Letter

Search for the decay $B_s^0 \to \pi^0\pi^0$ at Belle

J. Borah, B. Bhuyan, I. Adachi, H. Aihara, D. M. Asner, V. Aulchenko, T. Aushev, R. Ayad, V. Babu, S. Bahinipati, Sw. Banerjee, P. Behera, K. Belous, J. Bennett, M. Bessner, V. Bhardwaj, T. Bilka, D. Biswas, D. Bodrov, A. Bozek, M. Bračko, P. Branchini, T. E. Browder, A. Budano, M. Campajola, D. Červenkov, M. C. Chang, V. Chekelian, C. Chen, B. G. Cheon, K. Chilikin, H. E. Cho, K. Cho, S. J. Cho, S. K. Choi, Y. Choi, S. Choudhury, D. Cinabro, S. Das, G. De Nardo, G. De Pietro, R. Dhamija, F. Di Capua, J. Dingfelder, Z. Doležal, T. V. Dong, D. Epifanov, D. Ferlewicz, A. Frey, B. G. Fulsom, R. Garg, V. Gaur, A. Garmash, A. Giri, P. Goldenzweig, E. Graziani, T. Gu, Y. Guan, K. Gudkova, C. Hadjivasiliou, S. Halder, T. Hara, K. Hayasaka, H. Hayashii, M. T. Hedges, D. Herrmann, M. Hernández Villanueva, W. S. Hou, C. L. Hsu, T. Iijima, K. Inami, G. Inguglia, N. Ipsita, A. Ishikawa, R. Itoh, M. Iwasaki, W. W. Jacobs, E. J. Jang, Q. P. Ji, Y. Jin, K. K. Joo, D. Kalita, A. B. Kaliyar, K. H. Kang, T. Kawasaki, H. Kichimi, C. Kiesling, C. H. Kim, D. Y. Kim, K. H. Kim, Y. K. Kim, K. Kinoshita, P. Kodyš, T. Konno, A. Korobov, S. Korpar, P. Križan, P. Krokovny, T. Kuhr, M. Kumar, R. Kumar, K. Kumara, Y. J. Kwon, K. Lalwani, T. Lam, J. S. Lange, M. Laurenza, S. C. Lee, D. Levit, L. K. Li, Y. Li, L. Li Gioi, J. Libby, Y. R. Lin, D. Liventsev, T. Luo, M. Masuda, D. Matvienko, S. K. Maurya, F. Meier, M. Merola, F. Metzner, K. Miyabayashi, R. Mizuk, G. B. Mohanty, R. Mussa, I. Nakamura, M. Nakao, D. Narwal, Z. Natkaniec, A. Natochii, L. Nayak, N. K. Nisar, S. Nishida, K. Ogawa, S. Ogawa, H. Ono, Y. Onuki, P. Oskin, P. Pakhlov, G. Pakhlova, T. Pang, S. Pardi, J. Park, S. H. Park, A. Passeri, S. Patra, S. Paul, T. K. Pedlar, R. Pestotnik, L. E. Piilonen, T. Podobnik, E. Prencipe, M. T. Prim, M. Röhrken, A. Rostomyan, G. Russo, Y. Sakai, S. Sandilya, L. Santelj, V. Savinov, G. Schnell, C. Schwanda, A. J. Schwartz, Y. Seino, K. Senyo, M. E. Sevior, W. Shan, M. Shapkin, C. Sharma, C. P. Shen, J. G. Shiu, F. Simon, J. B. Singh, A. Sokolov, E. Solovieva, M. Starič, Z. S. Stottler, M. Sumihama, T. Sumiyoshi, M. Takizawa, U. Tamponi, K. Tanida, F. Tenchini, K. Trabelsi, M. Uchida, T. Uglov, Y. Unno, S. Uno, P. Urquijo, S. E. Vahsen, R. van Tonder, G. Varner, K. E. Varvell, A. Vinokurova, A. Vossen, D. Wang, M. Z. Wang, M. Watanabe, E. Won, X. Xu, B. D. Yabsley, W. Yan, S. B. Yang, J. Yelton, J. H. Yin, Y. Yook, C. Z. Yuan, L. Yuan, Y. Yusa, Z. P. Zhang, V. Zhilich, and V. Zhukova

(Belle Collaboration)

(Received 23 January 2023; accepted 2 March 2023; published 28 March 2023)

We report the results of the first search for the decay $B_s^0 \to \pi^0\pi^0$ using 121.4 fb^{-1} of data collected at the $\Upsilon(5S)$ resonance with the Belle detector at the KEKB asymmetric-energy e^+e^- collider. We observe no signal and set a 90% confidence level upper limit of 7.7×10^{-6} on the $B_s^0 \to \pi^0\pi^0$ decay branching fraction.

