Evidence for the suppressed decay $B^- \to DK^-$, $D \to K^+\pi^-\pi^0$

M. Nayak,[16] J. Libby,[16] K. Trabelsi,[12] I. Adachi,[12] H. Aihara,[55] D. M. Asner,[42] T. Aushev,[20] A. M. Bakich,[49] A. Bala,[43] P. Behera,[16] K. Belous,[18] V. Bhardwaj,[34] G. Bonvicini,[60] A. Bozek,[38] M. Bračko,[27,21] T. E. Browder,[11] D. Červenkov,[5] M.-C. Chang,[8] P. Chang,[37] V. Chekelian,[28] A. Chen,[35] B. G. Cheon,[10] R. Chistov,[20] I.-S. Cho,[62] K. Cho,[24] V. Chobanova,[28] Y. Choi,[48] D. Cinabro,[60] J. Dalseno,[28,51] M. Danilov,[20,30] Z. Doležal,[5] Z. Drásal,[5] D. Dutta,[15] S. Eidelman,[4] S. Esen,[6] H. Farhat,[60] J. E. Fast,[42] T. Ferber,[7] V. Gaur,[50] N. Gabyshev,[4] S. Ganguly,[60] R. Gillard,[60] Y. M. Goh,[10] B. Golob,[63,21] J. Haba,[12] H. Hayashii,[34] Y. Horii,[33] Y. Hoshi,[53] W.-S. Hou,[37] H. J. Hyun,[26] T. Iijima,[33,32] A. Ishikawa,[54] T. Iwashita,[34] I. Jaegle,[11] T. Julius,[29] D. H. Kah,[26] E. Kato,[54] D. Y. Kim,[47] H. J. Kim,[26] J. B. Kim,[25] M. J. Kim,[26] Y. J. Kim,[24] K. Kinoshita,[6] J. Klucar,[21] B. R. Ko,[25] P. Kodyš,[5] S. Korpar,[27,21] P. Krishnan,[16] P. Križan,[63,21] P. Krokovny,[4] T. Kuhr,[23] T. Kumita,[57] A. Kuzmin,[4] Y.-J. Kwon,[62] S.-H. Lee,[25] J. Li,[46] Y. Li,[59] L. Li Gioi,[28] Y. Liu,[6] D. Liventsev,[12] P. Lukin,[4] H. Miyake,[12] R. Mizuk,[20,30] G. B. Mohanty,[50] A. Moll,[28,51] T. Mori,[32] N. Muramatsu,[45] R. Mussa,[19] Y. Nagasaka,[13] M. Nakao,[12] E. Nedelkovska,[28] K. Negishi,[54] C. Ng,[55] N. K. Nisar,[50] O. Nitoh,[58] S. Ogawa,[52] S. Okuno,[22] Y. Onuki,[55] P. Pakhlov,[20,30] G. Pakhlova,[20] C. W. Park,[48] H. Park,[26] T. K. Pedlar,[64] M. Petrič,[21] L. E. Piilonen,[59] M. Ritter,[28] M. Röhrken,[23] A. Rostomyan,[7] H. Sahoo,[11] T. Saito,[54] Y. Sakai,[12] S. Sandilya,[50] L. Santelj,[21] T. Sanuki,[54] V. Savinov,[44] O. Schneider,[65] G. Schnell,[1,14] C. Schwanda,[17] A. J. Schwartz,[6] K. Senyo,[61] O. Seon,[32] M. E. Sevior,[29] M. Shapkin,[18] C. P. Shen,[2] T.-A. Shibata,[56] J.-G. Shiu,[37] B. Shwartz,[4] A. Sibidanov,[49] F. Simon,[28,51] Y.-S. Sohn,[62] A. Sokolov,[18] E. Solovieva,[20] M. Starič,[21] M. Steder,[7] Z. Suzuki,[54] U. Tamponi,[19,66] G. Tatishvili,[42] Y. Teramoto,[41] M. Uchida,[56] T. Uglov,[20,31] Y. Unno,[10] S. Uno,[12] P. Urquijo,[3] S. E. Vahsen,[11] C. Van Hulse,[1] P. Vanhoefer,[28] G. Varner,[11] K. E. Varvell,[49] M. N. Wagner,[9] C. H. Wang,[36] M.-Z. Wang,[37] Y. Watanabe,[22] K. M. Williams,[59] E. Won,[25] Y. Yamashita,[39] S. Yashchenko,[7] Y. Yusa,[40] V. Zhilich,[4] V. Zhulanov,[4] and A. Zupanc[23]

