Observation of $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ Using e^+e^- Collisions and a Determination of the B_s-$\bar{B}_s$ Width Difference $\Delta\Gamma_s$

S. Esen,[3] A. J. Schwartz,[3] I. Adachi,[9] H. Aihara,[42] K. Arinstein,[1,33] V. Aulchenko,[1,33] T. Aushev,[20,13] T. Aziz,[39] A. M. Bakich,[38] V. Balagura,[13] E. Barberio,[24] A. Bay,[20] M. Bischofberger,[26] A. Bondar,[1,33] A. Bozek,[30] M. Bračko,[22,14] T. E. Browder,[8] M.-C. Chang,[4] P. Chang,[29] A. Chen,[27] P. Chen,[29] B. G. Cheon,[7] C.-C. Chiang,[29] Y. Choi,[37] J. Dalseno,[23,40] M. Dash,[45] Z. Doležal,[2] Z. Drásal,[2] A. Drutskoy,[3] S. Eidelman,[1,33] P. Goldenzweig,[3] B. Golob,[21,14] H. Ha,[18] J. Haba,[9] T. Hara,[9] K. Hayasaka,[25] T. Higuchi,[9] Y. Hoshi,[41] W.-S. Hou,[29] Y. B. Hsiung,[29] H. J. Hyun,[19] T. Iijima,[25] K. Inami,[25] R. Itoh,[9] M. Iwabuchi,[46] N. J. Joshi,[39] T. Julius,[24] J. H. Kang,[46] T. Kawasaki,[32] H. Kichimi,[9] H. J. Kim,[19] H. O. Kim,[19] J. H. Kim,[17] Y. J. Kim,[6] K. Kinoshita,[3] B. R. Ko,[18] P. Kodyš,[2] S. Korpar,[22,14] P. Križan,[21,14] P. Krokovny,[9] T. Kuhr,[16] T. Kumita,[43] Y.-J. Kwon,[46] S.-H. Kyeong,[46] J. S. Lange,[5] S.-H. Lee,[18] Y. Liu,[29] D. Liventsev,[13] R. Louvot,[20] A. Matyja,[30] S. McOnie,[38] H. Miyata,[32] R. Mizuk,[13] G. B. Mohanty,[39] E. Nakano,[34] M. Nakao,[9] H. Nakazawa,[27] Z. Natkaniec,[30] S. Neubauer,[16] S. Nishida,[9] O. Nitoh,[44] T. Ohshima,[25] S. Okuno,[15] S. L. Olsen,[36,8] P. Pakhlov,[13] C. W. Park,[37] H. Park,[19] H. K. Park,[19] M. Petrič,[14] L. E. Piilonen,[45] M. Röhrken,[16] S. Ryu,[36] H. Sahoo,[8] Y. Sakai,[9] O. Schneider,[20] C. Schwanda,[11] K. Senyo,[25] M. E. Sevior,[24] M. Shapkin,[12] C. P. Shen,[8] J.-G. Shiu,[29] P. Smerkol,[14] E. Solovieva,[13] M. Starič,[14] K. Sumisawa,[9] T. Sumiyoshi,[43] Y. Teramoto,[34] K. Trabelsi,[9] S. Uehara,[9] Y. Unno,[7] S. Uno,[9] P. Urquijo,[24] Y. Usov,[1,33] G. Varner,[8] K. E. Varvell,[38] K. Vervink,[20] C. H. Wang,[28] M.-Z. Wang,[29] P. Wang,[10] Y. Watanabe,[15] R. Wedd,[24] J. Wicht,[9] E. Won,[18] B. D. Yabsley,[38] Y. Yamashita,[31] Z. P. Zhang,[35] V. Zhilich,[1,33] and A. Zupanc[16]

(The Belle Collaboration)

