Observation of $B_s^0 \rightarrow D_s^{(*)} + D_s^{(*)*}$ using $e^+e^-$ collisions and a
determination of the $B_s - \bar{B}_s$ width difference $\Delta \Gamma_s$

S. Esen, A. J. Schwartz, I. Adachi, H. Aihara, K. Arinstein, V. Aulchenko, 
T. Aushev, T. Aziz, A. M. Bakich, V. Balagura, E. Barberio, A. Bay, 
M. Bischofberger, A. Bondar, A. Bozek, M. Bračko, T. E. Browder, 
M.-C. Chang, P. Chang, A. Chen, P. Chen, B. G. Cheon, C.-C. Chiang, Y. Choi, 
J. Dalseno, M. Dash, Z. Doležal, Z. Drášal, A. Drutskoy, S. Eidelman, 
P. Goldenzweig, B. Golob, H. Ha, J. Haba, K. Hayasaka, T. Higuchi, 
Y. Hoshi, W.-S. Hou, Y. B. Hsiung, H. J. Hyun, T. Iijima, K. Inami, R. Itoh, 
M. Iwabuchi, N. J. Joshi, T. Juliuss, J. H. Kang, T. Kawasaki, H. Kichimi, 
H. J. Kim, H. O. Kim, J. H. Kim, Y. J. Kim, K. Kinoshita, B. R. Ko, 
P. Kodyš, S. Korpar, P. Križan, P. Krokovny, T. Kuhr, T. Kumita, 
Y.-J. Kwon, S.-H. Kyeong, J. S. Lange, S.-H. Lee, Y. Liu, D. Liventsev, 
R. Louvot, A. Matyja, S. McOnie, H. Miyata, R. Mizuk, G. B. Mohanty, 
E. Nakano, M. Nakao, H. Nakazawa, Z. Natkaniec, S. Neubauer, S. Nishida, 
O. Nitoh, T. Ohshima, S. Okuno, S. L. Olsen, P. Pakhlov, C. W. Park, 
H. Park, H. K. Park, M. Petrič, L. E. Piilonen, M. Röhrken, S. Ryu, 
H. Sahoo, Y. Sakai, O. Schneider, C. Schwanda, K. Senyo, M. E. Sevior, 
M. Shapkin, C. P. Shen, J.-G. Shiu, P. Smerkol, E. Solovieva, M. Starič, 
K. Sumisawa, T. Sumiyoshi, Y. Teramoto, K. Trabelsi, S. Uehara, Y. Unno, 
S. Uno, P. Urquijo, Y. Usov, G. Varner, K. E. Varvell, K. Vervink, 
C. H. Wang, M.-Z. Wang, P. Wang, Y. Watanabe, R. Wedd, J. Wicht, E. Won, 
B. D. Yabsley, Y. Yamashita, Z. P. Zhang, V. Zhilich, and A. Zupanc

(The Belle Collaboration)

1Budker Institute of Nuclear Physics, Novosibirsk
2Faculty of Mathematics and Physics, Charles University, Prague
3University of Cincinnati, Cincinnati, Ohio 45221
4 Department of Physics, Fu Jen Catholic University, Taipei
5 Justus-Liebig-Universität Gießen, Gießen
6 The Graduate University for Advanced Studies, Hayama
7 Hanyang University, Seoul
8 University of Hawaii, Honolulu, Hawaii 96822
9 High Energy Accelerator Research Organization (KEK), Tsukuba
10 Institute of High Energy Physics, Chinese Academy of Sciences, Beijing
11 Institute of High Energy Physics, Vienna
12 Institute of High Energy Physics, Protvino
13 Institute for Theoretical and Experimental Physics, Moscow
14 J. Stefan Institute, Ljubljana
15 Kanagawa University, Yokohama
16 Institut für Experimentelle Kernphysik, Karlsruher Institut für Technologie, Karlsruhe
17 Korea Institute of Science and Technology Information, Daejeon
18 Korea University, Seoul
19 Kyungpook National University, Taegu
20 École Polytechnique Fédérale de Lausanne (EPFL), Lausanne
21 Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana
22 University of Maribor, Maribor
23 Max-Planck-Institut für Physik, München
24 University of Melbourne, School of Physics, Victoria 3010
25 Nagoya University, Nagoya
26 Nara Women’s University, Nara
27 National Central University, Chung-li
28 National United University, Miao 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 Novosibirsk State University, Novosibirsk
Abstract