(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 and Novosibirsk State University, Novosibirsk 630090*
[5]*Faculty of Mathematics and Physics, Charles University, 121 16 Prague*
[6]*University of Cincinnati, Cincinnati, Ohio 45221*
[7]*Deutsches Elektronen–Synchrotron, 22607 Hamburg*
[8]*Department of Physics, Fu Jen Catholic University, Taipei 24205*
[9]*Justus-Liebig-Universität Gießen, 35392 Gießen*
[10]*Hanyang University, Seoul 133-791*
[11]*University of Hawaii, Honolulu, Hawaii 96822*
[12]*High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801*
[13]*Hiroshima Institute of Technology, Hiroshima 731-5193*
[14]*Ikerbasque, 48011 Bilbao*
[15]*Indian Institute of Technology Guwahati, Assam 781039*
[16]*Indian Institute of Technology Madras, Chennai 600036*
[17]*Institute of High Energy Physics, Vienna 1050*
[18]*Institute for High Energy Physics, Protvino 142281*
[19]*INFN—Sezione di Torino, 10125 Torino*
[20]*Institute for Theoretical and Experimental Physics, Moscow 117218*
[21]*J. Stefan Institute, 1000 Ljubljana*
[22]*Kanagawa University, Yokohama 221-8686*
[23]*Institut für Experimentelle Kernphysik, Karlsruher Institut für Technologie, 76131 Karlsruhe*
[24]*Korea Institute of Science and Technology Information, Daejeon 305-806*
[25]*Korea University, Seoul 136-713*
[26]*Kyungpook National University, Daegu 702-701*
[27]*University of Maribor, 2000 Maribor*
[28]*Max-Planck-Institut für Physik, 80805 München*
[29]*School of Physics, University of Melbourne, Victoria 3010*
[30]*Moscow Physical Engineering Institute, Moscow 115409*
[31]*Moscow Institute of Physics and Technology, Moscow Region 141700*
[32]*Graduate School of Science, Nagoya University, Nagoya 464-8602*

^{33}*Kobayashi-Maskawa Institute, Nagoya University, Nagoya 464-8602*
^{34}*Nara Women's University, Nara 630-8506*
^{35}*National Central University, Chung-li 32054*
^{36}*National United University, Miao Li 36003*
^{37}*Department of Physics, National Taiwan University, Taipei 10617*
^{38}*H. Niewodniczanski Institute of Nuclear Physics, Krakow 31-342*
^{39}*Nippon Dental University, Niigata 951-8580*
^{40}*Niigata University, Niigata 950-2181*
^{41}*Osaka City University, Osaka 558-8585*
^{42}*Pacific Northwest National Laboratory, Richland, Washington 99352*
^{43}*Panjab University, Chandigarh 160014*
^{44}*University of Pittsburgh, Pittsburgh, Pennsylvania 15260*
^{45}*Research Center for Electron Photon Science, Tohoku University, Sendai 980-8578*
^{46}*Seoul National University, Seoul 151-742*
^{47}*Soongsil University, Seoul 156-743*
^{48}*Sungkyunkwan University, Suwon 440-746*
^{49}*School of Physics, University of Sydney, New South Wales 2006*
^{50}*Tata Institute of Fundamental Research, Mumbai 400005*
^{51}*Excellence Cluster Universe, Technische Universität München, 85748 Garching*
^{52}*Toho University, Funabashi 274-8510*
^{53}*Tohoku Gakuin University, Tagajo 985-8537*
^{54}*Tohoku University, Sendai 980-8578*
^{55}*Department of Physics, University of Tokyo, Tokyo 113-0033*
^{56}*Tokyo Institute of Technology, Tokyo 152-8550*
^{57}*Tokyo Metropolitan University, Tokyo 192-0397*
^{58}*Tokyo University of Agriculture and Technology, Tokyo 184-8588*
^{59}*CNP, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061*
^{60}*Wayne State University, Detroit, Michigan 48202*
^{61}*Yamagata University, Yamagata 990-8560*
^{62}*Yonsei University, Seoul 120-749*
^{63}*Faculty of Mathematics and Physics, University of Ljubljana, 1000 Ljubljana*
^{64}*Luther College, Decorah, Iowa 52101*
^{65}*École Polytechnique Fédérale de Lausanne (EPFL), Lausanne 1015*
^{66}*University of Torino, 10124 Torino*

We report a study of the suppressed decay $B^- \to DK^-$, $D \to K^+\pi^-\pi^0$, where D denotes either a D^0 or a $\bar{D}^0$ meson. The decay is sensitive to the CP-violating parameter ϕ_3. Using a data sample of 772×10^6 $B\bar{B}$ pairs collected at the $\Upsilon(4S)$ resonance with the Belle detector, we measure the ratio of branching fractions of the above suppressed decay to the favored decay $B^- \to DK^-$, $D \to K^-\pi^+\pi^0$. Our result is $R_{DK} = [1.98 \pm 0.62(\text{stat}) \pm 0.24(\text{syst})] \times 10^{-2}$, which indicates the first evidence of the signal for this suppressed decay with a significance of 3.2 standard deviations. We measure the direct CP asymmetry between the suppressed B^- and B^+ decays to be $A_{DK} = 0.41 \pm 0.30(\text{stat}) \pm 0.05(\text{syst})$. We also report measurements for the analogous quantities $R_{D\pi}$ and $A_{D\pi}$ for the decay $B^- \to D\pi^-$, $D \to K^+\pi^-\pi^0$.

Several hadronic weak decays related by the combined charge-conjugation and parity (CP) transformations exhibit different behavior. Such violation of CP symmetry is described by the Standard Model of particle physics via an irreducible complex phase in the 3×3 Cabibbo–Kobayashi–Maskawa (CKM) quark mixing matrix [1], which has elements $V_{qq'}$, with $q = u$, c, t and $q' = d$, s, b. The unitarity triangle (UT) is used to represent the amount of CP violation parametrized by the CKM matrix. The UT angle $\phi_3 = \gamma \equiv \arg(-V_{ud}V_{ub}^*/V_{cd}V_{cb}^*)$ is less precisely measured compared to the other two angles $\phi_1(\equiv \beta)$ and $\phi_2(\equiv \alpha)$. The particular importance of improving the determination of ϕ_3 lies in the fact that it is the only CP-violating parameter that describes the UT that can be measured solely in tree-level processes. As a result, such measurements provide a benchmark to search for new physics contributions in loop-dominated processes that would otherwise constrain the UT.

