Journal of Physics: Conference Series **631** (2015) 012038 doi:10.1088/1742-6596/631/1/012038

Exclusive $b \to s\gamma$ decays at Belle

D Dutta and B Bhuyan on behalf of the Belle Collaboration

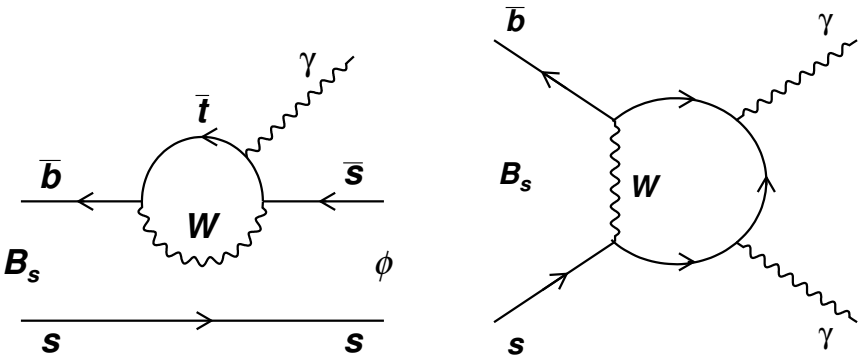
Indian Institute of Technology Guwahati, Assam, India

E-mail: d.dutta@iitg.ac.in; bhuyan@iitg.ac.in

Abstract. We report results on two exclusive $b \to s\gamma$ decays, namely $B_s^0 \to \phi\gamma$ and $B_s^0 \to \gamma\gamma$ at Belle. For this study, we have used 121.4 fb^{-1} of $\Upsilon(5S)$ data collected with the Belle detector at the KEKB asymmetric-energy e^+e^- collider. Calculations within the Standard Model framework predict the $B_s^0 \to \phi\gamma$ branching fraction to be in the range $(3.9 - 4.3) \times 10^{-5}$ and the $B_s^0 \to \gamma\gamma$ branching fraction to be in the range $(0.18 - 2.45) \times 10^{-6}$. The $B_s^0 \to \phi\gamma$ branching fraction is measured as $(3.6 \pm 0.5 \pm 0.3 \pm 0.6) \times 10^{-5}$, where the first uncertainty is statistical, the second is systematic and the last one is due to the uncertainty in the fraction of $B_s^{(*)}\bar{B}_s^{(*)}$ in $b\bar{b}$ events. This result agrees with the theoretical predictions and other experimental results. We observe no statistically significant signal for $B_s^0 \to \gamma\gamma$ and have thus set a 90% confidence-level upper limit on its branching fraction at 3.1×10^{-6}. This improves on the previous result by a factor of about 3 and is the most stringent limit till date.

1. Introduction

$B_s^0 \to \phi\gamma$ and $B_s^0 \to \gamma\gamma$ are exclusive $b \to s\gamma$ and $b \to s\gamma\gamma$ decay processes that are sensitive to physics beyond the Standard Model (SM). The leading order Feynman diagrams for these processes are shown in Figure 1. Precise measurements of the branching fraction of these decays along with other $b \to s\gamma$ modes may help to put stringent limits on the permitted values of Wilson coefficients $\mathcal{C}_7$, which in turn may constrain various new-physics model parameters. The present constraints on $\mathcal{C}_7$ obtained using various decay modes are summarized in Refs. [1, 2].



(a) (b)

Figure 1. Leading order Feynman diagrams for the decays (a) $B_s^0 \to \phi\gamma$ and (b) $B_s^0 \to \gamma\gamma$.

Calculations of the $B_s^0 \to \phi\gamma$ branching fraction within the SM framework yields a value between (3.9 - 4.3) $\times 10^{-5}$ with about 30% uncertainty [3, 4]. This decay was first observed by the Belle experiment using 23.6 fb^{-1} of data collected at the $\Upsilon(5S)$ resonance with the branching fraction (BF) measured to be $(5.7^{+2.2}_{-1.9}) \times 10^{-5}$ [5]. Recently, the LHCb Collaboration has measured the ratio $\mathcal{B}(B^0 \to K^{*0}\gamma)/\mathcal{B}(B_s^0 \to \phi\gamma)$ and using the world-average value of $\mathcal{B}(B^0 \to K^{*0}\gamma)$ [6], the BF for $B_s^0 \to \phi\gamma$ is estimated to be (3.5 $\pm$ 0.4) $\times 10^{-5}$ [7]. These experimental results are in good agreement with the theoretical ones. The $B_s^0 \to \phi\gamma$ BF is further constrained by a good agreement between theory and experimental results on exclusive $b \to s\gamma$ modes [3, 4, 6, 8] and on inclusive $B^0 \to X_s\gamma$ rates [6, 10, 11, 12]. Small contributions from various new physics effects may still remain hidden within the large uncertainties of this mode [13, 14]. The decay $B_s^0 \to \gamma\gamma$ is yet to be observed and its upper limit at 90% confidence level (CL) is estimated to be 8.7 $\times 10^{-6}$ [5]. This result is an order of magnitude larger than the SM predictions, which lie in the range of $(0.18 - 2.45) \times 10^{-6}$ [15, 16, 17]. Results from $B^0 \to X_s\gamma$ decays constrain the $B_s^0 \to \gamma\gamma$ BF in the R-parity conserving scenario [13]. However, presence of processes mediated by λ-irreducible diagrams [18] may increase its BF by more than an order of magnitude in R-parity violating (RPV) SUSY scenarios [13].

