Search for the Decay $B_s^0 \to \eta \eta$

N. K. Nisar, V. Savinov, I. Adachi, H. Alihara, S. Al Said, D. M. Asner, A. H. Atmacan, T. Aushev, R. Ayad, V. Babu, S. Bahinipati, P. Behera, J. Bennett, M. Bessner, V. Bhardwaj, B. Bhuyan, T. Bilka, J. Biswal, G. Bonvini, A. Bozek, M. Bračko, T. E. Browder, M. Campajola, D. Červenkov, M.-C. Chang, V. Chekelian, A. Chen, B. G. Cheon, K. Chilikin, H. E. Cho, K. Cho, S.-K. Choi, Y. Choi, S. Choudhury, D. Cinabro, S. Cunliffe, S. Das, N. Dash, G. De Nardo, R. Dhamija, F. Di Capua, Z. Doležal, T. V. Dong, S. Dubey, S. Eidelman, D. Epifanov, T. Ferber, D. Ferlevicz, A. Frey, B. G. Fulsom, R. Garg, V. Gaur, N. Gabyshev, A. Garmash, A. Giri, P. Goldenzweig, Y. Guan, K. Gudikova, C. Hadjivasiliou, S. Halder, T. Hará, O. Hartbrich, K. Hayasaka, H. Hayashii, M. T. Hedges, C.-L. Hsu, T. Iijima, K. Inami, A. Ishikawa, R. Itoh, M. Iwasaki, Y. Iwasaki, W. W. Jacobs, S. Jia, Y. Jin, C. W. Joo, K. K. Joo, J. Kahn, A. B. Kaliyar, K. H. Kang, K. K. Kang, T. Kawasaki, H. Kichimi, C. Kiesling, C. H. Kim, D. Y. Kim, S. H. Kim, Y.-K. Kim, K. Kinoshita, P. Kodyš, K. Konno, A. Korobov, S. Korpar, E. Kovalenko, P. Križan, R. Kroeger, P. Krokovny, T. Kuhr, M. Kumer, K. Kumara, A. Kuzmin, Y.-J. Kwon, K. Lalwani, J. S. Lange, S. C. Lee, B. Li, L. Li Gioi, J. Libby, K. Lieret, D. Liventsev, C. MacQueen, M. Masuda, T. Matsuda, D. Matvienko, M. Merola, R. Mizuk, G. B. Mohanty, S. Mohanty, M. Nakao, A. Natochii, L. Nayak, M. Nayak, S. Nishida, K. Nishimura, S. Ogawa, H. Ono, Y. Onuki, P. Oskin, P. Pakhlov, G. Pakhlova, T. Pang, S. Pardi, H. Park, S. Patra, S. Paul, T. K. Pedlar, R. Pestotnik, L. E. Piilonen, T. Podolnok, E. Prencipe, M. T. Prim, M. Rührer, A. Rostomyan, N. Routh, G. Russo, D. Sahoo, S. Sandilya, A. Sangal, L. Santelj, T. Sanuki, G. Schmell, J. Schueler, C. Schwanda, Y. Seino, K. Senyo, M. E. Sevior, M. Shapkin, C. Sharma, C. P. Shen, J.-G. Shiu, B. Shwartz, F. Simon, E. Solovieva, S. Stanič, Z. S. Stottler, M. Sumihama, T. Sumiyoshi, M. Takizawa, J. S. Lange, S. C. Lee, F. Tenchini, K. Trabelsi, M. Uchida, T. Uglov, Y. Unno, S. Uno, Y. Urquijo, K. Chilikin, S. V. Vassen, E. Waheed, C. H. Wang, M. Z.-W. Wang, P. Wang, X. L. Wang, S. Watanuki, X. Xu, B. D. Yabsley, W. Yan, S. B. Yang, H. Ye, Z. F. Zhang, V. Zhilich, V. Zhukova

