Study of the Suppressed Decays $B^- \rightarrow [K^+\pi^-]_D K^-$ and $B^- \rightarrow [K^+\pi^-]_D \pi^-$

M. Saigo,1 K. Abe,7 K. Abe,40 H. Aihara,42 M. Akatsu,20 Y. Asano,16 V. Aulchenko,1 T. Aushev,11 S. Bahinipati,4 A. M. Bakich,37 Y. Ban,31 S. Banerjee,38 I. Bedny,1 U. Bitenc,12 I. Bizjak,12 S. Blyth,24 A. Bondar,1 A. Bozek,25 M. Bračko,7,18,12 J. Brodzicka,25 T. E. Browder,6 Y. Chao,24 A. Chen,22 K.-F. Chen,24 W. T. Chen,22 B. G. Cheon,3 R. Chistov,11 S.-K. Choi,5 Y. Choi,36 Y. K. Choi,36 A. Chuvikov,32 S. Cole,37 J. Dalseno,19 M. Danilov,11 M. Dash,47 L. Y. Dong,9 A. Drutskoy,4 S. Eidelman,1 V. Eiges,11 Y. Enari,20 S. Fratina,12 N. Gabyshev,1 A. Garmash,32 T. Gershon,7 G. Gokhroo,38 B. Golob,17,12 J. Habu,7 K. Hayasaka,20 H. Hayashii,21 M. Hazumi,7 T. Higuchi,7 L. Hinz,16 T. Hokuue,20 Y. Hoshi,40 S. Hou,22 W.-S. Hou,24 Y. B. Hsiung,24 T. Iijima,20 A. Imoto,21 K. Inami,20 A. Ishikawa,7 R. Itoh,7 M. Iwasaki,42 Y. Iwasaki,7 J. H. Kang,48 J. S. Kang,14 P. Kapusta,25 N. Katayama,7 H. Kawai,2 T. Kawasaki,27 N. Kent,6 H. R. Khan,43 H. Kichimi,7 H. J. Kim,15 S. K. Kim,35 S. M. Kim,36 K. Kinoshita,4 P. Koppenburg,7 S. Korpar,18,12 P. Križan,17,12 P. Krokovny,1 R. Kulasi,4 C. C. Kuo,22 A. Kuzmin,1 Y.-J. Kwon,48 G. Leder,10 S. E. Lee,35 T. Lesiak,25 J. Li,34 S.-W. Lin,24 D. Liventsev,11 J. MacNaughton,10 G. Majumder,38 F. Mandl,10 D. Marlow,32 T. Matsumoto,44 A. Matyja,25 Y. Mikami,41 W. Mitaloff,10 H. Miyake,29 R. Mizuk,11 D. Mohapatra,47 T. Mori,43 Y. Nagasaka,8 E. Nakano,28 M. Nakao,7 Z. Nektaniec,25 S. Nishida,7 O. Nitoh,45 S. Ogawa,39 T. Ohshima,20 T. Okabe,20 S. Okuno,13 S. L. Olsen,6 W. Ostrowicz,25 H. Ozaki,7 P. Pakhlov,11 H. Palka,25 H. Park,15 K. S. Park,36 N. Parslow,37 L. S. Peak,35 R. Pestotnik,12 L. E. Piilonen,47 A. Poluektov,1 F. J. Ronga,7 H. Sagawa,7 Y. Sakai,7 N. Sato,20 T. Schietinger,16 O. Schneider,16 P. Schönherr,16 J. Schimani,24 A. J. Schwartz,4 S. Semenov,11 R. Seuster,6 M. E. Sevior,19 H. Shibuya,39 J. B. Singh,30 A. Somov,4 N. Soni,30 R. Stamen,7 S. Stanić,46 M. Starić,12 T. Suniyoshi,44 S. Suzuki,43 Y. Suzuki,7 O. Tajima,7 F. Takasaka,7 K. Tanai,7 N. Tamura,27 M. Tanaka,7 Y. Teramoto,28 X. C. Tian,31 K. Trabelsi,6 T. Tsukamoto,7 S. Uehara,7 T. Uglow,11 S. Uno,7 Y. Ushiroda,7 G. Varner,6 K. E. Varvell,37 S. villa,16 C. C. Wang,24 C. H. Wang,23 M.-Z. Wang,24 W. Watanabe,27 Y. Watanabe,43 A. Yamaguchi,41 H. Yamamoto,41 Y. Yamashita,26 M. Yamauchi,7 Y. Yusa,41 L. M. Zhang,34 Z. P. Zhang,34 V. Zhilich,1 D. Žontar,17,12 and D. Zürcher,16

