Study of the $B \to X(3872) (\to D^{*0}D^0)K$ decay

T. Aushev,18,12 N. Zwahlen,18 I. Adachi,8 H. Aihara,42 A. M. Bakich,37 V. Balagura,12 A. Bay,18 K. Belous,11 V. Bhardwaj,33 M. Bischofberger,24 A. Bondar,1,31 A. Bozek,28 J. Brodzicka,28 T. E. Browder,7 Y. Chao,27 A. Chen,25 P. Chen,27 B. G. Cheon,8 C.-C. Chiang,27 R. Chistov,12 I.-S. Cho,46 S.-K. Choi,5 Y. Choi,36 J. Dalseno,21,39 M. Danilov,12 A. Drutskoy,2 S. Eidelman,1,31 N. Gabyshhev,1,31 P. Goldenzweig,2 H. Ha,16 J. Haba,8 B.-Y. Han,16 H. Hayashii,24 Y. Hoshi,41 W.-S. Hou,27 H. J. Hyun,17 T. Iijima,23 K. Inami,23 R. Itoh,8 M. Iwabuchi,46 M. Iwasaki,42 Y. Iwasaki,8 N. J. Joshi,38 T. Julius,22 D. H. Kah,17 J. H. Kang,46 P. Kapusta,28 T. Kawasaki,30 H. J. Kim,17 H. O. Kim,17 J. H. Kim,36 Y. I. Kim,17 Y. J. Kim,4 B. R. Ko,16 S. Korpar,20,13 P. Križan,19,13 P. Krokovny,8 T. Kuhr,15 R. Kumar,33 Y.-J. Kwon,46 J. S. Lange,3 S.-H. Lee,16 J. Li,7 C. Liu,34 D. Liventsev,12 R. Louvot,18 A. Matyja,28 S. McOnie,37 T. Medvedeva,12 K. Miyabayashi,24 H. Miyata,30 Y. Miyazaki,23 R. Mizuk,12 E. Nakano,32 M. Nakao,8 Z. Natkaniec,28 S. Nishida,8 K. Nishimura,7 O. Nito,44 S. Ogawa,40 T. Ohshima,23 S. Okuno,14 S. L. Olsen,35,7 P. Pakhlov,12 G. Pakhlova,12 H. Palka,28 C. W. Park,36 H. Park,17 H. K. Park,17 R. Pestotnik,13 M. Petrič,13 L. E. Piilonen,45 S. Ryu,35 H. Sahoo,7 K. Sakai,30 Y. Sakai,8 O. Schneider,18 C. Schwanda,10 K. Senyo,23 M. Shapkin,11 C. P. Shen,7 J.-G. Shiu,27 B. Shwartz,1,31 J. B. Singh,33 P. Smerkol,13 A. Sokolov,14 E. Solovieva,12 M. Starič,13 T. Sumiyoshi,43 Y. Teramoto,32 I. Tikhomirov,12 K. Trabelsi,8 S. Uehara,8 T. Uglow,12 Y. Unno,6 S. Uno,8 Y. Usov,1,31 G. Varner,7 K. Vervink,18 C. H. Wang,26 P. Wang,9 Y. Watanabe,14 R. Wedd,22 J. Wicht,8 E. Won,16 B. D. Yabsley,37 Y. Yamashita,29 M. Yamauchi,8 C. Z. Yuan,9 Z. P. Zhang,34 V. Zhulanov,1,31 T. Zivko,13 A. Zupanc,13 and O. Zvyukova1,31 (The Belle Collaboration)