DOI: 10.1103/PhysRevD.107.L051101

The study of heavy-flavored hadrons decaying to hadronic final states provides an important input for understanding the interplay between strong and weak interactions. These types of decays involving weak annihilation amplitudes can be a promising place to look for disagreement between theoretical predictions and experimental observations. These decays are highly suppressed and often neglected in theoretical calculations. However, the inclusion of rescattering effects into the theoretical framework naturally enhances their contribution [1]. Recently it was observed that the predicted branching fraction for the decay $B_s^0 \to \pi^+\pi^-$, which involves topological annihilation diagrams, was substantially smaller than its measured value by the LHCb experiment [2]. This discrepancy between theoretical prediction and experimental measurement may require some models of strong interaction processes to be revisited [3]. In these aspects, searches for decays involving weak annihilation amplitudes become important and necessary.

Within the standard model (SM), the decay $B_s^0 \to \pi^0\pi^0$ proceeds via the W-exchange and "penguin" annihilation amplitudes, as shown in Fig. 1. Theoretical calculations based on the flavor diagram approach [4],

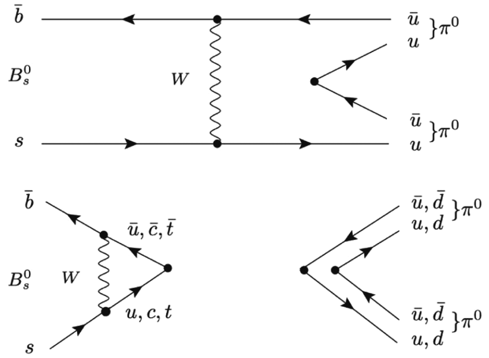

FIG. 1. W-exchange (top) and "penguin" annihilation (bottom) Feynman diagrams for $B_s^0 \to \pi^0\pi^0$.

perturbative quantum chromodynamics [5], and QCD factorization [6] predict the branching fraction ($\mathcal{B}$) to be $(0.40 \pm 0.27) \times 10^{-6}$, $(0.28 \pm 0.09) \times 10^{-6}$, and $(0.13 \pm 0.05) \times 10^{-6}$, respectively. The only measurement for this decay was made by the L3 experiment in 1995, which reported an upper limit (UL) of $\mathcal{B} < 2.1 \times 10^{-4}$ at 90% confidence level (CL) [7]. The search for the decay $B_s^0 \to \pi^0\pi^0$ [8] described in this paper is based on a data sample of 121.4 fb^{-1} collected at the $\Upsilon(5S)$ resonance using the Belle detector.

The Belle detector at the KEKB [9] asymmetric-energy e^+e^- collider is a large-solid-angle magnetic spectrometer that consists of a silicon vertex detector, a 50-layer central drift chamber, an array of aerogel threshold Cherenkov counters, a barrel-like arrangement of time-of-flight scintillation counters, and a CsI(Tl) crystal-based electromagnetic calorimeter (ECL) located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux return outside the coil is instrumented to detect K_L^0 mesons and identify muons. A detailed description of the Belle detector can be found elsewhere [10,11]. The analysis relies on the ECL component of the detector for the reconstruction of the photons in the $B_s^0 \to \pi^0\pi^0$ decay final state.

The production cross section of the $\Upsilon(5S)$ resonance at the e^+e^- center of mass (c.m.) energy of 10.86 GeV is $\sigma_{b\bar{b}}^{\Upsilon(5S)} = (0.340 \pm 0.016)$ nb [12], and the fraction of $b\bar{b}$ events giving rise to B_s^0 production modes, $B_s^{(*)0}\bar{B}_s^{(*)0}$, is measured to be $f_s = (0.201 \pm 0.031)$ [13]. There are three kinematically allowed modes of production of B_s^0 mesons: $B_s^{*0}\bar{B}_s^{*0}$, $B_s^0\bar{B}_s^{*0}$ or $B_s^{*0}\bar{B}_s^0$, and $B_s^0\bar{B}_s^0$. The production fractions from the former two are $(87.0 \pm 1.7)\%$ and $(7.3 \pm 1.4)\%$, respectively [12], while the remaining fraction is from the $B_s^0\bar{B}_s^0$ mode. The B_s^{*0} decays to B_s^0 by radiating a low-energy photon that is usually not identified due to its poor reconstruction efficiency. The number of events with $B_s^0\bar{B}_s^0$ is, therefore, estimated to be $N_{B_s^0\bar{B}_s^0} = 121.4 \text{ fb}^{-1} \cdot \sigma_{b\bar{b}}^{\Upsilon(5S)} \cdot f_s = (8.30 \pm 1.34) \times 10^6$.

We employ a "blind" analysis procedure to leave out the experimenter's biases and develop our analysis strategy with Monte-Carlo (MC) samples. In a "blind" analysis, the signal region is kept hidden until the selection criteria are finalized. The signal MC samples are generated with EvtGen [14] and simulated with GEANT3 [15] to model all possible detector effects. Background studies are performed with MC samples 6 times larger than the integrated luminosity of data. The analysis procedure is validated with a control sample of $B_d^0 \to \pi^0\pi^0$ decays produced at the $\Upsilon(4S)$ resonance, which closely resembles the signal.