2. The Belle Detector and the Dataset

The results reported here are based on 121.4 fb^{-1} of $\Upsilon(5S)$ data collected with the Belle detector [19, 20] at the KEKB [21] asymmetric-energy B-factory at KEK in Japan. The Belle detector consists of a 4-layer silicon vertex detector (SVD) primarily for an efficient vertex reconstruction, a central drift chamber (CDC) for measuring the momenta of charged particles, Aerogel Cherenkov counters (ACC) to distinguish between charged hadrons, mainly kaons and pions, plastic scintillation counters (TOF) for measuring the time of flight of a particle from the interaction point, and an electromagnetic calorimeter (ECL) for efficient detection of photons. All these detector components are located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux return located outside of the coil is instrumented to detect K_L^0 mesons and muons. The detector is described in details elsewhere [19, 20].

The Υ(5S) mass is substantially larger than the $B_s^0\bar{B}_s^0$ production threshold. So, several hadronic processes are possible at the Υ(5S) resonance. The $b\bar{b}$ production cross-section $[\sigma_{b\bar{b}}^{\Upsilon(5S)} = \sigma(e^+e^- \to b\bar{b})]$ at the Υ(5S) resonance is measured to be (0.340 $\pm$ 0.016) nb, by subtracting $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$) component obtained just above the Υ(4S) resonance [22]. The $b\bar{b}$ pairs can hadronize to produce three categories of events:

- events containing 2 strange B i.e., B_s mesons ($e^+e^- \to B_s^{(*)}\bar{B}_s^{(*)}$)
- events containing 2 non-strange B mesons ($e^+e^- \to B^{(*)}\bar{B}^{(*)}(X)$)
- events containing no $B_{(s)}$ mesons ($e^+e^- \to non - B_{(s)}\bar{B}_{(s)}$)

Charge conjugate modes are implied throughout the proceedings. The fraction of $b\bar{b}$ events that hadronize to $B_s^{(*)}\bar{B}_s^{(*)}$ $\left(f_s = \dfrac{\sigma(e^+e^- \to B_s^{(*)}\bar{B}_s^{(*)})}{\sigma(e^+e^- \to b\bar{b})}\right)$ is determined by the following relations:

$$\frac{\mathcal{B}(\Upsilon(5\mathrm{S}) \to D_sX)}{2} = f_s\mathcal{B}(B_s \to D_sX) + (1 - f_s)\mathcal{B}(B \to D_sX) \quad (1)$$

$$\frac{\mathcal{B}(\Upsilon(5\mathrm{S}) \to D^0X)}{2} = f_s\mathcal{B}(B_s \to D^0X) + (1 - f_s)\mathcal{B}(B \to D^0X) \quad (2)$$

via a measurement of the inclusive rates $\mathcal{B}(\Upsilon(5\mathrm{S}) \to D_s(D^0)X)$ and information on the inclusive BF of the $B_s(B) \to D_s(D^0)X$ decays [23]. The value of f_s is measured to be $(17.2 \pm 3.0)\%$ [22].

The $B_s^{(*)}\bar{B}_s^{(*)}$ pairs include $B_s^{*0}\bar{B}_s^{*0}$, $B_s^{*0}\bar{B}_s^0$ and $B_s^0\bar{B}_s^0$. Their relative abundances, measured using $B_s \to D_s\pi$ [22], are:

$$f_{B_s^{*0}\bar{B}_s^{*0}} = \frac{\sigma(e^+e^- \to B_s^{*0}\bar{B}_s^{*0})}{\sigma(e^+e^- \to B_s^{(*)}\bar{B}_s^{(*)})} = (87.0 \pm 1.7)\%, \tag{3}$$

$$f_{B_s^{*0}\bar{B}_s^0} = \frac{\sigma(e^+e^- \to B_s^{*0}\bar{B}_s^0)}{\sigma(e^+e^- \to B_s^{(*)}\bar{B}_s^{(*)})} = (7.3 \pm 1.4)\% and \tag{4}$$

$$f_{B_s^0\bar{B}_s^0} = 1 - f_{B_s^{*0}\bar{B}_s^{*0}} - f_{B_s^{*0}\bar{B}_s^0} \tag{5}$$

The B_s can thus be produced directly from $B_s^0\bar{B}_s^0$, or from the excited states $B_s^{*0}\bar{B}_s^{*0}$ and $B_s^{*0}\bar{B}_s^0$, where the B_s^{*0} mesons decay to the ground state B_s^0 mesons through the emission of a photon.

3. Event Generation

MC simulated events for both signal and background processes are generated to optimize the signal selection criteria, estimate the signal reconstruction efficiency, study background sources and to parameterize the signal and background shapes. This is done in two steps: event generation using the EvtGen package [24] and modeling of the detector response using GEANT3 [25]. Beam related backgrounds are added to the simulated samples during detector simulation. Events are generated according to various physics processes using a decay table, which specifies the decay models, modes, BFs, etc for all possible particles involved in the decay chain. MC simulated events for the decay $B^0 \to K^{*0}\gamma$ are also generated for the calibration purposes.

