Measurement of the Decay $B_s^0 \to D_s^- \pi^+$ and Evidence for $B_s^0 \to D_s^- K^\pm$ in $e^+e^-$ Annihilation at $\sqrt{s} \sim 10.87$ GeV

R. Louvot, J. Wicht, O. Schneider, I. Adachi, H. Aihara, K. Arinstein, V. Aulchenko, T. Ashevi, A. Bakich, V. Balagura, A. Bay, V. Bhardwaj, U. Bitenc, A. Bondar, A. Bozek, M. Bračko, T. Browder, A. Chen, B. G. Cheon, R. Chistov, I.-S. Cho, Y. Choi, J. Dalseno, M. Danilov, M. Dash, A. Drutskoy, W. Dungel, S. Eidelman, N. Gabyshev, P. Goldenzweig, B. Golob, H. Ha, J. Haba, K. Hayasaka, H. Hayashii, M. Hazumi, Y. Hoshi, W.-S. Hou, H. J. Hyun, T. Iijima, K. Inami, A. Ishikawa, H. Ishino, R. Itoh, M. Iwasaki, N. J. Joshi, D. H. Kah, J. H. Kang, N. Katayama, H. Kawai, T. Kawasaki, H. Kichimi, S. K. Kim, Y. I. Kim, Y. J. Kim, K. Kinoshita, S. Korpar, P. Križan, P. Krovkov, R. Kumar, A. Kuzmin, Y.-J. Kwon, S.-H. Kyeong, J. S. Lange, J. S. Lee, M. J. Lee, S. E. Lee, T. Lesiak, J. Lii, A. Limosani, S.-W. Lin, D. Liventsev, F. Mandl, A. Matyja, S. McOnie, T. Medvedeva, K. Miyabayashi, H. Miyake, H. Miyata, Y. Miyazaki, R. Mizuk, T. Mori, E. Nakano, M. Nakao, S. Nishida, O. Nitoh, S. Ogawa, T. Oshtima, S. Okuno, H. Ozaki, P. Pakhlova, C. W. Park, H. K. Park, R. Pestotnik, L. E. Piilonen, H. Sahoo, S. Takai, J. Schümann, A. J. Schwartz, A. Sekiya, K. Senyo, M. E. Sevior, M. Shapkin, J.-G. Shiu, J. B. Singh, A. Somov, S. Stanic, M. Staric, K. Sumisawa, T. Sumiyoshi, M. Tanaka, G. N. Taylor, Y. Teramoto, I. Tikhomirov, K. Trabelski, S. Uehara, T. Uglow, Y. Unno, S. Uno, Y. Usov, G. Varner, K. Vervink, C. C. Wang, C. H. Wang, P. Wang, X. L. Wang, Y. Watanabe, R. Wedd, E. Won, B. D. Yabsley, Y. Yamashita, M. Yamashita, Z. P. Zhang, V. Zhilich, Z. Zhulanov, T. Zivko, A. Zupanc, N. Zwahlen, and O. Zyukova

(The Belle Collaboration)
We have studied $B^0 \rightarrow D_s^- \pi^+$ and $B^0 \rightarrow D_s^+ K^\mp$ decays using 23.6 fb$^{-1}$ of data collected at the $\Upsilon(5S)$ resonance with the Belle detector at the KEKB $e^+e^-$ collider. This highly pure $B^0 \rightarrow D_s^- \pi^+$ sample is used to measure the branching fraction, $B(B^0 \rightarrow D_s^- \pi^+) = [3.67^{+0.25}_{-0.33} \text{(stat.)}^{+0.42}_{-0.34} \text{(syst.)}] \pm 0.49(f_s) \times 10^{-3}$ ($f_s = N_{B^0}^{b(*)} / N_{B^0}$) and the fractions of $B^0$ event types at the $\Upsilon(5S)$ energy, in particular $N_{B^0}^{b(*)} / N_{B^0}^{(*)} = (90.3^{+4.8}_{-4.0} \pm 2.0)\%$. We also determine the masses $M(B^0) = (5364.4 \pm 1.3 \pm 0.7)$ MeV/c$^2$ and $M(B_s^\mp) = (5416.4 \pm 0.4 \pm 0.5)$ MeV/c$^2$. In addition, we observe $B^0 \rightarrow D_s^\mp K^\pm$ decays with a significance of 3.5$\sigma$ and measure $B(B^0 \rightarrow D_s^\mp K^\pm) = [2.4^{+1.2}_{-1.0} \text{(stat.)} \pm 0.1 \text{(syst.)}] \pm 0.3(f_s) \times 10^{-4}$.