(The Belle Collaboration)
The charmedless hadronic decay $B^0_s \rightarrow \eta'\eta$ is suppressed in the Standard Model (SM) and proceeds only through transitions sensitive to Beyond-the-Standard-Model (BSM) physics [1]. BSM scenarios, such as a fourth generation of fermions, supersymmetry with broken R-parity, and a two-Higgs doublet model with flavor-changing neutral currents, could affect the branching fraction and CP asymmetry of this decay [2]. The expected branching fraction for $B^0_s \rightarrow \eta'\eta$ in the SM spans a range of $(2 - 4) \times 10^{-5}$ [3,7]. Once branching fractions for two-body decays $B^0_{s,d,*} \rightarrow \eta\eta, \eta'\eta$, and $\eta'\eta'$ are measured, it would be possible to extract CP-violating parameters using a formalism based on SU(3)/U(3) symmetry [3]. To achieve this goal, at least four of these six branching fractions need to be measured. Only the branching fraction for $B^0_s \rightarrow \eta'\eta'$ has been measured so far [8].

In this Letter, we report the results of the first search for the decay $B^0_s \rightarrow \eta'\eta$ using the full Belle data sample of 121.4 fb$^{-1}$ collected at the $\Upsilon(5S)$ resonance. The inclusion of the charge-conjugate decay mode is implied throughout. The Belle detector was a large-solid-angle magnetic spectrometer that operated at the KEKB asymmetric-energy $e^+e^-$ collider [9]. The detector components relevant to our study include a tracking system comprising a silicon vertex detector (SVD) and a central drift chamber (CDC), a particle identification (PID) system that consists of a barrel-like arrangement of time-of-flight scintillation counters (TOF) and an array of aerogel threshold Cherenkov counters (ACC), and a CsI(Tl) crystal-based electromagnetic calorimeter (ECL). All these components were located inside a superconducting solenoid coil that provided a 1.5 T magnetic field. A detailed description of the Belle detector can be found elsewhere [10].

The $\Upsilon(5S)$ resonance decays into $B^{*0}\bar{B}^{*0}$, $B^0_s\bar{B}^{*0}$, and $B^0_s\bar{B}^0$ pairs, where the relative fractions of the two former decays are $f_{B^{*0}\bar{B}^{*0}} = (87.0 \pm 1.7)\%$ and $f_{B^0_s\bar{B}^{*0}} = (7.3 \pm 1.4)\%$ [11], respectively. Signal $B^0_s$ mesons originate from the direct decays of $\Upsilon(5S)$ or from radiative decays of the excited vector state $B^{*0}_s$. The $\Upsilon(5S)$ production cross section is $340 \pm 16$ pb [11]. To present our nominal result for $B(B^0_s \rightarrow \eta'\eta)$ we use the world average value for the fraction of $B^{(*)0}_sB^{(*)0}_s$ in $b\bar{b}$ events $f_s = 0.201 \pm 0.031$ [12], the data sample is therefore estimated to contain $(16.60 \pm 2.68) \times 10^6$ $B^0_s$ mesons. We also report the results for $f_s \times B(B^0_s \rightarrow \eta'\eta)$.

To maximize discovery potential of the analysis and to validate the signal extraction procedure, we use a sample of background Monte Carlo (MC) simulated events equivalent to six times the data statistics. In addition, to estimate the overall reconstruction efficiency we use a high-statistics signal MC sample, where the other $B^0_s$ meson decays according to known branching fractions [12]. Both samples are used to develop a model implemented in the unbinned extended maximum-likelihood (ML) fit to data. The MC-based model is validated with a control sample of $B^0 \rightarrow \eta'K^0_S$ decays reconstructed from 711 fb$^{-1}$ of $\Upsilon(4S)$ data.

We reconstruct $\eta$ candidates using pairs of electromagnetic showers not matched to the projections of charged tracks to the ECL and therefore identified as photons. We require that the reconstructed energies of these showers exceed 50 (100) MeV in the barrel (endcap) region of the ECL. The larger energy threshold for the endcaps is due to the larger beam-related background in these regions. To reject hadronic showers mimicking photons, the ratio of the energies deposited by a photon candidate in the $(3 \times 3)$ and $(5 \times 5)$ ECL crystal arrays centered on the crystal with the largest deposited energy is required to exceed 0.75. The reconstructed invariant mass of the $\eta$ candidates is required to be $515 \leq M(\gamma\gamma) \leq 580$ MeV/c$^2$, which corresponds, approximately, to a $\pm 3\sigma$ Gaussian resolution window around the nominal $\eta$ mass [12]. To suppress combinatorial background arising due to low-energy photons, the magnitude of the cosine of the helic-
ity angle ($\cos \theta_{\text{beam}}$) is required to be less than 0.97, where $\theta_{\text{beam}}$ is the angle in the $\eta$ candidate’s rest frame between the directions of its Lorentz boost from the laboratory frame and one of the photons.