Various methods to determine ϕ_3 in the tree decay $B^- \to DK^-$, where D is a D^0 or $\bar{D}^0$ decaying to a common final state [2], have been proposed [3–5]. In this paper, we focus on the Atwood–Dunietz–Soni (ADS) method [4] using the decay $B^- \to DK^-$ followed by $D \to K^+\pi^-\pi^0$. Several ADS measurements have been made using

$D \to K^+ \pi^-$ [6–10]. However, given a significantly larger branching fraction for $\bar{D}^0 \to K^+ \pi^- \pi^0$ [(13.9 ± 0.5)%] than $\bar{D}^0 \to K^+ \pi^-$ [(3.89 ± 0.05)%] [11], the former mode is potentially more sensitive to ϕ_3 despite a reduced acceptance owing to the presence of a π^0 meson in the final state. Herein, we search for $B^- \to [K^+ \pi^- \pi^0]_D K^-$ events for the first time in Belle, where the favored $B^- \to D^0 K^-$ decay followed by the doubly Cabibbo-suppressed (DCS) $D^0 \to K^+ \pi^- \pi^0$ decay interferes with the suppressed $B^- \to \bar{D}^0 K^-$ decay followed by the Cabibbo-favored (CF) $\bar{D}^0 \to K^+ \pi^- \pi^0$ decay. The interference between the two amplitudes can lead to a large direct CP asymmetry between the suppressed decays. We use $B^- \to D\pi^-$ as a control channel because of the kinematic similarity to $B^- \to DK^-$ and its much larger branching fraction.

One observable measured is the ratio of the suppressed to favored branching fractions,

$$R_{DK} = \frac{\mathcal{B}([K^+ \pi^- \pi^0]_D K^-) + \mathcal{B}([K^- \pi^+ \pi^0]_D K^+)}{\mathcal{B}([K^- \pi^+ \pi^0]_D K^-) + \mathcal{B}([K^+ \pi^- \pi^0]_D K^+)}$$
$$= r_B^2 + r_D^2 + 2r_B r_D R_{K\pi\pi^0} \cos\phi_3 \cos(\delta_B + \delta_D^{K\pi\pi^0}); \tag{1}$$

the second is the direct CP asymmetry,

$$A_{DK} = \frac{\mathcal{B}([K^+ \pi^- \pi^0]_D K^-) - \mathcal{B}([K^- \pi^+ \pi^0]_D K^+)}{\mathcal{B}([K^+ \pi^- \pi^0]_D K^-) + \mathcal{B}([K^- \pi^+ \pi^0]_D K^+)}$$
$$= \frac{2r_B r_D R_{K\pi\pi^0} \sin\phi_3 \sin(\delta_B + \delta_D^{K\pi\pi^0})}{r_B^2 + r_D^2 + 2r_B r_D R_{K\pi\pi^0} \cos\phi_3 \cos(\delta_B + \delta_D^{K\pi\pi^0})}, \tag{2}$$

where r_B and δ_B are the absolute ratio and strong-phase difference between the suppressed $B^- \to \bar{D}^0 K^-$ decay and the favored $B^- \to D^0 K^-$ decay amplitudes. Furthermore, the ratio of DCS and CF D decays r_D is defined via

$$r_D^2 \equiv \frac{\Gamma(D^0 \to K^+ \pi^- \pi^0)}{\Gamma(D^0 \to K^- \pi^+ \pi^0)} = \frac{\int d\vec{\mathbf{m}} A_{\text{DCS}}^2(\vec{\mathbf{m}})}{\int d\vec{\mathbf{m}} A_{\text{CF}}^2(\vec{\mathbf{m}})} \tag{3}$$

and the coherence factor $R_{K\pi\pi^0}$ and average strong-phase difference $\delta_D^{K\pi\pi^0}$ [12] via

$$R_{K\pi\pi^0} e^{i\delta_D^{K\pi\pi^0}} \equiv \frac{\int d\vec{\mathbf{m}} A_{\text{DCS}}(\vec{\mathbf{m}}) A_{\text{CF}}(\vec{\mathbf{m}}) e^{i\delta(\vec{\mathbf{m}})}}{\sqrt{\int d\vec{\mathbf{m}} A_{\text{DCS}}^2(\vec{\mathbf{m}}) \int d\vec{\mathbf{m}} A_{\text{CF}}^2(\vec{\mathbf{m}})}}. \tag{4}$$

Here, $A_{\text{CF}}(\vec{\mathbf{m}})$ and $A_{\text{DCS}}(\vec{\mathbf{m}})$ are the magnitudes of the CF and DCS amplitudes, respectively; $\delta(\vec{\mathbf{m}})$ is the relative strong phase; and $\vec{\mathbf{m}} \equiv [m_{K\pi}^2, m_{K\pi^0}^2]$ indicates a point in the Dalitz plane.

The definition of $R_{K\pi\pi^0}$ is such that its value is bounded between zero and one. Sensitivity to ϕ_3 through measurements of R_{DK} and A_{DK} is maximal when $R_{K\pi\pi^0}$ is unity. The measured value of $R_{K\pi\pi^0}$ is 0.84 ± 0.07 [13], which means that these observables are suitable to obtain information about ϕ_3. The previous measurement of this

channel [14] has constrained R_{DK} to be less than 2.1×10^{-2} at the 90% confidence level; no limit on A_{DK} is presented.