PACS numbers: 13.25.Hw, 13.25.Gv, 14.40.Gx, 14.40.Nd

The decay $B^0 \rightarrow D_s^- \pi^+$ has a relatively large branching fraction and is a primary normalization mode at hadron colliders, where the absolute production rate of $B^0_s$ mesons is difficult to measure directly. It proceeds dominantly via a Cabibbo-favoured tree process. The decay $B^0 \rightarrow D^- \pi^+$ proceeds through the same tree process but may also have additional contributions from $W$-exchange, so a comparison of the partial width of the two decays can give insight into the poorly known $W$-exchange process. The Cabibbo-suppressed mode $B^0 \rightarrow D_s^\mp K^\pm$ is mediated by $b \rightarrow c$ and $b \rightarrow u$ tree transitions of similar order ($\sim \lambda^3$, in the Wolfenstein parameterization), which raises the possibility of measuring time-dependent CP-violating effects. It has recently become possible to produce $B_s^0$ events from $e^+e^-$ collisions at the $\Upsilon(5S)$ resonance in sufficiently large numbers to achieve interesting and competitive measurements. $\Upsilon(5S)$ events may also be used to determine precisely the $D_s^\mp$ and $B_s^0$ masses, as the mass difference can be compared with that of $B^{0\pm}$ and $B^0$ to test heavy quark symmetry, which predicts equality between them. Properties of the $\Upsilon(5S)$ such as the fraction of events containing a $B_{s0}$ and the relative proportions of $B_s^0 B_s^0$, $B_s^+ B_s^-$, and $B_s^+ B_s^*$ provide additional tests of heavy quark theories.

In this Letter, we report measurements performed with fully reconstructed $B^0_s \rightarrow D_s^- \pi^+$ and $B^0_s \rightarrow D_s^+ K^\mp$ decays in $L_{\text{int}} = (23.6 \pm 0.3)$ fb$^{-1}$ of data collected with the Belle detector at the KEKB asymmetric-energy (3.6 GeV on 8.2 GeV) $e^+e^-$ collider operated at the $\Upsilon(5S)$ resonance. The beam energy in the center-of-mass (CM) frame is measured to be $E_B^\text{CM} = \sqrt{s}/2 = 5433.5 \pm 0.5$ MeV with $\Upsilon(5S) \rightarrow \Upsilon(1S) \pi^+ \pi^-$, $\Upsilon(1S) \rightarrow \mu^+ \mu^-$ decays. The total $b\bar{b}$ cross section at the $\Upsilon(5S)$ energy has been measured to be $\sigma_{bb}^{\Upsilon(5S)} = (0.302 \pm 0.014) \text{ nb}$, which includes $B^0$, $B^+$, and $B^0_s$ events. Three $B^0_s$ production modes are kinematically allowed: $B^0_s B^0_s$, $B^0_s B^\mp$, and $B^* \rightarrow B_s$. The $B^0_s$ decays electromagnetically to $B^0$, emitting a photon with energy $E_\gamma \sim 53$ MeV. The fraction of $b\bar{b}$ events containing a $B_s^{b(*)} B_s^{(*)}$ pair has been measured to be $f_s = N_{B_s^{b(*)} B_s^{(*)}} / N_{b\bar{b}} = (19.5^{+3.0}_{-2.3}) \%$. The number of $B^0_s$ mesons in the sample is thus $N_{B^0_s} = 2 \times L_{\text{int}} \times \sigma_{bb}^{\Upsilon(5S)} \times f_s = (2.78^{+0.45}_{-0.36}) \times 10^6$. The $B^0_s$ production mode ratios are defined as $f_{B^0_s} B_s^\mp = N_{B^0_s} B_s^\mp / N_{b\bar{b}}$, $f_{B^0_s} B_s^{(*)} = N_{B^0_s} B_s^{(*)} / N_{b\bar{b}}$, and $f_{B^0_s} B_s^* = N_{B^0_s} B_s^* / N_{b\bar{b}}$. Belle previously measured $f_{B^0_s} B_s^\mp = (93.3^{+7.9}_{-6.0}) \%$.

