We study the rare hadronic transitions $\Upsilon(2S) \to \Upsilon(1S)\eta$ and $\Upsilon(2S) \to \Upsilon(1S)\pi^0$ using a sample of $158 \times 10^6 \Upsilon(2S)$ decays collected with the Belle detector at the KEKB asymmetric-energy $e^+e^-$ collider. We measure the ratios of branching fractions $\mathcal{B}(\Upsilon(2S) \to \Upsilon(1S)\eta)/\mathcal{B}(\Upsilon(2S) \to \Upsilon(1S)\pi^+\pi^-) = (1.99\pm0.14\text{ (stat)}\pm0.11\text{ (syst)}) \times 10^{-3}$ and $\mathcal{B}(\Upsilon(2S) \to \Upsilon(1S)\pi^+\pi^-)/\mathcal{B}(\Upsilon(2S) \to \Upsilon(1S)\pi^0) < 2.3 \times 10^{-4}$ at the 90% confidence level (CL).

Assuming the value $\mathcal{B}(\Upsilon(2S) \to \Upsilon(1S)\pi^+\pi^-) = (17.92\pm0.26\%)$, we obtain $\mathcal{B}(\Upsilon(2S) \to \Upsilon(1S)\eta) = (3.57\pm0.25\text{(stat)} \pm0.21\text{(syst)}) \times 10^{-4}$ and $\mathcal{B}(\Upsilon(2S) \to \Upsilon(1S)\pi^0) < 4.1 \times 10^{-5}$ (90% CL).

PACS numbers: 14.40.Pq, 13.25.Gv

In recent years, hadronic transitions between quarkonia have led to an impressive series of discoveries \cite{1}: $X(3872)$, $Y(4260)$, as well as $h_c$ and $h_b$ were observed in transitions either from or to the $\psi$ and $\Upsilon$ states. The phenomenology of these transitions is commonly described with the QCD Multipole Expansion formalism (QCDME) \cite{2,3}, which allows one to classify the transitions in a series of chromoelectric and chromomagnetic multiplets. In particular, theoretical predictions for $\eta$ and $\pi^0$ transitions \cite{4,5} among states are being challenged by experimental measurements. The $\eta$ and $\pi^0$ transitions between vector bottomonia should be mediated either by two M1 gluons or by one E1 and one M2 gluon: both cases imply a spin flip of the $b$ quark. The corresponding amplitude should scale as $1/m_b$, and its measurement yields information about the chromomagnetic moment of the $b$ quark.

By scaling from the $\psi(2S) \to J/\psi\eta$ transition, one expects a transition width of: $\Gamma[\psi(2S) \to \Upsilon(1S)\eta] = 0.0025 \times \Gamma[\psi(2S) \to J/\psi\eta]$, and therefore a ratio of branching fractions $\mathcal{R}_{\eta,\pi^+\pi^-} = \mathcal{B}(\Upsilon(2S) \to \Upsilon(1S)\eta)/\mathcal{B}(\Upsilon(2S) \to \Upsilon(1S)\pi^+\pi^-) \approx 4.4 \times 10^{-3}$ \cite{6}. From \cite{4}, one can calculate $\mathcal{R}_{\eta,\pi^+\pi^-} = 2.3 \times 10^{-3}$ assuming the $b$ quark mass to be $m_b = 4.67$ GeV/$c^2$ \cite{7}; the prediction from \cite{5} is $\mathcal{R}_{\eta,\pi^+\pi^-} = 1.7 \times 10^{-3}$. A further suppression is expected for the $\pi^0$ transition, which violates isospin; here, theory predicts $\Gamma[\Upsilon(2S) \to \Upsilon(1S)\pi^0] = 0.16 \times \Gamma[\psi(2S) \to \Upsilon(1S)\eta]$.