We have made the first observation of $B_s^0 \to D_s^{(*)}+D_s^{(*)}-$ decays using 23.6 fb$^{-1}$ of data recorded by the Belle experiment running on the $\Upsilon(5S)$ resonance. The branching fractions are measured to be $B(B_s^0 \to D_s^+D_s^-) = (1.03^{+0.39}_{-0.32}^{+0.26})\%$, $B(B_s^0 \to D_s^{(*)}D_s^{(*)}) = (2.75^{+0.83}_{-0.71} \pm 0.69)\%$, and $B(B_s^0 \to D_s^+D_s^-) = (3.08^{+1.22}_{-1.04}^{+0.85})\%$; the sum is $B(B_s^0 \to D_s^{(*)}D_s^{(*)}) = (6.85^{+1.53}_{-1.30}^{+1.79})\%$. Assuming $B_s^0 \to D_s^{(*)}D_s^{(*)}$ saturates decays to $CP$-even final states, the branching fraction determines the ratio $\Delta\Gamma_s/\cos\varphi$, where $\Delta\Gamma_s$ is the difference in widths between the two $B_s-B_s$ mass eigenstates, and $\varphi$ is a $CP$-violating weak phase. Taking $CP$ violation to be negligibly small, we obtain $\Delta\Gamma_s/\Gamma_s = 0.147^{+0.036}_{-0.036} \text{(stat.)}^{+0.042}_{-0.041} \text{(syst.)}$, where $\Gamma_s$ is the mean decay width.
Decays of $B_s$ mesons help elucidate the weak Cabibbo-Kobayashi-Maskawa structure of the Standard Model (SM). Because they are not produced in $\Upsilon(4S)$ decays, $B_s$ mesons are much less studied than their $B^0_s$ and $B^\pm$ counterparts. Most $B_s$ data comes from the hadron collider experiments CDF and DØ. Recently, another method to study $B_s$ decays has been exploited: that of running an $e^+e^-$ collider at a center-of-mass (CM) energy corresponding to the $\Upsilon(5S)$ resonance, which subsequently decays to $B_s^{(*)}\overline{B}_s^{(*)}$ pairs. Both the CLEO [1] and Belle [2–4] Collaborations have used this method to measure inclusive and exclusive $B_s^0$ decays. In this paper we use $L_{\text{int}} = 23.6 \text{ fb}^{-1}$ of data recorded by Belle at the $\Upsilon(5S)$ ($\sqrt{s} = 10.87 \text{ GeV}$) to make the first observation ($>5\sigma$) of $B_s^0 \rightarrow D_s^{(*)}+D_s^{(*)}-$. First measurements of the $D_s^{(*)}+D_s^{(*)}-$ final state were made by the ALEPH [6] and DØ [8] Collaborations using inclusive $\phi\phi X$ and $D_s^+D_s^-X$ samples. CDF [7] measured the single decay $B_s^0 \rightarrow D_s^+D_s^-$. Here we exclusively reconstruct all three final states: $D_s^+D_s^-$, $D_s^{*\pm}D_s^{*\mp}$, and $D_s^{*+}D_s^{*-}$.

These final states are expected to be predominantly $CP$-even [9], and the (Cabibbo-favored) partial widths should dominate the difference in decay widths $\Delta\Gamma_s^{CP}$ between the two $B_s\overline{B}_s$ $CP$ eigenstates [9]. This parameter equals $\Delta\Gamma_s/\cos\varphi$, where $\Delta\Gamma_s$ is the width difference between the mass eigenstates, and $\varphi$ is a $CP$-violating weak phase [13]. Thus the branching fraction gives a constraint in $\Delta\Gamma_s-\varphi$ parameter space. Both of these parameters can receive contributions from new physics [10, 11]. The values favored by current measurements [12] differ somewhat from the SM prediction [10].