The observables for the $B^- \to D\pi^-$ mode are $R_{D\pi}$ and $A_{D\pi}$. They can be defined using Eqs. (1) and (2) with the following substitutions: $K \to \pi$ for the B daughter, $r_B \to r_B^{D\pi}$, and $\delta_B \to \delta_B^{D\pi}$. Here, $r_B^{D\pi}$ and $\delta_B^{D\pi}$ are the absolute ratio and strong-phase difference between the suppressed and favored $B^- \to D\pi^-$ decay amplitudes. The sensitivity to ϕ_3 is reduced in this mode because $r_B^{D\pi}$ is approximately an order of magnitude smaller than r_B. There have been no previous measurements of $R_{D\pi}$ and $A_{D\pi}$. In Ref. [15] it was shown that the corrections due to D mixing on $R_{D\pi}$ and $A_{D\pi}$ are potentially large; therefore, such corrections would need to be taken into account if these measurements are used in the determination of ϕ_3.

Our measurement uses a data sample of $772 \times 10^6 B\bar{B}$ pairs, collected with the Belle detector [16] located at the KEKB asymmetric energy $e^+ e^-$ (3.5 on 8 GeV) collider [17] operating near the $\Upsilon(4S)$ resonance. The principal detector elements used in this analysis are a silicon vertex detector, a 50-layer central drift chamber (CDC), an array of aerogel threshold Cherenkov counters (ACC), a barrel-like arrangement of time-of-flight scintillation counters (TOF), and an electromagnetic calorimeter comprised of CsI(Tl) crystals located inside a superconducting solenoid coil that provides a 1.5 T magnetic field.

We reconstruct π^0 candidates from photon pairs that have a momentum greater than 400 MeV/c in the $e^+ e^-$ center-of-mass (CM) frame and an invariant mass between 120 and 145 MeV/c^2, which corresponds to approximately $\pm 3.2\sigma$ in resolution around the nominal π^0 mass [11]. Each photon candidate is required to have an energy greater than 50 MeV. We apply a mass-constrained fit to the π^0 candidate to improve its momentum resolution.

Neutral D meson candidates are reconstructed from a pair of oppositely charged tracks and a π^0 candidate. Each track must have a distance of closest approach to the interaction point of less than 0.2 cm in the plane transverse to the positron beam direction and less than 1.5 cm along the positron beam axis. We also define L_K (L_π), the likelihood of a track being a kaon (pion), based on particle identification (PID) information [18] from the ACC and the TOF, combined with specific ionization measured in the CDC. We apply likelihood-ratio requirements of $L(K/\pi) = \frac{L_K}{L_K + L_\pi} > 0.6$ for a kaon candidate and $L(K/\pi) < 0.4$ for a pion candidate. The efficiency to identify a kaon (pion) is approximately 83% (88%) averaged over momentum, and the probability of misidentifying a pion (kaon) as a kaon (pion) is approximately 8% (7%). The invariant mass of $K\pi\pi^0$ candidates is required to satisfy $1.804 \text{ GeV}/c^2 < M_{K\pi\pi^0} < 1.885 \text{ GeV}/c^2$, which corresponds to approximately $\pm 2.5\sigma$ in resolution around the nominal D mass [11]. To improve the four-momentum resolution of the daughters, we apply a D mass constrained fit.

A B meson candidate is reconstructed by combining the D candidate with a charged hadron. The same set of $L(K/\pi)$ requirements is applied for the prompt track as that used for D meson reconstruction. The signal is identified with the beam-energy-constrained mass $M_{\mathrm{bc}} = c^{-2}\sqrt{E_{\mathrm{beam}}^2 - |\vec{p}_B|^2 c^2}$ and the energy difference $\Delta E = E_B - E_{\mathrm{beam}}$, where E_{beam} is the beam energy and $\vec{p}_B$ (E_B) is the momentum (energy) of the B meson candidates in the CM frame. For $B \to DK$ decays, M_{bc} peaks at the nominal mass of the B meson [11], and ΔE peaks at zero. We select candidates in the ranges 5.27 GeV/$c^2 <$ $M_{\mathrm{bc}} < 5.29$ GeV/c^2 and -0.1 GeV $< \Delta E < 0.2$ GeV.

To suppress background coming from the $D^{*\pm} \to D\pi^\pm$ decays in $e^+e^- \to c\bar{c}$, we use the mass difference between the $D^{*\pm}$ and D candidates (ΔM). We reconstruct $D^{*\pm}$ candidates from the D meson used for B reconstruction and a $\pi^\pm$ candidate not used in the B reconstruction. No PID requirement is applied to the $\pi^\pm$ because of its low momentum when coming from the $D^{*\pm}$ decay. After requiring $\Delta M > 0.15$ GeV/c^2, we remove 99% of $D^{*\pm}$ backgrounds and 17% of all $c\bar{c}$ backgrounds. The relative loss of signal efficiency is 3.4%.