1Budker Institute of Nuclear Physics, Novosibirsk
2University of Cincinnati, Cincinnati, Ohio 45221
3Justus-Liebig-Universität Gießen, Gießen
4The Graduate University for Advanced Studies, Hayama
5Gyeongsang National University, Chinju
6Hanyang University, Seoul
7University of Hawaii, Honolulu, Hawaii 96822
8High Energy Accelerator Research Organization (KEK), Tsukuba
9Institute of High Energy Physics, Chinese Academy of Sciences, Beijing
10Institute of High Energy Physics, Vienna
11Institute of High Energy Physics, Protvino
12Institute for Theoretical and Experimental Physics, Moscow
13J. Stefan Institute, Ljubljana
14Kanagawa University, Yokohama
15Institut für Experimentelle Kernphysik, Karlsruhe Institut für Technologie, Karlsruhe
16Korea University, Seoul
17Kyungpook National University, Taegu
18École Polytechnique Fédérale de Lausanne (EPFL), Lausanne
19Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana
20University of Maribor, Maribor
21Max-Planck-Institut für Physik, München
22University of Melbourne, School of Physics, Victoria 3010
23Nagoya University, NagoYa
24Nara Women’s University, Nara
25National Central University, Chung-li
26National United University, Miaoli
27Department of Physics, National Taiwan University, Taipei
28H. Niewodniczanski Institute of Nuclear Physics, Krakow
29Nippon Dental University, Niigata
30Niigata University, Niigata
31Novosibirsk State University, Novosibirsk
32Osaka City University, Osaka
33Panjab University, Chandigarh
34University of Science and Technology of China, Hefei
35Seoul National University, Seoul
We present a study of $B \to X(3872)K$ with $X(3872)$ decaying to $D^{*0}\bar{D}^0$ using a sample of 657 million $B\bar{B}$ pairs recorded at the $\Upsilon(4S)$ resonance with the Belle detector at the KEKB asymmetric-energy $e^+e^-$ collider. Both $D^{*0} \to D^0\pi^+$ and $D^{*0} \to D^0\pi^0$ decay modes are used. We find a peak of 50.1$^{+14.8}_{-11.1}$ events with a mass of $(3872.9^{+0.6+0.1}_{-0.4-0.1})$ MeV/c$^2$, a width of $(3.9^{+2.8+0.2}_{-1.4-1.1})$ MeV/c$^2$ and a product branching fraction $B(B \to X(3872)K) \times B(X(3872) \to D^{*0}\bar{D}^0) = \left(\frac{0.80 \pm 0.20}{0.10}\right) \times 10^{-4}$, where the first errors are statistical and the second ones are systematic. The significance of the signal is 6.4$\sigma$. The difference between the fitted mass and the $D^{*0}\bar{D}^0$ threshold is calculated to be $(1.1^{+0.6+0.1}_{-0.4-0.3})$ MeV/c$^2$. We also obtain an upper limit on the product of branching fractions $B(B \to Y(3940)K) \times B(Y(3940) \to D^{*0}\bar{D}^0)$ of $0.67 \times 10^{-4}$ at 90% CL.

PACS numbers: 14.40.Rt, 13.25.Hw, 12.39.Mk

The $X(3872)$ was discovered by Belle in 2003 in $B^\pm \to J/\psi\pi^+\pi^-K^\pm$ \cite{1} with a mass of $(3872.0 \pm 0.6 \pm 0.5)$ MeV/c$^2$, and was later confirmed by CDF \cite{2}, DØ \cite{3} and BaBar \cite{4}. It is one of the many new and unexpected hidden-charm states states recently discovered with masses around 4 GeV/c$^2$. So far it remains unclassified; it does not seem to be a pure $c\bar{c}$ charmonium state, but may be a $D^{*0}\bar{D}$ deuon \cite{5,6}, a tetraquark state \cite{7} or a charmonium-gluon hybrid \cite{8}. The current average mass in the $J/\psi\pi^+\pi^-$ channel is $(3871.50 \pm 0.19)$ MeV/c$^2$ \cite{9}.

An important feature of the $X(3872)$ is that its mass is very close to the $D^{*0}\bar{D}^0$ threshold $\left(M_{D^{*0}\bar{D}^0} = (3871.81 \pm 0.36)\text{ MeV}/c^2\right)$. The $X(3872)$ was also observed by Belle \cite{11} as a near-threshold enhancement in the $D^{0}\bar{D}^{0}\pi^{0}$ invariant mass spectrum of the $B \to D^{0}\bar{D}^{0}\pi^{0}K$ channel, with a peak at $(3875.2 \pm 0.7^{+0.3}_{-1.0} \pm 0.8)$ MeV/c$^2$, where the third error is from the uncertainty on the $D^{*0}$ mass \cite{10}, a Gaussian width of $(2.42 \pm 0.55)$ MeV/c$^2$, and a branching fraction $B(B \to D^{0}\bar{D}^{0}\pi^{0}K) = (1.22 \pm 0.31^{+0.23}_{-0.30}) \times 10^{-4}$. This initial study did not distinguish between decays via the $D^{*0}$, and more general $D^{*0}\pi^{0}$ final states. Looking in addition for the $D^{*0} \to D^{0}\gamma$ decay is crucial to demonstrate the presence of $X(3872)$ decay through a $D^{*0}$.

The BaBar collaboration recently published an observation of the decay $B \to X(3872)\to (D^{*0}\bar{D}^0)K$ with a $4.9\sigma$ significance \cite{12}. The observed mass is $(3875.1^{+0.3}_{-0.5} \pm 0.5)$ MeV/c$^2$ and the width is $(3.0^{+1.2}_{-1.3} \pm 0.9)$ MeV/c$^2$, with a product branching fraction $B(B^+ \to X(3872)K^+) \times B(X(3872) \to D^{*0}\bar{D}^0) = (1.67 \pm 0.36 \pm 0.47) \times 10^{-4}$. In the BaBar analysis, $D^{*0}$ candidates were subjected to a mass-constrained fit.