The Belle detector is a large-solid-angle magnetic spectrometer that consists of a silicon vertex detector, a central drift chamber (CDC), an array of aerogel threshold Cherenkov counters (ACC), a barrel-like arrangement of acid-drift chambers, and two large CsI(Tl) electromagnetic calorimeters that cover to a pseudorapidity of 1.5. The detector is described in detail elsewhere.

Reconstructed charged tracks are required to have a maximum impact parameter with respect to the nominal interaction point of 0.5 cm in the radial direction and 3 cm in the beam-axis direction. A likelihood ratio $R_{K_{S^0}} = L_{K} / (L_{K^\mp} + L_{K})$ is built using ACC, TOF, and CDC $(dE/dx)$ measurements. A track is identified as a pion if $R_{K_{S^0}} < 0.6$ or as a kaon otherwise. With this selection, the identification efficiency for pions (kaons) is about 91% (85%), while the fake rate is about 9% (14%).
Neutral kaons are reconstructed via the decay \( K_S^0 \to \pi^+\pi^- \) with no identification requirements for the two charged pions. The \( K_S^0 \) candidates are required to have an invariant mass within \( \pm 7.5 \text{ MeV/c}^2 \) (\( \pm 4 \sigma \)) of the nominal \( K_S^0 \) mass (all nominal mass values are taken from Ref. [12]). Requirements on the \( K_S \) vertex displacement from the interaction point and on the difference between vertex and \( K_S \) flight directions are applied. The criteria are described in detail elsewhere [13].

The \( K^0 \) (\( \phi \)) candidates are reconstructed via the decay \( K^{\prime 0} \to K^+\pi^- (\phi \to K^+K^-) \) with an invariant mass within \( \pm 50 \text{ MeV/c}^2 \) (\( \pm 12 \text{ MeV/c}^2 \)) of the nominal mass.

Candidates for \( D_s^0 \) are reconstructed in the three modes \( D_s^+ \to \phi\pi^- \), \( D_s^- \to K^{*0}K^- \), and \( D_s^- \to K_S^0K^- \) and required to have mass within \( \pm 15 \text{ MeV/c}^2 \) (\( \pm 3 \sigma \)) of the nominal \( D_s^- \) mass for \( B_s^0 \to D_s^- \pi^+ \) and within \( \pm 8 \text{ MeV/c}^2 \) for \( B_s^0 \to D_s^+ K^\pm \). Following Ref. [10], the signals for \( B_s^0 \to D_s^- \pi^+ \) and \( B_s^0 \to D_s^+ K^\pm \) are observed using two variables: the beam-constrained mass of the \( B_s^0 \) candidate \( M_{bc} = \sqrt{E_{B_s^0}^2 - p_{B_s^0}^2} \) and the energy difference \( \Delta E = E_{B_s^0} - E_s \), where \((E_{B_s^0}, p_{B_s^0})\) is the four-momentum of the \( B_s^0 \) candidate expressed in the CM frame. We select candidates with \( M_{bc} > 5.3 \text{ GeV/c}^2 \) and \(-0.3 \text{ GeV} < \Delta E < 0.4 \text{ GeV} \). In each event the \( B_s^0 \) candidate with the \( D_s^- \) mass closest to its nominal value is selected for further analysis; only \( \approx 1\% \) of events have more than one candidate.

Further selection criteria are developed using Monte Carlo (MC) samples based on EvtGen [14] and GEANT [15] detector simulation. The most significant source of background is continuum events, \( e^+e^- \to u\bar{u}, d\bar{d}, s\bar{s}, c\bar{c} \).