A possible source of peaking background is the favored $B^- \to [K^-\pi^+\pi^0]_D h^-$ $(h = K$ or $\pi)$ decay, which can contribute to the signal region of the respective suppressed decay, due to misidentification of both the K^- and π^+ mesons in the D decay. To reject this background, we veto events satisfying 1.804 GeV/$c^2 < M_{K\pi\pi^0} < 1.885$ GeV/c^2 when the mass assignments of the K^- and π^+ are exchanged. This criterion reduces the background to a negligible level with a relative loss of signal efficiency of around 17%. About 6% of events have multiple B candidates; the candidate with $M_{K\pi\pi^0}$ and M_{bc} most consistent with the corresponding nominal values is retained for further analysis.

The dominant remaining background for both the favored and the suppressed Dh decays comes from $e^+e^- \to q\bar{q}$ $(q = u, d, s,$ or $c)$ continuum events. The daughters from $B\bar{B}$ events tend to emerge isotropically in the CM frame, whereas the particles from continuum events are collimated into back-to-back jets. We exploit this difference in event topology by using a neural network [7,19] to combine shape variables that describe the particle distribution with other properties of the event that differentiate between $q\bar{q}$ and $B\bar{B}$ events.

The neural network utilizes the following nine input variables: 1) the likelihood ratio of the Fisher discriminant formed from 17 modified Fox–Wolfram moments [20]; 2) the absolute value of the cosine of the angle in the CM frame between the thrust axis of the B decay and that of the remaining particles in the event; 3) the vertex separation between the B candidate and the remaining charged tracks along the beam direction; 4) the cosine of the angle between the direction of the K candidate from the D decay and the direction opposite the flight of the B

candidate measured in the D rest frame; 5) the absolute value of the B flavor tagging dilution factor [21]; 6) the cosine of the angle between the B flight direction and the beam axis in the CM frame; 7) the cosine of the angle between the D and $\Upsilon(4S)$ directions in the rest frame of the B; 8) the product of the charge of the B candidate and the sum of the charges of all kaons not used for the reconstruction of the B candidate; and 9) the difference between the sum of the charges of particles in the D hemisphere and the sum of charges in the opposite hemisphere, excluding the particles used in the B meson reconstruction.

The neural network output $\mathcal{C}_{\mathrm{NB}}$ is in the range -1 to 1, where events at $\mathcal{C}_{\mathrm{NB}} = 1$ (-1) are signal- (continuum-) like. The training and optimization of the neural network are carried out with signal and $q\bar{q}$ Monte Carlo (MC) samples after event-selection requirements are imposed. We require $\mathcal{C}_{\mathrm{NB}} > -0.6$, which rejects 70% of the $q\bar{q}$ continuum background and only 3% of the signal. The selection efficiency after all criteria have been applied is 10.9% (11.2%) for $B \to DK$ $(B \to D\pi)$ decays.

The $\mathcal{C}_{\mathrm{NB}}$ distribution peaks strongly at $|\mathcal{C}_{\mathrm{NB}}| \sim 1$ and is therefore difficult to model with a simple analytic function. Therefore, to improve this modeling, we transform $\mathcal{C}_{\mathrm{NB}}$ to a new variable $\mathcal{C}'_{\mathrm{NB}}$:

$$\mathcal{C}'_{\mathrm{NB}} = \log\left(\frac{\mathcal{C}_{\mathrm{NB}} - \mathcal{C}_{\mathrm{NB,min}}}{\mathcal{C}_{\mathrm{NB,max}} - \mathcal{C}_{\mathrm{NB}}}\right). \tag{5}$$

Here, $\mathcal{C}_{\mathrm{NB,min}} = -0.6$ and $\mathcal{C}_{\mathrm{NB,max}} = 1$ are the minimum and maximum values of $\mathcal{C}_{\mathrm{NB}}$ for the events used for the signal extraction. The distribution of $\mathcal{C}'_{\mathrm{NB}}$ can be modeled by Gaussian or asymmetric Gaussian functions.

We extract the signal yield using an unbinned extended maximum likelihood fit to ΔE and $\mathcal{C}'_{\mathrm{NB}}$ distributions. We perform separate fits to the suppressed and favored $B \to DK$ $(B \to D\pi)$ modes. The total probability density function (PDF) for each component is formed by multiplying the individual PDFs for ΔE and $\mathcal{C}'_{\mathrm{NB}}$, as they have negligible correlation. The ΔE and $\mathcal{C}'_{\mathrm{NB}}$ PDF for each fit component are described as follows.

For signal, the ΔE distribution is parametrized by a sum of two Gaussian functions of common mean. The $\mathcal{C}'_{\mathrm{NB}}$ distribution is parametrized by the sum of a symmetric Gaussian and an asymmetric Gaussian having different means. The PDF shape parameters used in the fit to the suppressed mode are fixed to the values obtained from the fit to the favored mode.

For $B \to DK$ decays, there is a background from $B \to D\pi$ decays where the π daughter of the B is misidentified as a K. This background peaks in ΔE at around 45 MeV and is modeled by the sum of a symmetric Gaussian and an asymmetric Gaussian. The distribution of $\mathcal{C}'_{\mathrm{NB}}$ is the same as for the signal, so the same PDF is used. For the fit to the suppressed DK data, the $D\pi$ background yield is fixed to that measured in the suppressed $D\pi$ signal fit multiplied by the misidentification rate; this procedure reduces the

statistical uncertainty on the signal yield. For this background component, all other PDF shape parameters in the suppressed mode are fixed to those measured in the fit to the favored mode.