Another new particle called $X(3940)$ was discovered by Belle in the $e^+e^- \to D\bar{D}^{*0}$, $J/\psi$ process \cite{13}. A state with the same mass, the $Y(3940)$ (also known as $X(3945)$ \cite{10}), was discovered by Belle in $B \to \omega J/\psi K$ \cite{14} and was later confirmed by BaBar, albeit with a smaller mass \cite{15}. The possibility that the $X(3940)$ and $Y(3940)$ are the same state has not yet been ruled out.

In this paper we measure the $X(3872) \to D^{*0}\bar{D}^0$ decay mode, followed either by $D^{*0} \to D^0\gamma$ or $D^{*0} \to D^{0}\pi^0$, in charged and neutral $B \to X(3872)K$ decays. Inclusion of charge-conjugate modes is implied throughout the paper. Furthermore, we use the notation $D^{*0}\bar{D}^0$ to indicate both $D^{0}\bar{D}^{0}$ and $D^{*0}\bar{D}^{0}$. The results are based on a 605 fb$^{-1}$ data sample, corresponding to 657 $\times 10^6$ $BB$ pairs, collected at the $\Upsilon(4S)$ resonance with the Belle detector \cite{10} at the KEKB asymmetric-energy $e^+e^-$ collider \cite{13}, which includes the statistics used in the previous Belle analysis \cite{11}.

The Belle detector is a general purpose spectrometer with a 1.5 T magnetic field provided by a superconducting solenoid. A silicon vertex detector and a 50-layer central drift chamber are used to measure the momenta of charged particles. Photons are detected in an electromagnetic calorimeter consisting of CsI(Tl) crystals. Particle identification likelihoods $L_k$ and $L_\pi$ are derived from information provided by an array of time-of-flight counters, an array of silica aerogel Cherenkov threshold counters and $dE/dx$ measurements in the central drift chamber.

Charged tracks are identified as kaons using a requirement on the likelihood ratio $L_K/(L_K + L_\pi)$, which has an efficiency of 88% for kaons and 10% for pions. Similarly, the charged pion selection has an efficiency of 98% for pions and 12% for kaons. $\pi^0$ candidates are reconstructed from pairs of photons with energies $E_\gamma > 50$ MeV and
with invariant mass in the range 118 MeV/c² < M_{\pi\gamma} < 150 MeV/c². A mass-constrained fit is applied to obtain the four-momenta of the π⁺ candidates. K⁰ candidates are reconstructed in the K⁰_S → π⁺π⁻ mode with the requirement |M_{π⁻} − m_{K⁰_S}| < 15 MeV/c², where m_{K⁰_S} is the K⁰_S mass [10]. Requirements on the K⁰_S vertex displacement from the interaction point and on the difference between vertex and K⁰_S flight directions are applied. The K⁰_S selection criteria are described in detail elsewhere [18]. A mass- and vertex-constrained fit is applied to improve the four-momenta measurements of the K⁰_S candidates.

D⁰ mesons are reconstructed in the modes D⁰ → K⁻π⁺, K⁻π⁺π⁰, K⁻π⁺π⁺π⁻, K⁰_Sπ⁺π⁻ and K⁻K⁺. The D⁰ candidates are required to have a mass within ±14 MeV/c² (±26 MeV/c² for the K⁻π⁺π⁰ mode) of the D⁰ mass, 1864.8 MeV/c² [10]. This mass window width corresponds to ±3σ (±3σ for K⁻π⁺π⁰). Mass- and vertex-constrained fits are applied to improve the resolution of the four-momenta of D⁰ candidates. D⁰ candidates are combined with a photon or a π⁰ to obtain D⁺π⁰ candidates. The photon candidate is required to have an energy in excess of 100 MeV and shower shape variables that are consistent with an electromagnetic shower; the ratio of energy deposition in the central 3 × 3 and 5 × 5 crystals of the cluster is required to be greater than 0.8. A mass window of ±27.5 MeV/c² for the D⁺π⁰ → D⁰γ channel and ±6 MeV/c² for the D⁺π⁰ → D⁰π⁰ channel is applied (±3σ), and a mass-constrained fit is used to improve the four-momenta of the D⁰ candidates; the mass is constrained to 2007.0 MeV/c² [10].