In addition, for the \( B_s^0 \to D_s^+ K^\pm \) mode there is also a large background from \( B_s^0 \to D_s^- \pi^+ \), where the \( \pi^+ \) is misidentified as a \( K^+ \). The expected continuum background, \( N_{bbk} \), is estimated using MC-generated continuum events representing three times the data. The expected signal, \( N_{sig} \), is obtained assuming \( \mathcal{B}(B_s^0 \to D_s^- \pi^+) = 3.0 \times 10^{-3} \) and \( f_{B_s^0} = 93\% \) for the \( B_s^0 \to D_s^- \pi^+ \) analysis and \( \mathcal{B}(B_s^0 \to D_s^\pm K^\mp) = 3.7 \times 10^{-4} \) for the \( B_s^0 \to D_s^\mp K^\mp \) analysis. For \( B_s^0 \to D_s^\mp K^\mp \), we assume the values of \( \mathcal{B}(B_s^0 \to D_s^- \pi^+) \) and \( f_{B_s^0} \) obtained in the \( B_s^0 \to D_s^- \pi^+ \) analysis.

To improve signal relative to background, criteria are chosen to maximize \( N_{sig}/\sqrt{N_{sig} + N_{bbk}} \), evaluated in the \( B_s^0 \) signal region (Fig. 1). Two topological variables are used. First, we use the ratio of the second and zeroth Fox-Wolfram moments \( R_2 \), which has a broad distribution between zero and one for jet-like continuum events and is concentrated in the range below 0.5 for the more spherical signal events. Candidates for \( B_s^0 \to D_s^- \pi^+ \) (\( B_s^0 \to D_s^+ K^\pm \)) are required to have \( R_2 < 0.5 \) (\( < 0.4 \)). We then use the helicity angle \( \theta_{hel} \) of the \( D_s^- \to \phi\pi^- (D_s^- \to K^{*0}K^-) \) decays, defined as the angle between the momentum of the positive daughter of the \( \phi \) (\( K^{*0} \)) and the momentum of the \( D_s^- \) in the \( (K^{*0}) \) rest frame; for signal decays consisting in a spin–0 particle decaying into a spin–1 particle and a spin–0 particle, the distribution is \( \propto \cos^2 \theta_{hel} \), while for combinatorial background under \( D_s \) signal it is flat. Candidates for \( D_s^- \to \phi\pi^- \) and \( D_s^- \to K^{*0}K^- \) are required to satisfy \( |\cos \theta_{hel}| > 0.2 \) (\( > 0.35 \)) for the \( B_s^0 \to D_s^- \pi^+ (B_s^0 \to D_s^+ K^\pm) \) mode. These two selections reject 43% (73%) of the continuum while retaining 95% (85%) of the \( B_s^0 \to D_s^- \pi^+ (B_s^0 \to D_s^+ K^\pm) \) signal. MC studies show that background from \( B^+ \) (\( B^0 \)) decays is small and flat enough to be described together with the continuum events for the \( B_s^0 \to D_s^- \pi^+ \) mode and is negligible for the \( B_s^0 \to D_s^+ K^\pm \) mode. The relevant background from \( B_s^0 \) decays is \( B_s^0 \to D_s^- \pi^+ \).

For each mode, a two-dimensional unbinned extended maximum likelihood fit \( 17 \) in \( M_{bc} \) and \( \Delta E \) is performed on the selected candidates, which are shown in Fig. 1. Each signal probability density function (PDF) is described by a sum of two Gaussians. For the \( B_s^0 \to D_s^- \pi^+ \) analysis, all three \( B_s^0 \) production modes \( (B_s^0 \to B_s^* \to B_s^0) \) and \( B_s^0 \) (\( B_s^0 \)) are fitted simultaneously. For the \( B_s^0 \to D_s^\mp K^\mp \) mode, only the \( B_s^0 \) component is taken into account. The resolutions for \( M_{bc} \) and \( \Delta E \) are estimated from the MC and scaled by a common factor (one for each variable) left free in the \( B_s^0 \to D_s^- \pi^+ \) fit. Approximating \( p_{B_s^0} \) with \( p_{B_s^0} \) in the \( B_s^0 \to D_s^\mp K^\mp \) decay, the mean values are parameterized, as shown in Table I as functions of the \( B_s^0 \) and \( B_s^0 \) masses, which are also left free in the \( B_s^0 \to D_s^- \pi^+ \) fit. The continuum (together with possible \( B^+ \) and \( B^0 \) background) is modeled with an ARGUS function \( 18 \) for \( M_{bc} \) and a linear function for \( \Delta E \). A non-parametric two-dimensional PDF, obtained from MC with the KEYS method \( 19 \), is used to describe the shape of the \( B_s^0 \to D_s^- \pi^+ \) background.