The $\eta'$ candidates are formed by combining pairs of oppositely charged pions with the $\eta$ candidates. We require the reconstructed $\eta'$ invariant mass to be in the range $920 \leq M(\pi^+\pi^-\eta) \leq 980$ MeV/$c^2$, which corresponds, approximately, to the range $[-10, +6] \sigma$ of the Gaussian resolution, after performing a kinematic fit constraining the reconstructed mass of the $\eta$ candidate to the nominal $\eta$ mass [12]. To identify charged pion candidates, the ratios of PID likelihoods, $R_{i/\pi} = L_i/(L_\pi + L_i)$, are used, where $L_\pi$ is the likelihood for the track being a pion, while $L_i$ is the corresponding likelihood for the kaon ($i = K$) or electron ($i = e$) hypotheses. We require $R_{K/\pi} \leq 0.6$ and $R_{e/\pi} \leq 0.95$ for pion candidates. The likelihood for each particle species is obtained by combining information from CDC, TOF and ACC [13], and (for electrons only) ECL [14]. According to MC studies, these requirements reject 28% of background, while the resulting efficiency loss is below 3%. Charged pion tracks are required to originate from near the interaction point (IP) by restricting their distance of closest approach to the $z$ axis to be less than 4.0 cm along the $z$ axis and 0.3 cm perpendicular to it, respectively. The $z$ axis is opposite to the direction of the $e^+$ beam. These selection criteria suppress beam-related backgrounds and reject poorly reconstructed tracks. To reduce systematic uncertainties associated with track reconstruction efficiency, the transverse momenta of charged pions are required to be greater than 100 MeV/c.

To identify $B_s^0 \rightarrow \eta'\eta$ candidates we use (shown here in natural units) the beam-energy-constrained $B_s^0$ mass, $M_{bc} = \sqrt{E_{\text{beam}}^2 - p_{\text{bc}}^2}$, the energy difference, $\Delta E = E_{B_s} - E_{\text{beam}}$, and the reconstructed invariant mass of the $\eta'$, where $E_{\text{beam}}$, $p_{\text{bc}}$, $E_{B_s}$ are the beam energy, the momentum and energy of the $B_s^0$ candidate, respectively. All these quantities are calculated in the $e^+e^-$ center-of-mass frame. To improve the $\Delta E$ resolution, the $\eta'$ candidates are further constrained to the nominal mass of $\eta'$, though most of the improvement comes from the $\eta$ mass constraint. Signal candidates are required to satisfy selection criteria $M_{bc} > 5.3$ GeV/$c^2$ and $-0.4 \leq \Delta E \leq 0.3$ GeV. In a Gaussian approximation, the $\Delta E$ resolution is approximately 40 MeV. Similarly, the $M_{bc}$ resolution is 4 MeV/$c^2$. To take advantage of all available information in case the data indicate signal presence, we include $M(\pi^+\pi^-\eta)$ in the three-dimensional (3D) ML fit used to statistically separate the signal from background. We define the signal region: $5.35 < M_{bc} < 5.43$ GeV/$c^2$, $-0.25 \leq \Delta E \leq 0.10$ GeV, and $0.94 < M(\pi^+\pi^-\eta) < 0.97$ GeV/$c^2$. The area outside the signal region is considered as sideband. To optimize sensitivity we use a narrower signal region $5.39 < M_{bc} < 5.43$ GeV/$c^2$ which would contain the largest signal contribution.