The $B\bar{B}$ background in the favored Dh modes has two components. The first is from $B^- \to D^* h^-$ and $B^- \to D\rho^-$ events and peaks at $\Delta E < -0.1$ GeV, so an upper tail is observed within the fit range. The second component is combinatorial. The peaking and combinatorial components are modeled by an exponential and first-order polynomial, respectively. The suppressed Dh has a much smaller peaking $B\bar{B}$ background contribution than the favored mode, so an exponential function is used to model the whole peaking and combinatorial background. The $\mathcal{C}'_{\rm NB}$ distribution for the $B\bar{B}$ background is parametrized by a Gaussian function, which is determined separately for suppressed and favored modes from the $B\bar{B}$ MC sample.

The ΔE and $\mathcal{C}'_{\rm NB}$ distributions for the $q\bar{q}$ continuum background are parametrized by a first-order polynomial and a sum of two Gaussian functions of common mean, respectively. The parameters for $\mathcal{C}'_{\rm NB}$ are determined using the $M_{\rm bc}$ sideband, given by 5.20 GeV/$c^2 < M_{\rm bc} <$ 5.24 GeV/c^2, for all modes. For the suppressed mode, the mean of one of the Gaussians is left free in the fit to the data; this minimizes the cross-feed between the $q\bar{q}$ and combinatorial $B\bar{B}$ backgrounds.
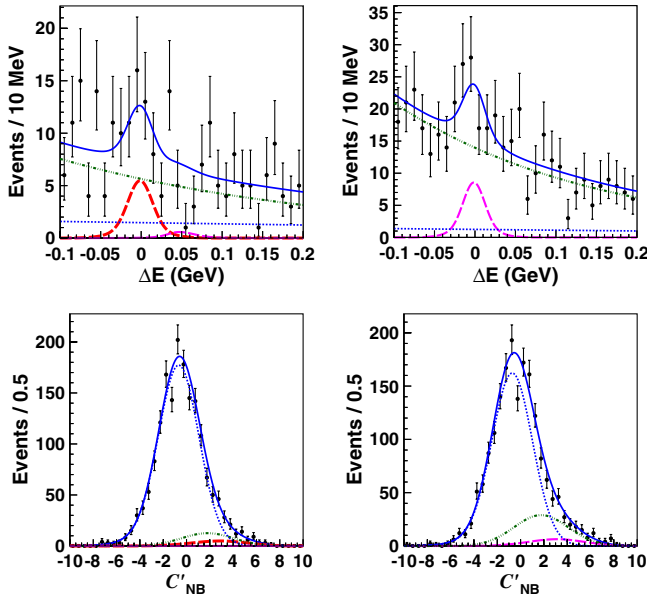

FIG. 1 (color online). ΔE ($\mathcal{C}'_{\rm NB} > 4$) and $\mathcal{C}'_{\rm NB}$ ($|\Delta E| <$ 0.02 GeV) distributions for $[K^+ \pi^- \pi^0]_D K^-$ (left) and $[K^+ \pi^- \pi^0]_D \pi^-$ (right). In these plots, points with error bars represent data while the total best-fit projection is shown with the solid blue curve, for which the components are shown with thicker dashed red (DK signal), thinner dashed magenta ($D\pi$), dashed dot green ($B\bar{B}$ background), and dotted blue ($q\bar{q}$ background) lines. To enhance the signal and suppress the dominant continuum background in the ΔE projection, a strict criterion on $\mathcal{C}'_{\rm NB}$ is applied.

The projections of the fits for the suppressed and favored Dh modes are shown in Figs. 1 and 2, respectively. Suppressed DK and $D\pi$ signal peaks are visible. The values of R_{Dh} are determined using the signal yields and efficiencies given in Table I:

$$R_{DK} = [1.98 \pm 0.62({\rm stat}) \pm 0.24({\rm syst})] \times 10^{-2}, \quad (6)$$

$$R_{D\pi} = [1.89 \pm 0.54({\rm stat})^{+0.22}_{-0.25}({\rm syst})] \times 10^{-3}. \quad (7)$$

The systematic uncertainties associated with R_{DK} and $R_{D\pi}$ are listed in Table II and estimated as follows. The uncertainties due to fixed PDF shape parameters that are obtained from data are estimated by varying each fixed parameter by $\pm 1\sigma$. The uncertainty due to the $B\bar{B}$ $\mathcal{C}'_{\rm NB}$ PDF is estimated by varying the mean and width of the Gaussian by the maximum differences observed between data and MC for the $\mathcal{C}'_{\rm NB}$ PDF from the favored signal. A possible bias related to the fit is checked with 10,000 simulated experiments. No bias is observed, and the systematic uncertainty due to the possible bias is taken to be the error on the mean residual. A small bias is observed in the yields of $B\bar{B}$ and $q\bar{q}$ backgrounds in the suppressed $B \to DK$ mode simulations. This is due to an imperfect modeling of the continuum $\mathcal{C}'_{\rm NB}$ distribution in the signal region by the fits to the $M_{\rm bc}$ sideband. The impact of this bias on the signal yield is estimated using simulated experiments to be at most 3%.

