Study of the lineshape of $X(3872)$ using B decays to $D^0\bar{D}^{*0}K$

H. Hirata, T. Iijima, Y. Kato, K. Tanida, I. Adachi, J. K. Ahn, H. Aihara, S. Al Said, D. M. Asner, H. Atmacan, T. Aushev, R. Ayad, V. Babu, Sw. Banerjee, P. Behera, K. Belous, J. Bennett, M. Bessner, V. Bhardwaj, B. Bhuyan, T. Bilka, D. Biswas, A. Bobrov, D. Bodrov, J. Borah, A. Bozek, M. Bračko, P. Branchini, T. E. Browder, A. Budano, M. Campajola, D. Červenkov, M.-C. Chang, P. Chang, B. G. Cheon, K. Chilikin, H. E. Cho, K. Cho, S.-J. Cho, S.-K. Choi, Y. Choi, S. Choudhury, D. Cinabro, S. Cunliffe, S. Das, G. de Marino, G. De Nardo, G. De Pietro, R. Dhamija, F. Di Capua, J. Dingfelder, Z. Doležal, T. V. Dong, D. Epifanov, T. Ferber, D. Ferlewicz, B. G. Fulsom, V. Gaur, A. Garmash, A. Giri, P. Goldenzweig, E. Graziani, K. Gudkova, C. Hadjivasiliou, S. Halder, T. Hara, K. Hayasaka, H. Hayashii, M. T. Hedges, D. Herrmann, W.-S. Hou, C.-L. Hsu, K. Inami, N. Ipsita, A. Ishikawa, R. Itoh, M. Iwasaki, W. W. Jacobs, E.-J. Jang, S. Jia, Y. Jin, D. Kalita, C. Kiesling, C. H. Kim, D. Y. Kim, K.-H. Kim, Y.-K. Kim, K. Kinoshita, P. Kodyš, T. Konno, A. Korobov, S. Korpar, E. Kovalenko, P. Križan, P. Krokovny, T. Kuhr, R. Kumar, K. Kumara, A. Kuzmin, Y.-J. Kwon, T. Lam, J. S. Lange, M. Laurenza, K. Lautenbach, S. C. Lee, L. K. Li, Y. Li, J. Libby, K. Lieret, Y.-R. Lin, D. Liventsev, T. Luo, Y. Ma, A. Martini, M. Masuda, T. Matsuda, D. Matvienko, S. K. Maurya, F. Meier, M. Merola, F. Metzner, K. Miyabayashi, G. B. Mohanty, M. Mrvar, R. Mussa, I. Nakamura, M. Nakao, Z. Natkaniec, A. Natochii, L. Nayak, M. Nayak, M. Niiyama, N. K. Nisar, S. Nishida, K. Ogawa, S. Ogawa, H. Ono, P. Oskin, P. Pakhlov, G. Pakhlova, T. Pang, S. Pardi, H. Park, J. Park, S.-H. Park, S. Patra, S. Paul, T. K. Pedlar, R. Pestotnik, L. E. Piilonen, T. Podobnik, E. Prencipe, M. T. Prim, N. Rout, G. Russo, S. Sandilya, A. Sangal, L. Santelj, V. Savinov, G. Schnell, C. Schwanda, A. J. Schwartz, Y. Seino, K. Senyo, M. E. Sevior, W. Shan, M. Shapkin, C. Sharma, J.-G. Shiu, B. Shwartz, A. Sokolov, E. Solovieva, M. Starič, Z. S. Stottler, M. Sumihama, W. Sutcliffe, M. Takizawa, U. Tamponi, S. Tanaka, F. Tenchini, R. Tiwary, K. Trabelsi, M. Uchida, T. Uglov, Y. Unno, K. Uno, S. Uno, P. Urquijo, Y. Usov, S. E. Vahsen, G. Varner, A. Vinokurova, A. Vossen, D. Wang, E. Wang, M.-Z. Wang, S. Watanuki, O. Werbycka, E. Won, X. Xu, B. D. Yabsley, W. Yan, S. B. Yang, J. Yelton, J. H. Yin, Y. Yook, C. Z. Yuan, L. Yuan, Y. Yusa, Z. P. Zhang, V. Zhilich, and V. Zhukova

(Belle Collaboration)

(Received 3 February 2023; accepted 15 May 2023; published 28 June 2023)

We present a study of the $X(3872)$ lineshape in the decay $B \to X(3872)K \to D^0\bar{D}^{*0}K$ using a data sample of 772×10^6 $B\bar{B}$ pairs collected at the $\Upsilon(4S)$ resonance with the Belle detector at the KEKB asymmetric-energy e^+e^- collider. The peak near the threshold in the $D^0\bar{D}^{*0}$ invariant mass spectrum is fitted using a relativistic Breit-Wigner lineshape. We determine the mass and width parameters to be $m_{\rm BW} = 3873.71^{+0.56}_{-0.50}(\text{stat}) \pm 0.13(\text{syst})$ MeV/c^2 and $\Gamma_{\rm BW} = 5.2^{+2.2}_{-1.5}(\text{stat}) \pm 0.4(\text{syst})$ MeV, respectively. The branching fraction is found to be $\mathcal{B}(B^+ \to X(3872)K^+) \times \mathcal{B}(X(3872) \to D^0\bar{D}^{*0}) = (0.97^{+0.21}_{-0.18}(\text{stat}) \pm 0.10(\text{syst})) \times 10^{-4}$. The signal from B^0 decays is observed for the first time with 5.2σ significance, and the ratio of branching fractions between charged and neutral B decays is measured to be $\mathcal{B}(B^0 \to X(3872)K^0)/\mathcal{B}(B^+ \to X(3872)K^+) = 1.34^{+0.47}_{-0.40}(\text{stat})^{+0.10}_{-0.12}(\text{syst})$. The peak is also studied using a Flatté lineshape. We determine the lower limit on the $D\bar{D}^*$ coupling constant g to be 0.075 at 95% credibility in the parameter region where the ratio of g to the mass difference from the $D^0\bar{D}^{*0}$ threshold is equal to -15.11 GeV^{-1}, as measured by LHCb.

DOI: 10.1103/PhysRevD.107.112011

I. INTRODUCTION

The charmoniumlike $X(3872)$ state, also known as $\chi_{c1}(3872)$ [1], was discovered by the Belle experiment as a narrow peak in the vicinity of the $D^0\bar{D}^{*0}$ threshold in the $J/\psi\pi^+\pi^-$ invariant mass distribution in exclusive $B^+ \to J/\psi\pi^+\pi^-K^+$ decays [2]. Its existence has been confirmed by multiple experiments: D0 [3], *BABAR* [4], CDF [5], LHCb [6], and BESIII [7]. In addition to the $J/\psi\pi^+\pi^-$ decay, other decays such as $J/\psi\omega$ [8], $J/\psi\gamma$, $\psi(2S)\gamma$ [9], $D^0\bar{D}^{*0}$ [10,11], $D^0\bar{D}^0\pi^0$ [12], and $\pi^0\chi_{c0}$ [13] have been observed. The $X(3872)$ quantum numbers J^{PC} have been determined to be 1^{++} [14,15]. Various interpretations such as a loosely bound state [16–19], an admixture of a molecular state and a pure charmonium resonance [20], a tetraquark [21], and a cusp at the $D^0\bar{D}^{*0}$ threshold [22–24] have been proposed, and the structure of the state remains uncertain. Measurement of the lineshape in various decay modes can help to discriminate among different choices for the structure. In this paper, we examine two models for the lineshape in the decay to $D^0\bar{D}^{*0}$: a Breit-Wigner, and a Flatté-inspired parametrization.

The $X(3872)$ peak has already been analyzed with the Breit-Wigner lineshape commonly used for resonance states. Based on the analyses of the decays including J/ψ, the mass is 3871.65 ± 0.06 MeV/c^2 and the width is 1.19 ± 0.21 MeV [1], with two measurements at the LHCb experiment [25,26] contributing significantly to these averages. Analyses of the decay to $D^0\bar{D}^{*0}$ based on the Breit-Wigner lineshape tend to yield a higher mass and a larger width, with the width measurement subject to large uncertainties [10,11]. Discrepancies in the lineshape between the decays to the $J/\psi\pi^+\pi^-$ and $D^0\bar{D}^{*0}$ final states can arise near the threshold due to coupled-channel effects [22]. This may be significant for the $X(3872)$, as the observed mass coincides with the $D^0\bar{D}^{*0}$ threshold of 3871.69 ± 0.10 MeV/c^2, and a 1^{++} state can couple to the $D^0\bar{D}^{*0}$ channel in S-wave. One model to account for coupled-channel effects is the Flatté-inspired parametrization [22,27], a Breit-Wigner model with an explicit expression for the energy-dependent partial width. At LHCb, an analysis of the $J/\psi\pi^+\pi^-$ invariant mass distribution was performed using this Flatté-inspired model [26]. It is difficult to determine all of the parameters using only this distribution, due to a scaling behavior in which the lineshape near the threshold does not change under a linear transformation of four of the five parameters [26,28]. To determine all the parameters, it is important to analyze not only the $J/\psi\pi^+\pi^-$ decay but also the $D^0\bar{D}^{*0}$ decay, as proposed in the theoretical analysis [27]. By analyzing the $D^0\bar{D}^{*0}$ decay, we aim to provide more information on the lineshape, and in particular on the coupling strength of $X(3872) \to D^0\bar{D}^{*0}$.

In this paper, we present a study of the $X(3872)$ lineshape using a sample of $X(3872) \to D^0\bar{D}^{*0}$ candidates produced in the exclusive decay $B \to D^0\bar{D}^{*0}K$ using the full Belle dataset. There have been three previous studies [10–12]. Reference [12] is an analysis of the $B \to D^0\bar{D}^0\pi^0K$ decay at Belle, and Refs. [10,11] are analyses of the $B \to D^0\bar{D}^{*0}K$ decays at *BABAR* and Belle, respectively. The latter two analyses apply a D^{*0} selection and a mass-constrained fit to the D^{*0} candidates. While this has the advantage of improving the signal-to-noise ratio, it has the disadvantage of disallowing entries below the $D^0\bar{D}^{*0}$ threshold, which is important for studying the structure. Given the limited size of our data sample, we adopt a similar technique to the latter analyses, i.e. subtracting the reconstructed D^{*0} mass and adding the nominal mass. The disadvantage of requiring the D^{*0} is partially compensated for by analyzing the Flatté model, in which we can obtain a lineshape reflecting poles of the scattering amplitude. Compared to Refs. [10,11], additional D^0 decay modes are included, increasing the efficiency to reconstruct $D^0\bar{D}^{*0}$ decays. Throughout this paper, charge conjugation is always included. We do not distinguish $D^0\bar{D}^{*0}$ from $\bar{D}^0D^{*0}$ unless otherwise indicated.