B mesons are reconstructed by combining a D⁰¯¯¯¯D⁰ candidate with invariant mass M_{D⁺D⁻} < 4.0 GeV/c² and a charged or neutral kaon candidate. To further reduce the background, at least one D⁰(¯¯¯¯D⁰) is required to decay to K⁻π⁺(K⁺π⁻). The beam-energy constrained mass

\[ M_{bc} = \sqrt{E_{beam}^2 - \left( \sum_i \vec{p}_i \right)^2}, \]

where \( \vec{p}_i \) is the momentum of the i-th daughter of the B candidate in the e⁺e⁻ center-of-mass (CM) system, is required to be larger than 5.2 GeV/c². The energy difference \( \Delta E = E_B - E_{beam} \), where \( E_{beam} \) is the CM energy of the B candidate and \( E_B \) is the CM beam energy, is restricted to the range \( |\Delta E| < 25 \) MeV/c². Continuum e⁺e⁻ → qq̅ background events (q = u, d, s, c) are suppressed by requiring the ratio of the second and zeroth Fox-Wolfram moments [19] to be smaller than 0.3.

The average B candidate multiplicity per event is 2.3 for D⁺π⁰ → D⁰γ and 2.7 for D⁺π⁰ → D⁰π⁰. We select the candidate with the smallest value of the quantity

\[
\left( \frac{\Delta M_{D^{0}}}{\sigma_{M_{D^{0}}}} \right)^2 + \left( \frac{\Delta M_{D^{0}}}{\sigma_{M_{D^{0}}}} \right)^2 + \left( \frac{\Delta (M_{D^{0}} - M_{D^{0}})}{\sigma_{(M_{D^{0}} - M_{D^{0}})}} \right)^2 + \left( \frac{\Delta E}{\sigma_{\Delta E}} \right)^2 + \left( \frac{\Delta M_{\pi^{0}}}{\sigma_{M_{\pi^{0}}}} \right)^2, \tag{1}
\]

where \( \Delta x \) is the deviation of the measured quantity \( x \) from its expected value and \( \sigma_x \) the uncertainty in its measurement obtained using a Monte Carlo (MC) method, and the last term applies to the D⁺π⁰ → D⁰π⁰ channel only.

MC samples are produced using the EVTGEN [20] generator. The X(3872) mass distribution is generated according to a relativistic Breit-Wigner function

\[
BW(m) = \frac{\mu m \Gamma(m)}{(m^2 - \mu^2)^2 + \mu^2 \Gamma(m)^2}, \tag{2}
\]

where \( \Gamma(m) = \Gamma_0 \frac{m p(m)}{m_0 p(\mu)}, \)

\[
p(m) = \frac{1}{2m} \frac{\sqrt{m^2 - (m_{D^{0}} + m_{D^{*}})^2}}{\sqrt{m^2 - (m_{D^{0}} - m_{D^{*}})^2}},
\]

and \( \mu \) and \( \Gamma_0 \) are the nominal mass and width of the resonance, respectively, \( p(m) \) is the momentum of one of the daughters in the rest frame of its parent. The term \( \Gamma_0(m) \) in the numerator of Eq. (2) behaves like a phase-space function, giving a smooth rise near the D⁺D⁰ threshold.

The response of the Belle detector is simulated using a GEANT3-based program [21]. Since the X(3872) is very close to the D⁺D⁰ threshold, its mass resolution varies rapidly with the D⁺D⁰ mass (\( M_{D^{0}} \)). This resolution was studied in detail using MC simulations. It is modeled as a single Gaussian with a width given by the function \( a\sqrt{M_{D^{0}} - m_0} \) (where \( a \) and \( m_0 \) are free parameters) shown in Fig. 1. At 3872 MeV/c², the resolution is 0.13 MeV/c².

FIG. 1: X(3872) mass resolution as a function of the X(3872) mass in the D⁺γ channel, obtained from MC with a X(3872) mass spectrum generated for a continuous range of masses from threshold to 4.0 GeV/c². Crosses are Gaussian resolutions for various generated D⁺D⁰ masses; the curve is the result of a fit with a square root function. Very similar results are obtained for the D⁺π⁰ channel.
A two-dimensional unbinned extended maximum likelihood fit to \(M_{bc}\) and \(M_{D^{+}D}\) is performed. The \(M_{bc}\) distribution is described by a single Gaussian function for the signal and an ARGUS function \(^{22}\) for the background; the \(M_{D^{+}D}\) distribution is described by a relativistic Breit-Wigner function convoluted with the mass-dependent Gaussian resolution for the signal and a square root function for the background. In the \(D^{0}\gamma\) channel, the signal function also includes a broad higher mass distribution, corresponding to \(D^{0}\bar{\pi}^{0}\gamma\) events incorrectly reconstructed as \(D^{0}\bar{D}^{0}\gamma(D\bar{D})\). The shape and fraction of this contribution is determined from MC. In the \(D^{0}\pi^{0}\) channel, however, the reflection shape is too similar to the signal one to be distinguished. This is taken into account as a systematic uncertainty. Additional components are the \(Y(3940)\) signal, described by a relativistic Breit-Wigner function, and the non-resonant \(B \rightarrow D^{*}DK\) background, which peaks in \(M_{bc}\) but not in \(M_{D^{+}D}\) and is therefore described by a single Gaussian function in \(M_{bc}\) and a square root function in \(M_{D^{+}D}\). The fitting procedure has been extensively tested using toy MC samples.