[1]*Budker Institute of Nuclear Physics, Novosibirsk, Russia*
[2]*Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic*
[3]*University of Cincinnati, Cincinnati, Ohio 45221, USA*
[4]*Department of Physics, Fu Jen Catholic University, Taipei, Taiwan*
[5]*Justus-Liebig-Universität Gießen, Gießen, Germany*
[6]*The Graduate University for Advanced Studies, Hayama, Japan*
[7]*Hanyang University, Seoul, South Korea*
[8]*University of Hawaii, Honolulu, Hawaii 96822, USA*
[9]*High Energy Accelerator Research Organization (KEK), Tsukuba, Japan*
[10]*Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China*
[11]*Institute of High Energy Physics, Vienna, Austria*
[12]*Institute of High Energy Physics, Protvino, Russia*
[13]*Institute for Theoretical and Experimental Physics, Moscow, Russia*
[14]*J. Stefan Institute, Ljubljana, Slovenia*
[15]*Kanagawa University, Yokohama, Japan*
[16]*Institut für Experimentelle Kernphysik, Karlsruher Institut für Technologie, Karlsruhe, Germany*
[17]*Korea Institute of Science and Technology Information, Daejeon, South Korea*
[18]*Korea University, Seoul, South Korea*
[19]*Kyungpook National University, Taegu, South Korea*
[20]*École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland*
[21]*Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana, Slovenia*
[22]*University of Maribor, Maribor, Slovenia*
[23]*Max-Planck-Institut für Physik, München, Germany*
[24]*University of Melbourne, School of Physics, Victoria 3010, Australia*
[25]*Nagoya University, Nagoya, Japan*
[26]*Nara Women's University, Nara, Japan*
[27]*National Central University, Chung-li, Taiwan*
[28]*National United University, Miao Li, Taiwan*
[29]*Department of Physics, National Taiwan University, Taipei, Taiwan*
[30]*H. Niewodniczanski Institute of Nuclear Physics, Krakow, Poland*
[31]*Nippon Dental University, Niigata, Japan*
[32]*Niigata University, Niigata, Japan*
[33]*Novosibirsk State University, Novosibirsk, Russia*
[34]*Osaka City University, Osaka, Japan*
[35]*University of Science and Technology of China, Hefei, China*

[36]Seoul National University, Seoul, South Korea
[37]Sungkyunkwan University, Suwon, South Korea
[38]School of Physics, University of Sydney, NSW 2006, Australia
[39]Tata Institute of Fundamental Research, Mumbai, India
[40]Excellence Cluster Universe, Technische Universität München, Garching, Germany
[41]Tohoku Gakuin University, Tagajo, Japan
[42]Department of Physics, University of Tokyo, Tokyo, Japan
[43]Tokyo Metropolitan University, Tokyo, Japan
[44]Tokyo University of Agriculture and Technology, Tokyo, Japan
[45]IPNAS, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061, USA
[46]Yonsei University, Seoul, South Korea

(Received 27 May 2010; published 10 November 2010)

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 $\mathcal{B}(B_s^0 \to D_s^+ D_s^-) = (1.03^{+0.39+0.26}_{-0.32-0.25})\%$, $\mathcal{B}(B_s^0 \to D_s^{*\pm} D_s^{\mp}) = (2.75^{+0.83}_{-0.71} \pm 0.69)\%$, and $\mathcal{B}(B_s^0 \to D_s^{*+} D_s^{*-}) = (3.08^{+1.22+0.85}_{-1.04-0.86})\%$; the sum is $\mathcal{B}[B_s^0 \to D_s^{(*)+} D_s^{(*)-}] = (6.85^{+1.53+1.79}_{-1.30-1.80})\%$. 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-$\bar{B}_s$ mass eigenstates, and φ is a CP-violating weak phase. Taking CP violation to be negligibly small, we obtain $\Delta\Gamma_s/\Gamma_s = 0.147^{+0.036}_{-0.030}(\text{stat})^{+0.042}_{-0.041}(\text{syst})$, where Γ_s is the mean decay width.

DOI: 10.1103/PhysRevLett.105.201802 PACS numbers: 13.25.Hw, 11.30.Er, 12.15.Ff, 14.40.Nd