The paper is organized as follows. In Sec. II, the Belle detector and dataset are described. In Secs. III and IV, the event selection and the fitted model are presented. In Sec. V, the results of fitting the data with the relativistic Breit-Wigner model and the Flatté model are presented. Section VI contains a discussion of the results, and the conclusions of the paper.

II. DETECTOR AND DATASET

We use a data sample of 772×10^6 $B\bar{B}$ pairs, collected at a center-of-mass energy of $\sqrt{s} = 10.58$ GeV, corresponding to the $\Upsilon(4S)$ resonance, with the Belle detector at the KEKB asymmetric-energy e^+e^- collider [29,30]. The Belle detector is a large-solid-angle magnetic spectrometer that consists of a silicon vertex detector (SVD), a 50-layer central drift chamber (CDC), an array of aerogel threshold Cherenkov counters (ACC), a barrel-like arrangement of time-of-flight scintillation counters (TOF), and an electromagnetic calorimeter comprised of CsI(Tl) crystals (ECL) located inside a super-conducting solenoid coil that provides a 1.5 T magnetic field. An iron flux-return located outside of the coil is installed to detect K_L^0 mesons and to identify muons (KLM). The detector is described in detail elsewhere [31,32].

To determine the event selection and the detector response, we use a sample of Monte Carlo (MC) simulated events generated using the `EvtGen` event generator [33]. The detector response is simulated using the GEANT3 package [34].

III. EVENT SELECTION

The event selection is determined using the MC samples in two steps. First, the selection criteria for the final-state

particles are determined based on our previous studies [11,12]. Second, the selection criteria for the intermediate-state particles are optimized by maximizing the figure-of-merit $S/\sqrt{S+B}$, where S and B are the estimated numbers of signal and background events, respectively. The resulting selection is described below.

Tracks are selected using vertex information measured by the tracking system. A track candidate is accepted if its distance along the detector axis from the point of closest approach to the interaction point is less than 4.0 cm, and its distance transverse to the detector axis is less than 1.0 cm. These requirements are not imposed for tracks in $K^0_S \to \pi^+\pi^-$ candidates. In addition, pion and kaon candidates are selected using likelihoods $\mathcal{L}_\pi$ and $\mathcal{L}_K$ based on the time-of-flight measured by the TOF, the number of Cherenkov photons detected by the ACC, and the ionization loss in the CDC. Tracks with a likelihood ratio $\mathcal{L}_\pi/(\mathcal{L}_\pi+\mathcal{L}_K) > 0.1$ are used as charged pion candidates, and tracks with $\mathcal{L}_\pi/(\mathcal{L}_\pi+\mathcal{L}_K) < 0.9$ are used as charged kaon candidates. The hadron identification efficiency is approximately 97% for both pions and kaons. Tracks satisfying $\mathcal{L}_e/(\mathcal{L}_e+\mathcal{L}_{\tilde{e}}) > 0.95$ are identified as electrons and eliminated. Here, $\mathcal{L}_e$ and $\mathcal{L}_{\tilde{e}}$ are distinct likelihoods for the electron and nonelectron hypotheses, based on ECL, tracking, and other information. The particle identification is described in detail elsewhere [35].

K^0_S candidates are reconstructed from charged pion pairs with opposite charges. The $\pi^+\pi^-$ invariant mass is required to agree with the known K^0_S mass [1] within 7 MeV/c^2 ($\approx 3.6\sigma$ of the resolution). Candidates are selected using a neural network classifier [36] with various kinematic variables as input. To improve the four-momentum resolution, a mass- and vertex-constrained fit is applied.

Photon candidates are reconstructed from ECL clusters with no matching charged tracks. Candidates are selected based on the ratio, E_9/E_{25}, of the energy deposited in the 3×3 array of crystals centered on the crystal with the highest energy deposition to that in the 5×5 array: we require $E_9/E_{25} > 0.8$.

Neutral pions are reconstructed from photon pairs. The photons are required to have energy greater than 30 MeV in the barrel region or 50 MeV in the endcaps. The $\gamma\gamma$ invariant mass is required to agree with the π^0 nominal mass [1] within 12 MeV/c^2. This mass window corresponds to 92% signal efficiency. A mass-constrained fit is applied to improve the momentum resolution.

D^0 candidates are reconstructed in six decay modes: $K^-\pi^+$, $K^-\pi^+\pi^0$, $K^-\pi^+\pi^-\pi^+$, $K^0_S\pi^+\pi^-$, $K^0_S\pi^+\pi^-\pi^0$, and K^+K^-. The π^0 candidates used in this reconstruction are required to have momentum in the center-of-mass system greater than 100 MeV/c, and energy in the laboratory system greater than 150 MeV. If a π^0 is included, the reconstructed D^0 invariant mass is required to be within 16 MeV/c^2 of the nominal mass [1] corresponding to 85% signal efficiency; otherwise, it is required to be within 8.5 MeV/c^2 corresponding to 91% efficiency. To improve the momentum resolution, a mass- and vertex-constrained fit is applied. Candidates where the χ^2 probability of the fit is less than 0.0001 are eliminated.

$\bar{D}^{*0}$ candidates are reconstructed in two decay modes: $\bar{D}^0\gamma$ and $\bar{D}^0\pi^0$. For the $\bar{D}^0\gamma$ mode, only γ candidates with an energy greater than 90 MeV in the laboratory system are used. For the $\bar{D}^0\pi^0$ mode, only π^0 candidates with a momentum in the center-of-mass system of less than 100 MeV/c and an energy in the laboratory system of less than 200 MeV are used. The difference in the reconstructed mass between $\bar{D}^{*0}$ and $\bar{D}^0$ is required to agree with the nominal value [1] within 9.0 MeV/c^2 and 2.0 MeV/c^2 for $\bar{D}^0\gamma$ and $\bar{D}^0\pi^0$, respectively, corresponding to 90% signal efficiency in each case.

B meson candidates are then reconstructed in the decay modes $D^0\bar{D}^{*0}K^+$ and $D^0\bar{D}^{*0}K^0_S$. To reduce wrong combinations, the daughter K^+ is required to have $\mathcal{L}_K/(\mathcal{L}_\pi+\mathcal{L}_K) > 0.6$, corresponding to an identification efficiency of 89%. The B candidates are selected based on the beam-energy constrained mass, $M_{\rm bc} \equiv \sqrt{(E^{\rm cms}_{\rm beam})^2 - (p^{\rm cms}_B)^2}$ and the difference of the energy in the center-of-mass system between the B candidate and the beam, $\Delta E \equiv E^{\rm cms}_B - E^{\rm cms}_{\rm beam}$, where $E^{\rm cms}_{\rm beam}$ is the beam energy in the center-of-mass system corresponding to half of $\sqrt{s}$, and $p^{\rm cms}_B$ and $E^{\rm cms}_B$ are the energy and momentum of B candidates in the center-of-mass system, respectively. We retain events with $M_{\rm bc} > 5.2$ GeV/c^2 and $|\Delta E| < 50$ MeV for later analysis. The $M_{\rm bc}$ signal region is defined as $|M_{\rm bc} - m_B| < 4.5$ MeV/c^2 ($\approx 2\sigma$) for $\bar{D}^{*0} \to \bar{D}^0\gamma$ and $|M_{\rm bc} - m_B| < 6.0$ MeV/c^2 ($\approx 2.5\sigma$) for $\bar{D}^{*0} \to \bar{D}^0\pi^0$, where m_B denotes the nominal B mass [1]. The ΔE signal region is defined as $|\Delta E| < 12$ MeV ($\approx 2\sigma$). For suppression of continuum events, we use a FastBDT classifier [37] trained on the simulation sample with the following event-shape information as input: modified Fox-Wolfram moments [38], the momentum flow in concentric cones around the thrust axis [39], and thrust-related quantities. Events for which the classifier output is less than 0.15 are eliminated. This requirement retains 96% of the signal candidates and rejects 49% of the candidates of continuum events.

After this selection, the average number of B candidates per event is 1.8, because $D^0\bar{D}^{*0}$ and $D^{*0}\bar{D}^0$ are often indistinguishable and double-counted. To avoid multiple counting of signal events, we select the candidate that has the highest value of the product of the following likelihood $\mathcal{L}$ and prior probability $\mathcal{P}$

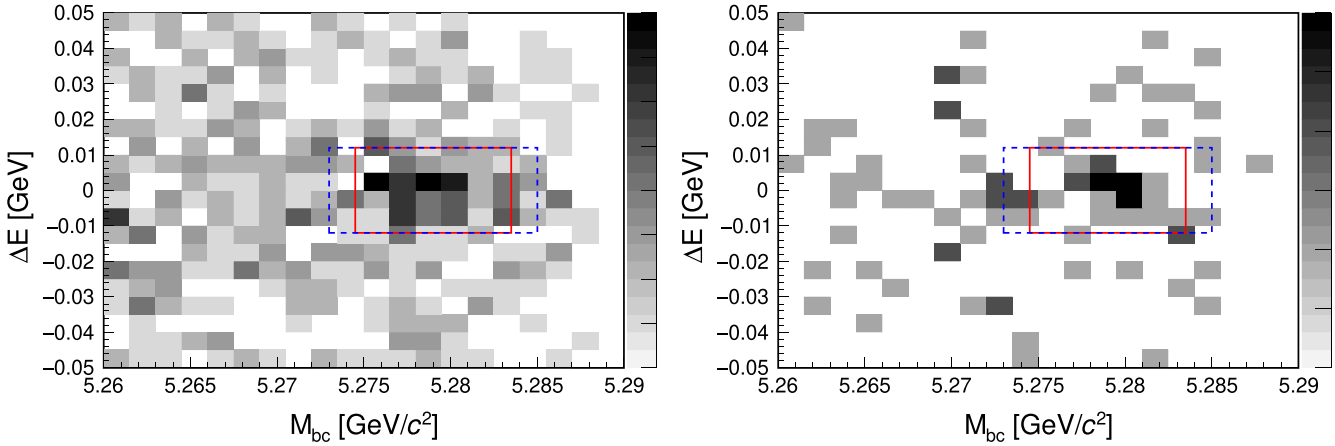

FIG. 1. Distributions of $(M_{\rm bc}, \Delta E)$ for B^+ (left) and B^0 (right) candidates in the $M(D^0\bar{D}^{*0}) < 3.88$ GeV/c^2 region, where the signal-to-background ratio for $X(3872)$ in data is relatively high. The red solid and blue dashed rectangles show the $(M_{\rm bc}, \Delta E)$ signal regions for $\bar{D}^{*0} \to \bar{D}^0\gamma$ and $\bar{D}^0\pi^0$, respectively.

$$\mathcal{L} = \mathcal{L}_{M(D^0)} \times \mathcal{L}_{M(\bar{D}^0)} \times \mathcal{L}_{M(\bar{D}^{*0})-M(\bar{D}^0)} \mathcal{L}_{\Delta E}[\times \mathcal{L}_{M(\pi^0)}],$$
$$\mathcal{P} = \frac{\varepsilon_{ijk}}{\zeta_{ijk}} \times \mathcal{B}(D^0 \to i) \times \mathcal{B}(\bar{D}^0 \to j) \times \mathcal{B}(\bar{D}^{*0} \to k), \quad (1)$$

where $\mathcal{L}$ is the product of the likelihoods of the measured D^0, $\bar{D}^0$, and $\bar{D}^{*0}$ masses, and ΔE; and, for the $\bar{D}^{*0} \to \bar{D}^0\pi^0$ mode, the likelihood of the measured π^0 mass. Each likelihood is obtained using probability density functions (PDFs) determined using the MC samples. The probability $\mathcal{P}$ is obtained from the probability that a signal event can be reconstructed ε_{ijk}, the average number of B candidates per event ζ_{ijk}, and the decay branching fraction, when D^0, $\bar{D}^0$, and $\bar{D}^{*0}$ are reconstructed in the i, j, and k modes, respectively. The values of ε_{ijk} and ζ_{ijk} are determined using the MC samples. The $(M_{\rm bc}, \Delta E)$ distribution of the selected B candidates is shown in Fig. 1. The red solid (blue dashed) rectangle shows the $(M_{\rm bc}, \Delta E)$ signal region for $\bar{D}^{*0} \to \bar{D}^0\gamma$ ($\bar{D}^0\pi^0$): B candidates used in the lineshape study are selected from this region.

