Recent Results from the Belle Experiment

Z King and A J Schwartz
(on behalf of the Belle Collaboration)

Physics Department, University of Cincinnati, P.O. Box 210011, Cincinnati, Ohio 45221 USA
E-mail: kingze@mail.uc.edu, alan.j.schwartz@uc.edu

Abstract. We review recent results from the Belle experiment, which took data at the KEKB asymmetric-energy $e^+e^-$ collider in Japan. The experiment recorded about 1000 fb$^{-1}$ of data running mainly at the $\Upsilon(4S)$ and $\Upsilon(5S)$ resonances. The results presented here are obtained from the full data set.

1. Introduction

The Belle experiment successfully operated for more than a decade until 2010 at the asymmetric-energy $e^+e^-$ collider KEKB [1]. The experiment took data at center-of-mass energies corresponding to several $\Upsilon(nS)$ resonances; the total data sample recorded exceeds 1 ab$^{-1}$. Here we present recent results based on the full data sample.

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 superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux-return located outside of the coil is instrumented to detect $K^0_L$ mesons and to identify muons (KLM). The detector is described in detail elsewhere [2].

2. Search for the Decay $B^0 \rightarrow \phi \gamma$

In the Standard Model (SM), the decay $B^0 \rightarrow \phi \gamma$ proceeds through electroweak and gluonic $b \rightarrow d$ penguin annihilation processes. These amplitudes are proportional to the small Cabibbo-Kobayashi-Maskawa [3] matrix element $V_{td}$ and thus are highly suppressed. The branching fraction has been estimated based on naive QCD factorization [4] and perturbative QCD [5] and found to be in the range $10^{-12}$ to $10^{-11}$. However, the internal loop can also be mediated by non-SM particles such as a charged Higgs boson or supersymmetric squarks, and thus the decay is sensitive to new physics (NP). It is estimated that such NP could enhance the branching fraction to the level of $10^{-9}$ to $10^{-8}$ [4]. Previously, no evidence for this decay has been found.

Here we describe a recent Belle search for this decay with much higher sensitivity than previously that uses the full data set of 711 fb$^{-1}$ recorded on the $\Upsilon(4S)$ resonance [6].

Candidate $\phi$ mesons are reconstructed via $\phi \rightarrow K^+K^-$ decays. The $K^+K^-$ invariant mass is required to be in the range [1.000, 1.039] GeV/c$^2$, which corresponds to 4.5$\sigma$ in resolution around the $\phi$ mass. Candidate $B$ mesons are identified using a modified beam-energy-constrained mass $M_{bc} = \sqrt{E_{beam}^2 - |\vec{p}_B|^2 c^2}$, and the energy difference $\Delta E = E_B - E_{beam}$, where $E_{beam}$
is the beam energy and $\vec{p}_B$ and $E_B$ are the momentum and energy, respectively, of the $B^0$ candidate. All quantities are evaluated in the center-of-mass frame. To improve the $M_{bc}$ resolution, the momentum $\vec{p}_B$ is calculated as $\vec{p}_B + (\vec{p}_\gamma/|{p}_\gamma|)\sqrt{(E_{\text{beam}} - E_\phi)^2/c}$, where $\vec{p}_\gamma$ is the photon momentum, and $E_\phi$ and $\vec{p}_\phi$ are the energy and momentum, respectively, of the $\phi$ candidate. We require that events satisfy $5.25 \text{ GeV}/c^2 < M_{bc} < 5.29 \text{ GeV}/c^2$ and $-0.30 \text{ GeV} < \Delta E < 0.15 \text{ GeV}$. The signal yield is calculated in a smaller “signal region” $5.27 \text{ GeV}/c^2 < M_{bc} < 5.29 \text{ GeV}/c^2$ and $-0.20 \text{ GeV} < \Delta E < 0.10 \text{ GeV}$.

Charless hadronic decays suffer from large backgrounds arising from continuum $e^+e^- \rightarrow q\bar{q} \,(q = u,d,s,c)$ production. To suppress this background, we use a multivariate analyzer based on a neural network (NN) [7]. The NN generates an output variable $C_{NN}$, which ranges from $-1$ for background-like events to $+1$ for signal-like events. We make a loose requirement $C_{NN} > 0.3$ and then translate $C_{NN}$ to $C_{NN}'$, defined as $C_{NN}' = \ln[(C_{NN} - C_{\text{min}})/(C_{\text{max}} - C_{\text{min}})]$ where $C_{\text{min}} = 0.3$ and $C_{\text{max}} = 1.0$. The variable $C_{NN}'$ is well-described by a sum of Gaussian functions.