(The Belle Collaboration)

1Budker Institute of Nuclear Physics, Novosibirsk 2Chiba University, Chiba 3Chonnam National University, Kwangju 4University of Cincinnati, Cincinnati, Ohio 45221 5Gyeongsang National University, Chinju 6University of Hawaii, Honolulu, Hawaii 96822 7High Energy Accelerator Research Organization (KEK), Tsukuba 8Hiroshima Institute of Technology, Hiroshima 9Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 10Institute of High Energy Physics, Vienna 11Institute for Theoretical and Experimental Physics, Moscow 12J. Stefan Institute, Ljubljana 13Kanagawa University, Yokohama 14Korea University, Seoul 15Kyungpook National University, Taegu 16Swiss Federal Institute of Technology of Lausanne, EPFL, Lausanne 17University of Ljubljana, Ljubljana 18University of Maribor, Maribor 19University of Melbourne, Victoria 20Nagoya University, Nagoya 21Nara Women’s University, Nara 22National Central University, Chung-li 23National United University, Miaoli 24Department of Physics, National Taiwan University, Taipei 25H. Niewodniczanski Institute of Nuclear Physics, Krakow 26Nihon Dental College, Niigata 27Niigata University, Niigata 28Osaka City University, Osaka 29Osaka University, Osaka
Precise measurements of the elements of the Cabibbo-Kobayashi-Maskawa matrix $\theta_C$ constrain the standard model and may reveal new physics. However, the extraction of the Unitarity Triangle angle $\phi_3$ is a challenging measurement even with modern high luminosity $B$ factories. Several methods for measuring $\phi_3$ use the interference between $B^- \to D^\ast K^-$ and $B^- \to D^0 K^-$, which occurs when $D^0$ and $D^\ast$ decay to common final states.\(^2\)\(^3\). As noted by Atwood, Dunietz and Soni (ADS)\(^4\), CP violation effects are enhanced if the final state is chosen so that the interfering amplitudes have comparable magnitudes; the archetype uses $B^- \to [K^+ \pi^-]_D K^-$, where $[K^+ \pi^-]_D$ indicates that the $K^+ \pi^-$ pair originates from a neutral $D$ meson. In this case, the color-allowed $B$ decay followed by the doubly Cabibbo-suppressed $D$ decay interferes with the color-suppressed $B$ decay followed by the Cabibbo-allowed $D$ decay (Fig.\(^1\)). Previous studies of this decay mode have not found any signals\(^5\). For the suppressed decay $B^- \to [K^+ \pi^-]_D \pi^-$, both topology and phenomenology are similar to $B^- \to [K^+ \pi^-]_D K^-$. In this analysis, the favored decays $B^- \to [K^- \pi^+]_D h^-$, where $h = \pi$ or $K$, are used as control samples to reduce systematic uncertainties. The same selection criteria for the suppressed decay modes are applied to the control samples whenever possible. Throughout this report, charge conjugate states are implied except where explicitly mentioned and we denote the analyzed decay modes as follows:

- Suppressed decay $B^- \to [K^+ \pi^-]_D h^- \to D_{\text{sup}} h^-$
- Favored decay $B^- \to [K^- \pi^+]_D h^- \to D_{\text{fav}} h^-$

The results are based on a data sample containing 275 million $B\bar{B}$ pairs, collected with the Belle detector at the KEKB asymmetric energy $e^+ e^-$ collider operating at the $Y(4S)$ resonance. The Belle detector is a large-solid-angle magnetic spectrometer that consists 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) comprised of CsI(Tl) crystals located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux-return located outside of the
respectively. We optimize the contributions of signal and continuum background events.