The Belle detector [14] running at the KEKB collider [15] includes a silicon vertex detector, a central drift chamber, an array of aerogel threshold Cherenkov counters, time-of-flight scintillation counters, and an electromagnetic calorimeter. At the $\Upsilon(5S)$ resonance, the $e^+e^- \rightarrow b\bar{b}$ cross section is $\sigma_{b\bar{b}} = 0.302 \pm 0.014 \text{ nb}$ [1, 2], and the fraction of $\Upsilon(5S)$ decays producing $B_s$ mesons is $f_s = 0.193 \pm 0.029$ [16]. Three production modes are kinematically allowed: $B_s\overline{B}_s$, $B_s\overline{B}_s^*$ or $B_s\overline{B}_s^*$, and $B_s^*\overline{B}_s^*$. In this analysis we use only the last (dominant) mode, for which the fraction is $f_{B_s^*\overline{B}_s^*} = 0.901 ^{+0.038}_{-0.040}$ [4]. The $B_s^*$ decays via $B_{s^*} \rightarrow B_s\gamma$, and the $\gamma$ is not reconstructed. Thus the number of $B_s\overline{B}_s$ pairs used in this analysis is $N_{B_s\overline{B}_s} = L_{\text{int}} \cdot \sigma_{b\bar{b}} \cdot f_s \cdot f_{B_s^*\overline{B}_s^*} = (1.24 \pm 0.20) \times 10^6$.

We select $B_s^0 \rightarrow D_s^+D_s^-$, $D_s^{*\pm}D_s^{*\mp}$, and $D_s^+D_s^-$ decays in which $D_s^+ \rightarrow \phi\pi^+$, $K_S^0 K^+$, $K_S^0 K^*$, and $K_S^0 K^{*+}$. We require that charged tracks originate from near the $e^+e^-$ interaction point. Charged kaons are selected by requiring that a kaon likelihood variable based on $dE/dx$ measured in the central drift chamber and information from the
aerogel threshold Cherenkov counters and time-of-flight scintillation counters be $>0.60$; this requirement is $\sim90\%$ efficient and has a $\pi^\pm$ misidentification rate of $\sim10\%$. Tracks having kaon likelihood $<0.60$ are identified as $\pi^\pm$. Neutral $K^0_S$ candidates are reconstructed from $\pi^+\pi^-$ pairs having an invariant mass within 10 MeV/$c^2$ of the $K^0_S$ mass and satisfying loose requirements on the decay vertex position. The momentum of tracks (except the $\pi^\pm$ from $K^0_S$ decay) must be $>100$ MeV/$c$.

Neutral $\pi^0$ candidates are reconstructed from $\gamma\gamma$ pairs having an invariant mass within 15 MeV/$c^2$ of the $\pi^0$ mass. The photons must have a laboratory energy greater than 100 MeV. Neutral $K^\star^0$ (charged $K^\star^+$) candidates are reconstructed from a $K^-$ ($K^0_S$) and $\pi^+$ having an invariant mass within 50 MeV/$c^2$ of $M_{K^\star^0}$ ($M_{K^\star^+}$). Neutral $\phi$ (charged $\rho^+$) candidates are reconstructed from a $K^+K^-$ ($\pi^+\pi^0$) pair having an invariant mass within 12 MeV/$c^2$ (100 MeV/$c^2$) of $M_\phi$ ($M_{\rho^+}$).

The invariant mass windows used for $D_s^+$ candidates are 10 MeV/$c^2$ ($2.5-3.2\sigma$) for the three final states containing $K^*$ candidates, 20 MeV/$c^2$ ($1.7\sigma$) for $\phi\rho^+$, and 15 MeV/$c^2$ ($\gtrsim4.0\sigma$) for the remaining two modes. For the three vector-pseudoscalar final states, we impose a loose requirement on the helicity angle $\theta_{\text{hel}}$, which is the angle between the momentum of the charged daughter of the vector particle and the direction opposite the $D_s$ momentum in the rest frame of the vector particle. We require $|\cos\theta_{\text{hel}}| > 0.20$, which retains 99\% of signal decays and rejects 18\% of remaining background.