First, each \(D^{*}\) decay channel sample is fitted separately. The yield, mass and width of the \(X(3872)\) peak are free parameters of the fit, as well as the \(Y(3940)\) yield and the number of background and non-resonant \(B \rightarrow D^{*}DK\) events. The \(Y(3940)\) mass and width and ratio are fixed to the values of Ref. \([14]\). The results of the fits are presented in Table I. Since all the results are consistent we proceed with a combined fit.

We subsequently perform a simultaneous fit to both \(D^{*}\) modes where the mass and width of the signal function are constrained to have the same values in both modes, but the ratio of the yields in the \(D^{0}\gamma\) and \(D^{0}\pi^{0}\) channels is left free. Table I shows the fit results. We obtain a yield ratio of \(N_{D^{0}\gamma}/N_{D^{0}\pi^{0}} = 1.16^{+0.53}_{-0.37}\) which is consistent with the value of 0.92 we expect from MC with no non-resonant \(D^{0}\bar{D}^{0}\pi^{0}\) contribution. We then fix the ratios of \(D^{0}\gamma/D^{0}\pi^{0}\) signal yields for the \((3872)\) and \((3940)\) from MC studies assuming that the \(D^{0}\pi^{0}\) and \(D^{0}\gamma\) come from a \(D^{0}\). This fits gives 50.1\(^{+14.8}_{-11.4}\) events with a statistical significance of 7.9\(\sigma\) (see Fig. 2 and Table I). We compute the significance from \(-2 \ln(L_{0}/L_{\text{max}})\), where \(L_{0}\) and \(L_{\text{max}}\) are the likelihood values returned by the fit with the signal yield fixed at zero and left free, respectively. This quantity should be distributed as \(χ^{2}(n_{\text{dof}} = 3)\), as three parameters are free for the signal. The results of the simultaneous fits are consistent with the results of the individual fits. The distributions for \(M_{bc}\) and \(X(3872)\) mass closer to the \((3872)\) signal region are presented in Fig. 3.

Next we fit the \(B^{+} \rightarrow X(3872)\bar{K}^{+}\) and \(B^{0} \rightarrow X(3872)K_{S}^{0}\) modes separately, fixing the \((3872)\) mass and width to the values obtained with the simultaneous fit. Table I shows the results of these fits. Assuming the \(B^{0} \rightarrow X(3872)K_{S}^{0}\) transition rate to be equal to twice the \(B^{0} \rightarrow X(3872)K_{S}^{0}\) rate, we obtain a ratio of branching fractions

\[
\frac{B(B^{0} \rightarrow X(3872)K_{S}^{0})}{B(B^{+} \rightarrow X(3872)\bar{K}^{+})} = 1.26 \pm 0.65 \pm 0.06,
\]

which is consistent with unity. In this ratio, most of the systematic uncertainties cancel out. The remaining uncertainties are MC statistics, particle identification and reconstruction efficiencies of the \(K^{+}\) and \(K_{S}^{0}\), which combine in quadrature to a 5% uncertainty.

The systematic uncertainties for the mass, width and signal yield are estimated by varying the fixed parameters of the simultaneous fit: the resolution function is scaled by factors 0.9 and 1.2 obtained from the inclusive \(D^{0}\) mass resolution study; the ratio between \(D^{0}\gamma\) and \(D^{0}\pi^{0}\) yields is varied according to the error in the relative branching ratio between \(D^{0}\rightarrow D^{0}\gamma\) and \(D^{0}\rightarrow D^{0}\pi^{0}\) \(^{10}\) and the uncertainty in their relative reconstruction efficiency; the \(Y(3940)\) mass and width are fixed instead to the values from Ref. \([15]\): the yield of \(Y(3940)\) is fixed to zero and to the upper limit of 40 events; the shape of the combinatorial background is changed from a square root to an inverted ARGUS function. Additional uncertainties due to possible biases of the fit parameters are obtained using a large ensemble of toy MC experiments generated for different values of \(X(3872)\) mass and width. We also consider systematic uncertainties due to the contribution from \(D^{0}\bar{D}^{0}\rightarrow D^{0}\bar{D}^{0}\) mis-reconstruction by changing the shape of the miss-reconstruction component to one obtained from MC generated with different values of \(X(3872)\) mass and width. Using MC we verify that the cross-sections between \(D^{0}\bar{D}^{0}\gamma\) and \(D^{0}\bar{D}^{0}\pi^{0}\) do not produce peaks in the signal region. Another systematic uncertainty on the mass is due to the \(D^{0}\bar{D}^{0}\) threshold mass, \(2m_{D^{0}} + \Delta(m_{D^{0} - m_{D^{0}}}) = (3871.80 \pm 0.41)\) MeV/c\(^2\). The systematic uncertainties for the branching fraction are estimated from the following sources: the number of \(B\bar{B}\) events, \(D^{0}\) branching fractions, track finding efficiencies, \(K/\pi\) identification efficiency, \(\gamma\) or \(\pi^{0}\) detection efficiency, \(K_{S}^{0}\) reconstruction efficiency, limited MC statistics and variation of the fixed parameters of the fit. Table I summarizes the systematic uncertainties. We obtain total systematic errors of \(+0.41_{-0.54}\) MeV/c\(^2\) for the mass, \(+0.22_{-1.07}\) MeV/c\(^2\) for the width and \(\pm 12\%\) for the branching fraction measurement. The significance of the signal including systematic uncertainties decreases to 6.4\(\sigma\).