Neutral $D$ mesons are reconstructed by combining two oppositely charged tracks. For each track, information from ACC, TOF and specific ionization measurements from the CDC are used to determine a $K/\pi$ likelihood ratio $P(K/\pi) = \mathcal{L}_K / (\mathcal{L}_K + \mathcal{L}_\pi)$, where $\mathcal{L}_K$ and $\mathcal{L}_\pi$ are kaon and pion likelihoods. We use the particle identification requirements $P(K/\pi) > 0.8$ for kaons and $P(K/\pi) < 0.2$ for pions. These requirements select kaons (pions) with momentum dependent efficiencies of $80 - 95\%$ ($90 - 95\%$) and pion (kaon) misidentification probabilities of $5 - 20\%$ ($15 - 20\%$). $D$ candidates are required to have an invariant mass within $\pm 2.5\sigma$ of the nominal $D$ mass: $1.850 \text{ GeV}/c^2 < M(K\pi) < 1.879 \text{ GeV}/c^2$. To improve the momentum determination, tracks from the $D$ candidate are refitted according to the nominal $D$ mass hypothesis and the reconstructed vertex position (a mass-and-vertex-constrained fit). $B$ mesons are reconstructed by combining $D$ candidates with primary charged hadron candidates. The signal is identified by two kinematic variables, the energy difference $\Delta E = E_D + E_h - E_{\text{beam}}$ and the beam-energy-constrained mass $M_{bc} = \sqrt{E_{\text{beam}}^2 - (\vec{p}_D + \vec{p}_h)^2}$, where $E_D$ is the energy of the $D$ candidate, $E_h$ is the energy of the $h^-$, and $E_{\text{beam}}$ is the beam energy, all evaluated in the center-of-mass (cm) frame; $\vec{p}_D$ and $\vec{p}_h$ are the momenta of the $D$ and $h^-$ in the cm frame. Event candidates are accepted if they have $5.2 \text{ GeV}/c^2 < M_{bc} < 5.3 \text{ GeV}/c^2$ and $|\Delta E| < 0.2 \text{ GeV}$. If there is more than one candidate in an event, we select the best candidate on the basis of a $\chi^2$ determined from the difference between the measured and nominal values of $M(K\pi)$ and $M_{bc}$.

To suppress the large background from the two-jet-like $e^+e^- \rightarrow q\bar{q}$ ($q = u, d, s, c$) continuum processes, variables that characterize the event topology are used. We construct a Fisher discriminant $\mathcal{F}$ of Fox-Wolfram moments called the Super-Fox-Wolfram (SFW) $\mathcal{F}$, where the Fisher coefficients are optimized by maximizing the separation between $B\bar{B}$ events and continuum events. Furthermore, $\cos \theta_B$, the cosine of the angle in the cm system between the $B$ flight direction with respect to the beam axis is also used to distinguish $B\bar{B}$ events from continuum events. These two independent variables, SFW and $\cos \theta_B$, are combined to form a likelihood ratio $R = \mathcal{L}_{\text{sig}} / (\mathcal{L}_{\text{sig}} + \mathcal{L}_{\text{cont}})$, where $\mathcal{L}_{\text{sig}}$ and $\mathcal{L}_{\text{cont}}$ are likelihoods calculated from the SFW and $\cos \theta_B$ distributions of signal and continuum background events, respectively. We optimize the $R$ requirement by maximizing $S/\sqrt{S+N}$, where $S$ and $N$ denote the expected numbers of signal and background events in the signal region.