To reconstruct $D_s^+\rightarrow D_s^+\gamma$ decays, we pair $D_s^+$ candidates with photon candidates and require that the mass difference $M_{\tilde{D}_s^+\gamma} - M_{D_s^+}$ be within 12.0 MeV/$c^2$ of the nominal value (143.8 MeV/$c^2$), where $\tilde{D}_s^+$ denotes the reconstructed $D_s^+$ candidate. This requirement (and that for $R$ discussed below) is determined by optimizing a figure-of-merit $S/\sqrt{S+B}$, where $S$ is the expected signal based on Monte Carlo (MC) simulation and $B$ is the expected background as estimated from a data sideband. We require that the photon energy in the CM system be greater than 50 MeV, and that the energy deposited in the central $3 \times 3$ array of cells of the electromagnetic calorimeter cluster be at least 85\% of the energy deposited in the central $5 \times 5$ array of cells.

Signal $B_s$ decays are reconstructed from $D_s^{(*)}D_s^{(*)}$ pairs using two quantities: the beam-energy-constrained mass $M_{bc} = \sqrt{E_{\text{beam}}^2 - p_B^2}$, and the energy difference $\Delta E = E_B - E_{\text{beam}}$, where $p_B$ and $E_B$ are the reconstructed momentum and energy of the $B^0_s$, and $E_{\text{beam}}$ is the beam energy. These quantities are evaluated in the $e^+e^-$ CM frame. When the $B^0_s$ is not
fully reconstructed, e.g., due to losing the $\gamma$ from $D_s^{*-} \rightarrow D_s^+ \gamma$, $\Delta E$ is shifted lower but $M_{bc}$ remains almost unchanged. We determine our signal yields by fitting events in the region $5.20 \text{ GeV}/c^2 < M_{bc} < 5.45 \text{ GeV}/c^2$ and $-0.15 \text{ GeV} < \Delta E < 0.10 \text{ GeV}$. The modes $\Upsilon(5S) \rightarrow B_s \overline{B}_s$, $B_s \overline{B}_s^*$ and $B_s^* \overline{B}_s^*$ are well-separated in $M_{bc}$-$\Delta E$ space. We see no evidence for $B_s \overline{B}_s$ and $B_s \overline{B}_s^*$ and thus do not fit for them. The expected yields based on Ref. [4] are less than one event for each of $D_s^+ / D_s^-$, $D_s^{+\ast} D_s^+$, $D_s^{+\ast} D_s^-$ final states.

Approximately half of the events have multiple $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ candidates, which usually arise from low momentum $\gamma$'s produced from $\pi^0$ decays. For these events we select the candidate that minimizes the quantity

$$
\chi^2 = \frac{1}{(2 + N)} \left\{ \sum_{D_s} \left[ (\bar{M}_{D_s} - M_{D_s})/\sigma_M \right]^2 + \sum_{\bar{D}_s} \left[ (\bar{\Delta}M - \Delta M)/\sigma_{\Delta M} \right]^2 \right\},
$$

where $\Delta M = M_{D_s} - M_{\bar{D}_s}$, $\bar{M}_{D_s}$ and $\bar{\Delta}M$ are reconstructed quantities; $\sigma_M$ and $\sigma_{\Delta M}$ are the uncertainties on $M_{D_s}$ and $\Delta M$; and the summations run over the two $D_s^+$ daughters and possible $D_s^{+\ast}$ daughters ($N = 0, 1, 2$) of a $B_s^0$ candidate. According to the MC simulation, this criterion selects the correct $B_s^0$ candidate 85%, 76%, and 75% of the time, respectively, for $D_s^+ D_s^-$, $D_s^{+\ast} D_s^+$, and $D_s^{+\ast} D_s^-$ final states.

We reject background from $e^+ e^- \rightarrow q\bar{q}$ ($q = u, d, s, c$) continuum events based on event topology: $q\bar{q}$ events tend to be collimated, while $B_{(s)} \overline{B}_{(s)}$ events tend to be spherical. We distinguish these topologies using a Fisher discriminant based on a set of modified Fox-Wolfram moments [18]. This discriminant is used to calculate a likelihood $L_s$ ($L_{q\bar{q}}$) for an event assuming the event is signal ($q\bar{q}$ background). We form the ratio $R = L_s / (L_s + L_{q\bar{q}})$ and require $R > 0.20$. This selection is 95% efficient for signal decays and removes $> 80\%$ of $q\bar{q}$ background.