The $\Upsilon(2S) \to \Upsilon(1S)\pi^0$ transitions have been studied by BaBar \cite{8,9} and CLEO \cite{6}; the measured branching fractions are either unexpectedly large ($\Upsilon(4S)$) or too small ($\Upsilon(2S)$ and $\Upsilon(3S)$). The parameters of the quark wave functions must be changed by more than 15% in order to account for these discrepancies \cite{10}. Searches for the $\pi^0$ transitions have only yielded upper limits \cite{6,9}.

We report here a new measurement of the transition $\Upsilon(2S) \to \Upsilon(1S)\eta$ and a search for $\Upsilon(2S) \to \Upsilon(1S)\pi^0$ using the Belle detector at the KEKB $e^+e^-$ collider \cite{11}. The $\Upsilon(1S)$ is reconstructed in both the $e^+e^-$ and $\mu^+\mu^-$ decay modes; we reconstruct the $\eta$ meson via its decay to $\gamma\gamma$ and $\pi^+\pi^-\pi^0$, and the $\pi^0$ in the $\gamma\gamma$ final state. As a normalization sample, we reconstruct the dominant transition $\Upsilon(2S) \to \Upsilon(1S)\pi^+\pi^-$, which has a branching fraction of $(17.92\pm0.26\%)$ \cite{7}. The data sample for this analysis includes an integrated luminosity of 24.7 fb$^{-1}$ at the $\Upsilon(2S)$ resonance peak, corresponding to $(158\pm4) \times 10^6 \Upsilon(2S)$ decays, and an additional 1.7 fb$^{-1}$ at $\sqrt{s}=9.993$ GeV used to study the QED continuum backgrounds.

The Belle detector is described in detail elsewhere \cite{12}. Here, we summarize the features that are relevant to the current analysis. The momentum of each charged track is measured using a four-layer double-sided Silicon Vertex Detector and a 50 layer Central Drift Chamber, embedded in a 1.5 T magnetic field, which allows tracking of charged particles with transverse momentum as low as 50 MeV/$c$. Electrons and photons are detected in a large array of CsI(Tl) crystals (ECL) also located inside the solenoid coil. An iron flux return located outside the coil is instrumented to detect $K_L^0$ mesons and to identify
Monte Carlo (MC) samples of the signal and of the dominant peaking backgrounds were generated using EvtGen \cite{13}. Dipion transitions were generated assuming that the amplitude is dominated by the S-wave contribution. QED continuum processes $e^+e^- \rightarrow (e^+e^- \text{ or } \mu^+\mu^-)$ + $n\gamma$ were generated using KKMC \cite{14}. The detector response was simulated using GEANT3 \cite{15}, and beam backgrounds were accounted for using random triggers taken during each period of data taking. Final state radiation effects are accounted for by using PHOTOS \cite{16} in EvtGen simulations.

Charged tracks with momentum $p^*$ in the center-of-mass frame of the colliding $e^+e^-$ pair (CM frame) greater than 4 GeV/c are selected as candidate leptons from $\Upsilon(1S)$ decay. In the following text, all the quantities computed in the CM frame are denoted with an asterisk. Electrons and muons are identified by the ratio $R_{e,\mu}$ between the electron (muon) hypothesis likelihood and the hadronic hypothesis likelihood. A track is identified as a lepton if $R_e$ or $R_\mu$ is above a threshold value of 0.2 and then as a muon if $R_\mu > R_e$. The identification efficiency is 93.2\% for electrons and 92.6\% for muons. Pairs of lepton candidates with opposite charge and an invariant mass in the range $9.0 \text{ GeV}/c^2 < M(\ell^+\ell^-) < 9.8 \text{ GeV}/c^2$ are then selected for further analysis.

In order to reduce the effect of final state radiation (FSR) and bremsstrahlung, the momentum of all photons detected in the ECL within 200 mrad of each leptonic track is added to its momentum. A mass-constrained kinematic fit performed on the $\Upsilon(1S)$ candidate lepton pair is required to have a confidence level $CL_{1S} > 10^{-5}$.