We reconstruct $B_s^0 \to \pi^0\pi^0$ with $\pi^0 \to \gamma\gamma$. Photon candidates are reconstructed from ECL clusters that do not match any charged track and have energy greater than 50 (100) MeV in the ECL's barrel (end caps) region. The forward end cap, barrel, and backward end cap regions of the ECL are given by $12° < \theta < 31.4°$, $32.2° < \theta < 128.7°$, and $131.5° < \theta < 157.2°$, respectively, where θ is the polar angle in the laboratory frame with respect to the detector axis, in the direction opposite to the e^+ beam. To remove the off-time (radiative) Bhabha and $e^+e^- \to \gamma\gamma$ events, a timing criterion based on the beam collision time is applied, which is determined at the trigger level for each candidate event. The invariant mass of the two-photon combination must lie in the range of $118 \text{ MeV}/c^2 < m(\gamma\gamma) < 152 \text{ MeV}/c^2$, corresponding to ± 2.4 standard deviations (σ) of the invariant mass resolution around the nominal π^0 mass [13]. A mass-constrained fit is subsequently performed to improve the π^0 momentum resolution.

To further select the B_s^0 candidates, we apply selection criteria on their beam-energy-constrained mass $M_{\rm bc} = \sqrt{(E_{\rm beam})^2 - |\vec{p}_{\rm reco}|^2c^2}/c^2$ and the energy difference $\Delta E' = E_{\rm reco} - E_{\rm beam} + M_{\rm bc}c^2 - m_{B_s^0}c^2$ in the e^+e^- c.m. frame, where $E_{\rm beam}$ is the beam energy, $\vec{p}_{\rm reco}$ and $E_{\rm reco}$ are the momentum and energy, respectively, of the reconstructed B_s^0 candidate. The world average value is used for the mass of the B_s^0 meson, $m_{B_s^0}$ [13]. In the $\Delta E'$ distribution, all the production channels ($B_s^{*0}\bar{B}_s^{*0}$, $B_s^0\bar{B}_s^{*0}$ or $B_s^{*0}\bar{B}_s^0$, and $B_s^0\bar{B}_s^0$) peak at zero, while in the $\Delta E = E_{\rm reco} - E_{\rm beam}$ variable, the peaks of $B_s^{*0}\bar{B}_s^{*0}$ and $B_s^0\bar{B}_s^{*0}$ or $B_s^{*0}\bar{B}_s^0$ are shifted to negative values by $(m_{B_s^{*0}}c^2 - m_{B_s^0}c^2)/2$ and $m_{B_s^{*0}}c^2 - m_{B_s^0}c^2$, respectively, where, $m_{B_s^{*0}}$ represents the mass of the B_s^{*0} meson. Therefore, to correct for these shifts in the ΔE variable, the $\Delta E'$ variable is used for this analysis. A B_s^0 candidate is retained for further analysis only if it satisfies the requirement that $5.300 \text{ GeV}/c^2 < M_{\rm bc} < 5.434 \text{ GeV}/c^2$ and $-0.60 \text{ GeV} < \Delta E' < 0.15$ GeV.

The backgrounds near the $\Upsilon(5S)$ resonance which can affect the analysis are: continuum ($e^+e^- \to q\bar{q}, q = u,d,s,c$), $B_s^{(*)}\bar{B}_s^{(*)}$ decays (referred as *bsbs*) and $B^*\bar{B}^*$, $B^*\bar{B}$, $B\bar{B}$, $B^*\bar{B}^*\pi$, $B^*\bar{B}\pi$, $B\bar{B}\pi$, and $B\bar{B}\pi\pi$ ($B = B^0, B^+$) decays (referred as *non-bsbs*). Additional background MC studies on the peaking background of the types $B_s^0 \to \rho^+\rho^-$ and $B_s^0 \to K_s^0\pi^0$ show that their contributions are negligible. We also find no *bsbs* and *non-bsbs* background after applying all of the aforementioned selection criteria. Background MC studies, therefore, reveal the dominance of continuum background over the other types of background. Their suppression requires topological variables, which classify the signal and the continuum background based on their event shape variables in the e^+e^- c.m. frame.

In signal events, B_s^0 pairs are produced with small momenta, and the distribution of their decay products tends to be spherical. In contrast, the quark pairs of the continuum background are produced with a significant amount of momentum; therefore, their decay product distribution has a jetlike topology. A neural network algorithm (NN) [16] is employed to suppress the continuum background. The input of the NN includes sixteen modified Fox-Wolfram moments [17], and thrust axis direction, $\cos\theta_T$ (see Sec. 9.3 in [18]) to provide additional discrimination between the signal and the continuum background. These modified Fox-Wolfram moments are calculated based on three categories of events: the B_s^0 candidate daughters, the rest of events, and the missing momentum of an event. The angle θ_T is defined as the angle between the thrust axis of the signal B_s^0 candidate and the thrust axis of the remainder of the events. The NN was trained on MC samples with consistency checks to ensure no overtraining.