Hadronic continuum events from $e^+e^- \rightarrow q\bar{q}$ ($q = u,d,c,s$) are the primary source of background. Because of large initial momenta of the light quarks, continuum events exhibit a “jetlike” event shape, while $B_s^{(*)0} \rightarrow \eta\pi^0$ events are distributed isotropically. We utilize modified Fox-Wolfram moments [15], used to describe the event topology, to discriminate between signal and continuum background. A likelihood ratio ($LR$) is calculated using Fisher discriminant coefficients obtained in an optimization based on these moments. We suppress the background using a discovery-optimized selection on $LR$ obtained by maximizing the value of Punzi’s figure of merit [16]:

$$FOM = \frac{\varepsilon(t)}{a/2 + \sqrt{B(t)}}, \quad (1)$$

where $t$ is the requirement on $LR$, $\varepsilon$ and $B$ are the signal reconstruction efficiency and the number of background events expected in the signal region for a given value of $t$, respectively. The quantity $a$ is the desired significance (which we vary between 3 and 5) in the units of standard deviation. To predict $B(t)$ we multiply the number of events in the background MC sample. We require signal candidates to satisfy the requirement $LR \geq 0.95$, which corresponds to $B(0.95) = 3.3$ and 48 background events in the signal region and sideband in the background MC sample. The background events containing real $\eta'$ mesons exhibit a peak in the $M(\pi^+\pi^-\eta)$ distribution, however, they are distributed smoothly in $M_{bc}$ and $\Delta E$. The fraction of this peaking background is a free parameter in our ML fits.

About 14% of the reconstructed signal MC events contain multiple candidates primarily arising due to misreconstructed $\eta$ mesons. In such events we retain the candidate with the smallest value of $\chi^2_{\eta}$, where $\chi^2_{\eta}$ denotes the $\eta$ mass-constrained fit statistic, the summation is over the two $\eta$ candidates, and $\chi^2_{\eta}$ quantifies the quality of the vertex fit for two pion tracks. Simulation shows that this procedure selects the correct $B_s^0$ candidate in 62% of such events. The overall reconstruction efficiency is 10%.

To extract the signal yield, we perform an unbinned extended ML fit to the 3D distribution of $M_{bc}$, $\Delta E$, and $M(\pi^+\pi^-\eta)$. The likelihood function is
\[
\mathcal{L} = \frac{e^{-\sum_{i} s_{i}}}{N!} \prod_{i=1}^{N} \left( \sum_{j} n_{i} \mathcal{P}_{j}[M_{bc}^{i}, \Delta E^{i}, M^{i}(\pi^{+}\pi^{-})] \right),
\]

where \( i \) is the event index, \( N \) is the total number of events, \( j \) denotes the fit component (the three components are background, correctly reconstructed signal, and misreconstructed signal described later), and the parameters \( n_{i} \) represent signal and background yields. Due to negligible correlations among fit variables for both background and correctly reconstructed signal events, the probability density function (PDF) for each fit component is assumed to factorize as 
\[
\mathcal{P}(M_{bc}^{i}, \Delta E^{i}, M^{i}(\pi^{+}\pi^{-})) = \mathcal{P}(M_{bc}^{i}) \cdot \mathcal{P}(\Delta E^{i}) \cdot \mathcal{P}(M^{i}(\pi^{+}\pi^{-})).
\]

The PDF signal is represented by a weighted sum of the three PDFs describing possible \( B^{0} \to \eta' K_{S}^{0} \) signal contributions from \( B^{0} \to \eta' K_{S}^{0} \) pairs, where the weights are fixed according to previous measurements [11].

To validate our fitting model and adjust the PDF shape parameters used to describe the signal, we use the control sample of \( B^{0} \to \eta' K_{S}^{0} \) decays. We reconstruct \( K_{S}^{0} \) candidates via secondary vertices associated with pairs of oppositely charged pions [17] using a neural network technique [18]. The following information is used in the network: the momentum of the \( K_{S}^{0} \) candidate in the laboratory frame; the distance along the \( z \) axis between the two track helices at the point of their closest approach; the flight length in the \( x \) - \( y \) plane; the angle between the \( K_{S}^{0} \) momentum and the vector joining the \( K_{S}^{0} \) decay vertex to the IP; the angle between the pion momentum and the laboratory-frame \( K_{S}^{0} \) momentum in the \( K_{S}^{0} \) rest frame; the distance-of-closest-approach in the \( x \) - \( y \) plane between the IP and the two pion helices; and the pion hit information in the SVD and CDC. The selection efficiency is 87% over the momentum range of interest. We also require that the reconstructed \( \pi^{+}\pi^{-} \) invariant mass is within 12 MeV/\( c^{2} \), which is about 3.5\( \sigma \) of the nominal \( K_{S}^{0} \) mass [12]. We require \( 5.20 < M_{bc} < 5.30 \) GeV/\( c^{2} \) for \( B^{0} \) candidates. The control-sample signal region is \( 5.27 < M_{bc} < 5.29 \) GeV/\( c^{2} \), \(-0.20 < \Delta E < 0.10 \) GeV, and \( 0.94 < M(\pi^{+}\pi^{-}) < 0.97 \) GeV/\( c^{2} \). All other selection criteria are the same as those used to select \( B^{0} \) candidates. This control sample is used to validate the \( \eta \) and \( \eta' \) reconstruction and its effect on the resolution functions and PDF shape parameters. The validation of \( K_{S}^{0} \) reconstruction was performed previously in a similar \( B^{0} \) analysis [19].