A requirement on the polar angle in the CM frame of the $e^-$ track with respect to the beam direction, $\cos(\theta_{e^-}^*) < 0.5$, is imposed in the reconstruction of the $\eta \rightarrow \gamma \gamma$, $\Upsilon(1S) \rightarrow e^+e^-$ final state in order to suppress singly- or doubly-radiative Bhabha events, which represent the dominant QED background for this channel. The Bhabha requirement is not included in the $\Upsilon(2S) \rightarrow \Upsilon(1S)\pi^0$ analysis since the $CL_{1S}$ requirement provides sufficient suppression.

Dipion candidates used both for $\eta \rightarrow \pi^+\pi^-\pi^0$ and the $\Upsilon(2S) \rightarrow \Upsilon(1S)\pi^+\pi^-$ transition are composed of oppositely charged pairs of tracks, each of which has a distance of closest approach of less than 1.5 cm (0.5 cm) in the axial (transverse) direction relative to the beam line. The cosine of the angle in the CM frame between these tracks is required to be less than 0.6 in order to reject events with photons that convert in the inner detectors.

In the $\eta \rightarrow \pi^+\pi^-\pi^0$ analysis, each photon produced in the $\pi^0$ decay must have $E_{\gamma\text{lab}} > 57 \text{ MeV}$, the optimal threshold for the rejection of photons arising from beam background, and $E^* < 220 \text{ MeV}$. The $\pi^0$ candidate is then selected as the $\gamma\gamma$ pair with invariant mass closest to the nominal $\pi^0$ mass \cite{7}. The threshold values are optimized by maximizing the figure of merit, defined as

$$FoM = \frac{s}{\sqrt{b}}$$

where $s$ is the signal yield and $b$ is the background contribution.

In the $\eta \rightarrow \gamma\gamma$ analysis, photons with $180 \text{ MeV} < E^* < 360 \text{ MeV}$ are subjected to the requirement on the opening angle in the CM frame of $\cos(\theta_{\gamma\gamma}^*) < -0.88$ in order to reject combinatorial background. Events with more than one $\eta$ candidate that satisfies these conditions are found to be a negligible fraction of the total Monte Carlo sample, and are rejected without introducing any further selection.

The planarity of the event is exploited in order to select the $\Upsilon(2S) \rightarrow \Upsilon(1S)\pi^0$ decay. We select the pair of photons with CM momentum $p_{\gamma\gamma}^*$ that minimizes the scalar product $(\hat{p}_{\gamma\gamma}^* \cdot \hat{u}_{\ell^+\ell^-})/|p_{\gamma\gamma}^*|$ (where $\hat{u}_{\ell^+\ell^-}$ is a vector normal to the plane formed by the dilepton pair in the CM frame) as the $\pi^0 \rightarrow \gamma\gamma$ candidate.

After $\eta$ or $\pi^0$ selection, the $\Upsilon(1S)$ and $\gamma\eta$ or $\pi^+\pi^-\pi^0$ are subject to a second kinematic fit, constraining them to have the $\Upsilon(2S)$ invariant mass. The minimum confidence level $CL_{2S}$ is optimized for each decay mode using MC samples. For the $\eta \rightarrow \gamma\gamma$, $CL_{2S}$ is required to be greater than $6 \times 10^{-4}$ ($2 \times 10^{-3}$) for $\mu^+\mu^-(e^+e^-)$ events; for $\eta \rightarrow \pi^+\pi^-\pi^0$, the event is accepted if the fit converges. Finally, in the analysis of $\Upsilon(2S) \rightarrow \Upsilon(1S)\pi^0$, $CL_{2S}$ must be greater than $10^{-5}$ for both $\Upsilon(1S)$ decay modes.