The choice of the selection criterion on the output of the NN, $\mathcal{C}_{\rm NN}$, is determined based on a Punzi's figure-of-merit optimization [19], where the significance level is set to three standard deviations. The $\mathcal{C}_{\rm NN}$ distributions for the continuum background and the signal lies in the range of $[-1,+1]$, where the continuum backgrounds peak at -1 and the signal candidates at $+1$. We require $\mathcal{C}_{\rm NN}$ to be greater than 0.90 for this analysis. This condition removes 99% of the continuum background with a signal loss of 53%. To facilitate the data modeling, $\mathcal{C}_{\rm NN}$ was transformed to another variable, $\mathcal{C}'_{\rm NN}$ using the following formula:

$$\mathcal{C}'_{\rm NN} = \log\left(\frac{\mathcal{C}_{\rm NN} - \mathcal{C}_{\rm NN(min)}}{\mathcal{C}_{\rm NN(max)} - \mathcal{C}_{\rm NN}}\right), \quad (1)$$

where $\mathcal{C}_{\rm NN(min)} = 0.90$ and $\mathcal{C}_{\rm NN(max)}$ is the maximum value of $\mathcal{C}_{\rm NN}$ obtained from the NN distribution.

After applying the selection criteria described above, 10.3% of signal MC events have more than one candidate. Most of the multiple candidates are due to a signal photon reconstructed as two ECL clusters. We select the best candidate based on the sum of the χ^2 of the mass-constrained fits to the two π^0s. The fraction of misreconstructed events after applying all the selection criteria is found to be negligible; hence, they are not treated separately. The overall signal reconstruction efficiency in this analysis is $(12.69 \pm 0.05)\%$.

To extract the signal yield, we perform a three-dimensional (3D) unbinned extended maximum likelihood (ML) fit to $M_{\rm bc}$, $\Delta E'$, and $\mathcal{C}'_{\rm NN}$. The likelihood function is defined as

$$\mathcal{L}_{\rm fit} = e^{-\sum_j n_j} \prod_i^N \left(\sum_j n_j P_j((M_{\rm bc})^i, (\Delta E')^i, (\mathcal{C}'_{\rm NN})^i)\right), \quad (2)$$

where $P_j(M_{\rm bc}, \Delta E', \mathcal{C}'_{\rm NN})$ is the probability distribution function (PDF) of the signal or background component (specified by index j), n_j is the yield of this component, i represents the event index, and N is the total number of events in the sample.

The linear correlation coefficients among $M_{\rm bc}$, $\Delta E'$, and $\mathcal{C}'_{\rm NN}$ are found to be below 3% in the signal region. Consequently, each of the 3D PDFs describing the signal and background contributions are assumed to factorize as $P_j \equiv P_j(M_{\rm bc})P_j(\Delta E')P_j(\mathcal{C}'_{\rm NN})$. These factorized PDFs are modeled using large signal and background MC samples. The signal $M_{\rm bc}$ PDF consists of three PDFs corresponding to the three B_s^0 production channels. Each of them is again separately modeled from large MC samples. They are then combined according to their production fractions [12] to produce the final signal PDF for the $M_{\rm bc}$ variable. The PDF used for parametrizing $B_s^0\bar{B}_s^0$ is a sum of two Gaussian distributions with a common mean, while each of $B_s^0\bar{B}_s^{*0}$ or $B^{*0}\bar{B}_s^0$, and $B_s^{*0}\bar{B}_s^{*0}$, are parametrized using a sum of a Gaussian function and an empirical PDF shape known as the Crystal Ball function [20]. The signal $\Delta E'$ variable, for all three B_s^0 channels, is modeled using the Crystal Ball function, which is modified for this analysis to include the asymmetric nature of the distribution about the mean position. The output from the NN is parametrized using a Gaussian and an asymmetric (bifurcated) Gaussian PDF for the signal $\mathcal{C}'_{\rm NN}$ variable. Unlike the signal PDF parameters for the $M_{\rm bc}$ variable, which is different for the three B_s^0 sources, $\Delta E'$ and $\mathcal{C}'_{\rm NN}$ variables take the same parameter values for the three B_s^0 production channels. The continuum background distribution of the $M_{\rm bc}$ variable is modeled through an empirically determined parametrized background shape referred to as the ARGUS function [21]. The continuum background is parametrized using a first-order Chebychev polynomial and a sum of two Gaussian distributions for the $\Delta E'$ and $\mathcal{C}'_{\rm NN}$ variables, respectively. All of the signal parameters and the background ARGUS end point are fixed to their best fit values obtained from 1D fits to the MC simulated events. In contrast, all other background parameter values and the signal and background yields are floated. The PDFs used for modeling the signal and continuum background are listed in Table I.

To validate our analysis, we use the Belle data sample collected at the Υ(4S) to reconstruct the decay $B_d^0 \to \pi^0\pi^0$ by applying similar event selection criteria. The results of

TABLE I. PDFs used to model the $M_{\rm bc}$, $\Delta E'$, and $\mathcal{C}'_{\rm NN}$ distributions. The notations G, BG, CB, ACB, *CP*, and A correspond to Gaussian, bifurcated Gaussian, Crystal Ball, asymmetric Crystal Ball, Chebyshev polynomial, and ARGUS functions, respectively.