The remaining background consists of $\Upsilon(5S) \rightarrow B_{(s)}^{(*)+} \overline{B}_{(s)}^{(*)-} \rightarrow D_s^{\pm} X$, $\Upsilon(5S) \rightarrow BBX$ (where $\bar{b}\bar{b}$ hadronizes into $B^0$, $\overline{B}^0$, or $B^\pm$), and $B_s \rightarrow D_{sJ}^{(*)}(2317) D_s^{(*)}$, $B_s \rightarrow D_{sJ}^{(*)}(2460) D_s^{(*)}$, and $B_s \rightarrow D_s^\pm D_s^\mp \pi^0$ decays. The last three processes peak at negative values of $\Delta E$, and their yields are estimated using analogous $B_{(s)} \rightarrow D_{sJ}^{(*)} D_s^{(*)}$ branching fractions. The total yields for all backgrounds within an $M_{bc}$-$\Delta E$ signal region spanning $3\sigma$ in $(M_{bc}, \Delta E)$ resolution are $0.25 \pm 0.03, 0.25 \pm 0.06$, and $0.15 \pm 0.13$ events, respectively, for $B_s \rightarrow D_s^+ D_s^-$, $D_s^{+\ast} D_s^+$, and $D_s^{+\ast} D_s^-$ decays. To check our background estimates, we count events in the sideband region $M_{bc} < 5.375 \text{ GeV}/c^2$ and find reasonable agreement with the yields predicted from MC simulation. All selection criteria are finalized before looking at events in the signal region.
regions. The final event samples are shown in Fig. 1.

To measure the signal yields, we perform a two-dimensional extended unbinned maximum-likelihood fit to the $M_{bc} - \Delta E$ distributions. For each sample, we include probability density functions (PDFs) for signal and $q\bar{q}$, $B_s^{(*)} \overline{B}_s^{(*)} \rightarrow D^+_s X$, and $\Upsilon(5S) \rightarrow BBX$ backgrounds. As these backgrounds have similar $M_{bc}$, $\Delta E$ shapes, we use a single PDF for them, taken to be an ARGUS function [20] for $M_{bc}$ and a third-order Chebyshev polynomial for $\Delta E$. All shape parameters are taken from MC simulation. Other backgrounds are very small and considered only when evaluating systematic uncertainties.

The signal PDFs have three components: correctly reconstructed (CR) decays; “wrong combination” (WC) decays in which a non-signal track or photon is included in place of a true daughter track or photon; and “cross-feed” (CF) decays in which a $D_s^{(*)}\pi^+$ or $D_{s0}^{(*)}\pi^0$ is reconstructed as $D_s^{(*)}\pi^+$ or $D_{s0}^{(*)}\pi^0$ respectively, or else a $D_s^{(*)}\pi^+$ or $D_{s0}^{(*)}\pi^0$ is reconstructed as $D_s^{(*)}\pi^+$ or $D_{s0}^{(*)}\pi^0$. In the former (latter) case the signal decay has lost (gained) a photon, and $\Delta E$ is typically shifted lower (higher) by 100-150 MeV. The PDF for CR events is modeled with a single Gaussian for $M_{bc}$ and a double-Gaussian with common mean for $\Delta E$. The means and widths are taken from MC simulation and calibrated using $B_0 \rightarrow D^{(*)}\pi^0$ and $B^0 \rightarrow D^{(*)}\pi^0$ control samples. The PDFs for WC and CF events are modeled from MC simulation using non-parametric PDFs with Kernel Estimation [19]. The fractions of WC and CF-down events are also taken from MC simulation. The fractions of CF-up events are difficult to simulate and thus floated in the fit, as the extraneous $\gamma$ usually originates from a $B$ decay chain and many $B, B_s$ partial widths are unmeasured. As the CF-down fractions are fixed, the three distributions ($D_s^+D_s^-, D_s^{(*)}\pi^0$, and $D_{s0}^{(*)}\pi^0$) are fitted simultaneously [21]. The CF fractions are typically 0.1–0.4.