For all events remaining in the selection, the following $D^0\bar{D}^{*0}$ invariant mass is calculated instead of applying a mass-constrained fit to improve the mass resolution,

$$M(D^0\bar{D}^{*0}) = \begin{cases} M(D^0\bar{D}^0\gamma) - M(\bar{D}^0\gamma) + m_{\bar{D}^{*0}} & \text{for } \bar{D}^{*0} \to \bar{D}^0\gamma, \\ M(D^0\bar{D}^0\pi^0) - M(\bar{D}^0\pi^0) + m_{\bar{D}^{*0}} & \text{for } \bar{D}^{*0} \to \bar{D}^0\pi^0, \end{cases} \quad (2)$$

where the reconstructed $\bar{D}^{*0}$ invariant mass, $M(\bar{D}^0\gamma)$ or $M(\bar{D}^0\pi^0)$, is subtracted, and the $\bar{D}^{*0}$ nominal mass, $m_{\bar{D}^{*0}}$, is added. The lineshape and signal yield are determined by fitting the distribution in the region below 4.0 GeV/c^2.

IV. FIT STRATEGY AND DETECTOR RESPONSE

In this work, the obtained $M(D^0\bar{D}^{*0})$ distributions are fitted with two lineshape models: the relativistic Breit-Wigner, and a Flatté-inspired model. The fits with these two models, shown in the next section, are performed with the following procedure.

When signal events are reconstructed correctly, the invariant mass distribution has a peak consisting of the natural lineshape convolved with the mass-dependent detector response. This response, i.e. the mass dependence of the signal efficiency and the mass resolution, is studied and parametrized using a set of $X(3872) \to D^0\bar{D}^{*0}$ MC samples generated with zero width, and a range of mass values from the $D^0\bar{D}^{*0}$ threshold to 4.0 GeV/c^2. Here, the $X(3872) \to D^0\bar{D}^{*0}$ decays are generated using a uniform phase space model; the D^{*0} width is assumed to be around 60 keV [23]. Since the signal-to-noise ratio depends on the D^{*0} decay mode, fits are performed separately for $D^{*0} \to D^0\gamma$ and $D^{*0} \to D^0\pi^0$. In addition, fits are performed separately for B^0 and B^+ candidates to determine the ratio of branching fractions between $B^0 \to X(3872)K^0$ and $B^+ \to X(3872)K^+$.

The fit function for correctly reconstructed $X(3872) \to D^0\bar{D}^{*0}$ decays, which we refer to as "signal", is constructed as follows. The signal efficiency varies depending on the mass by a few tens of percent, especially around the threshold, as shown in Fig. 2(a). It is parametrized by the threshold function $p_0\{1 - p_1 e^{p_2(M-m_{D^0}-m_{D^{*0}})} + p_3(M - m_{D^0} - m_{D^{*0}})\}$ with parameters p_0–p_3 in the low-mass region, which is continuously connected to a constant value in the high-mass region. The mass resolution for the signal is modeled as the sum of a Gaussian and a reversed Crystal Ball function [40] with a common mean. Figure 2(b) shows the $M(D^0\bar{D}^{*0})$ spread due to the resolution, and the resolution function used in this work,

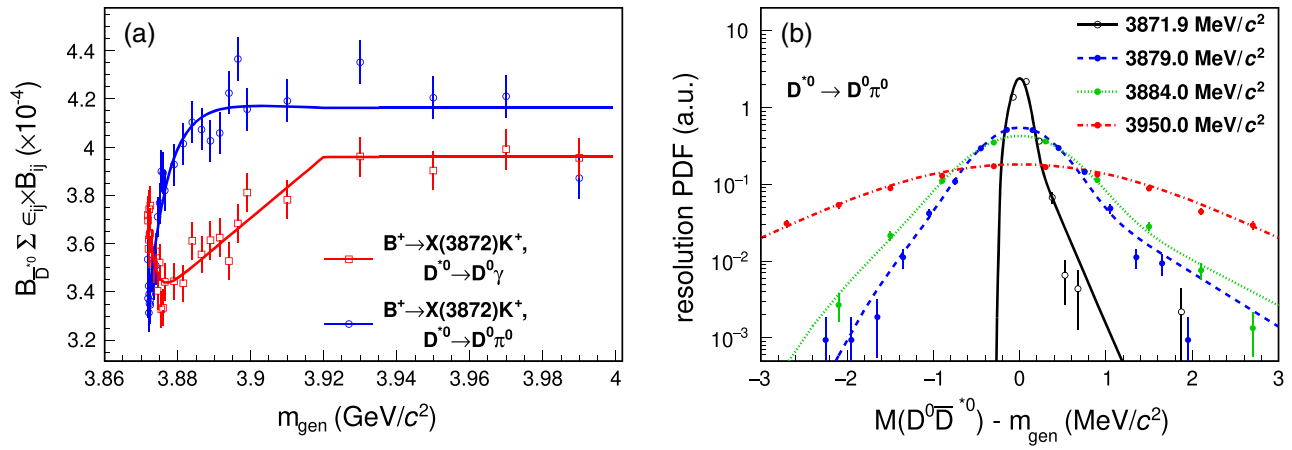

FIG. 2. The detector response for the signal component. (a) The sum of products of the signal efficiency and the branching fraction of the intermediate states $\mathcal{B}_{\bar{D}^{*0}} \sum \epsilon_{ij} \times \mathcal{B}_{ij}$ as a function of the $X(3872)$ mass generated in the MC samples for $B^+ \to X(3872)K^+$; the blue circles and the red squares are for $D^{*0} \to D^0\pi^0$ and $D^{*0} \to D^0\gamma$, respectively. The lines represent the parametrized efficiency functions. For $B^0 \to X(3872)K^0$, similar structures are obtained with a ratio of $B^+ \to X(3872)K^+$ to $B^0 \to X(3872)K^0$ of almost 4:1. (b) The $M(D^0\bar{D}^{*0})$ spread due to the detector response for the $X(3872)$ lineshape generated with zero width and masses of 3871.9 MeV/c^2, 3879.0 MeV/c^2, 3884.0 MeV/c^2, and 3950.0 MeV/c^2 for the $D^{*0} \to D^0\pi^0$ decay mode. The circles show normalized distributions obtained from the MC sample. The curves show the parametrized resolution functions. Similar results are obtained for $D^{*0} \to D^0\gamma$.

for several choices of the $X(3872)$ mass. As noted in the previous Belle study, the resolution degrades with the square root of the difference between the mass and the threshold [11]. The convolution with the mass-dependent resolution function entails longer computation times. The effect of smearing due to the resolution is small at masses away from the peak, since the natural lineshape is broad [10,11]. For example, the full width at half maximum (FWHM) of the natural lineshape is a few MeV, while the FWHM of the mass resolution near the peak is only about 220 keV. Therefore, instead of convolution with the mass-dependent resolution function, convolution with the specific resolution function at the mass of 3871.9 MeV/c^2 (near the peak) is adopted as an approximation. To reproduce the behavior near threshold, the signal function is multiplied by a soft threshold function that rises from zero to one at the threshold using an error function. The procedure is validated on $X(3872) \to D^0\bar{D}^{*0}$ MC samples generated with a broad lineshape. The effect of the approximation is negligible.

The ratios of signal yields among the decay modes are fixed in the fit using the product of the expected total signal efficiency and the branching fraction of each decay mode. Here the expected signal efficiency depends on the lineshape because of the mass dependence of the signal efficiency. The total signal efficiency is obtained by averaging the signal efficiencies as a function of mass weighted by the values of the lineshape function. It is then corrected by taking account of the signal which may leak out of the fit range, depending on the lineshape, and by taking the ratio of the area of the signal function in the fit range to that from the threshold to $m_B - m_K$. Here, a mass-dependent resolution function is convolved with the signal function, because smearing due to the resolution is important at higher masses. The calculation of the total signal efficiency is validated on MC samples for a broad range of lineshape parameters.

A separate fit function is used for "broken signal": cases where the D^0 is reconstructed incorrectly, a wrong π^0 or γ is combined in the $\bar{D}^{*0}$ reconstruction, or a $D^{*0}\bar{D}^0$ signal event is misinterpreted as $D^0\bar{D}^{*0}$ by combining π^0 or γ from D^{*0} incorrectly with the $\bar{D}^0$ to make a fake $\bar{D}^{*0}$. For the $\bar{D}^{*0} \to \bar{D}^0\pi^0$ mode, such events produce a broad peak in the $M(D^0\bar{D}^{*0})$ signal region and possibly distort the lineshape of the signal. The fit function for the broken signal therefore takes account of the mass dependencies of the resolution and the efficiency as in the case of correctly reconstructed signal events. The mass dependence of the efficiency is parametrized using the same threshold function used for the signal. The resolution is reproduced by a triple Gaussian multiplied by a soft threshold function at the $D^0\bar{D}^{*0}$ threshold, and its mass dependence is studied and parametrized using zero-width MC samples. Since the resolution for the broken signal is several times worse than that for the signal, we do not use approximated convolution: we instead use a discrete convolution with the mass-dependent resolution function. In the fit, the yield of the broken signal relative to the signal is fixed to the value expected from the lineshape and the ratio of the total efficiency of the broken signal to that of the signal.

The broken signal peak due to the $\bar{D}^{*0} \to \bar{D}^0\gamma$ mode has little sensitivity to the natural lineshape. To reduce the systematic uncertainty due to the shape, we use a histogram PDF depending on the lineshape. This PDF is obtained by

plotting the broken-signal histogram for each of the zero-width MC samples, scaling it by the value of the assumed lineshape at the generated mass, and summing up all of the scaled histograms. Here, the bin widths are adjusted to increase as the mass increases to suppress statistical fluctuations.