Decays of B_s mesons help elucidate the weak Cabibbo-Kobayashi-Maskawa structure of the standard model. Because they are not produced in $\Upsilon(4S)$ decays, B_s mesons are much less studied than their B_d^0 and $B^{\pm}$ counterparts. Most B_s data comes from the hadron collider experiments CDF and D0. Recently, another method to study B_s decays has been exploited: that of running an e^+e^- collider at a center-of-mass (c.m.) energy corresponding to the $\Upsilon(5S)$ resonance, which subsequently decays to $B_s^{(*)}\bar{B}_s^{(*)}$ 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 Letter we use $L_{\text{int}} = 23.6$ fb^{-1} of data recorded by Belle at the $\Upsilon(5S)$ ($\sqrt{s} = 10.87$ GeV) to make the first observation ($> 5\sigma$) of $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ [5]. First measurements of the $D_s^{(*)+} D_s^{(*)-}$ final state were made by the ALEPH [6] and D0 [7] Collaborations using inclusive $\phi\phi X$ and $D_s^+ D_s^- X$ samples. CDF [8] measured the single decay $B_s^0 \to D_s^+ D_s^-$. 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-$\bar{B}_s$ CP eigenstates [9]. This parameter equals $\Delta\Gamma_s/\cos\varphi$, where $\Delta\Gamma_s$ is the width difference between the mass eigenstates, and φ is a CP-violating weak phase [10]. Thus the branching fraction gives a constraint in $\Delta\Gamma_s - \varphi$ parameter space. Both of these parameters can receive contributions from new physics [11,12]. The values favored by current measurements [13] differ somewhat from the standard model prediction [11].

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 $e^+e^- \to b\bar{b}$ cross section is $\sigma_{b\bar{b}} = 0.302 \pm 0.014$ 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: $B_s\bar{B}_s$, $B_s\bar{B}_s^*$ or $B_s^*\bar{B}_s$, and $B_s^*\bar{B}_s^*$. In this analysis we use only the last (dominant) mode, for which the fraction is $f_{B_s^*\bar{B}_s^*} = 0.901^{+0.038}_{-0.040}$ [4]. The B_s^* decays via $B_s^* \to B_s\gamma$, and the γ is not reconstructed. Thus the number of $B_s\bar{B}_s$ pairs used in this analysis is $N_{B_s\bar{B}_s} = L_{\text{int}}\sigma_{b\bar{b}} f_s f_{B_s^*\bar{B}_s^*} = (1.24 \pm 0.20) \times 10^6$.

We select $B_s^0 \to D_s^{*+} D_s^{*-}$, $D_s^{*\pm} D_s^{\mp}$, and $D_s^+ D_s^-$ decays in which $D_s^+ \to \phi\pi^+$, $K_S^0 K^+$, $\bar{K}^{*0} K^+$, $\phi\rho^+$, $K_S^0 K^{*+}$, and $\bar{K}^{*0} K^{*+}$. 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 $\sim 90\%$ efficient and has a $\pi^{\pm}$ misidentification rate of $\sim 10\%$. Tracks having kaon likelihood <0.60 are identified as $\pi^{\pm}$. Neutral K_S^0 candidates are reconstructed from $\pi^+\pi^-$ pairs having an invariant mass within 10 MeV/c^2 of the K_S^0 mass [16] and satisfying loose requirements on the decay vertex position [17]. The momentum of tracks (except the $\pi^{\pm}$ from K_S^0 decay) must be >100 MeV/c.

Neutral π^0 candidates are reconstructed from $\gamma\gamma$ pairs having an invariant mass within 15 MeV/c^2 of the π^0 mass. The photons must have a laboratory energy greater than 100 MeV. Neutral $\bar{K}^{*0}$ (charged K^{*+}) candidates are reconstructed from a K^- (K_S^0) and π^+ having an invariant mass within 50 MeV/c^2 of $M_{K^{*0}}(M_{K^{*+}})$. Neutral ϕ (charged ρ^+) 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 (2.5–3.2σ) for the three final states containing K^* candidates, 20 MeV/c^2 (1.7σ) for $\phi\rho^+$, and 15 MeV/c^2 ($\gtrsim 4.0\sigma$) for the remaining two modes. For the three vector-pseudoscalar final states, we impose a loose requirement on the helicity angle $\theta_{\rm hel}$, 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 $|\cos\theta_{\rm hel}| > 0.20$, which retains 99% of signal decays and rejects 18% of remaining background.

To reconstruct $D_s^{*+} \to D_s^+\gamma$ decays, we pair D_s^+ candidates with photon candidates and require that the mass difference $M_{\tilde{D}_s^+\gamma} - M_{\tilde{D}_s^+}$ 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 $\mathcal{R}$ 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 c.m. system be greater than 50 MeV, and that the energy deposited in the central 3×3 array of cells of the electromagnetic calorimeter cluster be at least 85% of the energy deposited in the central 5×5 array of cells.