Fit component	$M_{\rm bc}$	ΔE	$\mathcal{C}'_{\rm NN}$
Signal	G + G ($B_s^0\bar{B}_s^0$) G + CB ($B_s^0\bar{B}_s^{*0}$) G + CB ($B_s^{*0}\bar{B}_s^{*0}$)	ACB	G + BG
Continuum	A	CP	G + G

TABLE II. Summary of systematic uncertainties.

Source	Value (%)
Fit bias	−3.3
Fixed PDF parametrization	$^{+3.5}_{-5.2}$
Fractions of $B_s^{*0}\bar{B}_s^{(*0)}$	$^{+5.2}_{-3.5}$
Reconstruction efficiency, $\epsilon_{\rm rec}$	±0.4
$\mathcal{C}'_{\rm NN}$ requirement	±3.0
$\pi^0 \to \gamma\gamma$ selection efficiency	±4.4
$\mathcal{B}(\pi^0 \to \gamma\gamma)$	±0.07
$b\bar{b}$ cross section, $\sigma_{b\bar{b}}$	±4.7
f_s	±15.4
Total	$^{+18.1}_{-18.4}$

the fit to $\Upsilon(4S)$ data are shown in Fig. 2, where each fit projection is plotted after additional selection criteria are applied as described in the caption. We calculate the branching fraction, $\mathcal{B}(B_d^0 \to \pi^0\pi^0) = (1.18 \pm 0.21) \times 10^{-6}$ (where only the statistical uncertainty is shown), which is in good agreement with our previous result [22].

The systematic uncertainties associated with the analysis are summarized in Table II. The systematic uncertainties due to the fit model are determined via ensemble investigations. To carry out an ensemble study, we generate and simulate 500,000 signal MC events. We randomly select signal events from this sample for different expected signal yields in data. In addition, background MC events are randomly extracted from the background PDFs based on the expected number of background events in the data. This MC sample that now has statistics equivalent to the expected yields in data is amplified by repeating the above procedure 1000 times. We then perform 3D unbinned extended ML fits on these 1000 pseudo-experiments to obtain pull distributions for each of the expected signal yields in data. The average deviation of a zeroth-order polynomial fit to the means of the pull values obtained for each expected signal yield from the no-bias condition is recognized as a fit bias.

We observe a fit bias of −3.3% and assign it as the corresponding systematic uncertainty. The uncertainty due to fixing the parameter values of the PDFs is determined by varying the best fit parameter values within $\pm 1\sigma$ of their statistical uncertainties and measuring the deviation of the signal yields in data. We find a fractional systematic uncertainty of $^{+3.5\%}_{-5.2\%}$ from this source. Apart from fixing the signal PDF parameters and the background PDF's ARGUS end point, we have also fixed the fractions of the B_s^0 production channels. We vary these fractions within $\pm 1\sigma$ of their measured values [12] and repeat the fit. The observed relative variation $^{+5.2\%}_{-3.5\%}$ of the signal yield is assigned as the systematic uncertainty. The systematic uncertainty of the signal reconstruction efficiency is 0.4% due to the finite number of signal MC events. The systematic uncertainty due to the efficiency of the $\mathcal{C}'_{\rm NN}$ requirement is estimated from the control sample using a parameter, $\mathcal{R}$. It is defined as the ratio between the efficiency of the $\mathcal{C}'_{\rm NN}$ requirement in data and MC. We assign a corresponding systematic uncertainty of ±3% due to the efficiency of the $\mathcal{C}'_{\rm NN}$ requirement.