In order to determine the reconstruction efficiency for the branching fraction measurement, MC samples of \(B^{+} \rightarrow X(3872)(\rightarrow D^{0}\bar{D}^{0})\bar{K}^{+}\) and \(B^{0} \rightarrow X(3872)(\rightarrow D^{0}\bar{D}^{0})K_{S}^{0}\) events are generated for \(D^{0}\rightarrow D^{0}\gamma\) and for \(D^{0}\rightarrow D^{0}\pi^{0}\). The \((3872)\) is generated with a mass of 3872.5 MeV/c\(^2\) and a width of 4.0 MeV/c\(^2\). The total MC efficiencies multiplied by subdecay branching fractions are presented in Table I.
FIG. 2: Distributions of $M_{D^*D}$ mass for $M_{bc} > 5.27$ GeV/c$^2$ for $D^{*0} \rightarrow D^0\gamma$ (left) and for $D^{*0} \rightarrow D^0\pi^0$ (right). The result of the simultaneous fit is shown by the superimposed lines. The points with error bars are data, the dotted curve is the signal, the dashed curve is the background, the dash-dotted curve (barely visible) is the $Y(3940)$ component, and the solid curve is the total fitting function.

FIG. 3: Distributions of $M_{bc}$ for $M_{D^*D} < 3.88$ GeV/c$^2$ (left) and of $M_{D^*D}$ for $M_{bc} > 5.27$ GeV/c$^2$ (right); the top row is for $D^{*0} \rightarrow D^0\gamma$ and the bottom row for $D^{*0} \rightarrow D^0\pi^0$. The result of the simultaneous fit is shown by the superimposed lines. The points with error bars are data, the dotted curve is the signal, the dashed curve is the background, the dash-dotted curve is the sum of the background and the $B \rightarrow D^*DK$ component, the dot-dot-dashed curve is the contribution from $D^0$–$\bar{D}^0$ reflections, and the solid curve is the total fitting function.
TABLE I: Summary of results: the fitted mass, width, and yield of the $X(3872)$ peak, and the total reconstruction efficiency, branching fraction and statistical significance for the various fits described in the text.

| Sample                  | $M_X$ (MeV/c^2) | $\Gamma$ (MeV/c^2) | Yield | $\epsilon \times B$ | $B$ (10^{-4}) | $\sigma$ |
|-------------------------|-----------------|--------------------|-------|---------------------|---------------|---------|
| $D^{*0} \rightarrow D^0 \gamma$ (XK^+ and XK^0) | 3873.4 ± 1.0    | 4.2^{+4.7}_{-1.8} | 26.2^{+7.0}_{-7.6} | 4.56 × 10^{-4} | 0.87 ± 0.28 ± 0.10 | 4.4σ |
| $D^{*0} \rightarrow D^{0} \pi^0$ (XK^+ and XK^0) | 3872.8 ± 0.7    | 3.4^{+4.1}_{-1.2} | 22.0^{+10.7}_{-6.4} | 4.93 × 10^{-4} | 0.68 ± 0.26 ± 0.09 | 6.8σ |

All (free $D^0 \gamma/D^0 \pi^0$ ratio) | 3872.9^{+0.6}_{-0.4} | 3.9^{+2.7}_{-1.4} | 50.6^{+14.2}_{-11.0} | 9.49 × 10^{-4} | 0.81 ± 0.20 ± 0.10 | 7.9σ |
All (fixed $D^0 \gamma/D^0 \pi^0$ ratio) | 3872.9^{+0.6}_{-0.4} | 3.9^{+2.8}_{-1.4} | 50.1^{+14.8}_{-11.3} | 9.49 × 10^{-4} | 0.80 ± 0.20 ± 0.10 | 7.9σ |