4. Selection Criteria

The background for these decay modes are light quark-antiquark pairs ($u\bar{u}$, $d\bar{d}$, $s\bar{s}$ and $c\bar{c}$), generic B_s ($B_s^{(*)}\bar{B}_s^{(*)}$) and non-B_s ($B^*\bar{B}^*\pi, \Upsilon(4S)\gamma$) decays of $\Upsilon(5S)$ and QED processes (Bhabha scattering or $e^+e^- \to \gamma\gamma$ events). Two-photon processes resulting from the π^0 and η decays can also contribute to the backgrounds. Loose selection criteria based on track multiplicity, photon energies, numbers of clusters and average cluster energy are applied to filter out the non-B events. Charged tracks are required to originate from the interaction point (IP) with impact parameters $|dz| < 3$ cm and $dr < 0.5$ cm, where $|dz|$ and dr are the distances of closest approach to the IP along the z axis (collinear with the e^+ beam) and in the transverse r-ϕ plane, respectively. Charged kaons are identified with about 85% efficiency by requiring:

$$\mathcal{L}(K/\pi) \equiv \frac{\mathcal{L}_K}{\mathcal{L}_K + \mathcal{L}_\pi} > 0.6$$

where $\mathcal{L}_K$ and $\mathcal{L}_\pi$ are the likelihoods of the track being due to a kaon and a pion, respectively. Tracks failing this requirement are assumed to be due to pions. The likelihoods are calculated based on the dE/dx measurement by the CDC, the Cherenkov light yield in the ACC and the time of flight information from TOF by taking a product of the likelihood functions for three discriminants:

$$\mathcal{L}_i = \mathcal{L}_i^{\mathrm{dE/dx}}.\mathcal{L}_i^{\mathrm{ACC}}.\mathcal{L}_i^{\mathrm{TOF}} (i = K, \pi)$$

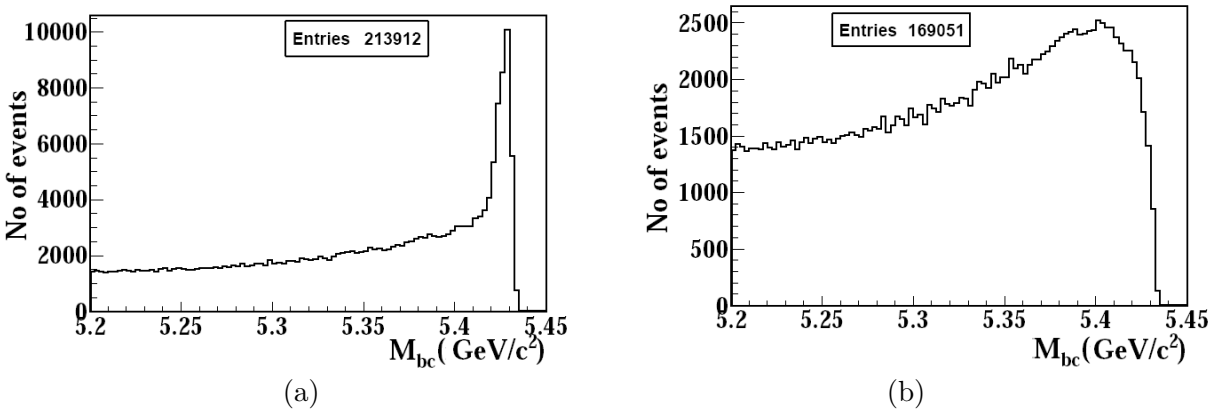
The ϕ meson candidates are reconstructed by combining a pair of oppositely charged kaons having an invariant mass $M(K^+K^-)$ within ± 12 MeV/c^2 ($\pm$ 2.5 σ) of the nominal ϕ mass. No events with multiple ϕ candidates are seen. Similarly, the K^{*0} candidates for the $B^0 \to K^{*0}\gamma$ control sample are reconstructed with oppositely charged kaon and pion candidates by requiring $|M_{K\pi} - m_{K^{*0}}| < 75$ MeV/c^2, where $M_{K\pi}$ and $m_{K^{*0}}$ are the invariant mass of the kaon-pion pair and the nominal K^{*0} mass, respectively. To remove the multiple K^{*0} candidates, K/π

vertex fit is performed and events with the least χ^2 value is chosen as the best K^{*0} candidate. Photons are reconstructed by identifying energy deposits in the ECL that are not matched to any charged track. Candidate photons are required to have a minimum energy (E_γ) of 100 MeV. Electromagnetic and hadronic showers have different shapes in both the transverse and longitudinal directions. Showers caused by hadrons grow wider than the electromagnetic ones due to multiple showers caused by hadrons. Thus, to reject neutral hadrons like K_L^0 mesons and neutrons and merged photons from π^0 we require the ratio of the energy deposited by a photon candidate in a (3×3) to that in (5×5) ECL crystal array (E_9/E_{25}) to be greater than 0.95. Daughter photons from π^0 and η decays are rejected on the basis of likelihood information [$P_{\pi^0}(\gamma)$ and $P_\eta(\gamma)$] obtained using energy and polar angle of the photons and the diphoton invariant mass, calculated by combining the candidate photon with any other photon in the event. Owing to the long decay time of the CsI crystals, the offtime QED events can leave two perfectly back-to-back clusters in the ECL and can very well mimic the $B_s^0 \to \gamma\gamma$ events. To remove these events, we require the timing characteristics of the clusters used for photon reconstruction to be consistent with the beam collision time, determined at the trigger level for the candidate event. As the timing information is not simulated, a selection criterion based on MC truth information is used to remove such background from the simulated samples with a matching efficiency. Figures 2 and 3 show the $M_{\rm bc}$ distributions before and after applying the MC truth information and timing criteria to background MC and real data, respectively, for the $B_s^0 \to \gamma\gamma$ analysis. For $B_s^0 \to \gamma\gamma$, the candidate photons are selected only from the barrel region ($33^\circ < \theta < 128^\circ$, θ being the photon polar angle), as there is a high level of beam background in the endcaps and it is very unlikely that one photon will be in the barrel and the other in the endcaps for a B_s^0 meson decaying at rest. The B_s^0 candidates are formed by combining a ϕ with a photon candidate for $B_s^0 \to \phi\gamma$ mode or by combining two photon candidates for the $B_s^0 \to \gamma\gamma$ mode. B_s^0 mesons are selected using the beam constrained mass ($M_{\rm bc}$) and energy difference (ΔE), defined as:

$$M_{\rm bc} = \sqrt{(E_{\rm beam}^{\rm CM})^2 - (p_{B_s}^{\rm CM})^2} \tag{6}$$

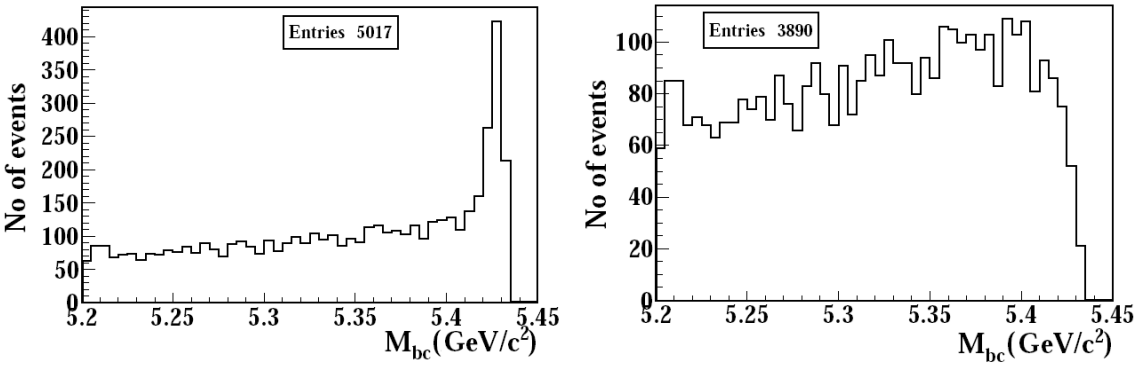
$$\Delta E = E_{B_s}^{\rm CM} - E_{\rm beam}^{\rm CM} \tag{7}$$

where, $E_{\rm beam}^{\rm CM}$ is the beam energy, $E_{B_s^0}^{\rm CM}$ and $p_{B_s^0}^{\rm CM}$ are the energy and momentum of the B_s^0 candidate, with all quantities evaluated at the e^+e^- center-of-mass (CM) frame. Signal candidates are required to satisfy $M_{\rm bc} > 5.3$ GeV/c^2 and $-0.4 < \Delta$E < 0.1 GeV for $B_s^0 \to \phi\gamma$



(a) (b)

Figure 2. $M_{\rm bc}$ distributions for background MC (a) before and (b) after the photonID criteria for $B_s^0 \to \gamma\gamma$.



(a) (b)

Figure 3. $M_{\rm bc}$ distributions for 1.86 fb^{-1} real data (a) before and (b) after the timing criteria for $B_s^0 \to \gamma\gamma$.

and $\mathrm{M_{bc}} > 5.3$ GeV/c^2 and $-0.7 < \Delta\mathrm{E} < 0.2$ GeV for the $B_s^0 \to \gamma\gamma$ mode. No events with multiple B_s^0 candidates are seen in signal MC samples, and negligible cases ($< 1.0\%$) of multiple B_s^0 are observed for background MC and real data for both $B_s^0 \to \phi\gamma$ and $B_s^0 \to \gamma\gamma$ analyses. The K^{*0} candidates are combined with photon candidates to form B^0 candidates by applying $M_{\rm bc} > 5.0$ GeV/c^2 and $-0.4 < \Delta\mathrm{E} < 0.2$ GeV for the $B^0 \to K^{*0}\gamma$ control sample. Multiple B_s^0 and B^0 candidates coming from a particular event ($< 1.0\%$) are removed by retaining the one with the more energetic photon.

5. Continuum Suppression

The dominant background for both the decay modes are from $e^+e^- \to q\bar{q}$ continuum events where, $q = u,\ d,\ s$ or c. The mass of a $B_s^{(*)}\bar{B}_s^{(*)}$ pair is less than that of the $\Upsilon(5\mathrm{S})$ resonance by only a few MeVs. So the B_s^0 produced at the $\Upsilon(5\mathrm{S})$ resonance has a very low momentum and its decay products have a random direction, i.e., they have a spherical topology. Since for continuum events, a smaller proportion of e^+e^- energy goes towards particle production, these events have a higher momentum and hence are topologically quite distinct from $\Upsilon(5\mathrm{S}) \to B_s^{(*)}\bar{B}_s^{(*)}$ events. Their decay products are distributed mostly along the beam direction. To reduce these backgrounds, event shape variables such as the modified Fox-Wolfram moments [26] and $|\cos(\theta_{\rm thrust})|$ are used as inputs to a Neural Network (NN) [27] based classifier, where the thrust axis is defined as the direction which maximises the sum of the longitudinal momenta of the particles in the decay. The classifier uses the information of all the discriminating variables and provides a single quantity (NN output) which can then be used as a discriminant instead of using all input variables separately. The NN output ($\mathcal{C}_{\rm NB}$) tends to peak at 1 for signal-like events and at -1 for background-like events. This quantity is used as an additional dimension in the unbinned extended maximum likelihood fit to extract the signal yield for $B_s^0 \to \phi\gamma$ mode. As $\mathcal{C}_{\rm NB}$ peaks sharply at 1 and -1, it is very difficult to model it with a simple analytic function. Thus, to improve the modeling, a modified NeuroBayes output is calculated as:

$$\mathcal{C}'_{\rm NB} = \log\left(\frac{\mathcal{C}_{\rm NB} - \mathcal{C}_{\rm NB_{min}}}{\mathcal{C}_{\rm NB_{max}} - \mathcal{C}_{\rm NB}}\right), \tag{8}$$

after rejecting events with $\mathcal{C}_{\rm NB} < \mathcal{C}_{\rm NB_{min}}$, where $\mathcal{C}_{\rm NB_{min}} = -0.6$ and $\mathcal{C}_{\rm NB_{max}} \sim 1$ are the lower and upper limits of $\mathcal{C}_{\rm NB}$ for the events used in the fit. For the $B_s^0 \to \gamma\gamma$ mode, an optimized criterion of $\mathcal{C}_{\rm NB} > 0.77$ is applied, since considerable correlations are observed between $\mathcal{C}_{\rm NB}$ with $M_{\rm bc}$ and $\Delta\mathrm{E}$ distributions. The optimal criterion is taken to be the value for which the signal

significance is maximum, where the signal significance defined as:

$$\text{Signal significance} = \frac{\text{S}}{\sqrt{\text{S} + \text{B}}} \tag{9}$$

after applying the cut at different $\mathcal{C}_{\text{NB}}$ values over the range -1 to $+1$. $\text{S} = w_s \times s$ and $\text{B} = w_b \times b$ are the expected signal and background events, w_s and w_b are the weight factors, determined using the signal and background expectations in data, s and b being the number of signal and background events remaining in the MC sample, after applying a particular selection criteria on $\mathcal{C}_{\text{NB}}$.

6. Extended Maximum Likelihood Fit

We perform a four-dimensional (two-dimensional) unbinned extended maximum likelihood fit involving M_{bc}, ΔE, $\cos\theta_{\text{hel}}$ and $\mathcal{C}'_{\text{NB}}$ (M_{bc} and ΔE) to extract the $B^0_s \to \phi\gamma$ ($B^0_s \to \gamma\gamma$) signal yields. The ϕ helicity angle (θ_{hel}) is the angle between B^0_s and any of the ϕ daughters in the ϕ rest frame. The total fit probability distribution function (PDF) consists of two components: signal and $q\bar{q}$ background. The signal further consists of three components: signal coming from $B^{*0}_s\bar{B}^{*0}_s$, $B^{*0}_s\bar{B}^0_s$ and $B^0_s\bar{B}^0_s$ decays, their relative fractions being fixed to the values reported in Ref. [22]. We combine the backgrounds arising from B_s and non-B_s decays with the $q\bar{q}$ continuum as they have a small contribution and have no peaking structure in the signal region. For the $B^0_s \to \phi\gamma$ mode, the fit regions chosen for different variables are listed below:

- M_{bc} $\longrightarrow$ (5.3 - 5.434) GeV/c^2
- ΔE $\longrightarrow$ (-0.4 - 0.1) GeV
- $\cos(\theta_{\text{hel}})$ $\longrightarrow$ (-1.0 - 1.0)
- $\mathcal{C}'_{\text{NB}}$ $\longrightarrow$ (-10.0 - 10.0)

For $B^0_s \to \gamma\gamma$ analysis, the fit regions chosen for M_{bc} and ΔE are listed below:

- M_{bc} $\longrightarrow$ (5.3 - 5.434) GeV/c^2
- ΔE $\longrightarrow$ (-0.7 - 0.4) GeV

The signal and background PDFs are determined using MC samples. The PDF for each component is represented by the product of one-dimensional functions since the correlations among the variables are negligible. For the $B^0_s \to \phi\gamma$ mode, the M_{bc}, ΔE, $\cos\theta_{\text{hel}}$ and $\mathcal{C}'_{\text{NB}}$ distributions of all three signal components are modeled with the sum of a Crystal Ball (CB) [28] and Gaussian function having a common mean, a CB function, a $\sin^2\theta_{\text{hel}}$ distribution and the sum of two Gaussian functions, respectively. The background PDFs are described by an ARGUS function [29] for M_{bc} having its endpoint fixed at 5.434 GeV/c^2, a first-order Chebychev polynomial for ΔE, a parabola for $\cos\theta_{\text{hel}}$ and a Gaussian function for $\mathcal{C}'_{\text{NB}}$. For the $B^0_s \to \gamma\gamma$ mode, the M_{bc} distributions of all three signal components are parameterized with a combination of CB and Gaussian functions having a common mean and the ΔE distributions are modeled with CB functions. The background is described by an ARGUS function having its endpoint fixed for M_{bc} and a first-order Chebychev polynomial for ΔE. For both analyses, the signal parameters are determined from MC events except for the means and widths of the M_{bc} and ΔE distributions describing the $B^*_s\bar{B}^*_s$ component. For both the decay modes, the widths of M_{bc} and ΔE for the $B^{*0}_s\bar{B}^{*0}_s$ component are calibrated using corrections obtained from the $B^0 \to K^{*0}\gamma$ control sample. As a cross-check, the branching fraction for $B^0 \to K^{*0}\gamma$ is measured and found to be in good agreement with its world average [6]. The M_{bc} mean of the $B^{*0}_s\bar{B}^{*0}_s$ component is similarly adjusted using M_{bc} mean corrections taken from the $B_s \to D_s\pi$ analysis [22]. No calibration is done for the other two components as their contribution is small. The ΔE mean of the $B^{*0}_s\bar{B}^{*0}_s$ component is allowed to float in the $B^0_s \to \phi\gamma$ analysis. For the $B^0_s \to \gamma\gamma$ mode,

we have fixed the ΔE mean to the signal MC value, as the correction to the ΔE mean obtained from the $B_s^0 \to \phi\gamma$ analysis is found to be within the statistical uncertainties. The uncertainty associated with this procedure is included as a systematic uncertainty for this mode. All the background parameters except the ARGUS endpoint is kept free. In total, we have nine free parameters for the $B_s^0 \to \phi\gamma$ fit: the ARGUS shape parameter, Chebychev polynomial coefficient, both parameters of $\cos(\theta_{\rm hel})$ background, mean and sigma of $\mathcal{C}'_{\rm NB}$ background, ΔE mean, signal and background yields. We have four free parameters for the $B_s^0 \to \gamma\gamma$ fit: the ARGUS shape parameter, Chebychev background parameter, signal and background yields.