The background from $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$) continuum events, and $e^+e^- \to \Upsilon(4S) \to B\bar{B}$ events other than signal, is studied using the background MC sample. The shape of the invariant mass distribution is reproduced using a threshold function, $\sqrt{M - (m_{D^0} + m_{D^{*0}})}$, where m_{D^0} and $m_{D^{*0}}$ are the nominal masses of D^0 and D^{*0} [1], respectively.

V. FIT TO DATA

A. Relativistic Breit-Wigner model

The relativistic Breit-Wigner lineshape function is defined as [1]

$$f_{\rm BW}(M) = \frac{m_{\rm BW} M \Gamma(M)}{(M^2 - m_{\rm BW}^2)^2 + m_{\rm BW}^2 \Gamma(M)^2}, \tag{3}$$

where M is the observed invariant mass, and $m_{\rm BW}$ is the mass of the resonance. The mass-dependent width $\Gamma(M)$ is defined as

$$\Gamma(M) = \Gamma_{\rm BW} \frac{m_{\rm BW}}{M} \left(\frac{p(M)}{p(m_{\rm BW})} \right)^{2L+1}, \tag{4}$$

where $\Gamma_{\rm BW}$ and L are the width of the resonance and the orbital angular momentum, respectively. Taking account of the closeness to the threshold, the decay is assumed to be pure S-wave ($L = 0$) with no D-wave ($L = 2$) admixture. The momentum of one of the daughters in the rest frame of $X(3872)$, $p(M)$, can be calculated as

$$p(M) = \frac{1}{2M} \sqrt{(M^2 - (m_{D^0} + m_{D^{*0}})^2)} \times \sqrt{(M^2 - (m_{D^0} - m_{D^{*0}})^2)}. \tag{5}$$

Figure 3 presents the $M(D^0\bar{D}^{*0})$ distributions obtained from the data. Here, unbinned maximum likelihood fits are performed simultaneously to the distributions for the $\bar{D}^{*0}$ decay modes, $\bar{D}^{*0} \to \bar{D}\pi^0$ and $\bar{D}\gamma$, and for the B^+ and B^0 samples, with common fit parameters m and Γ_0. The PDFs are convolved with the detector response as described in the previous section. Table I summarizes the parameters obtained from the fit. The significance is determined from the log-likelihood ratio $-2\ln(\mathcal{L}_0/\mathcal{L})$ accounting for the difference in the number of degrees of freedom, where $\mathcal{L}_0$ and $\mathcal{L}$ are the fit likelihood without and with the peak component; i.e., the yield is constrained to be zero for the significance of each B mode, and the parameters m and Γ_0 are additionally dropped for the combined significance. Here the likelihood is smeared to take account of the systematic uncertainties on the signal yields as described below. The significance is found to be 5.9σ for $B^+ \to X(3872)K^+$, and 5.2σ for $B^0 \to X(3872)K^0$. The absence of peaks in the $M(D^0\bar{D}^{*0})$ distribution in the $(M_{\rm bc}, \Delta E)$ sideband region confirms that any contribution from peaking background is small; here the sideband region is defined as $12\ {\rm MeV}/c^2 < |M_{\rm bc} - m_B| < 20\ {\rm MeV}/c^2$ or $30\ {\rm MeV} < |\Delta E| < 50\ {\rm MeV}$.

The lineshape parameters are determined to be

$$m_{\rm BW} = 3873.71^{+0.56}_{-0.50}({\rm stat}) \pm 0.13({\rm syst})\ {\rm MeV}/c^2,$$
$$\Gamma_{\rm BW} = 5.2^{+2.2}_{-1.5}({\rm stat}) \pm 0.4({\rm syst})\ {\rm MeV}.$$

The difference between $m_{\rm BW}$ and the $D^0\bar{D}^{*0}$ threshold is found to be

$$m_{\rm BW} - (m_{D^0} + m_{D^{*0}}) = 2.02^{+0.56}_{-0.50}({\rm stat}) \pm 0.08({\rm syst})\ {\rm MeV}/c^2.$$

The systematic uncertainties are listed in Table II. We consider the following nine sources of uncertainty on the mass, the width, and the signal yield: (i) The uncertainty due to the assumed shape of the generic background is estimated by performing a fit after changing the PDF from the threshold function with a square root to an inverted ARGUS function [41]. (ii) The mass resolution is validated by comparing the data and MC ΔE resolution in the $B^+ \to \bar{D}^{*0}\pi^+\pi^-\pi^+$ control sample, which has a similar decay topology to $B \to X(3872)(\to D^0\bar{D}^{*0})K$. The ratios of the mass resolution obtained for MC and data are 1.01 ± 0.10 for $\bar{D}^{*0} \to \bar{D}^0\gamma$ and 1.08 ± 0.13 for $\bar{D}^{*0} \to \bar{D}^0\pi^0$. This resolution is consistent in data and MC, so no correction is applied, and the associated uncertainty is assigned by performing fits with the resolution varied by the precision, $\pm 1\sigma \equiv \pm 13\%$. (iii, iv) The uncertainties arising from the mass dependence of the efficiency and the ratio of the broken-signal to the signal are evaluated by summing in quadrature the changes induced by $\pm 1\sigma$ variations of the relevant parameters. (v) The uncertainty due to possible bias in the fit is evaluated by performing pseudoexperiments. The input value of a parameter subtracted from the median of the parameter distribution is regarded as the corresponding uncertainty. (vi) For the $m_{\rm BW}$ measurement only, the uncertainty arising from the finite precision of the D^0 mass and the $\Delta(m_{D^{*0}} - m_{D^0})$ mass difference is taken as the $\pm 1\sigma$ uncertainty of $2m_{D^0} + \Delta(m_{D^{*0}} - m_{D^0}) = 3871.69 \pm 0.10\ {\rm MeV}/c^2$ following Ref. [1]. (vii) The nonzero D^{*0} width ($\Gamma_{D^{*0}}$) leads to three potential sources of bias: a bias arising from the mass difference technique, a bias arising from the consideration of the D^{*0} width in the lineshape model, and a bias due to the interference between $X(3872) \to D^0\bar{D}^{*0}$ and $\bar{D}^0D^{*0}$. Biases from these three sources are evaluated as follows. For the first bias, two

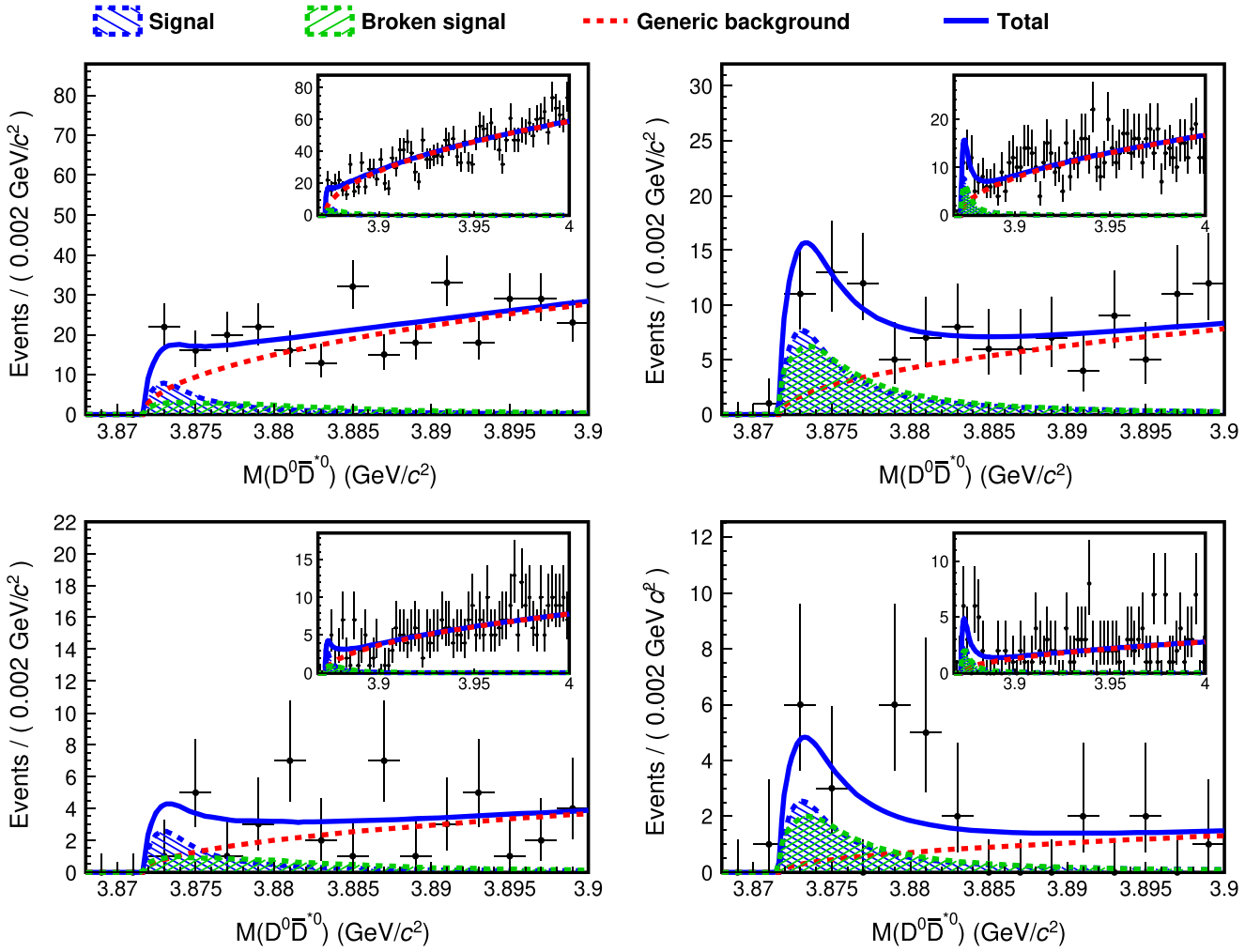

FIG. 3. The $M(D^0\bar{D}^{*0})$ distributions with the fit result with the relativistic Breit-Wigner lineshape for $B^+ \to X(3872)K^+$ (top) and $B^0 \to X(3872)K^0$ (bottom). The left and right rows are for $\bar{D}^{*0} \to \bar{D}^0\gamma$ and $\bar{D}^{*0} \to \bar{D}^0\pi^0$, respectively. The points with error bars represent data. The blue solid line shows the total fit result. The blue and green dashed lines show the signal contributions and broken signal contributions, respectively. The red dashed line shows the contribution of generic background.