For $B^- \rightarrow D_{\text{sup}}K^-$, there can be a contribution from $B^- \rightarrow D^0\pi^-$, $D^0 \rightarrow K^+K^-$, which has the same final state and can peak under the signal. In order to reject these events, we veto events that satisfy $1.843 \text{ GeV}/c^2 < M(K\bar{K}) < 1.894 \text{ GeV}/c^2$. The favored decay $B^- \rightarrow D_{\text{fav}}h^-$ can also cause a peaking background for the suppressed decay modes if both the kaon and pion from the $D_{\text{fav}}$ decay are misidentified. Therefore, we veto events for which the invariant mass of the $K\pi$ pair is inside the $D$ mass window when the mass assignments are exchanged. After applying this veto, the residual background from $K\pi$ misidentification is found to be negligible. The three-body charmless decay $B^- \rightarrow K^+K^-\pi^-$ can peak inside the signal regions of $\Delta E$ and $M_{bc}$ for $B^- \rightarrow D_{\text{sup}}K^-$. This background is estimated from the $\Delta E$ distribution of events in a $D$ mass sideband, defined as $1.637 \text{ GeV}/c^2 < M(K\pi) < 1.836 \text{ GeV}/c^2$ and $1.893 \text{ GeV}/c^2 < M(K\pi) < 2.093 \text{ GeV}/c^2$. The estimated peaking background inside the $\Delta E$ signal region is $1.7 \pm 0.9$ events, which we subtract from the observed $B^- \rightarrow D_{\text{sup}}K^-$ yield. As a check, we estimate the expected background level from the measured $B^- \rightarrow K^+K^-\pi^-$ branching fraction $\mathcal{B}$. Using this result, the expected background in our signal region is $2.1 \pm 0.6$ events, assuming that the $B^- \rightarrow K^+K^-\pi^-$ yield is uniformly distributed in phase space. For $B^- \rightarrow D_{\text{sup}}\pi^-$, the peaking background estimated from the $D$ mass sideband is consistent with zero.

The signal yields are extracted using binned maximum likelihood fits to the $\Delta E$ distributions of events in the $M_{bc}$ signal region, $(5.27 \text{ GeV}/c^2 < M_{bc} < 5.29 \text{ GeV}/c^2)$. Backgrounds from decays such as $B^- \rightarrow D\rho^-$ and $B^- \rightarrow D\pi^-$ are distributed in the negative $\Delta E$ region and make a small contribution to the signal region. The shape of this $B\bar{B}$ background is modeled with a smoothed histogram obtained from generic Monte Carlo (MC) samples. The continuum background populates the entire $\Delta E$ region. The shape of the continuum background is modeled with a linear function. The slope is determined from the $\Delta E$ distribution of the $M_{bc}$ sideband $(5.20 \text{ GeV}/c^2 < M_{bc} < 5.26 \text{ GeV}/c^2)$. The $\Delta E$ fitting function is the sum of two Gaussians for the signal, the linear function for the continuum, and the smoothed histogram for the $B\bar{B}$ background distribution.