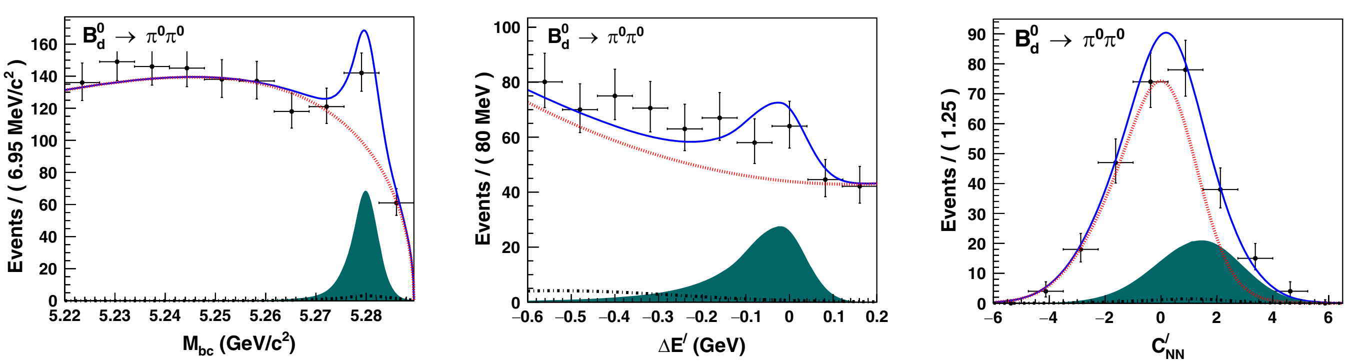

FIG. 2. Signal enhanced projections of $M_{\rm bc}$ (left), $\Delta E'$ (middle), and $\mathcal{C}'_{\rm NN}$ (right) for the control sample, $B_d^0 \to \pi^0\pi^0$. Each plot is generated by applying the signal region selection criteria on the two variables other than the plotted variable. The signal regions for the three variables are as follows: 5.2700 GeV$/c^2 < M_{\rm bc} <$ 5.2895 GeV$/c^2$, −0.23 GeV $< \Delta E' <$ 0.15 GeV, and $-3.10 < \mathcal{C}'_{\rm NN} < 7.61$. The dark-filled, red (dotted), black (dash-dotted), and blue (solid) color distributions represent the signal, continuum background, rare B_d^0 background (backgrounds arising due to $b \to u$ transitions) and total fit function, respectively. Points with error bars represent data.

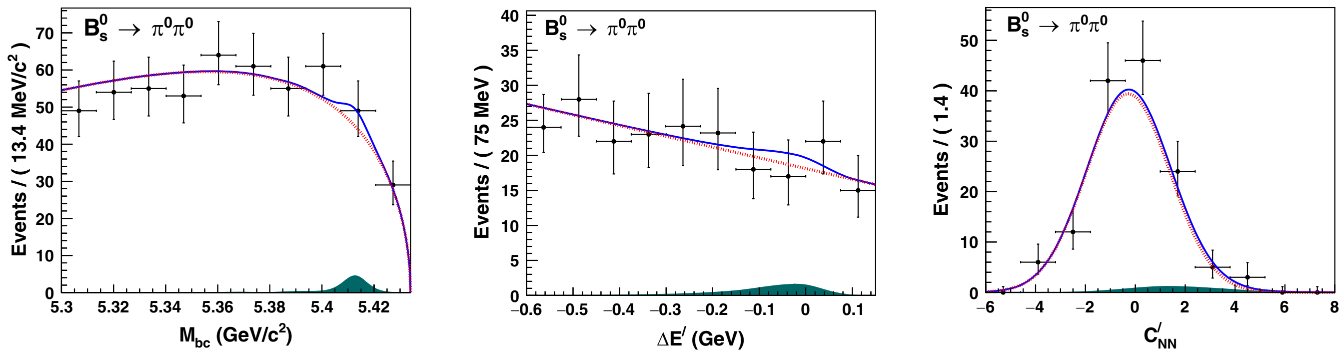

FIG. 3. Signal enhanced projections of M_{bc} (left), $\Delta E'$ (middle), and $\mathcal{C}'_{NN}$ (right) for the analysis, $B_s^0 \to \pi^0\pi^0$. Each plot is generated by applying the signal region selection criteria on the two variables other than the plotted variable. The signal regions for the three variables are as follows: $5.395\ \mathrm{GeV}/c^2 < M_{bc} < 5.434\ \mathrm{GeV}/c^2$, $-0.310\ \mathrm{GeV} < \Delta E' < 0.140\ \mathrm{GeV}$, and $-3.901 < \mathcal{C}'_{NN} < 7.451$. The dark-filled, red (dotted), and blue (solid) color distributions represent the signal, continuum background and total fit function, respectively. Points with error bars represent data. The peak in the M_{bc} distribution is due to the dominant B_s^0 production channel, $B_s^{*0}\bar{B}_s^{*0}$ (87.0%). The other two production channels, $B_s^{*0}\bar{B}_s^0$ (7.3%) and $B_s^0\bar{B}_s^0$ (5.7%), are present but do not contain a significant number of signal events.

The systematic uncertainty for the π^0 selection efficiency is determined to be 2.2% per π^0 using the decay $\tau^- \to \pi^-\pi^0\nu_\tau$. Since this uncertainty is completely correlated for the two π^0s, a total systematic uncertainty of 4.4% is assigned. We assign a fractional systematic uncertainty of 0.034% on the branching fraction of $\pi^0 \to \gamma\gamma$ [13]. Therefore, for the two π^0 branching fractions, a total systematic uncertainty of 0.07% is assigned.

The systematic uncertainty due to the $b\bar{b}$ production cross section at $\Upsilon(5S)$ resonance, $\sigma_{b\bar{b}}$ is estimated to be $\pm 4.7\%$ [12]. In addition, the systematic uncertainty due to the three charmless production processes arising from $b\bar{b}$ events, f_s is assumed to be $\pm 15.4\%$ [13]. This uncertainty on f_s is the dominant systematic uncertainty associated with any B_s^0 measurement at Belle.