The fit results are listed in Table I and projections of the fit are shown in Fig. 2. The branching fraction for channel $i$ is calculated as $B_i = Y_i / (\varepsilon_{MC}^i \cdot N_{B_i\overline{B}_s} \cdot 2)$, where $Y_i$ is the fitted CR yield, and $\varepsilon_{MC}^i$ is the MC efficiency with intermediate branching fractions included. The efficiencies $\varepsilon_{MC}^i$ include small correction factors to account for differences between MC simulation and data for kaon identification. Inserting all values gives the branching fractions listed in Table I. The statistical significance is calculated as $\sqrt{-2 \ln(L_0/L_{\text{max}})}$, where $L_0$ and $L_{\text{max}}$ are the values of the likelihood function when the signal yield $Y_i$ is fixed to zero and when it is the fitted value, respectively. We include systematic uncertainty in the significance by smearing the likelihood function by a Gaussian having a width equal to the
total systematic error related to the signal yield.

TABLE I: Signal yields (Y), efficiencies including intermediate branching fractions (ε), branching fractions (B), and signal significance (S) including systematic uncertainty. The first error listed is statistical, the second is from systematics due to the analysis procedure, and the third is from systematics due to external inputs.

| Mode       | Y (events) | ε (×10⁻⁴) | B (%) | S       |
|------------|------------|-----------|-------|---------|
| $D_s^+D_s^-$ | $8.5^{+3.2}_{-2.6}$ | 3.31 | $1.03^{+0.39+0.15}_{-0.32-0.13}$ | ± 0.21 | 6.2 |
| $D_s^{±}D_s^{±}$ | $9.2^{+2.8}_{-2.4}$ | 1.35 | $2.75^{+0.83}_{-0.71}$ | ± 0.40 | ± 0.56 | 6.6 |
| $D_s^{+}D_s^{-}$ | $4.9^{+1.9}_{-1.7}$ | 0.643 | $3.08^{+1.22+0.57}_{-1.04-0.58}$ | ± 0.63 | 3.1 |
| Sum        | $22.6^{+4.7}_{-3.9}$ | $6.85^{+1.53}_{-1.30}$ | $1.11^{+1.40}_{-1.41}$ |         |

The systematic errors are listed in Table II. The error due to PDF shapes is evaluated by varying shape parameters by ±1σ for backgrounds, and by trying different parameterizations for the WC and CF components. The systematic error for the fixed CF-down fractions is evaluated by fitting a $B^0_s \rightarrow D_s^\pm \pi^\mp$ control sample and comparing the fraction of $B^0_s \rightarrow D_s^\pm \pi^\mp$ contamination with that predicted by MC simulation. The difference is taken as the range of variation for the CF-down fractions. We vary the fractions over this range and take the r.m.s. variation in B as the systematic error. The error due to the fixed WC fractions is evaluated in the same way, but the range of variation for the WC fractions is ±20%. The uncertainty due to $K^{±}$ identification depends on momentum but is ~2.5% per track; as our final states typically have four charged kaons, this error is 10-11%. The error due to tracking efficiency is 1.0% per track. Uncertainty due to the unknown $B^0_s \rightarrow D_s^+D_s^-$ longitudinal polarization fraction ($f_L$) affects all three modes due to the CF components. For our nominal result, we take $f_L$ to be the world average (WA) value for the analogous spectator decay $B^0_d \rightarrow D^+_sD_s^-$: $0.52 \pm 0.05$ [16]. The systematic error is taken as the change in B when $f_L$ is varied by twice the error on the WA value. Significant uncertainties arise from $D_s^+$ branching fractions, $\sigma_{Y(5S)}$, $f_s$, and $f_{B_s^0D_s^0}$, which are external factors that should be measured more precisely in the future. We list separately the systematic error due to
FIG. 1: $M_{bc}$ vs. $\Delta E$ scatter plots. The signal ellipses correspond to $3\sigma$ in resolution for $\Upsilon(5S) \rightarrow B_s^* B_s^*$ decays; the number of candidates within the ellipses is listed. The top, middle, and bottom plots correspond to $B_s^0 \rightarrow D_s^+ D_s^-$, $B_s^0 \rightarrow D_s^{*\pm} D_s^{\mp}$, and $B_s^0 \rightarrow D_s^{*+} D_s^{-}$, respectively.
FIG. 2: $M_{bc}$ and $\Delta E$ projections of the fit result. The rows correspond to $B_s^0 \rightarrow D_{s}^{+}D_{s}^{-}$ (top), $B_s^0 \rightarrow D_{s}^{+}D_{s}^{-}$ (middle), and $B_s^0 \rightarrow D_{s}^{+}D_{s}^{-}$ (bottom). The red dashed (dotted) curves show RC (WC) signal, the green and blue dash-dotted curves show CF signal, the grey solid curve shows background, and the black solid curves show the total.