| Signal | Mean of \( (M_{bc}, \Delta E) \) |
|--------|---------------------|
| \( B_s^0 \to B_s^* \) | \( \left( m_{B_s^*}, \sqrt{E_{B_s^*}^2 - (m_{B_s^*}^2 + m_{B^0}^2)} - E_b \right) \) |
| \( B_s^0 \to B_s^0 \) | \( \left( m_{B^0}, 0 \right) \) |

For the \( B_s^0 \to D_s^- \pi^+ \) mode, the three signal yields are expressed as a function of three free parameters, \( \mathcal{B}(B_s^0 \to D_s^- \pi^+) \), \( f_{B_s^0}B_s^* \), and \( f_{B_s^0}B_s^0 \), with the relations \( N_M = N_{B_s^0}B_s^0 \mathcal{B}(B_s^0 \to D_s^- \pi^+) \sum_k \varepsilon_k M_{B_k} \) where \( M \) is one of the three \( B_s^{(*)}B_s^{(*)} \) pair production modes and \( k \) runs over the \( D_s^- \) modes; the third fraction is defined as \( f_{B_s^0}B_s^0 = 1 - f_{B_s^0}B_s^* - f_{B_s^0}B_s^0 \). The values of \( \sum_k \varepsilon_k M_{B_k} \), which are the total \( D_s^- \) branching fractions weighted
by the reconstruction efficiencies, are listed in Table II.

Figure 2 shows the $M_{bc}$ and $\Delta E$ projections in the $B_s^* \bar{B}_s^*$ and in the $B_s^* B_0^0$ regions of the data, together with the fitted function. In the $M_{bc}$ distribution, the three signal components are present due to overlap of the signal boxes; the peak on the right (middle, left) is due to $B_s^* B_s^*$ ($B_s^* B_0^0$, $B_s^* \bar{B}_s^*$) production. Table III presents the fitted signal yields as well as the significance defined by $S = \sqrt{2 \ln (L_{\text{max}} / L_0)}$ where $L_{\text{max}}$ ($L_0$) is the value at the maximum (with the corresponding yield set to zero) of the likelihood function convolved with a Gaussian distribution that represents the systematic errors.

Systematic uncertainties on the branching fractions are shown in Table III. Those on $f_{B_s^* B_s^*}$ and $f_{B_s^* \bar{B}_s^*}$ are mainly due to PDF uncertainties. Those due to the beam energy, the momentum calibration and the $p_T(B_s^*) \approx p_T(B_0^0)$ approximation are propagated as systematics on the $B_s^*$ mass and $B_0^0$ mass. The momentum normalization uncertainties are much more important in the latter case because the measured energy of the $B_0^0$ candidate is used instead of the beam energy.

We measure the branching fraction $\mathcal{B}(B_s^0 \to D_s^- \pi^+) = [3.67^{+0.90}_{-0.33}(\text{stat.})^{+0.42}_{-0.40}(\text{syst.}) \pm 0.49(f_s)] \times 10^{-3}$ where the largest systematic uncertainty, due to $f_s$, is quoted separately, the fraction $f_{B_s^* \bar{B}_s^*} = (90.1^{+3.8}_{-4.0} \pm 0.2)$ % and the two fitted masses $m_{B_0^0} = (5364.4 \pm 1.3 \pm 0.7)$ MeV/c² and $m_{D_s} = (5416.4 \pm 0.4 \pm 0.5)$ MeV/c². These four measurements supersede the previous Belle values [10]. We obtain for the first time values for the two fractions $f_{B_s^* B_0^0} = (7.3^{+3.3}_{-3.0} \pm 0.1)$ % and $f_{B_s^* \bar{B}_s^*} = (2.6^{+2.6}_{-2.3})$ %, using the correlation ($-0.77$) between $f_{B_s^* B_0^0}$ and $f_{B_s^* \bar{B}_s^*}$.