The presence of four photons in the final state gives rise to a sizable misreconstruction probability for the signal events. We study these self-crossfeed (SCF) events using the signal MC sample. A large correlation between \( M_{bc} \) and \( \Delta E \) for such signal events is taken into account by describing the correctly reconstructed signal and SCF components separately with two different PDF sets. The latter comprise approximately 14% of the reconstructed signal and are excluded from the estimate of its efficiency. The Pearson correlation coefficient for the region with largest correlations for SCF signal events is 27%.

A sum of a Gaussian and a Crystal Ball [20] function is used to model the correctly reconstructed signal in each of the three fit variables. For \( M_{bc} \) and \( M(\pi^{+}\pi^{-}) \) we use a sum of these two functions with the same mean but different widths, while for \( \Delta E \) both the mean and width are different. A Bubin function [21] and an asymmetric Gaussian are used to model the SCF contribution in \( M_{bc} \) and \( \Delta E \), respectively. For \( M(\pi^{+}\pi^{-}) \), we use a sum of a Gaussian and a first-order Chebyshev polynomial. In our nominal fit to data the fraction of correctly reconstructed signal is fixed to its MC value. The signal PDF shape parameters for \( M_{bc} \) and \( \Delta E \) are validated using the \( B^{0} \to \eta' K_{S}^{0} \) control sample.

We use an ARGUS [22] function to describe the background distribution in \( M_{bc} \) and a first-order Chebyshev polynomial for \( \Delta E \). To model the peaking part in \( M(\pi^{+}\pi^{-}) \) we use the signal PDF, because the peak is due to real \( \eta' \) mesons, while an additional first-order Chebyshev polynomial is used for the non-peaking contribution. The projections of the fit to the \( B^{0} \to \eta' K_{S}^{0} \) control sample are shown in Fig. 4.

To further test and validate our fitting model, ensemble tests are carried out by generating MC pseudoexperiments. In these experiments we use PDFs obtained from full detector simulation and the \( B^{0} \to \eta' K_{S}^{0} \) data. We perform 1000 pseudoexperiments for each assumed number of signal events. An ML fit is executed for each sample prepared in these experiments. The signal yield distribution obtained from these fits exhibits good linearity. We use the results of pseudoexperiments to construct classical confidence intervals (without ordering) using a procedure due to Neyman [23]. For each ensemble of pseudoexperiments, the lower and upper ends of the respective confidence interval represent the values of fit signal yields for which 10% of the results lie below and above these values, respectively. These intervals are then combined to prepare a classical confidence belt [24, 25] used to make a statistical interpretation of the results obtained from data. The confidence intervals prepared using this statistical method are known to slightly “overcover” for the number of signal events [26], therefore resulting in a conservative upper limit.

We apply the 3D model to the data and obtain 2.7\( \pm \)2.5 signal and 57.3\( \pm \)7.8 background events. The signal-region projections of the fit are shown in Fig. 2. We observe no significant signal and estimate a 90% confidence-level (CL) upper limit on the branching fraction for the decay \( B^{0} \to \eta' \eta \) using the following formula:
\[ \mathcal{B}(B_s^0 \rightarrow \eta' \eta) < \frac{N_{UL}^{90\%}}{N_{B_s^0} \times \varepsilon \times B} \]  