$B^+ \rightarrow XK^+$ | 3872.9 (fixed) | 3.9 (fixed) | 41.3^{+9.1}_{-8.1} | 8.17 × 10^{-4} | 0.77 ± 0.16 ± 0.10 | 7.6σ |
$B^0 \rightarrow XK^0$ | 3872.9 (fixed) | 3.9 (fixed) | 8.4^{+4.5}_{-4.6} | 1.32 × 10^{-4} | 0.97 ± 0.46 ± 0.13 | 2.8σ |

The branching fraction, assumed to be equal for charged and neutral $B$ modes, is

$$B(B \rightarrow X(3872)K) \times B(X(3872) \rightarrow D^{*0} \bar{D}^0) = (0.80 \pm 0.20 \pm 0.10) \times 10^{-4},$$

(3)

where $B(X(3872) \rightarrow D^{*0} \bar{D}^0)$ stands for the sum of branching fractions for $X(3872) \rightarrow D^{0} \bar{D}^0$ and $X(3872) \rightarrow D^{*0} \bar{D}^0$.

We obtain a mass of $(3872.9^{+0.6}_{-0.4} \pm 0.4)$ MeV/c^2 and a width of $(3.9^{+2.7}_{-1.4} \pm 0.7)$ MeV/c^2. The difference in mass between the peak and the $D^{*0} \bar{D}^0$ threshold is

$$\delta M = M_X - m_{D^{*0}} - m_{\bar{D}^0} = (1.1^{+0.6}_{-0.4} \pm 0.1) \text{ MeV/c}^2. \quad (4)$$

As a cross-check, different shapes are used as a signal function to fit the data. Using a non-relativistic Breit-Wigner function truncated at the threshold, we obtain a mass of $(3873.4^{+0.6}_{-0.9} \pm 0.9)$ MeV/c^2, a width of $(4.3^{+2.5}_{-1.4})$ MeV/c^2 and a signal yield of $39.6^{+9.3}_{-8.1}$ events with a statistical significance of 8.0σ. Using a Flatté-like parameterization [23], with $g = 0.3$ and $f_B = 0.007$, we obtain $E_f = (-14.9 \pm 2.0)$ MeV/c^2, consistent with expected $E_f = -11$ MeV/c^2 [23], and a signal of 65 ± 12 events with a statistical significance of 8.8σ. The data statistics are not sufficient to distinguish between other fitting functions suggested in the literature [24].

The peak mass in this mode, $M_X$, like that in the BaBar analysis [12], should not be directly compared to the mass of the peak seen in $J/\psi \pi^+ \pi^-$, or to the mass of the $X(3872)$ state itself. In this analysis, to improve signal/background separation, $D^0 \gamma$ and $D^0 \pi^0$ combinations are each subjected to a mass window selection, and then a mass-constrained fit. This procedure returns masses above $D^0 \bar{D}^0$ threshold by construction, so the distribution (and the mass and width of the peak) should be interpreted accordingly. Efforts in this direction have already appeared in the literature [24, 25], addressing a preliminary version of the results in this paper. It should be noted that the mass measurement in our earlier paper [11], while lacking the mass constraint, is also non-trivial to interpret, due to the role of the $D^*$ width, and interference, in the decay amplitudes [24, 25].

For the $Y(3940)$ state, the simultaneous fit yields $7 \pm 21 \pm 4$ signal events and we set an upper limit of

$$B(B \rightarrow Y(3940)K) \times B(Y(3940) \rightarrow D^{*0} \bar{D}^0) < 0.67 \times 10^{-4}$$

(5)
at 90% CL. By averaging the branching fractions of Refs [14] and [15], we obtain $B(B \rightarrow Y(3940)K) \times B(Y(3940) \rightarrow \omega J/\psi) = (0.51 \pm 0.11) \times 10^{-4}$; combining this with the upper limit (5) we get $B(Y(3940) \rightarrow D^{*0} \bar{D}^0) > 0.71$ at 90% CL, to be compared with the 90% CL limits from Ref. [13], $B(X(3940) \rightarrow \omega J/\psi) < 0.26$ and $B(X(3940) \rightarrow D^{*0} \bar{D}^0) > 0.45$, thus $B(Y(3940) \rightarrow D^{*0} \bar{D}^0) < 0.58$ with more than 90% CL, an incompatibility that suggests that the $X(3940)$ and the $Y(3940)$ are different states.

