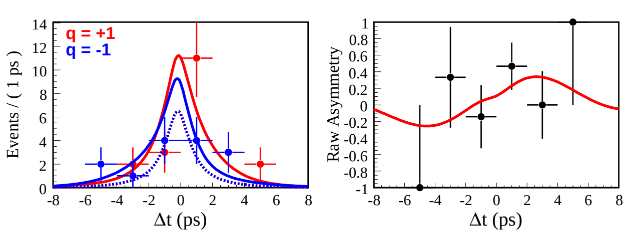

FIG. 3: Δt distributions for $q = +1$ and $q = -1$ (left) and the raw asymmetry (right) for well-tagged events. The dashed curves in the Δt plot are the sum of backgrounds while the solid curves are the sum of signal and backgrounds. The solid curve in the asymmetry plot shows the result of the UML fit.

The only free parameters in the CP fit are $\mathcal{S}$ and $\mathcal{A}$. The results of the fit are $\mathcal{S} = +0.74^{+0.72}_{-1.05}({\rm stat})^{+0.10}_{-0.24}({\rm syst})$ and $\mathcal{A} = +0.35 \pm 0.58({\rm stat})^{+0.23}_{-0.10}({\rm syst})$, where the uncertainties are obtained as described below. We define the raw asymmetry in each Δt bin by $(N_+ - N_-)/(N_+ + N_-)$, where N_+ (N_-) is the number of events with $q = +1$ (-1). Figure 3 shows the Δt distributions and raw asymmetry for events with good tagging quality ($r > 0.5$, 48% of the total).

We find that the error on $\mathcal{S}$ in the MINUIT minimization [24] is much smaller than the expectation from MC simulations and has a probability of only 0.6% [25]. This is due to low statistics and the presence of a single special event (with $\Delta t = -3.64$ ps and $r = 0.96$). A similar effect was found in our early time-dependent analyses of $B^0 \to \pi^+\pi^-$ [26]. Instead of the errors from MINUIT, we use the $\pm 68\%$ confidence intervals in the residual distributions of $\mathcal{S}$ and $\mathcal{A}$, determined from toy MC simulations as the statistical uncertainties on the result.

We evaluate the systematic uncertainties from the following sources. A significant contribution is from the vertex reconstruction (0.08 on $\mathcal{S}$, 0.04 on $\mathcal{A}$). We refit the data with each fixed parameter varied by its error to evaluate the uncertainties due to signal and background fractions (0.03, 0.07), resolution function (0.02, 0.03), ΔE-$M_{\rm bc}$ shapes (0.01, 0.01), continuum Δt PDF (0.01, 0.02), flavor tagging (0.01, 0.01) and effects of tagside interference [27] (0.004, 0.030). The uncertainty from physics parameters (τ_{B^0}, Δm_d), effective lifetime and CP asymmetry of the $B\overline{B}$ background, is (0.05, 0.03). We also include a possible fit bias due to low statistics and the proximity of the central value to the physical boundary ($^{+0.00}_{-0.22}$, $^{+0.21}_{-0.00}$). MC simulations show that this bias decreases to 0.04 with twice the signal yield. Adding all these contributions in quadrature, we obtain a systematic error of $^{+0.10}_{-0.24}$ on $\mathcal{S}$ and $^{+0.23}_{-0.10}$ on $\mathcal{A}$.

In summary, we report the first observation of a new radiative decay mode, $B^0 \to \phi K^0 \gamma$ using a data sample of 772×10^6 $B\overline{B}$ pairs. The observed signal yield is 37 ± 8 with a significance of $5.4\,\sigma$ including systematic uncertainties, and the measured branching fraction is $\mathcal{B}(B^0 \to \phi K^0 \gamma) = (2.74 \pm 0.60 \pm 0.32) \times 10^{-6}$. We also measure $\mathcal{B}(B^+ \to \phi K^+ \gamma) = (2.48 \pm 0.30 \pm 0.24) \times 10^{-6}$ with a significance of $9.6\,\sigma$. Furthermore, we measure the charge asymmetry $\mathcal{A}_{CP} = [N(B^-) - N(B^+)]/[N(B^-) + N(B^+)] = -0.03 \pm 0.11 \pm 0.08$, where $N(B^-)$ and $N(B^+)$ are the signal yields for B^- and B^+ decays, respectively. The signal events are mostly concentrated at low ϕK mass near threshold. The branching fractions and ϕK mass spectra are in agreement with the theoretical prediction of Ref. [20]. We also report the first measurements of time-dependent CP violation parameters in the neutral mode: $\mathcal{S} = +0.74^{+0.72+0.10}_{-1.05-0.24}$ and $\mathcal{A} = +0.35 \pm 0.58^{+0.23}_{-0.10}$. We have established that the mode $B^0 \to \phi K^0_S \gamma$ can be used at future high luminosity e^+e^- [2, 3] and hadronic facilities [4] to perform time-dependent CP violation measurements and to carry out sensitive tests for NP.