(3)

where \( N_{B_s^0} \) is the number of \( B_s^0 \) mesons in the full Belle data sample, \( \varepsilon \) is the overall reconstruction efficiency for the signal \( B_s^0 \) decay, and \( B \) is the product of the sub-decay branching fractions for \( \eta \) and \( \eta' \) reconstructed in our analysis. Further, \( N_{UL}^{90\%} \) is the expected signal yield of approximately 6.6 events at 90\% CL obtained from the confidence belt constructed using the frequentist approach \[23\]. Using Eq. (3) we estimate a 90\% CL upper limit on the branching fraction \( \mathcal{B}(B_s^0 \rightarrow \eta' \eta) < 6.2 \times 10^{-5} \). We also estimate a 90\% CL upper limit on the product \( \varepsilon \times \mathcal{B}(B_s^0 \rightarrow \eta' \eta) < 1.2 \times 10^{-5} \). The systematic uncertainties are not included in these estimates.

Sources of systematic uncertainties and their relative contributions are listed in Table I. The relative uncertainties on \( f_s \) and \( \sigma(Y(5\,S)) \) are 15.4\% and 4.7\%, respectively. The systematic uncertainty due to \( \eta \) reconstruction is 2.1\% per \( \eta \) candidate \[27\]. Track reconstruction \[28\] and PID systematic uncertainties are 0.35\% and 2\% per track, respectively. We estimate the systematic uncertainty due to the \( \mathcal{LR} \) requirement to be 10\%, which represents the relative change in efficiency when this requirement is varied by \( \pm 0.02 \) about the nominal value of 0.95. This range of variation is defined by the statistics of the control sample which is used to validate the efficiency and its dependence on the \( \mathcal{LR} \) requirement. Systematic uncer-

FIG. 1: Signal-region projections of the fit results on \( M_{bc}, \Delta E, \) and \( M(\pi^+ \pi^- \eta) \) for the \( B^0 \rightarrow \eta' K_S^0 \) control sample. Points with error bars are data, blue solid curves are the results of the fit, black dashed curves are the background component, and cyan-filled regions show the signal component.

FIG. 2: Signal-region projections of the fit results on \( M_{bc}, \Delta E, \) and \( M(\pi^+ \pi^- \eta) \) for \( B^0 \rightarrow \eta' \eta \). The \( M_{bc} \) signal region of the dominant signal contribution, \( 5.39 < M_{bc} < 5.43 \text{ GeV}/c^2 \), is used to plot the \( \Delta E \) and \( M(\pi^+ \pi^- \eta) \) projections. Points with error bars are data, blue solid curves are the results of the fit, black dashed curves are the background component, and pink-filled regions show the signal component. The three \( M_{bc} \) peaks in the signal component (from right to left) correspond to contributions from \( B_s^0 \bar{B_s}^0, B_s^0 \bar{B_s}^0, \) and \( B_s^0 \bar{B_s}^0 \) pairs.
tainty due to signal PDF shape is estimated by varying the fixed parameters within their statistical uncertainties determined with \( B^0 \rightarrow \eta' K_S^0 \) data. When varying these parameters, we observe an 11% change in the signal yield obtained from the data and use this number as an estimate of PDF parametrization systematics. Systematic uncertainty due to \( f_{B^{(*)0} \to \eta'} \) is evaluated by varying relative fractions of possible contributions to signal PDF and is 1.3%. When varying the SCF contribution by ±50% of itself, we observe a 4% change in the results of the fit to data, which we use as an estimate of SCF PDF systematic uncertainty. The relative uncertainties on \( \eta \) and \( \eta' \) branching fractions are 1% and 1.2%, respectively. The statistical uncertainty due to MC statistics is estimated to be 0.1%. The overall systematic uncertainties for \( B(\eta \rightarrow \eta') \) and \( f_s \times B(\eta' \rightarrow \eta') \) are estimated by adding the individual contributions in quadrature and are 23.1% and 17.2%, respectively. These systematic uncertainties are included in the \( N_{\text{90\%}} \) estimates of approximately 7.0 and 6.9 events by smearing the fit yield distributions while constructing the confidence belt used to extract the results. We estimate the upper limits on the branching fraction \( B(\eta \rightarrow \eta') < 6.5 \times 10^{-5} \) and on the product \( f_s \times B(\eta' \rightarrow \eta') < 1.3 \times 10^{-5} \) at 90% CL. Finally, using the number of signal events obtained from the fit we estimate \( B(\eta \rightarrow \eta') = (2.5 \pm 2.2 \pm 0.6) \times 10^{-5} \) and \( f_s \times B(\eta' \rightarrow \eta') = (0.51 \pm 0.44 \pm 0.09) \times 10^{-5} \), where, for each of the two estimates, the first uncertainty is statistical and the second is systematic. We summarize the results in Table I.