The fit projections obtained from a 3D unbinned extended maximum likelihood fit in the signal regions are shown in Fig. 3. We obtain 5.7 ± 5.8 signal events and 989 ± 32 continuum background events in our fit to the data. The branching fraction is calculated using

$$\mathcal{B}(B_s^0 \to \pi^0\pi^0) = \frac{N_{\mathrm{yield}}^{\mathrm{sig}}}{2 \times N_{B_s^0\bar{B}_s^0} \times \epsilon^{\mathrm{rec}} \times \mathcal{B}}, \qquad (3)$$

where $N_{B_s^0\bar{B}_s^0}$ is the number of $B_s^0\bar{B}_s^0$ pairs; ϵ^{rec} and $N_{\mathrm{yield}}^{\mathrm{sig}}$ are the signal selection efficiency obtained from MC simulation and the signal yield obtained from the fit, respectively; and $\mathcal{B}$ is the product of the two π^0-decay branching fractions [13].

Incorporating the signal yield, $N_{\mathrm{yield}}^{\mathrm{sig}} = (5.7 \pm 5.8)$, number of $B_s^0\bar{B}_s^0$ pairs, $N_{B_s^0\bar{B}_s^0} = (8.30 \pm 1.34) \times 10^6$, the signal reconstruction efficiency, $\epsilon_{\mathrm{rec}} = (12.69 \pm 0.05)\%$, and branching fraction, $\mathcal{B}(\pi^0 \to \gamma\gamma) = (98.82 \pm 0.03)\%$ in Eq. (3), the branching fraction for $B_s^0 \to \pi^0\pi^0$ and its product with f_s are calculated to be

$$\mathcal{B}(B_s^0 \to \pi^0\pi^0) = (2.8 \pm 2.8 \pm 0.5) \times 10^{-6},$$
$$f_s \times \mathcal{B}(B_s^0 \to \pi^0\pi^0) = (0.6 \pm 0.6 \pm 0.1) \times 10^{-6}.$$

The first uncertainty is statistical, and the second uncertainty is systematic.

Without significant signal yield, we calculate the UL on the branching fraction using a Bayesian approach. The UL on the branching fraction is estimated by integrating the likelihood function obtained from the maximum likelihood fit procedure from 0% to 90% of the area under the likelihood curve. The systematic uncertainties are incorporated by convolving the likelihood curve with a Gaussian distribution with a mean of zero and width equivalent to the total systematic uncertainty listed in Table II. The UL on the branching fraction, $\mathcal{B}(B_s^0 \to \pi^0\pi^0)$ at 90% CL [23] and the product of the branching fraction with f_s, $f_s \times \mathcal{B}(B_s^0 \to \pi^0\pi^0)$, is found to be

$$\mathcal{B}(B_s^0 \to \pi^0\pi^0) < 7.7 \times 10^{-6},$$
$$f_s \times \mathcal{B}(B_s^0 \to \pi^0\pi^0) < 1.5 \times 10^{-6}.$$

The total systematic uncertainties associated with $\mathcal{B}(B_s^0 \to \pi^0\pi^0)$ and $f_s \times \mathcal{B}(B_s^0 \to \pi^0\pi^0)$ are $^{+18.1\%}_{-18.4\%}$ and $^{+9.5\%}_{-10.0\%}$, respectively. The results are summarized in Table III.

TABLE III. Summary of results on branching fractions and UL for $\mathcal{B}(B_s^0 \to \pi^0\pi^0)$ and $f_s \times \mathcal{B}(B_s^0 \to \pi^0\pi^0)$.

Quantity	Value
$\mathcal{B}(B_s^0 \to \pi^0\pi^0)$	$(2.8 \pm 2.8 \pm 0.5) \times 10^{-6}$ $< 7.7 \times 10^{-6}$ at 90% CL
$f_s \times \mathcal{B}(B_s^0 \to \pi^0\pi^0)$	$(0.6 \pm 0.6 \pm 0.1) \times 10^{-6}$ $< 1.5 \times 10^{-6}$ at 90% CL

To summarize, we search for the decay $B_s^0 \to \pi^0\pi^0$ using the final Belle data sample available at $\Upsilon(5S)$ resonance, which corresponds to an integrated luminosity of 121.4 fb^{-1}. We do not observe a significant signal yield, and thus set a 90% CL upper limit on the $B_s^0 \to \pi^0\pi^0$ branching fraction of 7.7×10^{-6}. This is the most stringent UL estimated for this decay representing an order-of-magnitude improvement over the previous result [7] by the L3 experiment in 1995.