Charmless $B^- \to K^- K^+ \pi^- \pi^0$ decay could result in an irreducible peaking background to the signal. The size of this background is bounded by fits to the sidebands of the reconstructed D mass: 1.45 GeV/$c^2 < M_D < 1.80$ GeV/c^2
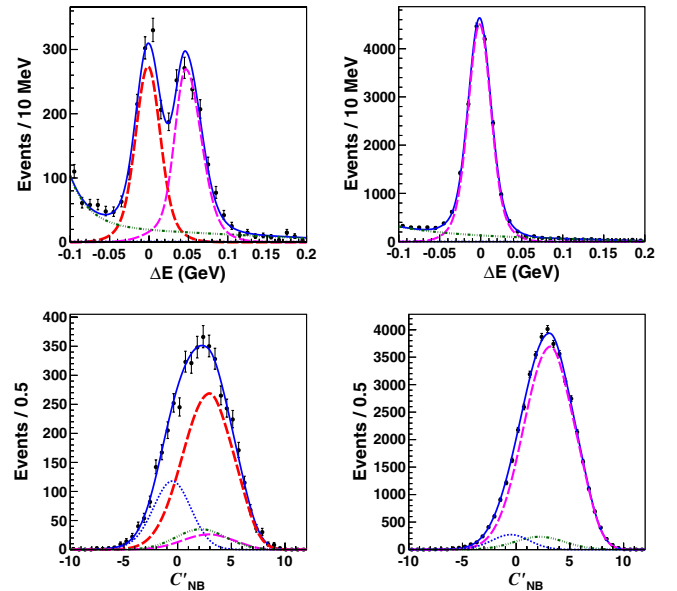

FIG. 2 (color online). ΔE ($\mathcal{C}'_{\rm NB} > 4$) and $\mathcal{C}'_{\rm NB}$ ($|\Delta E| <$ 0.02 GeV) distributions for $[K^- \pi^+ \pi^0]_D K^-$ (left) and $[K^- \pi^+ \pi^0]_D \pi^-$ (right). The color legend and fit components are the same as those in Fig. 1.

TABLE I. Signal yields, reconstruction efficiencies for signals after PID calibration for any data-MC discrepancy, and significances ($\mathcal{S}$) including systematic uncertainties. The uncertainties listed for the signal yield are statistical only, and those on efficiency are from MC statistics and the PID correction.

Mode	Yield	Efficiency (%)	$\mathcal{S}$
$B^- \to [K^+\pi^-\pi^0]_D K^-$	77 ± 24	10.9 ± 0.1	3.2σ
$B^- \to [K^-\pi^+\pi^0]_D K^-$	3844 ± 125	10.8 ± 0.1	
$B^- \to [K^+\pi^-\pi^0]_D \pi^-$	94 ± 27	11.2 ± 0.1	3.3σ
$B^- \to [K^-\pi^+\pi^0]_D \pi^-$	49668 ± 338	11.2 ± 0.1	

and $1.90~\text{GeV}/c^2 < M_D < 2.25~\text{GeV}/c^2$. We apply the same fitting method used in the signal extraction to the sideband sample to obtain an expected yield of -9 ± 7 and -11 ± 8 events for suppressed DK and $D\pi$, respectively. Since the yields are consistent with zero, we include the uncertainty on the obtained yield as a systematic uncertainty. This is the dominant source of systematic uncertainty on the measurement of R_{DK}.

There are also uncertainties on the efficiency coming from the limited statistics of the MC sample and the calibration of the PID efficiency for potential data-MC differences. The uncertainty due to fixing the $B \to D\pi$ yield in the fit to the suppressed $B \to DK$ sample is found to be negligible.

The signal significance is calculated as $\mathcal{S} = \sqrt{-2\ln(L_0/L_{\max})}$, where $L_{\max}$ is the maximum likelihood and L_0 is the likelihood when the signal yield is constrained to be zero. In order to include systematic uncertainty in the significance, we convolve the fit likelihood with a Gaussian for which the width is equal to the systematic uncertainty for R_{DK} and with an asymmetric Gaussian for which the widths are the negative and positive systematic uncertainties for $R_{D\pi}$. The significance of R_{DK} ($R_{D\pi}$) is 3.2σ (3.3σ).

We measure A_{Dh} in a separate fit to the suppressed candidates, including the charge of the kaon or pion from the B decay as an additional observable and A_{Dh} as a new free parameter. Since asymmetries associated with $B\bar{B}$ and $q\bar{q}$ parameters are expected to be negligible, they are fixed to zero in the A_{Dh} fit. The measured values are

TABLE II. Summary of the systematic uncertainties for R_{Dh} and A_{Dh}. Negligible contributions are denoted by "$\cdots$."

Source	R_{DK} (%)	$R_{D\pi}$ (%)	A_{DK}	$A_{D\pi}$
ΔE and $\mathcal{C}'_{\text{NB}}$ PDFs	$^{+6.5}_{-7.1}$	$^{+8.3}_{-10.3}$	$^{+0.03}_{-0.02}$	$^{+0.02}_{-0.03}$
Fit bias	$+0.1$	$+0.4$	$\cdots$	$\cdots$
Due to $B\bar{B}$ and $q\bar{q}$ bias	± 3.0	$\cdots$	$\cdots$	$\cdots$
Peaking background	± 9.5	± 8.2	± 0.04	± 0.01
Efficiency	± 0.1	± 0.1	$\cdots$	$\cdots$
Detector asymmetry	$\cdots$	$\cdots$	± 0.02	± 0.02
Total	$^{+11.9}_{-12.2}$	$^{+11.7}_{-13.2}$	± 0.05	$^{+0.03}_{-0.04}$