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

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

$$\chi^2 = \frac{1}{(2+N)}\left\{\sum_{\#D_s}[(\tilde{M}_{D_s} - M_{D_s})/\sigma_M]^2 + \sum_{\#D_s^*}[(\widetilde{\Delta M} - \Delta M)/\sigma_{\Delta M}]^2\right\}, \quad (1)$$

where $\Delta M = M_{D_s^*} - M_{D_s}$; $\tilde{M}_{D_s}$ and $\widetilde{\Delta M}$ are reconstructed quantities; σ_M and $\sigma_{\Delta M}$ are the uncertainties on $\tilde{M}_{D_s}$ and $\widetilde{\Delta M}$; and the summations run over the two D_s^+ daughters and possible D_s^{*+} 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 $D_s^+D_s^-$, $D_s^{*\pm}D_s^{\mp}$, and $D_s^{*+}D_s^{*-}$ final states.

We reject background from $e^+e^- \to q\bar{q}(q = u, d, s, c)$ continuum events based on event topology: $q\bar{q}$ events tend to be collimated, while $B_{(s)}\bar{B}_{(s)}$ 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 $\mathcal{L}_s$ ($\mathcal{L}_{q\bar{q}}$) for an event assuming the event is signal ($q\bar{q}$ background). We form the ratio $\mathcal{R} = \mathcal{L}_s/(\mathcal{L}_s + \mathcal{L}_{q\bar{q}})$ and require $\mathcal{R} > 0.20$. This selection is 95% efficient for signal decays and removes >80% of $q\bar{q}$ background.

The remaining background consists of $\Upsilon(5S) \to B_s^{(*)}\bar{B}_s^{(*)} \to D_s^+X$, $\Upsilon(5S) \to BBX$ (where $b\bar{b}$ hadronizes into B^0, $\bar{B}^0$, or $B^\pm$), and $B_s \to D_{sJ}^\pm(2317)D_s^{(*)}$, $B_s \to D_{sJ}^\pm(2460)D_s^{(*)}$, and $B_s \to D_s^\pm D_s^\mp\pi^0$ decays. The last three processes peak at negative values of ΔE, and their yields are estimated using analogous $B_d \to D_{sJ}^\pm D^{(*)}$ branching fractions. The total yields for all backgrounds within an $M_{\rm bc} - \Delta E$ signal region spanning 3σ in $(M_{\rm bc}, \Delta E)$ resolution are 0.25 ± 0.03, 0.25 ± 0.06, and 0.15 ± 0.13 events, respectively, for $B_s \to D_s^+D_s^-$, $D_s^{*\pm}D_s^{\mp}$, and $D_s^{*+}D_s^{*-}$ decays. To check our background estimates, we count events in the sideband region $M_{\rm bc} < 5.375$ 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 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 $M_{\rm bc} - \Delta E$ distributions. For each sample, we include probability density functions (PDFs) for signal and $q\bar{q}$, $B_s^{(*)}\bar{B}_s^{(*)} \to D_s^+X$, and $\Upsilon(5S) \to BBX$ backgrounds. As these backgrounds have similar $M_{\rm bc}$, ΔE shapes, we use a single PDF for them, taken to be an ARGUS function [19] for $M_{\rm bc}$ and a third-order Chebyshev polynomial for ΔE. 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 nonsignal track or photon is included in place of a true daughter track or photon; and "cross-feed" (CF) decays in which a $D_s^{*\pm}D_s^{\mp}$ or $D_s^{*+}D_s^{*-}$ is reconstructed as $D_s^+D_s^-$ or $D_s^{*\pm}D_s^{\mp}$, respectively, or else a $D_s^+D_s^-$ or $D_s^{*\pm}D_s^{\mp}$ is reconstructed as $D_s^{*\pm}D_s^{\mp}$ or $D_s^{*+}D_s^{*-}$. In the former (latter) case the signal decay has lost (gained) a photon, and ΔE is typically shifted lower (higher) by 100–150 MeV. The PDF for CR events is modeled with a single Gaussian for $M_{\rm bc}$ and a double Gaussian with common mean for ΔE. The means and widths are taken from MC simulation and calibrated using