$M(D^0\bar{D}^{*0})$ distributions are formed in MC with a broad lineshape: one where $m_{D^{*0}}$ in Eq. (2) is fixed to the nominal value (as in our analysis), and the other where $m_{D^{*0}}$ is replaced by the true D^{*0} mass generated by EvtGen, where $\Gamma_{D^{*0}} = 65.5$ keV [23] is assumed. Each distribution is fitted with the PDF of the signal component, and the largest difference is regarded as the associated uncertainty. For the second bias, the distribution for data is fitted after smearing the assumed lineshape with a Breit-Wigner function of $\Gamma_{D^{*0}} = 65.5$ keV, and the change from the original result is regarded as the associated uncertainty. The third bias is ignored since the interference effect is negligible above the threshold [42]. The uncertainties associated with the first and second biases are added in quadrature. (viii) Limited MC statistics lead to uncertainty on the shape of the broken-signal for $\bar{D}^{*0} \to \bar{D}^0\gamma$. This is

TABLE I. Results using the relativistic Breit-Wigner lineshape: the fitted mass, width and signal yield, the total signal efficiency, and the significance.

Mode	m (MeV/c^2)	Γ_0 (MeV)	$N_{\rm sig}$	$\mathcal{B}_{\bar{D}^{*0}} \sum \epsilon_{ij} \times \mathcal{B}_{ij}$	Significance
Combined	$3873.71^{+0.56}_{-0.50}$	$5.2^{+2.2}_{-1.5}$	$70.5^{+13.6}_{-11.5}$	8.70×10^{-4}	7.5σ
$X(3872)K^+$	⋯	⋯	$53.2^{+11.6}_{-9.8}$	6.92×10^{-4}	5.9σ
$X(3872)K^0$	⋯	⋯	$17.3^{+4.7}_{-4.1}$	1.78×10^{-4}	5.2σ

TABLE II. Summary of systematic uncertainty for the mass, width, and branching fractions measurements using the relativistic Breit-Wigner lineshape.

Source	m (MeV/c^2)	Γ_0 (MeV)	$X(3872)K^+$ (%)	$X(3872)K^0$ (%)	Ratio (K^0/K^+) (%)
(i) Generic BG PDF	±0.07	±0.38	±8.2	±1.4	±6.7
(ii) Mass resolution	±0.02	$-0.11/+0.13$	$-0.2/+0.4$	$-0.3/+0.4$	$-0.1/+0.0$
(iii) Mass dependence of efficiency	±0.02	$-0.08/+0.07$	$-2.7/+2.0$	$-2.3/+1.7$	$-0.5/+0.6$
(iv) Ratio of broken-signal BG to signal	±0.01	±0.02	±2.1	±0.6	±2.1
(v) Fit bias	$-0.02/+0.00$	$-0.02/+0.00$	$-1.3/+0.0$	$-7.3/+0.0$	$-4.5/+0.0$
(vi) D^{*0} and D^0 masses	±0.10	...	...	...	...
(vii) D^{*0} width	$-0.01/+0.02$	±0.02	±0.0	±0.0	±0.0
(viii) Broken-signal shape for $\bar{D}^{*0} \to \bar{D}^0\gamma$	±0.00	±0.01	±0.1	±0.1	±0.1
(ix) Signal ratio of $\bar{D}^{*0} \to \bar{D}^0\gamma$ to $\bar{D}^0\pi^0$	±0.01	±0.05	±0.8	±0.2	±0.6
(x) Tracking efficiency	...	...	±2.1	±2.4	±0.3
(xi) PID efficiency	...	...	±2.9	±2.4	±0.4
(xii) K^0_S efficiency	...	...	±0.2	±1.0	±0.8
(xiii) π^0 reconstruction	...	...	±1.9	±1.9	...
(xiv) γ reconstruction	...	...	±1.5	±1.5	...
(xv) $\sum \epsilon_{ij} \times \mathcal{B}_{ij}$	...	...	±1.4	$-3.1/+2.3$	$-1.7/+0.9$
(xvi) $N_{B\bar{B}}$	...	...	±1.4	±1.4	...
(xvii) $\mathcal{B}(\Upsilon(4S) \to B\bar{B})$	...	...	±1.2	±1.2	±2.4
Total	±0.13	±0.4	±10	$-9.6/+5.7$	$-9.0/+7.6$

evaluated by repeating the fit while varying each bin entry of the MC PDF histogram assuming Poisson distributions. The 68% interval of the distributions of the resulting fit values is used to assign the uncertainty. (ix) The uncertainty arising from the fixed ratio of the signal yields for $\bar{D}^{*0} \to \bar{D}^0\gamma$ to $\bar{D}^0\pi^0$ is evaluated by performing new fits, and varying the ratio of branching fractions between $D^{*0} \to D^0\gamma$ and $D^{*0} \to D^0\pi^0$ by $\pm1\sigma$ [1]. The difference from the original result is treated as the uncertainty.

The product branching fraction is calculated as follows:

$$\mathcal{B}(B \to X(3872)K) \times \mathcal{B}(X(3872) \to D^0\bar{D}^{*0}) = \frac{N_{\rm sig}}{2N_{B\bar{B}}\mathcal{B}(\Upsilon(4S) \to B\bar{B})\mathcal{B}_{\bar{D}^{*0}} \sum \epsilon_{ij} \times \mathcal{B}_{ij}}, \quad (6)$$

where $\mathcal{B}_{\bar{D}^*}$ is the appropriate $\bar{D}^{*0}$ branching fraction, and $\sum \epsilon_{ij} \times \mathcal{B}_{ij}$ is the sum of efficiencies multiplied by the product of branching fractions for the various $D^0 \to i$ and $\bar{D}^0 \to j$ decay modes used. For $\mathcal{B}(\Upsilon(4S) \to B\bar{B})$, 0.514 and 0.486 are assigned for the B^+B^- and $B^0\bar{B}^0$ modes, respectively [1]. The results are

$$\begin{aligned}
&\mathcal{B}(B^+ \to X(3872)K^+) \times \mathcal{B}(X(3872) \to D^0\bar{D}^{*0}) \\
&= (0.97^{+0.21}_{-0.18}(\text{stat}) \pm 0.10(\text{syst})) \times 10^{-4}, \\
&\mathcal{B}(B^0 \to X(3872)K^0) \times \mathcal{B}(X(3872) \to D^0\bar{D}^{*0}) \\
&= (1.30^{+0.36}_{-0.31}(\text{stat})^{+0.12}_{-0.07}(\text{syst})) \times 10^{-4}.
\end{aligned}$$

Here we consider the following eight sources of systematic uncertainties in addition to those previously described for the lineshape parameters; (x) The uncertainty of the tracking efficiency is estimated using a $D^{*+} \to \pi^+D^0(\to \pi^+\pi^-K^0_S)$ sample for tracks with high momentum. The efficiency is consistent in data and MC; the precision of the test, 0.35% per track, is taken as a systematic uncertainty. For tracks with low momentum, the sample of soft π^- in $D^{*-} \to \bar{D}^0\pi^-$ in the $B^0 \to D^{*-}\pi^+$ decay is used. The ratio of tracking efficiency obtained for MC and data is applied as a correction factor. The uncertainty in the correction factor is regarded as a systematic uncertainty. (xi, xii, xiii) Efficiencies for hadron identification, K^0_S selection, and π^0 detection are evaluated using control samples: $D^{*+} \to D^0(\to K^-\pi^+)\pi^+$, $D^{*+} \to D^0(\to K^0_S\pi^0)\pi^+$, and $\tau^- \to \pi^-\pi^0\nu_\tau$, respectively. In each case a correction factor is applied to the signal efficiency based on the ratio of efficiencies obtained for MC and data, and the uncertainty on the correction factor is taken as the associated systematic uncertainty. (xiv) The uncertainty of the efficiency of γ detection is evaluated using a $B^+ \to \chi_{c1}(\to J/\psi\gamma)K^+$ sample: 3.0% is assigned for the $D^{*0} \to D^0\gamma$ decay mode. (xv) The uncertainty on $\sum \epsilon_{ij} \times \mathcal{B}_{ij}$ mainly arises from the uncertainties on the D^0 branching fractions, and the limited size of the signal MC sample. In addition, validation of the calculation method for the total signal efficiency shows input-output differences in the B^0 decay mode larger than expected from statistical fluctuations: the largest of these is assigned as a systematic uncertainty. The uncertainties from these sources are added. (xvi) The number of $B\bar{B}$ pairs in the dataset is measured to be $(772 \pm 11) \times 10^6$: the associated uncertainty is set to 1.4%. (xvii) The uncertainties on the branching fractions $\mathcal{B}(\Upsilon(4S) \to B^+B^-) = (51.4 \pm 0.6)\%$ and $\mathcal{B}(\Upsilon(4S) \to B^0\bar{B}^0) = (48.6 \pm 0.6)\%$ [1] are also included.

The ratio of branching fractions between $B^0 \to X(3872)K^0$ and $B^+ \to X(3872)K^+$ is found to be

$$\frac{\mathcal{B}(B^0 \to X(3872)K^0)}{\mathcal{B}(B^+ \to X(3872)K^+)} = 1.34^{+0.47}_{-0.40}(\text{stat})^{+0.10}_{-0.12}(\text{syst}),$$

with the same sources of systematic uncertainty as for the branching fractions; some sources cancel, or partially cancel, in the ratio (see Table II).

B. Flatté model

The Flatté-inspired parametrization is defined as follows using the energy from the $D^0\bar{D}^{*0}$ threshold, $E = M - (m_{D^0} + m_{D^{*0}})$ [22,27,43]:

$$f_{\text{Flatte}}(E) = \frac{g k_{D^0\bar{D}^{*0}}}{|D(E)|^2}, \tag{7}$$

$$D(E) = \begin{cases} E - E_f - \frac{1}{2}g\kappa_{D^+D^{*-}} + \frac{i}{2}[g k_{D^0\bar{D}^{*0}} + \Gamma(E)] & \text{for } 0 < E < \delta, \\ E - E_f + \frac{i}{2}[g(k_{D^0\bar{D}^{*0}} + k_{D^+D^{*-}}) + \Gamma(E)] & \text{for } E > \delta, \end{cases} \tag{8}$$

where $E_f = m_0 - (m_{D^0} + m_{D^{*0}})$ is the mass difference of this state (m_0) from the threshold, and g is the coupling constant for the $D\bar{D}^*$ channels; we assume the coupling constants for the $D^0\bar{D}^{*0}$ and D^+D^{*-} channels are the same due to isospin symmetry. The momenta k_a and κ_a for the channel a are measured in the rest frame of the $X(3872)$. They are expressed using the reduced mass μ as

$$\begin{aligned} k_{D^0\bar{D}^{*0}} &= \sqrt{2\mu_{D^0\bar{D}^{*0}}E}, \\ k_{D^+D^{*-}} &= \sqrt{2\mu_{D^+D^{*-}}(E-\delta)}, \\ \kappa_{D^+D^{*-}} &= \sqrt{2\mu_{D^+D^{*-}}(\delta-E)}, \\ \delta &= (m_{D^+} + m_{D^{*-}}) - (m_{D^0} + m_{D^{*0}}). \end{aligned} \tag{9}$$