We obtain the signal yield using an unbinned extended maximum likelihood fit to the observables $M_{bc}$, $\Delta E$, $C_{NN}'$, and $\cos \theta_\phi$. The helicity angle $\theta_\phi$ is the angle between the $K^+$ momentum and the opposite of the $B$ flight direction in the $\phi$ rest frame. The resulting branching fraction is calculated as

$$B\left(B^0 \rightarrow \phi \gamma\right) \equiv \frac{N_{\text{sig}}}{N_{BB} \cdot \epsilon \cdot B(\phi \rightarrow K^+ K^-)} \quad (1)$$

where $N_{\text{sig}} = 3.4^{+4.6}_{-3.8}$ is the signal yield in the signal region, $N_{BB} = (772 \pm 11) \times 10^6$ is the number of $B\bar{B}$ events, $\epsilon = 0.296 \pm 0.001$ is the signal efficiency as calculated from Monte Carlo (MC) simulation, and $B(\phi \rightarrow K^+ K^-) = (48.9 \pm 0.5)\%$ [8]. We find no evidence for this decay and set an upper limit on the branching fraction of $B(B^0 \rightarrow \phi \gamma) < 1.0 \times 10^{-7}$ at 90% confidence level (C.L.). This limit is almost an order of magnitude lower than the previous most stringent result [9].

3. Angular Analysis of $B^0 \rightarrow K^*(892)^0 \ell^+ \ell^-$

The rare decay $B^0 \rightarrow K^*(892)^0 \ell^+ \ell^-$ involves the quark transition $b \rightarrow s \ell^+ \ell^-$, a flavor changing neutral current that is forbidden at tree level in the SM. Higher order SM processes such as penguin or $W^+W^-$ box diagrams allow for such transitions, leading to branching fractions below $10^{-6}$. Various extensions to the SM predict NP amplitudes that can interfere with the SM amplitudes and lead to enhanced or suppressed branching fractions and modified angular distributions.

Belle performed an angular analysis using the decay modes $B^0 \rightarrow K^*(892)^0 e^+e^-$ and $B^0 \rightarrow K^*(892)^0 \mu^+\mu^-$. $K^{*0}$ candidates are reconstructed in the channel $K^{*0} \rightarrow K^+ \pi^-$ [10]. Large combinatoric background is suppressed by applying requirements on kinematic variables. Large backgrounds also arise from charmonium decays $B \rightarrow K^{(*)}J/\psi$ and $B \rightarrow K^{(*)}\psi(2S)$, in which $\psi \rightarrow \ell^+ \ell^-$. To maximize signal efficiency and purity, neural networks are applied sequentially beginning from the end of the decay chain. Signal and background yields are determined from an unbinned extended maximum likelihood fit to $M_{bc}$ in bins of $q^2$. The bin ranges used and the resulting fitted yields are listed in Table 1. In total, $117.6 \pm 12.4$ signal candidates are obtained for $B^0 \rightarrow K^*(892)^0 \mu^+\mu^-$, and $69.4 \pm 12.0$ candidates for $B^0 \rightarrow K^*(892)^0 e^+e^-$. We subsequently perform the angular analysis. The decay is kinematically described by three angles $(\theta_\ell, \theta_K$, and $\phi)$ and the invariant mass squared of the lepton pair $(q^2)$. The definitions of the angles and the full angular distribution are described in Ref. [12]. For this fit we require $M_{bc} > 5.27 \text{ GeV}/c^2$, and the number of signal ($N_{\text{sig}}$) and background events ($N_{\text{bkg}}$) are fixed to the values in Table 1.
Table 1. Fitted yields in bins of $q^2$ for signal ($N_{\text{sig}}$) and background ($N_{\text{bkg}}$), for electron and muon channels combined.

| Bin | $q^2$ range in GeV$^2$/c$^4$ | $N_{\text{sig}}$ | $N_{\text{bkg}}$ |
|-----|-----------------------------|-----------------|-----------------|
| 0   | 1.00 – 6.00                 | 49.5 ± 8.4      | 30.3 ± 5.5      |
| 1   | 0.10 – 4.00                 | 30.9 ± 7.4      | 26.4 ± 5.1      |
| 2   | 4.00 – 8.00                 | 49.8 ± 9.3      | 35.6 ± 6.0      |
| 3   | 10.09 – 12.90               | 39.6 ± 8.0      | 19.3 ± 4.4      |
| 4   | 14.18 – 19.00               | 56.5 ± 8.7      | 16.0 ± 4.0      |

Figure 1. Results for $P'_5$ as compared to SM predictions and to previous measurements by LHCb [15][16].