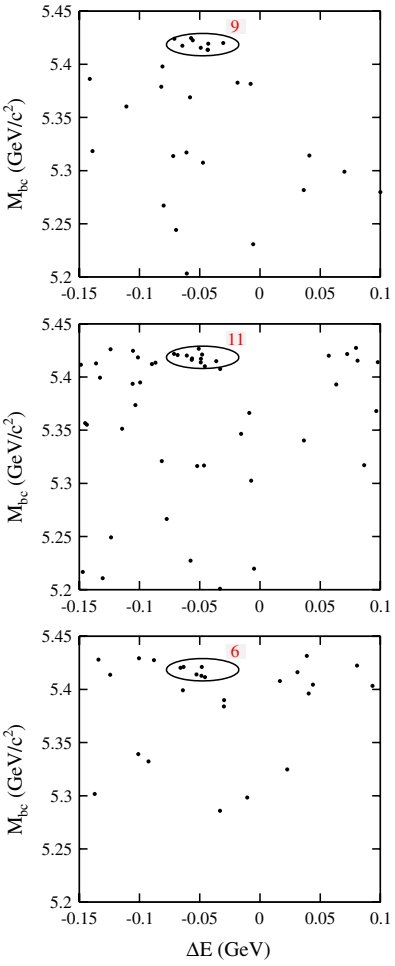

FIG. 1 (color online). M_{bc} vs ΔE scatter plots. The signal ellipses correspond to 3σ in resolution for $\Upsilon(5S) \to B_s^* \bar{B}_s^*$ decays; the number of candidates within the ellipses is listed. The top, middle, and bottom plots correspond to $B_s^0 \to D_s^+ D_s^-$, $B_s^0 \to D_s^{*\pm} D_s^{\mp}$, and $B_s^0 \to D_s^{*+} D_s^{*-}$, respectively.

$B_s^0 \to D_s^{(*)-} \pi^+$ and $B^0 \to D_s^{(*)+} D^-$ control samples. The PDFs for WC and CF events are modeled from the MC simulation using nonparametric PDFs with Kernel estimation [20]. 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 γ 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 ($D_s^+ D_s^-$, $D_s^{*\pm} D_s^{\mp}$, and $D_s^{*+} D_s^{*-}$) 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 $\mathcal{B}_i = Y_i/(\varepsilon^i_{\rm MC} N_{B_s \bar{B}_s} 2)$, where Y_i is the fitted CR yield, and $\varepsilon^i_{\rm MC}$ is the MC efficiency with intermediate branching fractions [16] included. The efficiencies $\varepsilon^i_{\rm MC}$ 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

TABLE I. Signal yields (Y), efficiencies including intermediate branching fractions (ε), branching fractions ($\mathcal{B}$), 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 ($\times 10^{-4}$)	$\mathcal{B}$ (%)	S
$D_s^+ D_s^-$	$8.5^{+3.2}_{-2.6}$	3.31	$1.03^{+0.39+0.15}_{-0.32-0.13} \pm 0.21$	6.2
$D_s^{*\pm} D_s^{\mp}$	$9.2^{+2.8}_{-2.4}$	1.35	$2.75^{+0.83}_{-0.71} \pm 0.40 \pm 0.56$	6.6
$D_s^{*+} D_s^{*-}$	$4.9^{+1.9}_{-1.7}$	0.643	$3.08^{+1.22+0.57}_{-1.04-0.58} \pm 0.63$	3.1
Sum	$22.6^{+4.7}_{-3.9}$		$6.85^{+1.53}_{-1.30} \pm 1.11^{+1.40}_{-1.41}$	

$\sqrt{-2\ln(\mathcal{L}_0/\mathcal{L}_{\rm max})}$, where $\mathcal{L}_0$ and $\mathcal{L}_{\rm max}$ 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.

The systematic errors are listed in Table II. The error due to PDF shapes is evaluated by varying shape parameters by $\pm 1\sigma$ for backgrounds, and by trying different parametrizations for the WC and CF components. The systematic