Since the signal events are fully reconstructed and the total momentum of all the charged tracks and photons in the CM frame is expected to be close to zero, a selection on $p^*_{\text{tot}} = |P_{\Upsilon(1S)} + P_{\eta,\pi^0}|$ is imposed, requiring $p^*_{\text{tot}} < 0.07 \text{ GeV}/c$ in $\Upsilon(1S) \rightarrow e^+e^-$, $\eta \rightarrow \gamma\gamma$ events, which are more contaminated by radiative Bhabha, and $p^*_{\text{tot}} < 0.1 \text{ GeV}/c$ when investigating the other modes.

The requirements described above result in an almost complete rejection of the QED and $\Upsilon(2S) \rightarrow \Upsilon(1S)\pi^0$ backgrounds.

The $\Upsilon(2S) \rightarrow \chi_{bJ}(1P)\gamma_1 \rightarrow \Upsilon(1S)\gamma_1\gamma_2$ decay has the same event topology as $\eta \rightarrow \gamma\gamma$ and $\pi^0$ transitions. The kinematic limit of the less energetic photon, which arises from $\Upsilon(2S) \rightarrow \chi_{bJ}(1P)\gamma_1$ decay, is 162 MeV in the CM frame. Therefore, this background is completely rejected for $\eta \rightarrow \gamma\gamma$ by the photon energy requirement mentioned earlier. In the search for the $\pi^0$ transition, this background is still larger than expected signal; hence, we require the energy of the less energetic photon to be higher than 170 MeV. This requirement rejects 99.5\% of the $\chi_{bJ}(1P)$ background and retains 34\% of the signal events.

The $\pi^+\pi^-$ transition represents a significant source of background only for the $\eta \rightarrow \pi^+\pi^-\pi^0$ channel; in this case, we require $\Delta M = M(\pi^+\pi^-\ell^+\ell^-) - M(\ell^+\ell^-) < 0.44\text{GeV}/c^2$, since this observable peaks at $\Delta M = M(\Upsilon(2S)) - M(\Upsilon(1S)) = 0.56 \text{ GeV}/c^2$ in $\pi^+\pi^-$ events; this cut rejects 99.92\% of the background and retains 99.1\% of the signal.

Signal efficiencies for the various final states are sum-
The efficiency in the $e^+e^-\gamma\gamma$ mode significantly differs from the one in $\mu^+\mu^-\gamma\gamma$ since this channel is affected by a Bhabha veto included at the trigger level. A trigger simulation is used in order to account for this effect.

The signal yield is extracted with a simultaneous, unbinned likelihood fit of the $\eta$ mass distribution in four different final states with a common branching fraction, as shown in Fig. 1. For the $\eta$ invariant mass peak, we use a double Gaussian with parameters, that differ from channel to channel and fixed at values determined by the simulation. The background probability density function (PDF) shape, a Crystal Ball function in $\eta \rightarrow \gamma\gamma$ and a Gaussian shape in $\eta \rightarrow \pi^+\pi^-\pi^0$, is chosen using the MC simulation; the parameters of the chosen PDFs, including the background yields in each channel, are left free in the fit of the four final states. The sum of the invariant mass distributions for the four independent final states is shown in Fig. 2.

| TABLE I: Signal efficiencies |
|-----------------------------|
| $\Upsilon(1S) \rightarrow \pi^+\pi^-$ | $\Upsilon(1S) \rightarrow \mu^+\mu^-$ |
| $\eta \rightarrow \gamma\gamma$ | 8.4% | 25.7% |
| $\eta \rightarrow \pi^+\pi^-\pi^0$ | 6.4% | 7.6% |
| $\pi^0 \rightarrow \gamma\gamma$ | 6.0% | 7.8% |