In summary, we have used the full data sample recorded by the Belle experiment at the \( \Upsilon(5S) \) resonance to search for the decay \( B_s^0 \rightarrow \eta' \eta \). We observe no statistically significant signal and set a 90% CL upper limit of \( 6.5 \times 10^{-5} \) on its branching fraction. To date, our result is the only experimental information on \( B_s^0 \rightarrow \eta' \eta \) and is twice as large as the most optimistic SM-based theoretical prediction. This decay can be probed further at the next-generation Belle II experiment [29] at the SuperKEKB collider in Japan.

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 5 (SINET5) for valuable network support. We acknowledge support from 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 DP180102629, DP170102389, DP170102204, DP150103061, 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. 11435013, No. 11475187, No. 11521505, No. 11575017, No. 11675166, No. 11705209; Key Research Program of Frontier Sciences, Chinese Academy of Sciences (CAS), Grant No. QYZDJ-SSW-SLH011; the CAS Center for Excellence in Particle Physics (CCEPP); the Shanghai Pujiang Program under Grant No. 18PJ1401000; the Shanghai Science and Technology Committee (STCSM) under Grant No. 19ZR1403000; the Ministry of Education, Youth and Sports of the Czech Republic under Contract No. LTT17020; Horizon 2020 ERC Advanced Grant No. 884719 and ERC Starting Grant No. 947006 “InterLeptons” (European Union); the Carl Zeiss Foundation, the Deutsche Forschungsgemeinschaft, 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 Grant Nos. 2016R1A1A1B01010135, 2016R1A1B0212900, 2018R1A2B3003643, 2018R1A6A1A06024970, 2018R1D1A1B07047294, 2019K1A3A7A09033840, 2019R1I1A3A01058933; Radiation Science Research Institute, Foreign Large-size

| Source            | Uncertainty (%) |
|-------------------|-----------------|
| \( f_s \)         | 15.4            |
| \( \sigma(\Upsilon(5S)) \) | 4.7            |
| \( \eta \)        | 4.2             |
| Tracking          | 0.7             |
| PID               | 4.0             |
| \( \mathcal{L}R \) selection | 10.0        |
| PDF parametrization | 11.0        |
| \( f_{B^{(*)0} \to \eta'} \) | 1.3            |
| SCF PDF           | 4.0             |
| Branching fraction of \( \eta \) | 1.0             |
| Branching fraction of \( \eta' \) | 1.2             |
| MC statistics     | 0.1             |

### Table II: Summary of the results for \( f_s \times B(\eta' \rightarrow \eta \eta) \) and \( B(\eta' \rightarrow \eta \eta) \).

| Quantity                                    | Value                              |
|---------------------------------------------|------------------------------------|
| \( f_s \times B(\eta' \rightarrow \eta \eta) \) | \((0.51 \pm 0.44 \pm 0.09) \times 10^{-5}\)  |
| \( B(\eta' \rightarrow \eta \eta) \)         | \(< 1.3 \times 10^{-5} @ \text{90\% CL}\)  |
|                                             | \(< 6.5 \times 10^{-5} @ \text{90\% CL}\)  |
Research Facility Application Supporting project, the Global Science Experimental Data Hub Center of the Korea Institute of Science and Technology Information and 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 14.W03.31.0026, and the HSE University Basic Research Program, Moscow; University of Tabuk research grants S-1440-0321, S-0256-1438, and S-0280-1439 (Saudi Arabia); the Slovenian Research Agency Grant Nos. J1-9124 and P1-0135; Ikerbasque, Basque Foundation for Science, Spain; 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.

[1] Ed. A. J. Bevan, B. Golob, Th. Mannel, S. Prell, and B. D. Yabsley, Eur. Phys. J. C 74, 3026 (2014); SLAC-PUB-15968; KEK Preprint 2014-3.

[2] E. Kou et al., Prog. Theor. Exp. Phys. 2019, 123C01 (2019).