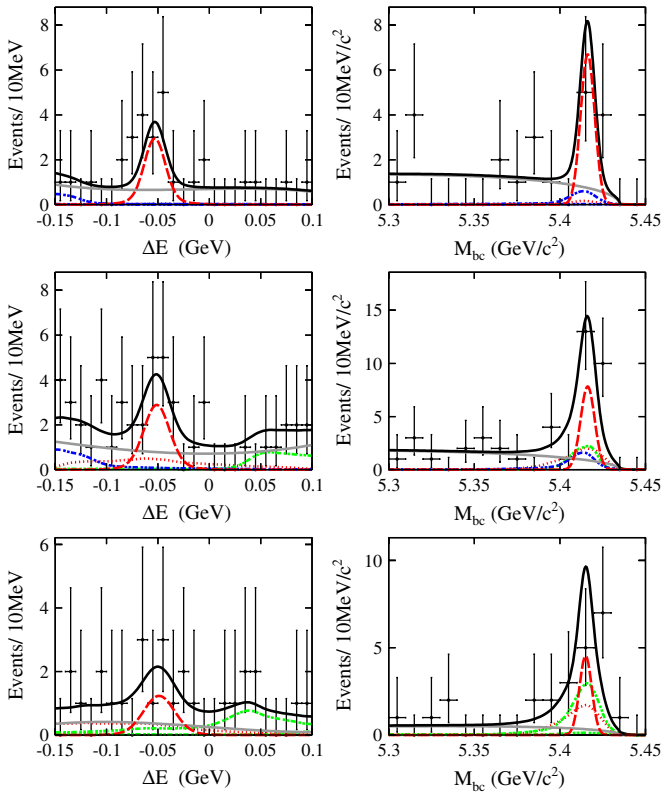

FIG. 2 (color online). M_{bc} and ΔE projections of the fit result. The rows correspond to $B_s^0 \to D_s^+ D_s^-$ (top plots), $B_s^0 \to D_s^{*\pm} D_s^{\mp}$ (middle plots), and $B_s^0 \to D_s^{*+} D_s^{*-}$ (bottom plots). The (red) dashed (dotted) curves show RC (WC) signal, the (green and blue) dash-dotted curves show CF signal, the grey solid curve shows background, and the black solid curves show the total.

TABLE II. Systematic errors (%). The first twelve sources affect the signal yield and thus the signal significance.

Source	$D_s^+ D_s^-$		$D_s^{*\pm} D_s^{\mp}$		$D_s^{*+} D_s^{*-}$	
	$+\sigma$	$-\sigma$	$+\sigma$	$-\sigma$	$+\sigma$	$-\sigma$
CR PDF Shape	0.8	0.8	0.3	0.3	0.5	0.4
Background PDF	1.1	1.3	1.9	2.0	3.0	6.1
WC + CF PDF	0.3	0.3	1.5	1.5	4.4	4.4
WC/CF Fractions	0.2	0.2	5.0	5.0	8.7	8.7
$\mathcal{R}$ Requirement ($q\bar{q}$ suppr.)	1.8	1.8	1.8	1.8	1.8	1.8
Best Candidate Selection	6.9	0.0	2.2	0.0	2.2	0.0
$K^{\pm}$ Identification	10.1	10.1	10.6	10.6	10.9	10.9
K_S^0 Reconstruction	2.1	2.1	2.1	2.1	2.2	2.2
π^0 Reconstruction	1.1	1.1	1.1	1.1	1.0	1.0
γ	-	-	3.8	3.8	7.6	7.6
Tracking	6.2	6.2	6.2	6.2	6.2	6.2
Polarization	0.2	0.0	0.8	0.5	0.7	0.3
MC Statistics for ε	1.1	1.1	0.9	0.8	1.0	1.0
$D_s^{(*)}$ BF's	12.4	12.4	12.4	12.4	12.5	12.5
Luminosity			± 1.3			
$\sigma_{\Upsilon(5S)}$			± 4.6			
f_s			± 15			
$f_{B_s^* \bar{B}_s^*}$			$+4.2 - 4.4$			
Total	24.9	24.0	25.1	25.1	27.5	28.0

error for the fixed CF-down fractions is evaluated by fitting a $B_s^0 \to D_s^- \pi^+$ control sample and comparing the fraction of $B_s^0 \to D_s^{*-} \pi^+$ contamination with that predicted by the 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 $\mathcal{B}$ 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 $\pm 20\%$. The uncertainty due to $K^{\pm}$ identification depends on momentum but is $\sim 2.5\%$ 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 $B_s^0 \to D_s^{*+} D_s^{*-}$ 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 value for the analogous spectator decay $B_d^0 \to D_s^{*+} D^{*-}$: 0.52 ± 0.05 [16]. The systematic error is taken as the change in $\mathcal{B}$ when f_L is varied by twice the error on the world average value. Significant uncertainties arise from D_s^+ branching fractions, $\sigma_{\Upsilon(5S)}$, f_s, and $f_{B_s^* \bar{B}_s^*}$, which are external factors that should be measured more precisely in the future. We list separately the systematic error due to these factors in Table I.