The energy-dependent width $\Gamma(E)$ is defined by

$$\Gamma(E) = \Gamma_{J/\psi\rho}(E) + \Gamma_{J/\psi\omega}(E) + \Gamma_0, \tag{10}$$

where Γ_a is the partial width for the channel a. For the $J/\psi\rho$ and $J/\psi\omega$ channels, the dependence on E is defined as follows using the phase space and effective coupling constants, f_ρ and f_ω [27]:

$$\Gamma_{J/\psi\rho}(E) = f_\rho \int_{2m_\pi}^{M(E)-m_{J/\psi}} \frac{dm'}{2\pi} \frac{q(m',E)\Gamma_\rho}{(m'-m_\rho)^2 + \Gamma_\rho^2/4}, \tag{11}$$

$$\Gamma_{J/\psi\omega}(E) = f_\omega \int_{3m_\pi}^{M(E)-m_{J/\psi}} \frac{dm'}{2\pi} \frac{q(m',E)\Gamma_\omega}{(m'-m_\omega)^2 + \Gamma_\omega^2/4}, \tag{12}$$

where Γ_ρ and Γ_ω are total widths for the ρ and ω resonances, respectively. The upper bound of the integral is set by the difference between

$$M(E) = E + (m_{D^0} + m_{D^{*0}}) \tag{13}$$

and $m_{J/\psi}$. In each case, $q(m', E)$ is the momentum of the two- or three-pion system in the rest frame of the $X(3872)$,

$$q(m',E) = \frac{1}{2M(E)}\sqrt{M^2(E) - (m' + m_{J/\psi})^2} \times \sqrt{M^2(E) - (m' - m_{J/\psi})^2}. \tag{14}$$

The parameter Γ_0 is the sum of the partial widths of other channels, such as radiative decays. In total, this model has five free parameters, E_f, g, f_ρ, f_ω, and Γ_0.

To obtain stable results in the fit, we apply two constraints, which were also used in the previous study at LHCb [26]. The first is to fix f_ω so that the branching fraction of the $J/\psi\pi^+\pi^-$ mode and that of the $J/\psi\omega$ mode are equal, consistent with experimental results to date [44–46]. Based on the feature that the area under the lineshape for a channel is proportional to the branching fraction, f_ω can be uniquely determined by calculating the ratio of the lineshape area in the $J/\psi\pi^+\pi^-$ channel to that in $J/\psi\omega$. The second is a soft constraint on the ratio of branching fractions between the $J/\psi\pi^+\pi^-$ and $D^0\bar{D}^{*0}$ decay modes: for each of the B^+ and B^0 modes, the $J/\psi\pi^+\pi^-$ product branching fraction is calculated as follows, and a Gaussian constraint to the measured value [47] is included in the fit,

$$\begin{aligned} &\mathcal{B}(B \to X(3872)K) \times \mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-) \\ &\quad = R_{D\bar{D}^*} \times \mathcal{B}(B \to X(3872)K) \times \mathcal{B}(X(3872) \to D^0\bar{D}^{*0}) \\ &\quad = \begin{cases} (8.61 \pm 0.32) \times 10^{-6} & \text{for the } B^+\text{mode} \\ (4.1 \pm 1.1) \times 10^{-6} & \text{for the } B^0\text{mode} \end{cases}, \end{aligned} \tag{15}$$

where $R_{D\bar{D}^*}$ is the ratio of the lineshape area in the $J/\psi\pi^+\pi^-$ channel to that in $D^0\bar{D}^{*0}$, and the $D^0\bar{D}^{*0}$ product branching fraction is given by Eq. (6).

There are too few events in our $X(3872) \to D^0\bar{D}^{*0}$ sample to simultaneously determine the four remaining parameters. Therefore, we focus on the parameter regions where scaling behavior was observed at LHCb [26]. We search for the best lineshape fitted to the $M(D^0\bar{D}^{*0})$ distribution when the following ratios of parameters are

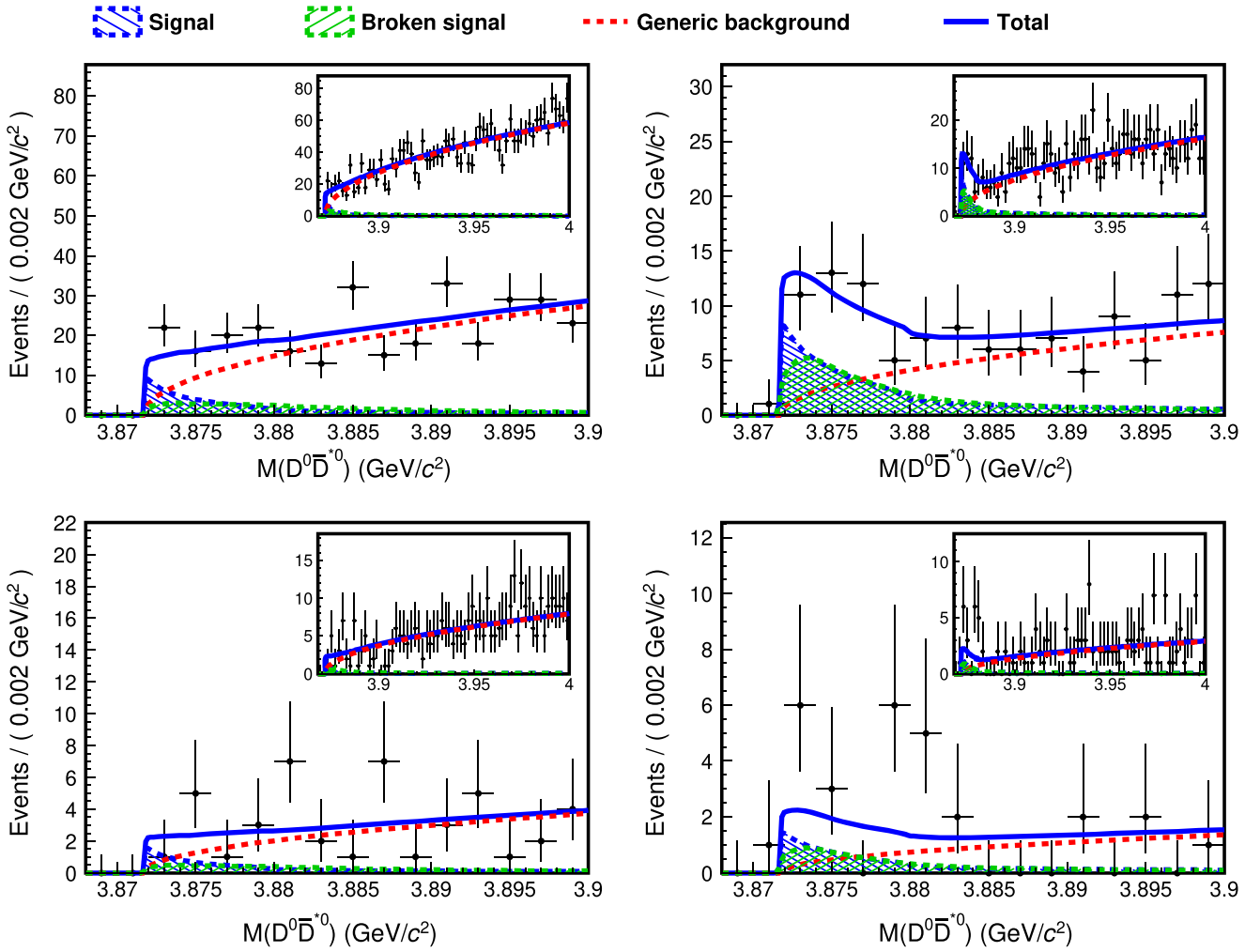

FIG. 4. The $M(D^0\bar{D}^{*0})$ distributions with the fit result with the Flatté lineshape for $B^+ \to X(3872)K^+$ (top) and $B^0 \to X(3872)K^0$ (bottom). The left and right rows are for $\bar{D}^{*0} \to \bar{D}^0\gamma$ and $\bar{D}^{*0} \to \bar{D}^0\pi^0$, respectively. The points with error bars represent data. The blue solid line shows the total fit result. The blue and green dashed lines show the signal contributions and broken-signal contributions, respectively. The red dashed line shows the contribution of generic background.

fixed to the values measured at LHCb: dg/dE_f is fixed to $-15.11\ \mathrm{GeV}^{-1}$, and f_ρ/E_f and Γ_0/E_f are fixed based on the measurements $f_\rho = 1.8 \times 10^{-3}$ and $\Gamma_0 = 1.4$ MeV, and the assumption $E_f = -7.2$ MeV. Thus, only g is floated as a free parameter.

We perform a simultaneous unbinned maximum likelihood fit under the above fit conditions. The fit results for the data are shown in Fig. 4 and Table III. The fitted g is $0.29^{+2.69}_{-0.15}$, where the uncertainty is statistical. Systematic uncertainties are summarized in Table IV. The method to evaluate the uncertainties due to the sources (i) to (ix) is the same as in the measurement of the relativistic Breit-Wigner

TABLE III. Results using the Flatté lineshape: the fitted coupling constant g, and the signal yield.

Mode	g	N_{sig}
Combined	$0.29^{+2.69}_{-0.15}$	$90.9^{+11.3}_{-15.9}$
$X(3872)K^+$	⋯	$77.9^{+9.6}_{-13.5}$
$X(3872)K^0$	⋯	$13.0^{+3.0}_{-2.9}$

TABLE IV. Summary of systematic uncertainties for the coupling constant g of the Flatté lineshape.

Source	g
(i) Generic BG PDF	$< O(0.001)$
(ii) Mass resolution	$-0.011/+0.003$
(iii) Mass dependence of efficiency	$-0.012/+0.024$
(iv) Ratio of broken-signal BG to signal	$-0.007/+0.020$
(v) Fit bias	$-0.000/+\infty$
(vi) D^{*0} and D^0 masses	⋯
(vii) D^{*0} width	$-0.006/+0.001$
(viii) Broken-signal shape for $\bar{D}^{*0} \to \bar{D}^0\gamma$	$-0.001/+0.002$
(ix) Signal ratio of $\bar{D}^{*0} \to \bar{D}^0\gamma$ to $\bar{D}^0\pi^0$	$-0.000/+0.004$
(x)–(xvii) Branching fraction	$-0.021/+0.042$
Total	$-0.029/+\infty$

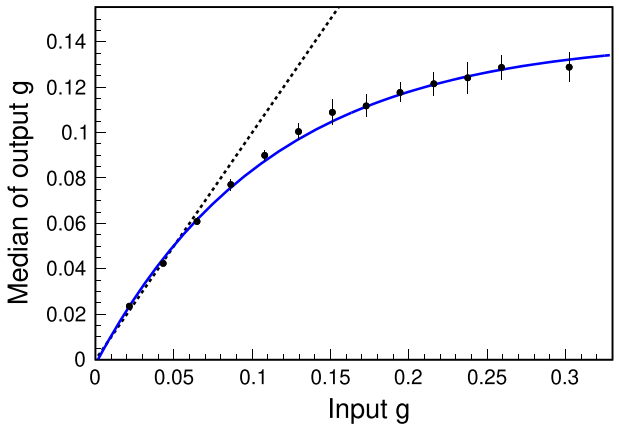

FIG. 5. The median of output values of the coupling constant g, as a function of the input g, evaluated using pseudoexperiments. The dotted black line represents perfect linearity $g_{\text{out}} = g_{\text{in}}$. The solid blue curve represents the threshold function $g_{\text{out}} = 0.14(1 - \exp(-9g_{\text{in}}))$.