The angular observables $P'_{i=4,5,6,8}$ introduced in Ref. [13] contain information about short-distance effects. These observables are considered to be largely free from form-factor uncertainties [14], and thus they are a promising place to search for NP. As the statistics in this analysis are insufficient to perform a full eight-dimensional fit, a “folding technique” is used as described in Ref. [15]. We determine $P'_{4,5,6,8}$ by performing a three-dimensional unbinned maximum likelihood fit in four bins of $q^2$ using the folded signal probability density function and fixed background yields and shapes. We also fit for the longitudinal polarization $F_L$ and the transverse polarization asymmetry $A_T^{(2)}$. The fit results are shown in Fig. 1. These results are consistent with SM predictions, although one value of $P_5'$ differs by 2.1σ from the SM value. It is notable that this deviation is for the same $q^2$ bin and in the same direction as a similar deviation reported by LHCb [15][16].

4. Measurement of the branching fraction of $\bar{B}^0 \to D^{*+} \tau^- \nu_\tau$ relative to that of $\bar{B}^0 \to D^{*+} \ell^- \nu_\ell$

Semi-tauonic $B$ decays of the type $b \to c\tau\nu_\tau$ are sensitive probes to search for NP. Charged Higgs bosons, which appear in supersymmetry and other models with at least two Higgs doublets, may induce measurable effects in the branching fraction due to the large mass of the $\tau$ lepton. Similarly, leptoquarks, which carry both baryon number and lepton number, can also contribute
to this process. The ratio of branching fractions

$$R(D^{(*)}) = \frac{B(B \rightarrow D^{(*)}\tau^-\nu_\tau)}{B(B \rightarrow D^{(*)}\ell^-\nu_\ell)} \quad (\ell = e, \mu)$$

(2)

is typically measured instead of the absolute branching fraction $B(B \rightarrow D^{(*)}\tau^-\nu_\tau)$ to reduce systematic uncertainties arising from reconstruction efficiencies, form factors, and the CKM matrix element $|V_{cb}|$. Standard Model calculations predict $R(D^*) = 0.252 \pm 0.003$ [17] and $R(D) = 0.297 \pm 0.017$ [18][19]. Semi-tauonic $B$ decays were first observed by Belle [20], with subsequent studies reported by Belle [19][21], BABAR [22], and LHCb [23]. The world average values [24] are $R(D^*) = 0.322 \pm 0.018 \pm 0.012$ and $R(D) = 0.391 \pm 0.041 \pm 0.028$, which exceed the SM predictions by 3.0σ and 1.7σ, respectively.

Thus far, measurements of $R(D^{(*)})$ at Belle and BABAR have been performed using either a hadronic [19][22] or an inclusive [20][21] tagging method. Here we report the first measurement of $R(D^*)$ using a semileptonic tagging method. We reconstruct signal events in modes in which one $B$ decays as $B_0 \rightarrow D^{(*)+}\ell^-\nu_\ell$ and the other $B$ decays as $B_0^* \rightarrow D^{(*)+}\tau^-\nu_\tau$. To reconstruct normalization events corresponding to the denominator in Eq. (2), we require that both $B$ mesons decay to $D^{(*)+}\ell^-\nu_\ell$.

Signal and normalization events are identified using a neural network. The dominant background contribution arises from events with falsely reconstructed $D^{(*)}$ mesons. To separate signal and normalization events from backgrounds, we use the energy variable $E_{ECL}$, which is defined as the sum of the energies of neutral clusters detected in the electromagnetic calorimeter (ECL) that are not associated with reconstructed particles.

We extract the signal and normalization yields using a two-dimensional extended maximumlikelihood fit to the neural network output NN and $E_{ECL}$. The resulting yields of signal and normalization events are $231 \pm 23$(stat) and $2800 \pm 57$(stat), respectively. The ratio $R(D^*)$ is $0.302 \pm 0.030 \pm 0.011$, where the first and second errors correspond to statistical and systematic uncertainties, respectively. Our measurement is 1.6σ higher than the SM prediction.