$$A_{DK} = 0.41 \pm 0.30(\text{stat}) \pm 0.05(\text{syst}), \tag{8}$$

$$A_{D\pi} = 0.16 \pm 0.27(\text{stat})^{+0.03}_{-0.04}(\text{syst}). \tag{9}$$

The ΔE projections for signal Dh^- and Dh^+ are shown in Fig. 3. The systematic uncertainties (see Table II) arise from the following sources. Uncertainties related to the fit parameters are obtained in the same way as those estimated for R_{Dh}. The uncertainty due to the yield of the peaking background is ± 0.04 (± 0.01) for A_{DK} ($A_{D\pi}$), which is estimated under the assumption of zero asymmetry in the peaking background. A possible bias in A_{Dh} due to any detLiector asymmetry is estimated by determining the asymmetry between B^+ and B^- in the favored mode, which is expected to be close to zero. No detector asymmetry is observed in the favored DK mode, so the uncertainty on the measurement is taken as a systematic uncertainty for the suppressed DK mode. An asymmetry is seen in the favored $D\pi$ mode, which is taken as a systematic uncertainty for the suppressed $D\pi$ mode. The remaining sources are found to give negligible contributions.

In summary, for the mode $B^- \to Dh^-$, $D \to K^+\pi^-\pi^0$ ($h = K, \pi$), we report the measurements R_{Dh} and A_{Dh}, using $772 \times 10^6~B\bar{B}$ pairs collected by the Belle detector. We obtain the first evidence for the suppressed $B \to DK$ signal with a significance of 3.2σ. In addition, we report the first measurements of A_{DK}, $R_{D\pi}$, and $A_{D\pi}$. The R_{DK} and A_{DK} results obtained can be used to constrain the UT angle ϕ_3 using the ADS method [4].
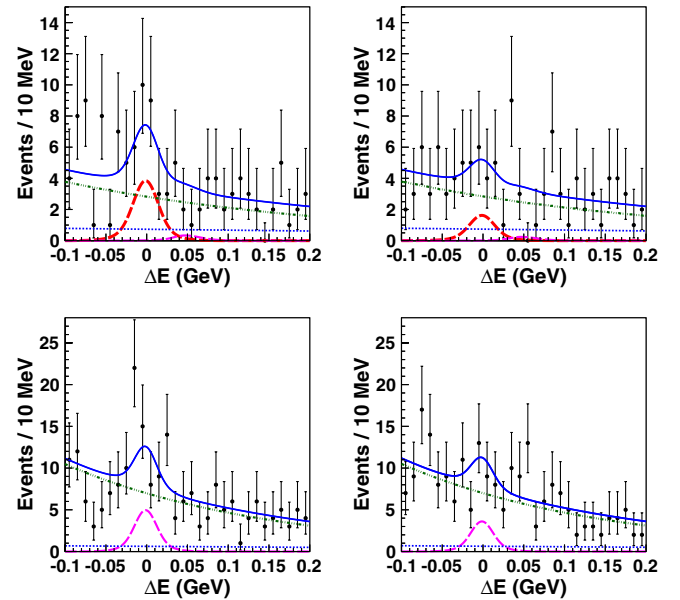
FIG. 3 (color online). ΔE distributions ($\mathcal{C}'_{\text{NB}} > 4$) for $[K^+\pi^-\pi^0]_D K^-$ (left upper), $[K^-\pi^+\pi^0]_D K^+$ (right upper), $[K^+\pi^-\pi^0]_D \pi^-$ (left lower), and $[K^-\pi^+\pi^0]_D \pi^+$ (right lower). The color legend and fit components are the same as those in Fig. 1.

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

[2] Thoughout this paper, the addition of the charge conjugate decay mode is implicit unless stated otherwise.

[3] M. Gronau and D. London, Phys. Lett. B **253**, 483 (1991); M. Gronau and D. Wyler, Phys. Lett. B **265**, 172 (1991).

[4] D. Atwood, I. Dunietz, and A. Soni, Phys. Rev. D **63**, 036005 (2001).

[5] A. Giri, Yu. Grossman, A. Soffer, and J. Zupan, Phys. Rev. D **68**, 054018 (2003).

[6] P. del Amo Sanchez *et al.* (*BABAR* Collaboration), Phys. Rev. D **82**, 072006 (2010).

[7] Y. Horii *et al.* (Belle Collaboration), Phys. Rev. Lett. **106**, 231803 (2011).

[8] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **84**, 091504(R) (2011).

[9] R. Aaij *et al.* (LHCb Collaboration), Phys. Lett. B **712**, 203 (2012).

[10] K. Negishi *et al.* (Belle Collaboration), Phys. Rev. D **86**, 011101(R) (2012).

[11] J. Beringer *et al.* (Particle Data Group), Phys. Rev. D **86**, 010001 (2012).

[12] D. Atwood and A. Soni, Phys. Rev. D **68**, 033003 (2003).

[13] N. Lowrey *et al.* (CLEO Collaboration), Phys. Rev. D **80**, 031105(R) (2009).

[14] J. P. Lees *et al.* (*BABAR* Collaboration), Phys. Rev. D **84**, 012002 (2011).

[15] M. Rama, arXiv:1307.4384.

[16] 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).

[17] 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**, 03A001 (2013) and following articles up to **2013**, 03A011 (2013).

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

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

[20] The Fox–Wolfram moments were introduced in 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).

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