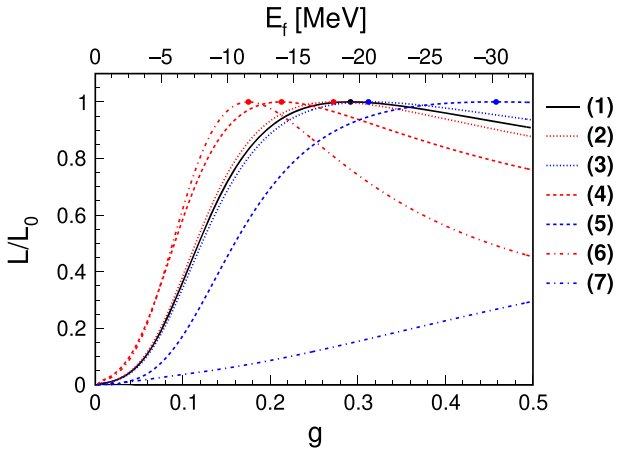

FIG. 6. For each of seven parameter sets, the likelihood ratio L/L_0 is shown, as a function of the coupling constant g, where $L = L(g)$ is the fitted likelihood and L_0 is the likelihood of the best fit for that parameter set. The solid black line shows the parameter set (1). The red and blue dotted lines show parameter sets (2) and (3), respectively. The red and blue dashed lines show sets (4) and (5), and the red and blue dot-dashed lines show sets (6) and (7), respectively. The parameter sets are described in Table V. Circles on the lines show the best fit g.

lineshape. Sources (x) to (xvii) also contribute through the constraint on the branching fraction applied in the fit. They are evaluated by performing fits after varying each parameter by $\pm 1\sigma$, and adding the resulting changes in quadrature. Regarding the fitter bias (v), the relationship between input values and medians of output values is evaluated using pseudoexperiments, as shown in Fig. 5. This study shows that for this sample size, g is likely to be underestimated as g increases, with the median of the output values converging to around 0.14. The main reason is that the lineshape converges to a fixed form for large g (given the assumed ratios for the other parameters), and fits fail, especially in determining an upper statistical uncertainty.

Since there is no input value for g for which the median of output values is 0.29, we cannot determine a central value for g. We can however set a lower limit. The likelihood including the systematic uncertainties listed in Table IV, $L(g)$, is shown as the black solid line in Fig. 6. Noting that the curve is asymmetric, with a larger integral above than below the best fit value, we conservatively set the lower limit g_{lower} from

$$\int_{g_{\text{lower}}}^{g_{\text{best}}} L(g)dg = = \begin{cases} 0.8\int_0^{g_{\text{best}}} L(g)dg & \text{for } 90\% \text{ credibility}, \\ 0.9\int_0^{g_{\text{best}}} L(g)dg & \text{for } 95\% \text{ credibility}, \end{cases} \tag{16}$$

where g_{best} denotes the coupling constant at the maximum likelihood. The effect of fixing dg/dE_f, f_ρ, and Γ_0 to the values measured by LHCb is studied by varying each parameter by $\pm 1\sigma$. Separate curves of the relative likelihood L/L_0 for each case are also shown in Fig. 6, where $L = L(g)$ is the likelihood of the fit and L_0 is the likelihood of the best fit for each parameter set. The corresponding fit results and lower limits are summarized in Table V. The L_0 values for the different parameter sets vary in a small range around the value for set (1): the best is favored by only 1.2σ relative to set (1), and the worst is disfavored by 3.4σ. The loosest lower limit is obtained for the parameter set (6), one of the disfavored scenarios, where f_ρ is changed by $+1\sigma$. We conservatively choose these as the final lower limits for this study:

$$g > 0.094 \text{ at } 90\% \text{ credibility},$$
$$g > 0.075 \text{ at } 95\% \text{ credibility}.$$

These correspond to upper limits of $E_f < -6.2$ MeV at 90% credibility and $E_f < -5.0$ MeV at 95% credibility, which are derived from $dg/dE_f = -15.11$ GeV^{-1}.

We investigate which lineshape model fits the $M(D^0\bar{D}^{*0})$ distribution better using the test statistic $t = -2\ln(\mathcal{L}_{\text{BW}}/\mathcal{L}_{\text{Flatte}})$. Here, $\mathcal{L}_{\text{BW}}$ is the best fit likelihood for the Breit-Wigner lineshape, and $\mathcal{L}_{\text{Flatte}}$ is the best likelihood without the $R_{D\bar{D}^*}$ constraint term, for the Flatté lineshape with parameter set (1). For data, we obtain $t = -8.5$; i.e. the Breit-Wigner lineshape is favored. Based on the t distribution obtained from pseudoexperiments, the exclusion level for the Flatté lineshape is only 2.2σ; this level decreases when the systematic uncertainties are taken into account. Thus, neither lineshape can be excluded.

Additionally, the consistency of the two lineshape measurements is confirmed using pseudoexperiments. The Breit-Wigner parameters measured for the data are

TABLE V. Summary of the seven parameter sets used in the evaluation of lower limits on the coupling constant g, showing the g of the best fit, the g lower limits, and corresponding E_f upper limits. The parameter sets are the center values of dg/dE_f, Γ_0, and f_ρ measured at LHCb [26] (1), changing dg/dE_f by $+1\sigma$ (2), changing dg/dE_f by -1σ (3), changing Γ_0 by $+1\sigma$ (4), changing Γ_0 by -1σ (5), changing f_ρ by $+1\sigma$ (6), and changing f_ρ by -1σ (7). For the parameter set (7), no lower limit is determined, because no best fit is found in the range $g < 50$.

Parameter set	(1)	(2)	(3)	(4)	(5)	(6)	(7)
dg/dE_f (GeV^{-1})	-15.11	$-14.95(+1\sigma)$	$-15.27(-1\sigma)$	-15.11	-15.11	-15.11	-15.11
Γ_0/E_f	-0.19	-0.19	-0.19	$-0.29(+1\sigma)$	$-0.09(-1\sigma)$	-0.19	-0.19
f_ρ/E_f (GeV^{-1})	-0.25	-0.25	-0.25	-0.25	-0.25	$-0.38(+1\sigma)$	$-0.12(-1\sigma)$
g of best fit	0.29	0.27	0.31	0.21	0.46	0.17	>50
g lower limit at 90% CL	>0.143	>0.136	>0.151	>0.105	>0.212	>0.094	⋯
at 95% CL	>0.113	>0.108	>0.119	>0.082	>0.167	>0.075	⋯
E_f upper limit at 90% CL (MeV)	<-9.5	<-9.0	<-10.0	<-6.9	<-14.0	<-6.2	⋯
at 95% CL (MeV)	<-7.6	<-7.2	<-7.9	<-5.5	<-11.1	<-5.0	⋯

consistent with those obtained from pseudoexperiments generated with the observed Flatté lineshape.

VI. DISCUSSION AND CONCLUSION

In this paper, we examine the $X(3872) \to D^0\bar{D}^{*0}$ lineshape using the full Belle dataset. When fitting with a relativistic Breit-Wigner lineshape, the mass and width parameters are measured to be

$$m_{\rm BW} = 3873.71^{+0.56}_{-0.50}(\text{stat}) \pm 0.13(\text{syst})\ \text{MeV}/c^2,$$
$$\Gamma_{\rm BW} = 5.2^{+2.2}_{-1.5}(\text{stat}) \pm 0.4(\text{syst})\ \text{MeV}.$$

The difference between $m_{\rm BW}$ and the $D^0\bar{D}^{*0}$ threshold is found to be $2.02^{+0.56}_{-0.50}(\text{stat}) \pm 0.08(\text{syst})\ \text{MeV}/c^2$. These values are in good agreement with those measured in previous studies of the $D^0\bar{D}^{*0}$ decay [10,11], and the precision of the measurement is improved by at least 22%. The measured branching fractions are as follows:

$$\mathcal{B}(B^+ \to X(3872)K^+) \times \mathcal{B}(X(3872) \to D^0\bar{D}^{*0}) = (0.97^{+0.21}_{-0.18}(\text{stat}) \pm 0.10(\text{syst})) \times 10^{-4},$$
$$\mathcal{B}(B^0 \to X(3872)K^0) \times \mathcal{B}(X(3872) \to D^0\bar{D}^{*0}) = (1.30^{+0.36}_{-0.31}(\text{stat})^{+0.12}_{-0.07}(\text{syst})) \times 10^{-4}.$$

This is the first measurement of $X(3872)$ production in B^0 decays with more than 5σ significance. The ratio of the branching fractions is determined to be

$$\frac{\mathcal{B}(B^0 \to X(3872)K^0)}{\mathcal{B}(B^+ \to X(3872)K^+)} = 1.34^{+0.47}_{-0.40}(\text{stat})^{+0.10}_{-0.12}(\text{syst}).$$

These results are in good agreement with those of previous studies [10–12].

We compare these results with the analysis of the Breit-Wigner lineshape using the $J/\psi\pi^+\pi^-$ decay mode. The measured Breit-Wigner mass is significantly higher than the $D^0\bar{D}^{*0}$ threshold, while the world-average mass with the $J/\psi\pi^+\pi^-$ decay is consistent with the threshold. The measured width and ratio $\mathcal{B}(B^0 \to X(3872)K^0)/\mathcal{B}(B^+ \to X(3872)K^+)$ are shifted from the average with the $J/\psi\pi^+\pi^-$ decay by 2.6σ and 2.0σ, respectively [50]. In previous studies of the $D^0\bar{D}^{*0}$ decay, it has also been seen that these properties in the $X(3872) \to D^0\bar{D}^{*0}$ decay mode differ from those in $J/\psi\pi^+\pi^-$.

We also fit the lineshape using a Flatté-inspired parametrization. With sufficient data, such a model could be used to simultaneously describe the lineshapes of the decays to the $J/\psi\pi^+\pi^-$ and $D^0\bar{D}^{*0}$ final states. Given the limited size of the $D^0\bar{D}^{*0}$ data sample at Belle, and the scaling behavior observed in the LHCb study of $J/\psi\pi^+\pi^-$, we set various ratios of parameters to their LHCb values, and fit with the coupling constant to the $D\bar{D}^*$ channel, g, as the undetermined parameter. We find that the fitted value of g is in a region that is relatively insensitive to the underlying value. We determine its lower limits to be

$$g > 0.094 \text{ at } 90\% \text{ credibility},$$
$$g > 0.075 \text{ at } 95\% \text{ credibility}.$$

These correspond to upper limits of $E_f < -6.2$ MeV at 90% credibility and $E_f < -5.0$ MeV at 95% credibility, which are slightly more stringent than the LHCb measurement, $-270\ \text{MeV} < E_f < -2.0\ \text{MeV}$ [26]. This suggests that analysis using $D^0\bar{D}^{*0}$ can indeed complement the study of the $J/\psi\pi^+\pi^-$ mode in this framework. The limit includes the solution $E_f = -7.2$ MeV assumed in the scattering amplitude analysis at LHCb. There is still uncertainty in the pole positions of the scattering amplitude, because the limit is not especially stringent.