We investigate the compatibility of the data with a Type II two-Higgs doublet model (2HDM) [25] and with $R_2$-type and $S_1$-type leptoquark models [26]. We find that the data allows for additional contributions from 2HDM scalar operators and also vector operators, but contributions from a tensor operator with $0.34 < C_T < 0.39$, an $R_2$-type leptoquark model with $0.34 < C_T < 0.38$, or an $S_1$-type leptoquark model with $0.22 < C_T < 0.28$ are disfavored.

5. Observation of the decay $B_s^0 \rightarrow K^0\bar{K}^0$

The two-body decays $B_s^0 \rightarrow hh'$, where $h^0$ is either a charged $\pi^\pm$ or $K^\pm$, have all been observed [8]. However, the neutral modes $\pi^0\pi^0$, $\pi^0K^0$, and $K^0\bar{K}^0$ have not. The decay $B_s^0 \rightarrow K^0\bar{K}^0$ is of particular interest because the branching fraction is predicted to be large: $(16 \sim 27) \times 10^{-6}$ [27]. The presence of non-SM particles or couplings could measurably affect this value [28]. The current upper limit, $B(B_s^0 \rightarrow K^0\bar{K}^0) < 6.6 \times 10^{-5}$ at 90% C.L., was set by Belle using only 23.6 fb$^{-1}$ of data recorded at the $\Upsilon(5S)$ [29]. Here we update that result using the full data set of 121.4 fb$^{-1}$ [30]. At the $\Upsilon(5S)$, $B^0_s\bar{B}^0_s$ pairs are produced in three decay channels: $B_s^0\bar{B}^0_s$, $B_s^0\bar{B}^0_s$, or $B_s^0\bar{B}^0_s$, and $B_0^0\bar{B}^0_s$. The latter two channels dominate, with production fractions of $f_{B^0_s\bar{B}^0_s} = (7.3 \pm 1.4)%$ and $f_{B^0_s\bar{B}^0_s} = (87.0 \pm 1.7)%$ [31].

Candidate $K^0$ mesons are reconstructed via the decay $K_S^0 \rightarrow \pi^+\pi^-$ using a neural network technique. To suppress background arising from $e^+e^- \rightarrow qq$ ($q = u, d, s, c$) production, a second NN is used. The output of the latter ($C_{NN}$) ranges from $-1$ for background-like events to $+1$ for signal-like events. We make a loose requirement $C_{NN} > -0.1$ and then translate $C_{NN}$ to the variable $C_{NN}'$ as described in Section 2.
The current world average branching fractions are $B(D^0 \rightarrow K^0\pi^0/s) = (69.20 \pm 0.05)$% is the branching fraction for $K^0_{S} \rightarrow \pi^+\pi^-$ [8]; and $\varepsilon = (46.3 \pm 0.1)$% is the reconstruction efficiency as determined from MC simulation. The factor 0.50 accounts for the 50% probability for $K^0_{S} \rightarrow K^0_{L}$ (since $K^0_{S}$ is CP even). Inserting these values gives $B(B_s^0 \rightarrow K^0\overline{K}^0) = (19.6^{+5.8}_{-5.1} \pm 1.0 \pm 2.0) \times 10^{-6}$, where the first uncertainty is statistical, the second is systematic, and the third reflects the uncertainty due to the total number of $B_s^0\overline{B}_s^0$ pairs. This result is in good agreement with the SM prediction [27].

The signal significance is calculated as $\sqrt{-2\ln(L_0/L_{\text{max}})}$, where $L_0$ is the likelihood value when the signal yield is fixed to zero, and $L_{\text{max}}$ is the likelihood value of the nominal fit. Systematic uncertainties are included in the significance by convolving the likelihood function with a Gaussian function whose width is equal to the systematic uncertainty associated with the signal yield. We obtain a signal significance of 5.1σ, and thus our measurement constitutes the first observation of this decay.

6. Measurement of the branching fraction and CP asymmetry for $D^0 \rightarrow V\gamma$

CP asymmetries in the charm sector are a promising area in which to search for new physics. The CP asymmetry ($A_{CP}$) of many two-body $D$ decays has been measured [24], but all results are consistent with zero. Here we present a first search for a $CP$ asymmetry in radiative $D^0 \rightarrow V\gamma$ decays, where $V$ is a neutral vector meson $\phi$, $K^0_{S}$, or $\rho^0$. Within the SM, such asymmetries are predicted to be small: $A_{CP} \sim 10^{-3}$. However, theoretical studies [33][34] indicate that extensions to the SM with chromomagnetic dipole operators can increase $A_{CP}$ up to a few percent. Several $D^0 \rightarrow V\gamma$ decay modes have been measured by Belle and BABAR [35][36]. The current world average branching fractions are $B(D^0 \rightarrow \phi\gamma) = (2.70 \pm 0.35) \times 10^{-5}$ and $B(D^0 \rightarrow K^{*0}\gamma) = (32.7 \pm 3.4) \times 10^{-5}$ [8]. The $D^0 \rightarrow \rho^0\gamma$ mode has not yet been observed, and the current upper limit on the branching fraction is $B(D^0 \rightarrow \rho^0\gamma) < 24 \times 10^{-5}$ at 90% C.L. [8].