7. Signal Significance, Branching Fraction and Upper Limit

In total, we have observed 91^{+14}_{-13} $B_s^0 \to \phi\gamma$ signal events coming from all 3 signal components with a significance of 10.7 σ including systematics. The signal significance is computed assuming a null or background only hypothesis as $\mathrm{sign}(N_{\rm signal})\sqrt{-2\ln(\mathcal{L}_0/\mathcal{L}_{\rm max})}$, where $N_{\rm signal}$ is the signal yield obtained from the fit, $\mathcal{L}_0$ is the likelihood value of the fit when the signal yield is fixed to 0 and $\mathcal{L}_{\rm max}$ is the maximum likelihood value of the likelihood fit. Additive systematic uncertainties are incorporated by convolving the statistical likelihood curve with a Gaussian function of width equal to the total additive systematics. The branching fraction for the decay $B_s^0 \to \phi\gamma$ is calculated as:

$$\mathcal{B}(B_s^0 \to \phi\gamma) \;=\; \frac{N(B_s^0 \to \phi\gamma)}{\mathcal{B}(\phi \to K^+K^-) \times \epsilon \times N(B_s^0)}, \;\; \text{with } N(B_s^0) \;=\; 2 \times f_s \times \sigma_{b\bar{b}}^{\Upsilon(5S)} \times \mathcal{L}_{\rm int} \quad (10)$$

where, $N(B_s^0 \to \phi\gamma)$ is the signal yield of $B_s^0 \to \phi\gamma$; $\mathcal{B}(\phi \to K^+K^-)$ is the branching fraction of $\phi \to K^+K^-$, ϵ is the signal selection efficiency of $B_s^0 \to \phi\gamma$ signal, f_s is the fraction of $B_s^{(*)}B_s^{(*)}$ events in the $b\bar{b}$ sample, $\sigma_{b\bar{b}}^{\Upsilon(5S)}$ and $\mathcal{L}_{\rm int}$ are the $b\bar{b}$ production cross-section and integrated luminosity at the $\Upsilon(5S)$ energy. The $B_s^0 \to \phi\gamma$ BF is measured as $(3.6 \pm 0.5 \pm 0.3 \pm 0.6) \times 10^{-5}$, where the first uncertainty is statistical, the second is systematic and the third is due to the uncertainty in f_s. We observe no statistically significant signal for the decay $B_s^0 \to \gamma\gamma$ and have rather measured the single-event sensitivity for $B_s^0 \to \gamma\gamma$ to be 0.5×10^{-6}. We use a Bayesian approach and integrate the likelihood curve from 0 to 90% of the total integral under the curve to obtain the 90% CL upper limit of 3.1×10^{-6} on the $B_s^0 \to \gamma\gamma$ branching fraction. The fit results are listed in Table 1, while the fit projections are shown in Figs. 4 and 5. The fits are projected onto the $B_s^{*0}\bar{B}_s^{*0}$ signal region defined as $M_{\rm bc} > 5.4$ GeV/c^2, $-0.2 < \Delta E < 0.02$ GeV, $|\cos\theta_{\rm hel}| < 0.8$ and $0.0 < \mathcal{C}'_{\rm NB} < 10.0$ for $B_s^0 \to \phi\gamma$ and $M_{\rm bc} > 5.4$ GeV/c^2 and $-0.3 < \Delta E < 0.05$ GeV for $B_s^0 \to \gamma\gamma$.

8. Systematic Uncertainties

The systematic uncertainties are classified into two types: additive and multiplicative. The multiplicative uncertainties do not affect the signal yield or signal significance of the decay channel but affects its BF. For both the modes, these uncertainties arise due to the uncertainty in signal reconstruction efficiency and on the number of B_s mesons. The former one comprises

Table 1. Results of the $B_s^0 \to \phi\gamma$ and $B_s^0 \to \gamma\gamma$ analyses.

	$B_s^0 \to \phi\gamma$	$B_s^0 \to \gamma\gamma$
ϵ (%)	36.1 ± 0.1	14.0 ± 0.1
N	91^{+14}_{-13}	$-3.9^{+3.7}_{-2.6}$
$\mathcal{B}(10^{-6})$	$36 \pm 5(\text{stat.}) \pm 3(\text{syst.}) \pm 6(f_s)$	< 3.1 (90% CL)