In the fit to the $\Delta E$ distribution of $B^- \rightarrow D_{\text{sup}}\pi^-$, the free parameters are the position, width and area of the signal peak, and the normalizations of continuum and $B\bar{B}$ backgrounds. For the signal, the relative widths and areas of the two Gaussians are fixed from the signal MC. For the $B^- \rightarrow D_{\text{sup}}K^-$ fit, the position and width of the signal peak are fixed from the $B^- \rightarrow D_{\text{sup}}\pi^-$ fit results.
The fit results are shown in Fig. 2. The numbers of continuum and feed-across from different widths since the shift caused by wrong mass assignment makes the shape asymmetric. The shape parameters of this function are fixed to values determined by the fit to the $B^− \to D_{s\text{fav}} \pi^−$ distribution using a kaon mass hypothesis for the prompt pion. The areas of signal and feed-across from $B^− \to D\pi^−$, and the normalizations of continuum and $B\bar{B}$ backgrounds are floated in the fit. For $B^− \to D_{s\text{sup}}K^−$ and $B^− \to D_{s\text{sup}}\pi^−$, the signal and $B\bar{B}$ background shapes are modeled using the fit results for the $B^− \to D_{\text{fav}}K^−$ and $B^− \to D_{\text{fav}}\pi^−$ modes, respectively. The area of the feed-across from $D_{s\text{sup}}\pi^−$ is estimated as the measured yield of $B^− \to D_{s\text{sup}}\pi^−$ multiplied by the $\pi$ to $K$ misidentification probability. However, the areas of the signal and the normalizations of continuum and $B\bar{B}$ backgrounds are floated. The fit results are shown in Fig. 2. The numbers of events for $B^− \to D_{s\text{sup}}h^−$ and $D_{\text{fav}}h^−$, and the statistical significances of the $B^− \to D_{s\text{sup}}h^−$ signals are given in Table I. The statistical significance is defined as $\sqrt{-2 \ln(L_0/L_{\text{max}})}$, where $L_{\text{max}}$ is the maximum likelihood in the $\Delta E$ fit and $L_0$ is the likelihood when the signal yield is constrained to be zero. The uncertainty in the peaking background contribution is taken into account in the statistical significance calculation.

The ratio of branching fractions is defined as

$$R_{Dh} \equiv \frac{B(B^− \to D_{s\text{sup}}h^-)}{B(B^− \to D_{\text{fav}}h^-)} = \frac{N_{D_{s\text{sup}}h^-}/\epsilon_{D_{s\text{sup}}h^-}}{N_{D_{\text{fav}}h^-}/\epsilon_{D_{\text{fav}}h^-}},$$

where $N_{D_{s\text{sup}}h^-}$ ($N_{D_{\text{fav}}h^-}$) and $\epsilon_{D_{s\text{sup}}h^-}$ ($\epsilon_{D_{\text{fav}}h^-}$) are the number of signal events and the reconstruction efficiency for the decay $B^− \to D_{s\text{sup}}h^−$ ($B^− \to D_{\text{fav}}h^−$), and are given in Table I. The ratios $R_{Dh}$ are calculated to be

$$R_{DK} = (2.3^{+1.6}_{-1}.6)\text{(stat)} \pm 0.1\text{(syst)} \times 10^{-2},$$
$$R_{D\pi} = (3.5^{+1.0}_{-0.9}\text{(stat)} \pm 0.2\text{(syst)}) \times 10^{-3}.$$  

Since the signal for $B^− \to D_{s\text{sup}}K^−$ is not significant, we set an upper limit at the 90% confidence level (C.L.) of $R_{DK} < 4.4 \times 10^{-2}$, where we take the likelihood function as a single Gaussian with width given by the quadratic sum of statistical and systematic errors, and the area is normalized in the physical region of positive branching fraction.

Most of the systematic uncertainties from the detection efficiencies and the particle identification cancel when taking the ratios, since the kinematics of the $B^− \to D_{s\text{sup}}h^−$ and $B^− \to D_{\text{fav}}h^−$ processes are similar. The systematic errors are due to uncertainties in the yield extraction (4.7%−5.4%) and the efficiency difference between $B^− \to D_{s\text{sup}}h^−$ and $B^− \to D_{\text{fav}}h^−$ (1.3%−1.7%). The uncertainties in the signal shapes and the $q\bar{q}$ background shapes are determined by varying the shape of the fitting function by $\pm 1\sigma$. The uncertainties in the $B\bar{B}$ background shapes are determined by fitting the $\Delta E$ distribution in the region $-0.07$ GeV < $\Delta E$ < 0.20 GeV ignoring the $B\bar{B}$ background contributions. The uncertainties in the efficiency differences are determined using signal MC. The total systematic error is the sum in quadrature of the above uncertainties.