Here we present new measurements of all three modes using the full Belle data set. Our results include the first measurement of the CP asymmetries. For our analysis, the $D^0$ is required to originate from $D^{*+} \rightarrow D^0\pi^+$ decays; this tags the $D^0$ flavor and suppresses combinatoric
background. The vector mesons in the final state are reconstructed via $\phi \to K^+K^-$, $\overline{K}^{*0} \to K^-\pi^+$, and $\rho^0 \to \pi^+\pi^-$. 

All branching fractions and $A_{CP}$ are determined by normalizing to decay channels $D^0 \to K^+K^-$ for $D^0 \to \phi \gamma$, $D^0 \to K^−\pi^+$ for $D^0 \to \overline{K}^{*0}\gamma$, and $D^0 \to \pi^+\pi^-$ for $D^0 \to \rho^0\gamma$. The branching fractions are calculated as

$$B_{\text{sig}} = B_{\text{norm}} \times \frac{N_{\text{sig}}}{N_{\text{norm}}} \times \varepsilon_{\text{norm}} \varepsilon_{\text{sig}}$$

where $B$ and $N$ are the branching fraction and fitted yield, respectively, of signal or normalization modes, and $\varepsilon$ is the reconstruction efficiency. For $B_{\text{norm}}$, the world-average value [8] is used.

The extracted raw asymmetry

$$A_{\text{raw}} = \frac{N(D^0) - N(\overline{D}^0)}{N(D^0) + N(\overline{D}^0)}$$

has several contributions: $A_{\text{raw}} = A_{CP} + A_{FB} + A_{\pm}^\pm$, where $A_{FB}$ is the forward-backward production asymmetry and $A_{\pm}^\pm$ is the detection asymmetry between positively and negatively charged particles. Both of these asymmetries are eliminated by measuring $A_{CP}$ of a signal mode relative to that of its normalization mode, which has the same final-state charged particles. The $CP$ asymmetry of a signal mode is then $A_{CP}^{\text{sig}} = A_{CP}^{\text{raw}} - A_{CP}^{\text{norm}} + A_{CP}^{\text{norm}}$, where $A_{CP}^{\text{norm}}$ is the PDG value of the $CP$ asymmetry for the normalization mode [8].

To extract the signal yields and $CP$ asymmetries, we perform a simultaneous two-dimensional unbinned extended maximum likelihood fit of the $D^0$ and $\overline{D}^0$ samples. The fit variables are the invariant mass of the $D^0$ and the cosine of the helicity angle $\theta_H$, defined as the angle between the $D^0$ and the positively or negatively charged hadron in the rest frame of the $V$ meson. For $D^0$ candidates we use the $K^+K^-\pi^+$ for $\phi/\overline{K}^{*0}/\rho^0$ decays, and we use the oppositely charged particles for $\overline{D}^0$ candidates. The fitted signal yields are $N_{\phi\gamma} = 524 \pm 35$, $N_{K^{*0}\gamma} = 9104 \pm 396$, and $N_{\rho\gamma} = 500 \pm 85$. The resulting branching fractions are

$$B(D^0 \to \phi \gamma) = (2.76 \pm 0.20 \pm 0.08) \times 10^{-5}$$
$$B(D^0 \to \overline{K}^{*0}\gamma) = (4.66 \pm 0.21 \pm 0.18) \times 10^{-4}$$
$$B(D^0 \to \rho^0\gamma) = (1.77 \pm 0.30 \pm 0.08) \times 10^{-5},$$

where the first uncertainty is statistical and the second is systematic. The result for $\phi\gamma$ is improved with respect to the previous result [35]. The result for $\overline{K}^{*0}\gamma$ is $3.3\sigma$ higher than the BABAR result [36]. The result for $\rho^0\gamma$ is close to that for $\phi\gamma$, which is consistent with theoretical expectations. The significance is calculated as $\sqrt{-2\ln(L_0/L_{\text{max}})}$ as described in Section 5. Systematic uncertainties are included in the significance by convolving the likelihood function with a Gaussian whose width is equal to the systematic uncertainty associated with the signal yield. The significance for $\rho^0\gamma$ is $5.5\sigma$, and thus our measurement constitutes the first observation of this decay.