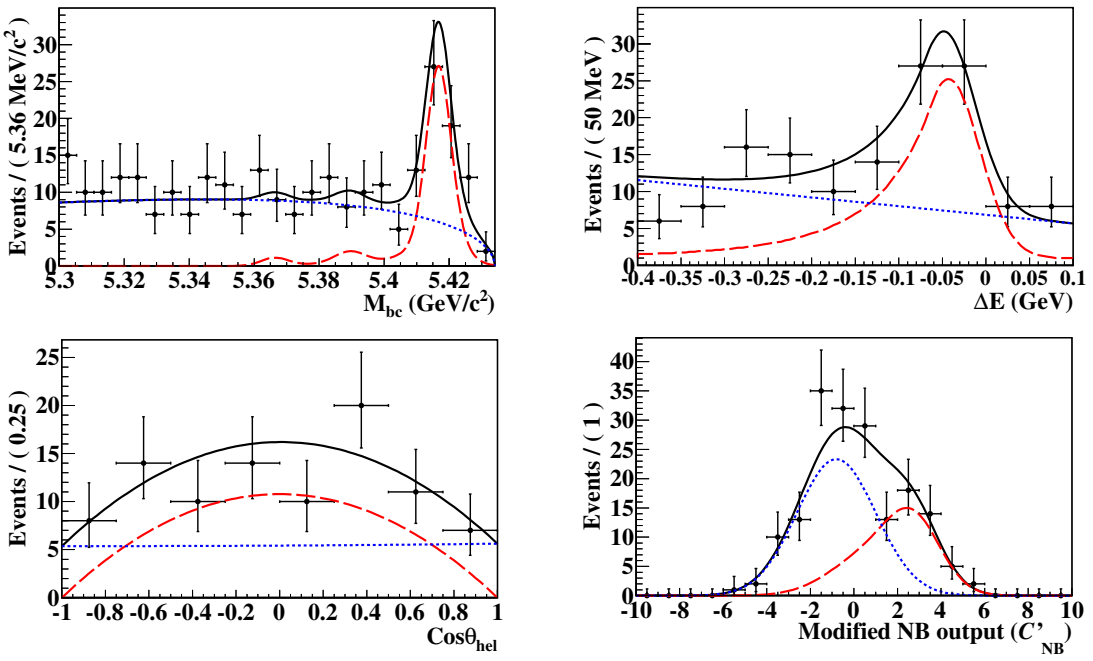


Figure 4. Data fits for the $B_s^0 \to \phi\gamma$ analysis. The projections are shown only for events inside the $B_s^{*0}\bar{B}_s^{*0}$ signal region except for the plotted variable. The points with error bars represent the data, the solid black curve represents the total fit function, the red dashed (blue dotted) curve represents the signal (continuum background) contribution.

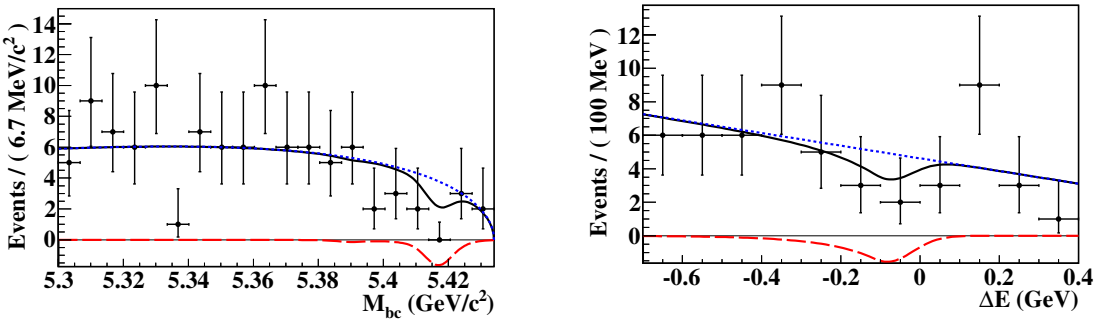


Figure 5. Data fits for the $B_s^0 \to \gamma\gamma$ analysis. The projections are shown only for events inside the $B_s^{*0}\bar{B}_s^{*0}$ signal region except for the plotted variable. The points with error bars represent the data, the solid black curve represents the total fit function, the red dashed (blue dotted) curve represents the signal (continuum background) contribution.

of the uncertainty associated with the photon reconstruction efficiency, kaon identification efficiency, tracking efficiency, the requirement on $\mathcal{C}_{\rm NB}$ that is estimated by comparing the efficiencies in data and MC simulations with the $B^0 \to K^{*0}\gamma$ control sample and limited MC statistics. The uncertainties due to kaon identification and tracking efficiency are 1.3% and 0.3%, respectively. These are obtained using control samples of $D^{*+} \to D^0\pi^+_{\rm slow} \to K^-\pi^+\pi^+_{\rm slow}$ and $D^{*+} \to D^0\pi^+_{\rm slow}, D^0 \to K^0_S\pi^+\pi^-, K^0_S \to \pi^+\pi^-$ decays, respectively. For $B_s^0 \to \phi\gamma$, the uncertainty due to the $\phi \to K^+K^-$ branching fraction also adds to the multiplicative systematics. This is taken from Ref. [9]. The uncertainty on the number of B_s mesons is the quadratic sum of the uncertainties in the integrated luminosity, $\sigma^{\Upsilon(5S)}_{b\bar{b}}$ and f_s. The additive

uncertainties reduce the signal significance of the observed peak and changes the BF of the decay. That arise due to the fit procedure, which includes the uncertainty due to PDF parametrization and fit bias, respectively. The uncertainty due to PDF parametrization is estimated by varying each fixed parameter of the fit PDF by $\pm 1\sigma$ and estimating the variation in the signal yield and then adding them in quadrature. The calibration uncertainties are also included in the calculation. To estimate the uncertainty due to fit bias, pseudo-experiments are generated according to the PDFs obtained from the fits to data i.e., by fixing all parameters including the background shape parameters and signal and background yields. Events generated from the pseudo-experiments are then fitted to obtain the signal yield and residual distributions. The observed biases of -0.28 ± 0.08 and -0.10 ± 0.07 for $B_s^0 \to \phi\gamma$ and $B_s^0 \to \gamma\gamma$ are corrected and their uncertainties are assigned as systematic uncertainties. The systematic uncertainties are listed in Table 2.