The ratio of the branching fraction for the $\eta$ transition to that for the dipion transition is given by:

$$R_{\eta,\pi^+\pi^-} = \frac{B(\Upsilon(1S)\eta)}{B(\Upsilon(1S)\pi^+\pi^-)} = \frac{N_{\eta,f}}{N_{\pi^+\pi^-}^{eff}} \times \frac{\epsilon_{\pi^+\pi^-}^{eff}}{\epsilon_{\eta,f}^{eff}},$$

where the $\eta$ final state $f$ is $\gamma\gamma$ or $\pi^+\pi^-\pi^0$. $N_{\eta}$ is the signal yield from the fit and $N_{\pi^+\pi^-}^{eff}$ is the number of detected $\Upsilon(2S) \rightarrow \Upsilon(1S)\pi^+\pi^-$ transitions for each $\Upsilon(1S) \rightarrow l^+l^-$ decay mode: $N_{\pi^+\pi^-}^{eff} = 228167$, $N_{\pi^+\pi^-}^{eff} = 276261$ and $N_{\pi^+\pi^-}^{total} = 504428$. The charged dipion transition is selected with the same cuts as in the $\eta \rightarrow \pi^+\pi^-\pi^0$ selection and requiring the event to have exactly one dipion and no $\pi^0$ or $\eta$ candidates. According to the Monte Carlo simulation the background contribution is negligible, and the efficiencies for the normalization channel are $\epsilon_{\pi^+\pi^-}^{eff} = 31.26\%$ and $\epsilon_{\eta,f}^{eff} = 37.94\%$. The number of $\Upsilon(2S) \rightarrow \Upsilon(1S)\eta$ events extracted from the simultaneous fit is $N_{\eta} = 241 \pm 17$. The resulting ratio is $R_{\eta,\pi^+\pi^-} = (1.99 \pm 0.14) \times 10^{-3}$, where the error is statistical.

In order to study systematic effects, the ratios $R_{\eta,\pi^+\pi^-}$ measured separately in four different final states are reported in Table II. The fitting procedure used for the individual channels is the same as that used for the fit to the full sample.

In the $\Upsilon(2S) \rightarrow \Upsilon(1S)\pi^0$ analysis, the $\gamma\gamma$ invariant mass distribution is fitted to a Gaussian function for the signal and a third-order polynomial for the background

| FIG. 1: Mass distribution of the $\eta$ candidates in $\gamma\gamma\mu^+\mu^-$ (top), $\gamma\gamma e^+e^-$ (middle), $\pi^+\pi^-\pi^0\mu^+\mu^-$, and $\pi^+\pi^-\pi^0 e^+e^-$ (bottom) final states. The fit function in blue represents the simultaneous fit of the four channels; the red curve shows the best-fit background component |
An upper limit on the number of $\Upsilon(1S)\pi^0$ candidates, $N_{\Upsilon(1S)\pi^0}$, is determined by generating 5000 pseudo-experiments for different values of the signal yield ranging from 0 to 30, using a Bayesian-frequentist hybrid approach based on the ratio of CLs between the $p$-value of the signal-background hypothesis and the $p$-value of the background-only hypothesis. The resulting upper limit is $N_{\Upsilon(1S)\pi^0} = 20.5$.

The upper limit on the ratio $R_{\Upsilon(1S)\pi^0}$ is then calculated from the relation:

$$R_{\Upsilon(1S)\pi^0} = \frac{N_{\Upsilon(1S)\pi^0} \epsilon^{\Upsilon(1S)\pi^0} \epsilon^{\pi^0}}{N_{\eta(1S)e^+e^-} \epsilon^\eta \epsilon^{\eta\pi^0} \epsilon^{\eta\pi^0}} = \frac{N_{\Upsilon(1S)\pi^0} \epsilon^{\Upsilon(1S)\pi^0} \epsilon^{\pi^0}}{N_{\eta(1S)e^+e^-} \epsilon^\eta \epsilon^{\eta\pi^0} \epsilon^{\eta\pi^0}}.$$

where $\epsilon^{\Upsilon(1S)\pi^0}$ is the efficiency reported in Table I reduced by a factor $1-\sigma_{\eta\pi^0}=0.922$ to account for systematic uncertainties. The contributions from each source, reported in Table III, are summed in quadrature in order to obtain the final uncertainty in each channel. The resulting upper limit is $R_{\eta\pi^0} < 2.3 \times 10^{-4}$ at 90% confidence level.