In summary, we find a significant near-threshold enhancement in the $D^{*0} \bar{D}^0$ invariant mass spectrum in $B \rightarrow D^{*0} \bar{D}^0 K$ decays. The significance of this en-
hancement including systematic uncertainties is $6.4\sigma$; significant signals are seen in both $D^{*0} \to D^0\gamma$ and $D^{0}\pi^0$ modes. Using a relativistic Breit-Wigner, we obtain a mass of $(3872.9^{+0.6+0.4}_{-0.4-0.5})$ MeV/$c^2$ and a width of $(3.9^{+2.8+0.2}_{-1.4-1.1})$ MeV/$c^2$. The difference between the fitted mass and the $D^{*0}D^0$ threshold is $(1.1^{+0.6+0.3}_{-0.4-0.5})$ MeV/$c^2$. Note that a $D^{*0}$ mass constraint has been used in this analysis; the fitted mass of the $X(3872)$ peak is $2.3\sigma$ lower than the value obtained by BaBar [12], where a similar constraint was used. For the $Y(3940)$ state, we set an upper limit on the $B(B \to Y(3940)K) \times B(Y(3940) \to D^{*0}D^0)$ branching fraction which suggests that the $X(3940)$ and the $Y(3940)$ are different states.

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 National Institute of Informatics for valuable computing and SINET3 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 and the Australian Department of Industry, Innovation, Science and Research; the National Natural Science Foundation of China under contract No. 10575109, 10775142, 10875115 and 10825524; the Department of Science and Technology Information; the Ministry of Education and Higher Education; the Polish Ministry of Science and Higher Education; the Ministry of Education and Science of the Russian Federation and the Russian Federal Agency for Atomic Energy; the Slovenian Research Agency; the Swiss National Science Foundation; the National Science Council and the Ministry of Education of Taiwan; and the U.S. Department of Energy. This work is supported by a Grant-in-Aid from MEXT for Science Research in a Priority Area (“New Development of Flavor Physics”), and from JSPS for Creative Scientific Research (“Evolution of Tau-lepton Physics”).

[1] S.-K. Choi et al. (Belle Collaboration), Phys. Rev. Lett. 91, 262001 (2003).
[2] D. Acosta et al. (CDF Collaboration), Phys. Rev. Lett. 93, 072001 (2004).
[3] V.M. Abazov et al. (DØ Collaboration), Phys. Rev. Lett. 93, 162002 (2004).
[4] B. Aubert et al. (BaBar Collaboration), Phys. Rev. D 71, 071103 (2005).
[5] E.S. Swanson, Phys. Lett. B 588, 189 (2004).
[6] N.A. Törnqvist, Phys. Lett. B 590, 209 (2004).
[7] L. Maiani, F. Piccinini, A.D. Polosa and V. Riquer, Phys. Rev. D 71, 014028 (2005).
[8] B.A. Li, Phys. Lett. B 605, 306 (2005).
[9] Our own average using the most recent measurements from Belle, BaBar, CDF and DØ [3, 26, 27, 28].
[10] C. Amsler et al. (Particle Data Group), Phys. Lett. B 667, 1 (2008).
[11] G. Gokhroo et al. (Belle Collaboration), Phys. Rev. Lett. 97, 162002 (2006).
[12] B. Aubert et al. (BaBar Collaboration), Phys. Rev. D 77, 011102 (2008).
[13] P. Pakhlov et al. (Belle Collaboration), Phys. Rev. Lett. 98, 082001 (2007).
[14] S.-K. Choi, S.L. Olsen et al. (Belle Collaboration), Phys. Rev. Lett. 94, 182002 (2005).
[15] B. Aubert et al. (BaBar Collaboration), Phys. Rev. Lett. 101, 082001 (2008).
[16] A. Abashian et al. (Belle Collaboration), Nucl. Instr. and Meth. A 479, 117 (2002).
[17] S. Kurokawa and E. Kikutani, Nucl. Instr. and Meth. A 499, 1 (2003).
[18] F. Fang, Ph.D. thesis, University of Hawaii (2003).
[19] G.C. Fox and S. Wolfram, Phys. Rev. Lett. 41, 1581 (1978).
[20] D.J. Lange, Nucl. Instr. and Meth. A 462, 152–155 (2001).
[21] R. Brun et al., GEANT 3.21, CERN Report DD/EE/84-1 (1984).
[22] H. Albrecht et al. (ARGUS Collaboration), Phys. Lett. B 185, 218 (1987).
[23] C. Hanhart, Yu.S. Kalashnikova, A.E. Kudryavtsev and A.V. Nefediev, Phys. Rev. D 76, 034007 (2007).
[24] E. Braaten and J. Stapleton, arXiv:0907.3167 [hep-ph].
[25] Yu.S. Kalashnikova and A.V. Nefediev, arXiv:0907.4901 [hep-ph].
[26] I. Adachi et al. (Belle Collaboration), arXiv:0809.1223 [hep-ex].
[27] B. Aubert et al. (BaBar Collaboration), Phys. Rev. D 77, 111101 (2008).
[28] T. Aaltonen et al. (CDF Collaboration), Phys. Rev. Lett. 103, 152001 (2009).