The product branching fractions for $B^− \to D_{s\text{sup}}h^−$ are determined as

$$B(B^− \to D_{s\text{sup}}h^-) = B(B^− \to D_{\text{fav}}h^-) \times R_{Dh},$$

and are given in Table I. A third uncertainty arises due

| Mode | Product branching fraction from [12] | Efficiency (%) | Signal Yield | Statistical significance | Measured product branching fraction | Upper limit (90% C.L.) |
|------|-----------------------------------|----------------|--------------|------------------------|-----------------------------------|----------------------|
| $B^− \to D_{s\text{sup}}K^−$ | − | $12.9 \pm 0.2$ | $8.5^{+0.1}_{-0.1}$ | $2.3\sigma$ | $(3.2 \pm 0.2 \pm 0.2 \pm 0.5) \times 10^{-1}$ | $6.3 \times 10^{-4}$ |
| $B^− \to D_{s\text{sup}}\pi^−$ | $(6.9 \pm 0.7) \times 10^{-7}$ | $20.1 \pm 0.2$ | $2.85 \pm 7.4 \times 10^{-7}$ | $6.4\sigma$ | $(6.6 \pm 1.9 \pm 0.4 \pm 0.3) \times 10^{-7}$ | − |
| $B^− \to D_{\text{fav}}K^−$ | $(1.4 \pm 0.2) \times 10^{-5}$ | $12.9 \pm 0.2$ | $376.0 \pm 21.8 \times 10^{-2}$ | − | − | − |
| $B^− \to D_{\text{fav}}\pi^−$ | $(1.9 \pm 0.1) \times 10^{-4}$ | $20.3 \pm 0.2$ | $8181.9 \pm 94.0 \times 10^{-2}$ | − | − | − |
to the error in the branching fraction of \( B^- \to D_{\text{sup}} h^- \), which is taken from [12]. The uncertainties are statistics-dominated. For the \( B^- \to D_{\text{sup}} K^- \) branching fraction, we set an upper limit at the 90\% C.L. of \( \mathcal{B}(B^- \to D_{\text{sup}} K^-) < 6.3 \times 10^{-7} \). For \( B^- \to D_{\text{sup}} \pi^- \), our measured branching fraction is consistent with expectation neglecting the contribution from \( B^- \to D^0 \pi^- \).

The ratio \( R_{DK} \) is related to \( \phi_B \) by

\[
R_{DK} = r_B^2 + r_D^2 + 2r_B r_D \cos \phi_B \cos \delta,
\]

where [12]

\[
r_B = \left| \frac{A(B^- \to D^0 K^-)}{A(B^- \to D^0 K^0)} \right|, \quad \delta = \delta_B + \delta_D,
\]

\[
r_D = \left| \frac{A(D^0 \to K^+ \pi^-)}{A(D^0 \to K^- \pi^+)} \right| = 0.060 \pm 0.003,
\]

and \( \delta_B \) and \( \delta_D \) are the strong phase differences between the two \( B \) and \( D \) decay amplitudes, respectively. Using the above result, we obtain a limit on \( r_B \). The least restrictive limit is obtained allowing \( \pm 1\sigma \) variation on \( r_D \) and assuming maximal interference \((\phi_B = 0^\circ, \delta = 180^\circ) \) or \((\phi_B = 180^\circ, \delta = 0^\circ) \) and is found to be \( r_B < 0.27 \).

We search for partial rate asymmetries \( A_{Dh} \) in \( B^\pm \to D_{\text{sup}} h^\mp \) decay, fitting the \( B^- \) and \( B^+ \) yields separately for each mode, where \( A_{Dh} \) is determined as

\[
A_{Dh} = \frac{B(B^- \to D_{\text{sup}} h^-) - B(B^+ \to D_{\text{sup}} h^+)}{B(B^- \to D_{\text{sup}} h^-) + B(B^+ \to D_{\text{sup}} h^+)}.
\]