We thank the KEKB group for excellent operation of the accelerator, the KEK cryogenics group for efficient solenoid operations, and the KEK computer group and the NII for valuable computing and SINET3 network support. We acknowledge support from MEXT, JSPS and Nagoya's TLPRC (Japan); ARC and DIISR (Australia); NSFC (China); MSMT (Czechia); DST (India); MEST, NRF, NSDC of KISTI, and WCU (Korea); MNiSW (Poland); MES and RFAAE (Russia); ARRS (Slovenia); SNSF (Switzerland); NSC and MOE (Taiwan); and DOE (USA).

[1] D. Atwood, M. Gronau and A. Soni, Phys. Rev. Lett. **79**, 185 (1997). D. Atwood, T. Gershon, M. Hazumi and A. Soni, Phys. Rev. D **71**, 076003 (2005).

[2] T. Aushev *et al.*, arXiv:1002.5012. S. Hashimoto *et al.*, Letter of intent for KEK Super B Factory, KEK Report No. KEK-REPORT-2004-4, 2004.

[3] B. O'Leary *et al.*, arXiv:1008.1541.

[4] P. Ball *et al.*, arXiv:hep-ph/0003238.

[5] Y. Ushiroda *et al.* (Belle Collaboration), Phys. Rev. Lett. **94**, 231601 (2005).

[6] B. Aubert *et al.* (BaBar Collaboration), Phys. Rev. D **78**, 071102 (2008).

[7] V. D. Orlovsky and V. I. Shevchenko, Phys. Rev. D **77**, 093003 (2008). D. Atwood, T. Gershon, M. Hazumi and A. Soni, hep-ph/0701021.

[8] A. Drutskoy *et al.* (Belle Collaboration), Phys. Rev. Lett. **92**, 051801 (2004).

[9] B. Aubert *et al.* (BaBar Collaboration), Phys. Rev. D **75**, 051102 (2007).

[10] Throughout this Letter, the inclusion of the charge-conjugate decay mode is implied unless otherwise stated.

[11] A. Abashian *et al.* (Belle Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 117 (2002).

[12] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods Phys. Res., Sect. A **499**, 1 (2003), and other papers included in this volume.

[13] K. Nakamura *et al.* (Particle Data Group), J. Phys. G **37**, 075021 (2010).

[14] K.-F. Chen *et al.* (Belle Collaboration), Phys. Rev. D **72**, 012004 (2005).

[15] P. Koppenburg *et al.* (Belle Collaboration), Phys. Rev. Lett. **93**, 061803 (2004).

[16] R. O. Duda, P. E. Hart and D. G. Stork, Pattern Classification, 2nd Edition (John Wiley and Sons, 2001). G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978). S. H. Lee *et al.* (Belle Collaboration), Phys. Rev. Lett. **91**, 261801 (2003).

[17] T. Skwarnicki, Ph.D. thesis, Institute for Nuclear Physics, Krakow (DESY Internal Report, DESY F31-86-02, 1986). The function is widely used to describe asymmetric distributions caused by shower leakage in crystal calorimeters.

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

[19] $f(\Delta E, M_{\rm bc}) = [G_1(\Delta E - E_1) + G_2(\Delta E - E_2)]\, G(M_{\rm bc} - M_0)$, where G_1, G_2 and G are Gaussian functions and E_1, E_2 and M_0 are constants.

[20] C. H. Chen and H.-n. Li, Phys. Rev. D **70**, 054006 (2004) and private communication with H.-n. Li. The pQCD model is in qualitative agreement with our data after including the kinematic effect of the kaon mass.

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

[22] H. Tajima *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **533**, 370 (2004).

[23] Private communication with A. Soni.

[24] F. James and M. Roos, Comput. Phys. Commun. **10**, 343 (1975).

[25] The MINOS errors are $^{+0.32}_{-0.45}$ on $\mathcal{S}$ and ± 0.45 on $\mathcal{A}$ for data, while the toy MC distributions with input values equal to those measured in data have a width of $^{+0.72}_{-1.05}$ for $\mathcal{S}$ and ± 0.58 for $\mathcal{A}$. The correlation between $\mathcal{S}$ and $\mathcal{A}$ is found to be 0.009 in data.

[26] K. Abe *et al.* (Belle Collaboration), Phys. Rev. D **68**, 012001 (2003).

[27] O. Long, M. Baak, R. N. Cahn and D. Kirkby, Phys. Rev. D **68**, 034010 (2003).