The systematic uncertainty arising from the Bhabha veto procedure is obtained comparing by the signal yields obtained with and without the veto. The systematic uncertainty introduced by the 2S kinematic fit is studied using a sample of $\Upsilon(2S) \rightarrow \chi_{b1,2}\gamma \rightarrow \Upsilon(1S)\gamma\gamma$ events that are identified by modifying the $\eta \rightarrow \gamma\gamma$ selection. To reconstruct these events, one photon in the range 90 MeV $< E^* < 180$ MeV and one with 380 MeV $< E^* < 700$ MeV are required. Three different threshold values for $CL_{2S}$ have been used to estimate the systematic uncertainties on $B(\Upsilon(2S) \rightarrow \chi_{b1,2}\gamma \rightarrow \Upsilon(1S)\gamma\gamma)$. The systematic uncertainty due to the choice of the parameters describing the background PDF is estimated by varying each within $\pm 1 \sigma$ from the MC value in the $\eta$ transition and changing the order of the polynomial fit in the $\pi^0$ channel. The uncertainties arising from the possible different signal PDF parameters between data and MC simulation are estimated by varying them within the errors and then comparing the obtained branching frac-
The upper limit for the ratio corresponding to the upper limit of stringency than that of CLEO [6], is obtained:

\[ R_{\eta,\pi^+\pi^-} = (1.99 \pm 0.14\text{(stat)} \pm 0.11\text{(syst)}) \times 10^{-3}. \]

This result is about 17% greater than the prediction of Ref. [4], 14% below the value extracted from Ref. [4] and less than half the value predicted by scaling from the \( \psi(2S) \to J/\psi \eta \) branching fraction [6]. Assuming the branching fraction \( \mathcal{B}(\Upsilon(2S) \to \Upsilon(1S)\pi^+\pi^-) = (17.92 \pm 0.26)\% \) [7], a new measurement of \( \mathcal{B}(\Upsilon(2S) \to \Upsilon(1S)\eta) \) is obtained:

\[ \mathcal{B}(\Upsilon(2S) \to \Upsilon(1S)\eta) = (3.57 \pm 0.25\text{(stat)} \pm 0.21\text{(syst)}) \times 10^{-4}, \]

where an additional systematic error of 1.4% is introduced in order to account for the uncertainties on \( \mathcal{B}(\Upsilon(2S) \to \Upsilon(1S)\pi^+\pi^-) \). This result is higher by about two standard deviations and more precise than those obtained by BaBar [9], and CLEO [6]. In addition, an upper limit for the \( R_{\eta^0,\pi^+\pi^-} \) ratio, a factor of four more stringent than that of CLEO [6], is obtained:

\[ R_{\eta^0,\pi^+\pi^-} < 2.3 \times 10^{-4}, \]

corresponding to the upper limit of

\[ \mathcal{B}(\Upsilon(2S) \to \Upsilon(1S)\pi^0) < 4.1 \times 10^{-5} \text{(90\% C.L.)}. \]

The upper limit for the ratio

\[ \frac{\mathcal{B}(\Upsilon(2S) \to \Upsilon(1S)\pi^0)}{\mathcal{B}(\Upsilon(2S) \to \Upsilon(1S)\eta)} < 0.13 \text{(90\% C.L.)} \]

is slightly below the expected value of \( 0.16 \pm 0.02 \) [6].

We thank the KEKB group for excellent operation of the accelerator, the KEK cryogenics group for efficient solenoid operations, and the KEK computer group and the NII for valuable computing and SINET4 network support. We acknowledge support from MEXT, JSPS and Nagoya’s TLSRC (Japan); ARC and DIISR (Australia); NSFC (China); MSDT (Czechia); DST (India); INFN (Italy); MEST, NRF, NSC of KISTI, and WCU (Korea); MNiSW (Poland); MES and RFAAE (Russia); ARRS (Slovenia); SNSF (Switzerland); NSC and MOE (Taiwan); and DOE and NSF (USA).