The peaking background for \( B^\pm \to D_{\text{sup}} K^\mp \) is subtracted assuming no \( CP \) asymmetry. The fit results are shown in Fig. 4 and Table III. We find

\[
A_{DK} = 0.88^{+0.77}_{-0.62}(\text{stat}) \pm 0.06(\text{syst}),
\]

\[
A_{D\pi} = 0.30^{+0.29}_{-0.25}(\text{stat}) \pm 0.06(\text{syst}),
\]

where the systematic uncertainties arise from possible biases in the analysis algorithms (estimated from the \( B^\pm \to D_{\text{sup}} \pi^\mp \) control sample to be 2.5\%); uncertainties in the extraction of the \( B^- \) and \( B^+ \) yields (estimated by varying fitting parameters by \( \pm 1\sigma \) to be 4.9\%); asymmetry in the particle identification efficiency of prompt kaons (estimated in [13] to be 0.6\%). We assume no \( CP \) asymmetry in the peaking background, and do not assign any systematic uncertainty from this source [14].

In summary, using 275 million \( BB \) pairs collected with the Belle detector, we report studies of the suppressed decay \( B^- \to D_{\text{sup}} h^- \) \((h = K, \pi)\). We observe \( B^- \to D_{\text{sup}} \pi^- \) for the first time, with a significance of 6.4\sigma. The size of the signal is consistent with expectation based on measured branching fractions [12]. The significance for \( B^- \to D_{\text{sup}} K^- \) is 2.3\sigma and we set an upper limit on the ratio of \( B \) decay amplitudes \( r_B < 0.27 \) at 90\% confidence level. This result is consistent with previous searches [3], and with the measurement of \( r_B \) in the decay \( B^- \to DK^-, D \to K_S^0 \pi^- \) [13], which provides the most precise current determination of \( \phi_B \).

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 and the NII for valuable computing and SuperSINET network support. We acknowledge support from MEXT and JSPS (Japan); ARC and DEST (Australia); NSFC (contract No. 10175071, China); DST (India); the BK21 program of MOEHRD and the CHEP SRC program of KOSEF (Korea); KBN (contract No. 2P03B 01324, Poland); MIST (Russia); MESS (Slovenia); NSC and MOE (Taiwan); and DOE (USA).

* on leave from Nova Gorica Polytechnic, Nova Gorica

[1] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. 49, 652 (1973).
[2] I.I. Bigi and A.I. Sanda, Phys. Lett. B 211, 213 (1988).
[3] M. Gronau and D. Wyler, Phys. Lett. B 165, 172 (1991); M. Gronau and D. London, Phys. Lett. B 253, 483 (1991).
[4] D. Atwood, I. Dunietz, and A. Soni, Phys. Rev. Lett. 78, 3257 (1997), Phys. Rev. D 63, 036005 (2001).
[5] B. Aubert et al. (BaBar collaboration), Phys. Rev. Lett. 93, 131804 (2004).
[6] S. Kurokawa and E. Kikutani, Nucl. Instr. and Meth. A 499, 1 (2003).
[7] A. Abashian et al. (Belle collaboration), Nucl. Instr. and Meth. A 479, 117 (2002).
[8] Y. Ushiroda (Belle SVD2 Group), Nucl. Instr. and Meth. A 511, 6 (2003).
[9] R.A. Fisher, Ann. Eugenics 7, 179 (1936).
[10] The Fox-Wolfram moments were introduced in G.C. Fox and S. Wolfram, Phys. Rev. Lett. 41, 1581 (1978). The Fisher discriminant used by Belle, based on modified Fox-Wolfram moments (SFW), is described in K. Abe et al. (Belle collaboration), Phys. Rev. Lett. 87, 101801 (2001).
[11] A. Garmash et al. (Belle collaboration), Phys. Rev. D 69, 012001 (2004).
[12] S. Eidelman et al. (Particle Data Group), Phys. Lett. B 592, 1 (2004).
[13] Y. Chao et al. (Belle collaboration), Phys. Rev. Lett. 93, 191802 (2004).
[14] An assumption of 30% $CP$ asymmetry in the peaking background would lead to a shift of 5.5% in $A_{DK}$.
[15] A. Poluektov et al. (Belle collaboration), Phys. Rev. D 70, 072003 (2004).