[3] Y.-K. Hsiao, C.-F. Chang, and X.-G. He, Phys. Rev. D 93, 114002 (2016).

[4] A. R. Williamson and J. Zupan, Phys. Rev. D 74, 014003 (2006).

[5] A. Ali et al., Phys. Rev. D 76, 074018 (2007).

[6] H.-Y. Cheng and C.-K. Chua, Phys. Rev. D 80, 114026 (2009).

[7] H.-Y. Cheng, C.-W. Chiang, and A.-L. Kuo, Phys. Rev. D 91, 014011 (2015).

[8] R. Aaij et al. (LHCb Collaboration), Phys. Rev. Lett. 115, 051801 (2015).

[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); also see Section 2 in J. Brodzicka et al., Prog. Theor. Exp. Phys. 2012, 04D001 (2012).

[11] S. Esen et al. (Belle Collaboration), Phys. Rev. D 87, 031101(R) (2013).

[12] P.A. Zyla et al. (Particle Data Group), Prog. Theor. Exp. Phys. 2020, 083C01 (2020).

[13] E. Nakano, Nucl. Instrum. Methods Phys. Res. Sect. A 494, 402 (2002).

[14] K. Hanagaki, H. Kakuno, H. Ikeda, T. Iijima and T. Tsukamoto, Nucl. Instrum. Methods Phys. Res. Sect. A 485, 490 (2002).

[15] The Fox–Wolfram moments were introduced in G. C. Fox and S. Wolfram, Phys. Rev. Lett. 41, 1581 (1978). The Fisher discriminant used by Belle, based on modified Fox-Wolfram moments, is described in K. Abe et al. (Belle Collaboration), Phys. Rev. Lett. 87, 101801 (2001) and K. Abe et al. (Belle Collaboration), Phys. Lett. B 511, 151 (2001).

[16] G. Punzi, eConf C 030908 (2003), arXiv:physics/0308063 [physics.data-an].

[17] H. Nakano, SEARCH FOR NEW PHYSICS BY A TIME-DEPENDENT CP VIOLATION ANALYSIS OF THE DECAY $B \rightarrow K_{S}\gamma\gamma$ USING THE BELLE DETECTOR. PhD thesis, Tohoku University (2014) Chapter 4, unpublished.

[18] M. Feindt and U. Kerzel, Nucl. Instrum. Methods Phys. Res. Sect. A 559, 190 (2006).

[19] B. Pal et al. (Belle Collaboration), Phys. Rev. Lett. 116, 161801 (2016).

[20] M. Oreglia, A STUDY OF THE REACTIONS $\psi' \rightarrow \gamma\gamma\psi$. PhD thesis, SLAC-0236, 1980; T. Skwarnicki, A STUDY OF THE RADIATIVE CASCADE TRANSITIONS BETWEEN THE UPSILON-PRIME AND UPSILON RESONANCES. PhD thesis, DESY-F31-86-02, 1986.

[21] A.D. Bukin, arXiv:0711.4449 [physics.data-an] (2007).

[22] H. Albrecht et al. (ARGUS Collaboration), Phys. Lett. B 241, 278-282 (1990).

[23] J. Neyman, Phil. Trans. Roy. Soc. Lond. A 236, 767, 333-380 (1937); Reprinted in A Selection of Early Statistical Papers of J. Neyman, (University of California Press, Berkeley, 1967).

[24] A. Stuart and J.K. Ord, Classical Inference and Relationship, 5th ed., Kendall’s Advanced Theory of Statistics, Vol. 2 (Oxford University Press, New York, 1991); see also earlier editions by Kendall and Stuart.

[25] W.T. Eadie, D. Drijard, F.E. James, M. Roos, and B. Sadoulet, Statistical Methods in Experimental Physics, (NorthHolland, Amsterdam, 1971).

[26] G. J. Feldman and R. D. Cousins, Phys. Rev. D 57, 3873-3889 (1998).

[27] J. Schumann et al. (ARGUS Collaboration), Phys. Rev. D 91, 061802 (2006).

[28] S. Ryu et al. (Belle Collaboration), Phys. Rev. D 89, 072009 (2014).

[29] T. Abe et al. (Belle II Collaboration), arXiv: 1011.0352 [physics.ins-det] (2010).