The results for $A_{CP}$ are

$$A_{CP}(D^0 \to \phi \gamma) = -0.094 \pm 0.066 \pm 0.001$$
$$A_{CP}(D^0 \to \overline{K}^{*0}\gamma) = -0.003 \pm 0.020 \pm 0.000$$
$$A_{CP}(D^0 \to \rho^0\gamma) = 0.056 \pm 0.151 \pm 0.006.$$

No $CP$ asymmetry is seen in any of these radiative decays.

Acknowledgements
The authors thank the BEACH 2016 organizers for a well-run workshop and excellent hospitality. This research is supported by the U.S. Department of Energy.
References

[1] Kurokawa S and Kikutani E 2003 *Nucl. Instrum. Methods Phys. Res., Sect. A* **499** 1 and other papers included in this volume

Abe T et al 2013 *Prog. Theor. Exp. Phys.* 03A001 and following articles up to 03A011

[2] Abashian A et al (Belle Collaboration) 2002 *Nucl. Instr. and Meth. A* **479** 117

[3] Cabibbo N 1963 *Phys. Rev. Lett.* **10** 531

Kobayashi M and Maskawa T 1973 *Prog. Theor. Phys.* **49** 652

[4] Li X-Q, Lu G-R, Wang R-M and Yang Y 2004 *Eur. Phys. J. C* **36** 97

Hua J, Kim C and Li Y 2010 *Eur. Phys. J. C* **69** 139

[5] Lu C-D, Shen Y-L and Wang W 2006 *Chin. Phys. Lett.* **23** 2684

[6] King Z et al (Belle Collaboration) 2016 *Phys. Rev. D* **93** 111101

[7] Olive K A et al (Particle Data Group) 2014 *Chin. Phys. C* **38** 090001

[8] Aubert B et al (BABAR Collaboration) 2005 *Phys. Rev. D* **72** 091103

[9] Abdesselem A et al (Belle Collaboration) 2016 arXiv:1604.04042

[10] Charge-conjugate modes are implicitly included unless noted otherwise

[11] Aaij R et al (LHCb Collaboration) 2013 *JHEP* **08** 131

[12] Descotes-Genon S, Matias J, Ramon M and Virto J 2013 *JHEP* **01** 048

Descotes-Genon S, Hurth T, Matias J and Virto J 2013 *JHEP* **1305** 137

[13] Aaij R et al (LHCb Collaboration) 2013 *Phys. Rev. Lett.* **111** 191801

[14] Fajfer S, Kamenik J F and Nisandzic I 2012 *Phys. Rev. D* **85** 094025

[15] Chin C H 2001 *Phys. Lett. B* **520** 33

Williamson A R and Zupan J 2006 *Phys. Rev. D* **74** 014003

Ali A, Kramer G, Li Y, Lu C-D, Shen Y-L, Wang W and Wang Y-M 2007 *Phys. Rev. D* **76** 074018

Chua C K 2008 *Phys. Rev. D* **78** 076002

Wang K and Zhu G 2013 *Phys. Rev. D* **88** 014043

Wang J J, Lin D T, Sun W, Ji Z J, Cheng S and Xiao Z J 2014 *Phys. Rev. D* **89** 074046

Chang Q, Sun J, Yang Y and Li X 2015 *Phys. Lett. B* **740** 56

Cheng H Y, Chiang C W and Kuo A L 2015 *Phys. Rev. D* **91** 014011

[16] Chang Q, Li X Q and Yang Y D 2014 *J. Phys. G* **41** 105002

[17] Peng C C et al (Belle Collaboration) 2010 *Phys. Rev. D* **82** 072007

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

[19] Eisen S et al (Belle Collaboration) 2013 *Phys. Rev. D* **87** 031101(R)

[20] Oswald C et al (Belle Collaboration) 2015 *Phys. Rev. D* **92** 072013

[21] Isidori G and Kamenik J F 2012 *Phys. Rev. Lett.* **109** 171801

[22] Lyon J and Zwicky R 2012 arXiv:1210.6546

[23] Tajima O et al (Belle Collaboration) 2014 *Phys. Rev. Lett.* **92** 101803

[24] Aubert B et al (BABAR Collaboration) 2008 *Phys. Rev. D* **78** 071101