Our branching fraction is compatible with the CDF result $[12, 20]$, and is slightly higher (1.3σ) than $\mathcal{B}(B_s^0 \to D_s^- \pi^+)$ [12]. The value of $f_{B_s^* B_s^*}$ is significantly larger than the theoretical expectation of $\approx 70\%$ [3, 6]. The

![FIG. 2](image-url)
$B_s^0$ mass is compatible with the world average value [12] while our value for the $B_s^*$ mass is 2.6σ larger than the result from CLEO [21]. The mass difference obtained, $m_{B_s^0} - m_{B_s^*} = 52.0 \pm 1.5 \text{ MeV}/c^2$, is 4.0σ larger than the world average of $m_{B_s^0} - m_{B_s^*}$ [12], while heavy-quark symmetry predicts equal values [4].

The distribution of the angle between the $B_s^0$ momentum and the beam axis in the CM frame is of theoretical interest [3] and is presented in Fig. 3 for the signal events in the $B_s^* B_s^*$ region, using the $\chi$Plot method [22]. A fit to a $1 + a \cos^2 \theta_{B_s^0}$ distribution returns $\chi^2$/n.d.f. = 8.74/8 and $a = -0.39^{+0.18}_{-0.16}$. It has been checked that the signal efficiency does not depend on this angle. We naively expect $a = -0.27$ by summing over all the possible polarization states.

![FIG. 3: Fitted distribution of the cosine of the angle between the $B_s^0$ momentum and the beam axis in the CM frame for the $\Upsilon(5S) \to B_s^* B_s^*$ signal.](image)

For the $B_s^0 \to D_s^- K^\pm$ mode, mean values and resolutions for $B_s^0 \to D_s^+ K^\pm$ and $B_s^0 \to D_s^- \pi^\mp$ components are calibrated using the results of the $B_s^0 \to D_s^- \pi^+$ fit. The four yields (signal, continuum, $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^- \pi^+$) are allowed to float, but due to the very small contribution of $B_s^0 \to D_s^- \pi^+$, the ratio between the yields of $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^- \pi^+$ is fixed from a fit to data without kaon identification.

The fit results are shown in Fig. 4 and Table I. Systematic errors are presented in Table II. We find $6.7^{+3.4}_{-2.7}$ signal events (3.5σ), corresponding to $B \to D_s^- K_s f_s$, is $2.4^{+1.0}_{-0.7}$ (stat.) ±0.3 (syst.) ±0.3(fit) × 10^{-4}. Using the previously fitted value of $f_{B_s^0} B_s^*$, the ratio $B \to D_s^- K_s f_s / B \to D_s^- K_s f_s$ is $(6.5^{+3.9}_{-2.5}) \%$, the errors are dominated by the low $B_s^0 \to D_s^- K^\pm$ statistics.

![FIG. 4: Left: $M_{bc}$ distribution of $B_s^0 \to D_s^- K^\pm$ candidates with $\Delta E$ in the $B_s^* B_s^*$ signal region. Right: $\Delta E$ distribution of the $B_s^0 \to D_s^- K^\pm$ candidates with $M_{bc}$ in the $B_s^* B_s^*$ signal region; the left (right) peak is the $B_s^0 \to D_s^- K^\pm$ ($B_s^0 \to D_s^- \pi^+$) component. The dashed, dotted and dash-dotted curves represent the signal, $B_s^0 \to D_s^- \pi^+$ backgrounds and continuum, respectively.](image)