In the limit $m_{c,b} \rightarrow \infty$ while $(m_b - 2m_c) \rightarrow 0$, the $b \rightarrow c\bar{c}s$ process saturates the decay width $[22]$. If also the number of colors $N_c \rightarrow \infty$, then $B_s^0 \rightarrow D_{s}^{+}D_{s}^{-}$, $D_{s}^{+}D_{s}^{-}$ (along with $D_{s}^{+}D_{s}^{-}$) are $CP$ even (+), and $\Gamma[B_s^0(CP \rightarrow D_{s}^{(*)}D_{s}^{(*)})]$ saturates $\Delta \Gamma_s^{CP}$ [9]. This gives the relationship $2B(B_s^0 \rightarrow D_{s}^{(*)}D_{s}^{(*)}) = (\Delta \Gamma_s^{CP}/2)[(1 + \cos \varphi)/\Gamma_L + (1 - \cos \varphi)/\Gamma_H]$. 

these factors in Table I.
TABLE II: Systematic errors (%). The first twelve sources affect the signal yield and thus the signal significance.

| Source                        | $D_s^+ D_s^-$ | $D_s^+ D_s^+$ | $D_s^+ D_s^-$ |
|-------------------------------|---------------|---------------|---------------|
|                               | $+\sigma$    | $-\sigma$    | $+\sigma$    |
| CR PDF Shape                  | 0.8           | 0.8           | 0.3           |
| Background PDF                | 1.1           | 1.3           | 1.9           |
| WC+CF PDF                     | 0.3           | 0.3           | 1.5           |
| WC/CF Fractions               | 0.2           | 0.2           | 5.0           |
| $\mathcal{R}$ Requirement ($q\bar{q}$ suppr.) | 1.8           | 1.8           | 1.8           |
| Best Candidate Selection      | 6.9           | 0.0           | 2.2           |
| $K^{\pm}$ Identification      | 10.1          | 10.1          | 10.6          |
| $K_S^0$ Reconstruction        | 2.1           | 2.1           | 2.1           |
| $\pi^0$ Reconstruction        | 1.1           | 1.1           | 1.1           |
| $\gamma$                     | -             | -             | 3.8           |
| Tracking                      | 6.2           | 6.2           | 6.2           |
| Polarization                  | 0.2           | 0.0           | 0.8           |
| MC Statistics for $\varepsilon$ | 1.1           | 1.1           | 0.9           |
| $D_s^{(\ast)}$ Branching Fractions | 12.4          | 12.4          | 12.4          |
| Luminosity                    | $\pm 1.3$    |               |               |
| $\sigma_{\Upsilon(5S)}$      | $\pm 4.6$    |               |               |
| $f_s$                         | $\pm 15$     |               |               |
| $f_{B_s^{0}\bar{B}_s^{0}}$   | $\pm 4.2$    |               |               |
| Total                         | 24.9          | 24.0          | 25.1          |

where $\Gamma_{L,H}$ are the decay widths of the light and heavy mass eigenstates $|13\rangle$. Substituting $\Gamma_{L,H} = \Gamma \pm \Delta \Gamma_s/2$ and $\Delta \Gamma_s^{CP} = \Delta \Gamma_s / \cos \varphi \,|13\rangle$ allows one to use the branching fraction $B$ to constrain $\Delta \Gamma_s$ and $\varphi$. If $CP$ violation is negligible, then $\cos \varphi \simeq 1$ and the above expression can be inverted to give $\Delta \Gamma_s / \Gamma_s = 2B/(1 - B)$. Inserting $B$ from Table I yields

$$\frac{\Delta \Gamma_s}{\Gamma_s} = 0.147 \pm 0.036^{+0.042}_{-0.041},$$

(2)

where the first error is statistical and the second is systematic. This result is $1.3\sigma$ higher than that of Ref. [8] but consistent with the theory prediction $|10\rangle$. There is theoretical uncertainty.
arising from the $CP$-odd component in $B^0 \to D_s^{\pm} D_s^{\mp}, D_s^{+} D_s^{-}$ and contributions from other two-body final states; the effect upon $\Delta \Gamma_s/\Gamma_s$ is estimated in Ref. [9] to be $\pm 3\%$. This is much smaller than the statistical/systematic errors on our measurement, but there may be additional contributions coming from three-body final states, which are neglected in [9].