Both Breit-Wigner and Flatté lineshapes fit the invariant mass distribution obtained from the data. Finally, we

examine which lineshape model best fits the invariant mass distribution. Based on the likelihood ratio from the fits, the Breit-Wigner lineshape is favored, but the Flatté lineshape is not excluded.

Analysis of the large dataset expected from Belle II will be important, because the statistical uncertainty dominates in both of the lineshape measurements. In the Flatté study, it is essential to reduce systematic uncertainty due to the fit bias and the measurement of parameters f_ρ and Γ_0. Increasing the size of the data sample will also reduce these uncertainties. In addition, a simultaneous fit of the $J/\psi\pi^+\pi^-$ and $D^0\bar{D}^{*0}$ decay modes will also be useful, because the ratio of branching fractions can further constrain the parameters. In such a fit, the most adequate samples would be an exclusive $B \to X(3872)(\to J/\psi\pi^+\pi^-)K$ sample at LHCb and a $B \to X(3872)(\to D^0\bar{D}^{*0})K$ sample at Belle II. Such an analysis could fully determine the lineshape in the coupled-channel framework, and greatly contribute to determining the internal structure.

ACKNOWLEDGMENTS

This work, based on data collected using the Belle detector, which was operated until June 2010, was supported by the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan, the Japan Society for the Promotion of Science (JSPS) including in particular the Grant-in-Aid for JSPS Fellows (No. 19J23314), the Tau-Lepton Physics Research Center of Nagoya University, and the Advanced Science Research Center of Japan Atomic Energy Agency; the Australian Research Council including Grants No. DP180102629, No. DP170102389, No. DP170102204, No. DE220100462, No. DP150103061, No. FT130100303; Austrian Federal Ministry of Education, Science, and Research (FWF) and FWF Austrian Science Fund No. P 31361-N36; the National Natural Science Foundation of China under Contracts No. 11675166, No. 11705209; No. 11975076; No. 12135005; No. 12175041; No. 12161141008; Key Research Program of Frontier Sciences, Chinese Academy of Sciences (CAS), Grant No. QYZDJ-SSW-SLH011; Project No. ZR2022JQ02 supported by Shandong Provincial Natural Science Foundation; the Ministry of Education, Youth and Sports of the Czech Republic under Contract No. LTT17020; the Czech Science Foundation Grant No. 22-18469S; Horizon 2020 ERC Advanced Grant No. 884719 and ERC Starting Grant No. 947006 "InterLeptons" (European Union); the Carl Zeiss Foundation, the Deutsche Forschungs gemeinschaft, the Excellence Cluster Universe, and the VolkswagenStiftung; the Department of Atomic Energy (Project Identification No. RTI 4002) and the Department of Science and Technology of India; the Istituto Nazionale di Fisica Nucleare of Italy; National Research Foundation (NRF) of Korea Grants No. 2016R1-D1A1B-02012900, No. 2018R1-A2B-3003643, No. 2018R1-A6A1A-06024970, No. RS-2022-00197659, No. 2019R1-I1A3A-01058933, No. 2021R1-A6A1A-03043957, No. 2021R1-F1A-1060423, No. 2021R1-F1A-1064008, No. 2022R1-A2C-1003993; Radiation Science Research Institute, Foreign Large-size Research Facility Application Supporting project, the Global Science Experimental Data Hub Center of the Korea Institute of Science and Technology Information and KREONET/GLORIAD; the Polish Ministry of Science and Higher Education and the National Science Center; the Ministry of Science and Higher Education of the Russian Federation, Agreement No. 14.W03.31.0026, and the HSE University Basic Research Program, Moscow; University of Tabuk Research Grants No. S-1440-0321, No. S-0256-1438, and No. S-0280-1439 (Saudi Arabia); the Slovenian Research Agency Grants No. J1-9124 and No. P1-0135; Ikerbasque, Basque Foundation for Science, Spain; the Swiss National Science Foundation; the Ministry of Education and the Ministry of Science and Technology of Taiwan; and the United States Department of Energy and the National Science Foundation. These acknowledgements are not to be interpreted as an endorsement of any statement made by any of our institutes, funding agencies, governments, or their representatives. We thank the KEKB group for the excellent operation of the accelerator; the KEK cryogenics group for the efficient operation of the solenoid; and the KEK computer group and the Pacific Northwest National Laboratory (PNNL) Environmental Molecular Sciences Laboratory (EMSL) computing group for strong computing support; and the National Institute of Informatics, and Science Information NETwork 6 (SINET6) for valuable network support.

[1] R. L. Workman *et al.* (Particle Data Group), Prog. Theor. Exp. Phys. **2022**, 083C01 (2022).
[2] S.-K. Choi *et al.* (Belle Collaboration), Phys. Rev. Lett. **91**, 262001 (2003).
[3] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **93**, 162002 (2004).
[4] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **71**, 071103 (2005).

[5] D. Acosta *et al.* (CDF II Collaboration), Phys. Rev. Lett. **93**, 072001 (2004).
[6] R. Aaij *et al.* (LHCb Collaboration), Eur. Phys. J. C **72**, 1947 (2012).
[7] M. Ablikim *et al.* (BESIII Collaboration), Phys. Rev. Lett. **112**, 092001 (2014).
[8] K. Abe *et al.* (Belle Collaboration), arXiv:hep-ex/0505037; P. del Amo Sanchez *et al.* (*BABAR* Collaboration), Phys. Rev. D **82**, 011101 (2010); M. Ablikim *et al.* (BESIII Collaboration), Phys. Rev. Lett. **122**, 232002 (2019).
[9] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. Lett. **102**, 132001 (2009); V. Bhardwaj *et al.* (Belle Collaboration), Phys. Rev. Lett. **107**, 091803 (2011); R. Aaij *et al.* (LHCb Collaboration), Nucl. Phys. **B886**, 665 (2014); M. Ablikim *et al.* (BESIII Collaboration), Phys. Rev. Lett. **124**, 242001 (2020).
[10] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **77**, 011102 (2008).
[11] T. Aushev *et al.* (Belle Collaboration), Phys. Rev. D **81**, 031103 (2010).
[12] G. Gokhroo *et al.* (Belle Collaboration), Phys. Rev. Lett. **97**, 162002 (2006).
[13] V. Bhardwaj *et al.* (Belle Collaboration), Phys. Rev. D **99**, 111101 (2019).
[14] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. Lett. **110**, 222001 (2013).
[15] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. D **92**, 011102 (2015).
[16] N. A. Törnqvist, Phys. Lett. B **590**, 209 (2004).
[17] E. S. Swanson, Phys. Lett. B **588**, 189 (2004).
[18] C.-Y. Wong, Phys. Rev. C **69**, 055202 (2004).
[19] D. Gamermann and E. Oset, Phys. Rev. D **80**, 014003 (2009).
[20] M. Takizawa and S. Takeuchi, Prog. Theor. Exp. Phys. **2013**, 093D01 (2013).
[21] L. Maiani, F. Piccinini, A. D. Polosa, and V. Riquer, Phys. Rev. D **71**, 014028 (2005).
[22] C. Hanhart, Y. S. Kalashnikova, A. E. Kudryavtsev, and A. V. Nefediev, Phys. Rev. D **76**, 034007 (2007).
[23] E. Braaten and M. Lu, Phys. Rev. D **76**, 094028 (2007).
[24] D. V. Bugg, J. Phys. G **35**, 075005 (2008).
[25] R. Aaij *et al.* (LHCb Collaboration), J. High Energy Phys. 08 (2020) 123.
[26] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. D **102**, 092005 (2020).
[27] Y. S. Kalashnikova and A. V. Nefediev, Phys. Rev. D **80**, 074004 (2009).
[28] V. Baru, J. Haidenbauer, C. Hanhart, A. Kudryavtsev, and Ulf-G. Meißner, Eur. Phys. J. A **23**, 523 (2005).
[29] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods Phys. Res., Sect. A **499**, 1 (2003), and other papers included in this volume.
[30] T. Abe *et al.*, Prog. Theor. Exp. Phys. **2013**, 03A001 (2013).
[31] A. Abashian *et al.* (Belle Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 117 (2002).
[32] J. Brodzicka *et al.*, Prog. Theor. Exp. Phys. **2012**, 04D001 (2012).
[33] D. J. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 152 (2001).
[34] R. Brun *et al.*, GEANT3, CERN Report No. DD/EE/84-1, 1987.
[35] E. Nakano, Nucl. Instrum. Methods Phys. Res., Sect. A **494**, 402 (2002).
[36] M. Feindt and U. Kerzel, Nucl. Instrum. Methods Phys. Res., Sect. A **559**, 190 (2006).
[37] T. Keck, Comput. Software Big Sci. **1**, 2 (2017).
[38] G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978); The modified moments used in this paper are described in S. H. Lee *et al.* (Belle Collaboration), Phys. Rev. Lett. **91**, 261801 (2003).
[39] D. M. Asner *et al.* (CLEO Collaboration), Phys. Rev. D **53**, 1039 (1996).
[40] T. Skwarnicki, Ph.D. thesis, Institute of Nuclear Physics, Krakow (1986), the original Crystal Ball function is a Gaussian with a tail on the lower side. The function defined as a Gaussian with a tail on the upper side is used in this work.
[41] H. Albrecht *et al.* (ARGUS Collaboration), Phys. Lett. B **241**, 278 (1990).
[42] C. Hanhart, Y. S. Kalashnikova, and A. V. Nefediev, Phys. Rev. D **81**, 094028 (2010).
[43] We always use $c = 1$ in this subsection.
[44] K. Abe *et al.* (Belle Collaboration), arXiv:hep-ex/0505037.
[45] P. del Amo Sanchez *et al.* (*BABAR* Collaboration), Phys. Rev. D **82**, 011101 (2010).
[46] M. Ablikim *et al.* (BESIII Collaboration), Phys. Rev. Lett. **122**, 232002 (2019).
[47] Our own average using the most recent measurements from Belle, *BABAR* and LHCb [25,48,49].
[48] S.-K. Choi *et al.* (Belle Collaboration), Phys. Rev. D **84**, 052004 (2011).
[49] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **77**, 111101 (2008).
[50] Our own average of $\mathcal{B}(B^0 \to X(3872)K^0)/\mathcal{B}(B^+ \to X(3872)K^+)$ using the most recent measurements from Belle and *BABAR* [48,49], 0.48 ± 0.13.