In summary, a large $B_s^0 \to D^- \pi^+$ signal is observed and six physics parameters are measured: the branching fraction $B \to D_s^- \pi^+$ = 3.67^{+0.35}_{-0.33} (stat.) \pm0.42 (syst.) 0.49 (f_s) \times 10^{-3}$, the fractions of the $B_s^0$ pair production modes at the $\Upsilon(5S)$ energy, $f_{B_s^0} f_{B_s^0} = (90.1^{+3.8}_{-4.0} \pm 0.2) \%$, $f_{B_s^0} f_{B_s^0} = (7.3^{+3.0}_{-3.0} \pm 0.1) \%$, $f_{B_s^0} f_{B_s^0} = (26.6^{+2.6}_{-2.5}) \%$, and the masses $m_{B_s^0} = (5416.4 \pm 0.4 \pm 0.5) \text{ MeV}/c^2$, $m_{B_s^0} = (5364.4 \pm 1.3 \pm 0.7) \text{ MeV}/c^2$. In addition, evidence (3.5σ) for the $B_s^0 \to D_s^- K^\pm$ decay is obtained, leading to a measurement $B \to D_s^- K^\pm$ = 2.4^{+1.2}_{-1.0} (stat.) ±0.3 (syst.) ±0.3 (fit) × 10^{-4}.

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 and JSPS (Japan); ARC and DEST (Australia); NSFC (China); DST (India); MOEHRD, KOSEF and KRF (Korea); KBN (Poland); MES and RFRAE (Russia); ARRS (Slovenia); SNSF (Switzerland); NSC and MOE (Taiwan); and DOE (USA).

* now at Okayama University, Okayama

[1] Unless specified otherwise, charge-conjugated modes are implied throughout the Letter.
[2] L. Wolfenstein, Phys. Rev. Lett. 51, 1945 (1983).
[3] R. Aleksan, I. Dunietz and B. Kayser, Z. Phys. C 54, 653 (1992). See also R. Fleicher, Nucl. Phys. B 671, 459 (2003) and S. Nandi and U. Nierste, Phys. Rev. D 77, 054010 (2008).
[4] W.A. Bardeen, E.J. Eichten and C.T. Hill, Phys. Rev. D 68, 054024 (2003).
[5] A.G. Grozin and M. Neubert, Phys. Rev. D 55, 272 (1997).
[6] N.A. Törnqvist, Phys. Rev. Lett. 53, 878 (1984).
[7] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods A 499, 1 (2003).
[8] K.-F. Chen et al. (Belle Collaboration), Phys. Rev. Lett. 100, 112001 (2008). We obtain $\sqrt{s} = m_{\Upsilon(1S)} + M \Delta m$ where $m_{\Upsilon(1S)}$ is the nominal $\Upsilon(1S)$ mass [12] and $\Delta m$ is the measured $M_{\mu^+\mu^-\pi^+\pi^-} - M_{\mu^+\mu^-}$.
[9] A. Drutskoy et al. (Belle Collaboration), Phys. Rev. Lett. 98, 052001 (2007), G. S. Huang et al. (CLEO Collaboration), Phys. Rev. D 75, 012002 (2007). These two published values of $\sigma_{\Upsilon(5S)}$ are averaged. Experimental $f_s$ values are also given by both of them; the average is given in Ref. 12.
[10] A. Drutskoy et al. (Belle Collaboration), Phys. Rev. D 76, 012002 (2007).
[11] A. Abashian et al. (Belle Collaboration), Nucl. Instrum. Methods A 479, 117 (2002).
[12] C. Ansler et al. (Particle Data Group), Phys. Lett. B 667, 1 (2008).
[13] F. Fang, Ph.D. thesis, University of Hawaii (2003).
[14] D.J. Lange, Nucl. Instrum. Methods A 462, 132 (2001).
[15] CERN Application Software Group (1993), CERN Program Library, W5013.
[16] G.C. Fox and S. Wolfram, Phys. Rev. Lett. 41, 1581 (1978).
[17] R. Barlow, Nucl. Instrum. Methods A 297, 496 (1990).
[18] H. Albrecht et al. (ARGUS Collaboration), Phys. Lett. B 185, 218 (1987).
[19] K. Cranmer, Comput. Phys. Commun. 136, 198 (2001).
[20] A. Abulencia et al. (CDF Collaboration), Phys. Rev. Lett. 98, 061802 (2007).
[21] O. Aquines et al. (CLEO Collaboration), Phys. Rev. Lett. 96, 152001 (2006).
[22] M. Pivk and F.R. Le Diberder, Nucl. Instrum. Methods A 555, 356 (2005).