In summary, we have measured the branching fractions for $B^0_s \to D_s^{(*)+} D_s^{(*)-}$ using $e^+ e^-$ data taken at the $\Upsilon(5S)$ resonance. Our results constitute the first observation of $B^0 \to D_s^{\pm} D_s^{\mp}$ ($6.6\sigma$ significance) and provide the first evidence for $B^0_s \to D_s^{(*)+} D_s^{(*)-}$ ($3.1\sigma$ significance). We use these measurements to determine the $B^0_s - \overline{B}_s$ decay width difference $\Delta \Gamma_s$ with improved fractional precision.

We thank R. Aleksan and L. Oliver for useful discussions. 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 (Korea); MNiSW (Poland); MES and RFAAE (Russia); ARRS (Slovenia); SNSF (Switzerland); NSC and MOE (Taiwan); and DOE (USA).

[1] G. S. Huang et al. (CLEO Collab.), Phys. Rev. D 75, 012002 (2007). M. Artuso et al. (CLEO Collab.), Phys. Rev. Lett. 95, 261801 (2005).
[2] A. Drutskoy et al. (Belle Collab.), Phys. Rev. Lett. 98, 052001 (2007).
[3] A. Drutskoy et al. (Belle Collab.), Phys. Rev. D 76, 012002 (2007); J. Wicht et al. (Belle Collab.), Phys. Rev. Lett. 100, 121801 (2008).
[4] R. Louvot et al. (Belle Collab.), Phys. Rev. Lett. 102, 021801 (2009).
[5] Charge-conjugate modes are implicitly included.
[6] R. Barate et al. (ALEPH Collab.), Phys. Lett. B 486, 286 (2000).
[7] T. Aaltonen et al. (CDF Collab.), Phys. Rev. Lett. 100, 021803 (2008).
[8] V. M. Abazov et al. (DØ Collab.), Phys. Rev. Lett. 102, 091801 (2009).
[9] R. Aleksan et al., Phys. Lett. B 316, 567 (1993).
[10] A. Lenz and U. Nierste, Jour. High Energy Phys. 0706, 072 (2007).
[11] See for example: A. J. Buras et al., arXiv:1005.5310; Z. Ligeti et al., arXiv:1006.0432.
[12] CDF Public Note 10206 (2010); DØ Note 6098-CONF (2010). CDF Public Note 9787 (2009).
[13] I. Dunietz, R. Fleischer, and U. Nierste, Phys. Rev. D 63, 114015 (2001); I. Dunietz, Phys. Rev. D 52, 3048 (1995).
[14] A. Abashian et al. (Belle Collab.), Nucl. Instr. Meth. Phys. Res. A 479, 117 (2002).
[15] S. Kurokawa and E. Kikutani, Nucl. Instrum. and Methods Phys. Res. A 499, 1 (2003), and other papers included in this volume.
[16] C. Amsler et al. (Particle Data Group), Phys. Lett. B 667, 1 (2008) and 2009 update for the 2010 edition.
[17] Y. Nakahama et al. (Belle Collab.), Phys. Rev. Lett. 100, 121601 (2008).
[18] 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 Collab.), Phys. Rev. Lett. 91, 261801 (2003).
[19] Kernel Estimation in High-Energy Physics, K. Cranmer, Comput. Phys. Commun. 136, 198 (2001).
[20] H. Albrecht et al. (ARGUS Collab.), Phys. Lett. B 241, 278 (1990).
[21] Thus the fit errors for yields are less than the square root of the yields due to the CF information.
[22] M. A. Shifman and M. B. Voloshin, Sov. J. Nucl. Phys. 47, 511 (1988).