In the limit $m_{c,b} \to \infty$ while $(m_b - 2m_c) \to 0$, the $b \to c\bar{c}s$ process saturates the decay width [22]. If also the number of colors $N_c \to \infty$, then $B_s^0 \to D_s^{*+} D_s^{*-}$, $D_s^{*\pm} D_s^{\mp}$ (along with $D_s^+ D_s^-$) are CP even ($+$), and $\Gamma[B_s^0(CP+) \to D_s^{(*)} D_s^{(*)}]$ saturates $\Delta\Gamma_s^{CP}$ [9]. This gives the relationship $2\mathcal{B}[B_s^0 \to D_s^{(*+)} D_s^{(*-)}] = (\Delta\Gamma_s^{CP}/2)[(1 + \cos\varphi)/\Gamma_L + (1 - \cos\varphi)/\Gamma_H]$, where $\Gamma_{L,H}$ are the decay widths of the light and heavy mass eigenstates [10]. Substituting $\Gamma_{L,H} = \Gamma \pm \Delta\Gamma_s/2$ and $\Delta\Gamma_s^{CP} = \Delta\Gamma_s/\cos\varphi$ [10] allows one to use the branching fraction $\mathcal{B}$ to constrain $\Delta\Gamma_s$ and φ. If CP violation is negligible, then $\cos\varphi \simeq 1$ and the above expression can be inverted to give $\Delta\Gamma_s/\Gamma_s = 2\mathcal{B}/(1 - \mathcal{B})$. Inserting $\mathcal{B}$ from Table I yields

$$\frac{\Delta\Gamma_s}{\Gamma_s} = 0.147^{+0.036+0.042}_{-0.030-0.041}, \tag{2}$$

where the first error is statistical and the second is systematic. This result is 1.3σ higher than that of Ref. [7] but consistent with the theory prediction [11]. 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 or 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 $B_s^0 \to D_s^{(*)+} D_s^{(*)-}$ using e^+e^- data taken at the $\Upsilon(5S)$ resonance. Our results constitute the first observation of $B^0 \to D_s^{*\pm} D_s^{\mp}$ (6.6σ significance) and provide the first evidence for $B_s^0 \to D_s^{*+} D_s^{*-}$ (3.1σ significance). We use these measurements to determine the $B_s^0 - \bar{B}_s$ 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 Collaboration), Phys. Rev. D **75**, 012002 (2007); M. Artuso *et al.* (CLEO Collaboration), Phys. Rev. Lett. **95**, 261801 (2005).
[2] A. Drutskoy *et al.* (Belle Collaboration), Phys. Rev. Lett. **98**, 052001 (2007).
[3] A. Drutskoy *et al.* (Belle Collaboration), Phys. Rev. D **76**, 012002 (2007); J. Wicht *et al.* (Belle Collaboration), Phys. Rev. Lett. **100**, 121801 (2008).
[4] R. Louvot *et al.* (Belle Collaboration), Phys. Rev. Lett. **102**, 021801 (2009).
[5] Charge-conjugate modes are implicitly included.
[6] R. Barate *et al.* (ALEPH Collaboration), Phys. Lett. B **486**, 286 (2000).
[7] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **102**, 091801 (2009).
[8] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **100**, 021803 (2008).
[9] R. Aleksan *et al.*, Phys. Lett. B **316**, 567 (1993).
[10] I. Dunietz, R. Fleischer, and U. Nierste, Phys. Rev. D **63**, 114015 (2001); I. Dunietz, Phys. Rev. D **52**, 3048 (1995).
[11] A. Lenz and U. Nierste, J. High Energy Phys. 06 (2007), 072.
[12] See for example: A. J. Buras *et al.*, Phys. Rev. Lett. **105**, 131601 (2010); Z. Ligeti *et al.*, arXiv:1006.0432.
[13] CDF Collaboration, Public Note 10206 2010, (unpublished); D0 Collaboration, Note 6098-CONF 2010 (unpublished); CDF Collaboration, Public Note 9787 2009 (unpublished).
[14] A. Abashian *et al.* (Belle Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 117 (2002).
[15] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods Phys. Res., Sect. 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 Collaboration), 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 Collaboration), Phys. Rev. Lett. **91**, 261801 (2003).
[19] H. Albrecht *et al.* (ARGUS Collaboration), Phys. Lett. B **241**, 278 (1990).
[20] K. Cranmer, Comput. Phys. Commun. **136**, 198 (2001).
[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).