Table 2. Summary of systematic uncertainties.

Additive systematic uncertainties (events)		
Source	$B_s^0 \to \phi\gamma$	$B_s^0 \to \gamma\gamma$
PDF parameterization	$^{+1.6}_{-1.7}$	± 0.4
Fit bias	± 0.1	± 0.1
Total (quadratic sum)	$^{+1.6}_{-1.7}$	± 0.4
Multiplicative systematic uncertainties (%)		
Source	$B_s^0 \to \phi\gamma$	$B_s^0 \to \gamma\gamma$
Photon reconstruction efficiency	2.2	2×2.2
Kaon identification efficiency	2.6	-
Tracking efficiency	0.7	-
$\mathcal{C}_{\rm NB}$ requirement	4.8	8.7
MC statistics	0.2	0.4
$\mathcal{B}(\phi \to K^+K^-)$	1.0	-
$\mathcal{L}_{\rm int}$	1.3	
$\sigma_{b\bar{b}}^{\Upsilon(5S)}$	4.7	
f_s	17.4	
Total (quadratic sum)	19.1	20.6

9. Results and Conclusions

To summarize, we have used the entire Belle $\Upsilon(5S)$ dataset in our study and have measured the $B_s^0 \to \phi\gamma$ branching fraction to be $\mathcal{B}(B_s^0 \to \phi\gamma) = [3.6 \pm 0.5(\text{stat.}) \pm 0.3(\text{syst.}) \pm 0.6(f_s)] \times 10^{-5}$. This measurement supersedes the earlier Belle result [5], and is consistent with theoretical predictions [3, 4] and a recent LHCb measurement [7]. For the $B_s^0 \to \gamma\gamma$ mode, in the absence of any statistically significant signal, we have set the 90% CL upper limit on its BF at 3.1×10^{-6}. This result improves on the previously published result by a factor of about three and is consistent with the expected sensitivity for our data sample. This also rules out large contributions from RPV SUSY in this channel. This decay will possibly be observed at upcoming flavor physics experiments like Belle II with a dedicated run at Υ(5S) resonance.

Acknowledgements
The authors acknowledge the funding received from Department of Science and Technology, Government of India to support this work.

References
[1] W. Altmannshofer and D. M. Straub, JHEP **8**, 1 (2012).
[2] W. Altmannshofer, P. Paradisi and D. M. Straub, JHEP **4**, 1 (2012).
[3] A. Ali, B. D. Pecjak, and C. Greub, Eur. Phys. J. C **55**, 577 (2008).
[4] P. Ball, G. W. Jones, and R. Zwicky, Phys. Rev. D **75**, 054004 (2007).
[5] J. Wicht *et al.* (Belle Collaboration), Phys. Rev. Lett. **100**, 121801 (2008).
[6] Y. Amhis *et al.* (Heavy Flavor Averaging Group), arXiv:1207.1158 (2012).
[7] R. Aaij *et al.* (LHCb Collaboration), Nucl. Phys. B **867**, 1 (2013).
[8] A. Ali and A. Y. Parkhomenko, Eur. Phys. J. C **23**, 89 (2002).
[9] K. A. Olive *et al.* (Particle Data Group), Chin. Phys. C**38**, 090001 (2014).
[10] M. S. Alam *et al.* (CLEO Collaboration), Phys. Rev. Lett. **74**, 2885 (1995).
[11] M. Misiak *et al.*, Phys. Rev. Lett. **98**, 022002 (2007).
[12] M. Misiak and M. Steinhauser, Nucl. Phys. B **764**, 62 (2007).
[13] A. Gemintern, S. Bar-Shalom, and G. Eilam, Phys. Rev. D **70**, 035008 (2004).
[14] P. Colangelo, F. De Fazio, R. Ferrandes, and T. Pham, Phys. Rev. D **77**, 055019 (2008).
[15] L. Reina, G. Ricciardi, and A. Soni, Phys. Rev. D **56**, 5805 (1997).
[16] S. W. Bosch and G. Buchalla, JHEP **2002**, 054 (2002).
[17] R. Mohanta and A. Giri, Phys. Rev. D **85**, 014008 (2012).
[18] A. Gemintern, S. Bar-Shalom, G. Eilam, and F. Krauss, Phys. Rev. D **67**, 115012 (2003).
[19] A. Abashian *et al.* (Belle Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 117 (2002).
[20] J. Brodzicka *et al.* (Belle Collaboration), PTEP **04D001**, (2012).
[21] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods Phys. Res., Sect. A **499**, 1 (2003).
[22] S. Esen *et al.* (Belle Collaboration), Phys. Rev. D **87**, 031101 (R) (2013).
[23] A. Drutskoy *et al.* (Belle Collaboration), Phys. Rev. Lett. **98**, 052001 (2007).
[24] D. J. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 152 (2001).
[25] R. Brun *et al.*, GEANT 3.21. Report No. CERN DD/EE/84-1 (1984).
[26] 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).
[27] M. Feindt and U. Kerzel, Nucl. Instrum. Methods Phys. Res., Sect. A **559**, 190 (2006).
[28] T. Skwarnicki, Ph.D. thesis, Institute for Nuclear Physics, Krakow 1986; DESY Internal Report, DESY F31-86-02 (1986).
[29] H. Albrecht *et al.* (ARGUS Collaboration), Phys. Lett. B **241**, 278 (1990).