This work, based on data collected using the Belle detector, which was operated until June 2010, was supported by the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan, the Japan Society for the Promotion of Science (JSPS), and the Tau-Lepton Physics Research Center of Nagoya University; the Australian Research Council including Grants No. DP180 102629, No. DP170102389, No. DP170102204, No. DE22 0100462, No. DP150103061, and No. FT130100303; Austrian Federal Ministry of Education, Science and Research (FWF) and FWF Austrian Science Fund No. P 31361-N36; the National Natural Science Foundation of China under Contracts No. 11675166, No. 11705209, No. 11975076, No. 12135005, No. 12175041, and No. 12161141008; Key Research Program of Frontier Sciences, Chinese Academy of Sciences (CAS), Grant No. QYZDJ-SSW-SLH011; Project No. ZR2022JQ02 supported by Shandong Provincial Natural Science Foundation; the Ministry of Education, Youth and Sports of the Czech Republic under Contract No. LTT17020; the Czech Science Foundation Grant No. 22-18469S; Horizon 2020 ERC Advanced Grant No. 884719 and ERC Starting Grant No. 947006 "InterLeptons" (European Union); the Carl Zeiss Foundation, the Deutsche Forschungs gemeinschaft, the Excellence Cluster Universe, and the VolkswagenStiftung; the Department of Atomic Energy (Project Identification No. RTI 4002) and the Department of Science and Technology of India; the Istituto Nazionale di Fisica Nucleare of Italy; National Research Foundation (NRF) of Korea Grants No. 2016R1D1A1B02012900, No. 2018R1A2B3003643, No. 2018R1A6A1A06024970, No. RS202200197659, No. 2019R1I1A3A01058933, No. 2021R1A6A1A03043957, No. 2021R1F1A1060423, No. 2021R1F1A1064008, and No. 2022R1A2C1003993; Radiation Science Research Institute, Foreign Large-size Research Facility Application Supporting project, the Global Science Experimental Data Hub Center of the Korea Institute of Science and Technology Information and Korea Research Environment Open Network/Global Research Network Collaborations (KREONET/GLORIAD); the Polish Ministry of Science and Higher Education and the National Science Center; the Ministry of Science and Higher Education of the Russian Federation, Agreement No. 14.W03.31.0026, and the HSE University Basic Research Program, Moscow; University of Tabuk research Grants No. S-1440-0321, No. S-0256-1438, and No. S-0280-1439 (Saudi Arabia); the Slovenian Research Agency Grants No. J1-9124 and No. P1-0135; Ikerbasque, Basque Foundation for Science, Spain; the Swiss National Science Foundation; the Ministry of Education and the Ministry of Science and Technology of Taiwan; and the United States Department of Energy and the National Science Foundation. These acknowledgements are not to be interpreted as an endorsement of any statement made by any of our institutes, funding agencies, governments, or their representatives. 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 Pacific Northwest National Laboratory (PNNL) Environmental Molecular Sciences Laboratory (EMSL) computing group for strong computing support; and the National Institute of Informatics, and Science Information NETwork 6 (SINET6) for valuable network support.

[1] M. Gronau, D. London, and J. L. Rosner, Phys. Rev. D **87**, 036008 (2013).
[2] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. Lett. **118**, 081801 (2017).
[3] Q. Chang, J. Sun, Y. Yang, and X. Li, Phys. Lett. B **740**, 56 (2015).
[4] H. Cheng, C. Chiang, and A. Kuo, Phys. Rev. D **91**, 014011 (2015).
[5] A. Ali, G. Kramer, Y. Li, C.-D. Lü, Y.-L. Shen, W. Wang, and Y.-M. Wang, Phys. Rev. D **76**, 074018 (2007).
[6] M. Beneke and M. Neubert, Nucl. Phys. **B675**, 333 (2003).
[7] M. Acciarri *et al.* (L3 Collaboration), Phys. Lett. B **363**, 127 (1995).
[8] Charge conjugation is implied throughout this paper.
[9] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods Phys. Res., Sect. A **499**, 1 (2003), and other papers included in this volume; T. Abe *et al.*, Prog. Theor. Exp. Phys. **2013**, 03A001 (2013), and references therein.
[10] A. Abashian *et al.* (Belle Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 117 (2002).
[11] J. Brodzicka *et al.* (Belle Collaboration), Prog. Theor. Exp. Phys. **2012**, 04D001 (2012).

[12] S. Esen *et al.* (Belle Collaboration), Phys. Rev. D **87**, 031101(R) (2013).
[13] R. L. Workman *et al.* (Particle Data Group), Prog. Theor. Exp. Phys. **2022**, 083C01 (2022).
[14] D. J. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 152 (2001).
[15] R. Brun *et al.*, GEANT3.21, Report No. CERN DD/EE/84-1, 1984.
[16] M. Feindt and U. Kerzel, Nucl. Instrum. Methods Phys. Res., Sect. A **559**, 190 (2006).
[17] G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978); S. H. Lee *et al.* (Belle Collaboration), Phys. Rev. Lett. **91**, 261801 (2003).
[18] A. J. Bevan, B. Golob, T. Mannel, S. Prell, and B. D. Yabsley, Eur. Phys. J. C **74**, 3026 (2014).
[19] G. Punzi, eConf **C030908**, MODT002 (2003), https://doi.org/10.48550/arXiv.physics/0308063.
[20] T. Skwarnicki, Ph.D. thesis, Institute for Nuclear Physics, Krakow, DESY Internal Report, Report No. DESY F31-86-02, 1986.
[21] H. Albrecht *et al.* (ARGUS Collaboration), Phys. Lett. B **241**, 278 (1990).
[22] T. Julius, M. E. Sevior, G. B. Mohanty *et al.* (Belle Collaboration), Phys. Rev. D **96**, 032007 (2017).
[23] As we use the Bayesian method, this is a “credible interval” but we use “confidence level” here following common convention.