

Angular analysis of $B \rightarrow K^* e^+ e^-$ in the low- q^2 region with new electron identification at Belle

D. Ferlewicz^{*}, P. Urquijo[✉], I. Adachi[✉], K. Adamczyk[✉], H. Aihara[✉], D. M. Asner[✉], H. Atmacan[✉], R. Ayad[✉], V. Babu[✉], Sw. Banerjee[✉], P. Behera[✉], K. Belous[✉], J. Bennett[✉], M. Bessner[✉], V. Bhardwaj[✉], B. Bhuyan[✉], T. Bilka[✉], D. Biswas[✉], D. Bodrov[✉], M. Bračko[✉], P. Branchini[✉], T. E. Browder[✉], A. Budano[✉], M. Campajola[✉], L. Cao[✉], D. Červenkov[✉], M.-C. Chang[✉], B. G. Cheon[✉], K. Chilikin[✉], H. E. Cho[✉], K. Cho[✉], S.-K. Choi[✉], Y. Choi[✉], S. Choudhury[✉], S. Das[✉], N. Dash[✉], G. de Marino[✉], G. De Pietro[✉], R. Dhamija[✉], F. Di Capua[✉], J. Dingfelder[✉], Z. Doležal[✉], T. V. Dong[✉], S. Dubey[✉], P. Ecker[✉], D. Epifanov[✉], T. Ferber[✉], B. G. Fulsom[✉], V. Gaur[✉], A. Garmash[✉], A. Giri[✉], P. Goldenzweig[✉], E. Graziani[✉], T. Gu[✉], Y. Guan[✉], K. Gudkova[✉], C. Hadjivasiliou[✉], T. Hara[✉], K. Hayasaka[✉], H. Hayashii[✉], S. Hazra[✉], M. T. Hedges[✉], D. Herrmann[✉], W.-S. Hou[✉], C.-L. Hsu[✉], T. Iijima[✉], A. Ishikawa[✉], R. Itoh[✉], M. Iwasaki[✉], W. W. Jacobs[✉], S. Jia[✉], Y. Jin[✉], A. B. Kaliyar[✉], 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. Kou[✉], E. Kovalenko[✉], P. Križan[✉], P. Krokovny[✉], T. Kuhr[✉], R. Kumar[✉], K. Kumara[✉], A. Kuzmin[✉], Y.-J. Kwon[✉], Y.-T. Lai[✉], T. Lam[✉], S. C. Lee[✉], D. Levit[✉], P. Lewis[✉], L. K. Li[✉], Y. B. Li[✉], L. Li Gioi[✉], J. Libby[✉], D. Liventsev[✉], Y. Ma[✉], D. Matvienko[✉], F. Meier[✉], M. Merola[✉], F. Metzner[✉], K. Miyabayashi[✉], R. Mizuk[✉], G. B. Mohanty[✉], I. Nakamura[✉], M. Nakao[✉], A. Natochii[✉], L. Nayak[✉], S. Nishida[✉], S. Ogawa[✉], H. Ono[✉], S. Pardi[✉], J. Park[✉], A. Passeri[✉], 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[✉], V. Savinov[✉], G. Schnell[✉], C. Schwanda[✉], Y. Seino[✉], K. Senyo[✉], M. E. Sevier[✉], W. Shan[✉], C. Sharma[✉], J.-G. Shiu[✉], B. Shwartz[✉], E. Solovieva[✉], M. Starič[✉], M. Sumihama[✉], M. Takizawa[✉], U. Tamponi[✉], K. Tanida[✉], F. Tenchini[✉], M. Uchida[✉], Y. Unno[✉], S. Uno[✉], Y. Ushiroda[✉], S. E. Vahsen[✉], K. E. Varvell[✉], E. Wang[✉], M.-Z. Wang[✉], S. Watanuki[✉], E. Won[✉], B. D. Yabsley[✉], W. Yan[✉], S. B. Yang[✉], J. Yelton[✉], J. H. Yin[✉], C. Z. Yuan[✉], L. Yuan[✉], Y. Yusa[✉], Z. P. Zhang[✉], V. Zhilich[✉], and V. Zhukova[✉]

(Belle Collaboration)



(Received 2 April 2024; accepted 11 September 2024; published 8 October 2024)

We perform an angular analysis of the $B \rightarrow K^* e^+ e^-$ decay for the dielectron mass squared, q^2 , range of 0.0008–1.1200 GeV²/c⁴ using the full Belle dataset in the $K^{*0} \rightarrow K^+ \pi^-$ and $K^{*+} \rightarrow K_S^0 \pi^+$ channels, incorporating new methods of electron identification to improve the statistical power of the dataset. This analysis is sensitive to contributions from right-handed currents from physics beyond the Standard Model by constraining the Wilson coefficients $C_7^{(\prime)}$. We perform a fit to the $B \rightarrow K^* e^+ e^-$ differential decay rate and measure the imaginary component of the transversality amplitude to be $A_T^{\text{Im}} = -1.27 \pm 0.52 \pm 0.12$, and the K^* transverse asymmetry to be $A_T^{(2)} = 0.52 \pm 0.53 \pm 0.11$, with F_L and A_T^{Re} fixed to the Standard Model values. The resulting constraints on the value of C_7' are consistent with the Standard Model within a 2σ confidence interval.

DOI: 10.1103/PhysRevD.110.072005

I. INTRODUCTION

The $b \rightarrow s \ell^+ \ell^-$ decay is a flavor-changing neutral current mediated by loop diagrams and is therefore suppressed in the Standard Model (SM). Studies of this decay

are a sensitive probe for contributions from physics beyond the SM, where angular analyses measuring parameters with reduced theoretical uncertainties, including P_5' , have in the past found disagreements with SM predictions [1,2]. Effective field theory is used to characterize the Hamiltonian and to describe the different processes that contribute to the decay [3]. The contributions from different operators across the range of the dilepton invariant mass squared, q^2 , are determined by their respective Wilson coefficients. In the q^2 region where the P_5' tension is found, the $C_9^{(\prime)}$ Wilson coefficients are dominant, with some contributions from $C_7^{(\prime)}$ and from charm loops that are

^{*}Contact author: dferlewicz@hep.phys.s.u-tokyo.ac.jp

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](#). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

not yet fully understood [4]. Studies focusing on the very low- q^2 region, $q^2 \lesssim 1.0 \text{ GeV}^2/c^4$, are accessible for the dielectron mode and can isolate the $C_7^{(\prime)}$ contribution, which measures the polarization of photons from $b \rightarrow s\gamma$ decays and is a probe for non-SM right-handed currents [5]. Angular analyses of $b \rightarrow s\ell^+\ell^-$ can be used to test other new physics scenarios, such as those detailed in Refs. [6–9]. The $B^0 \rightarrow K^{*0}(\rightarrow K^+\pi^-)e^+e^-$ decay at low- q^2 can be used to measure the ratio of the right- and left-handed Wilson coefficients, C_7'/C_7 , which constrains non-SM contributions. Such angular analyses have been performed by LHCb for $0.002 < q^2 < 1.120 \text{ GeV}^2/c^4$ [10] and $0.0008 < q^2 < 0.257 \text{ GeV}^2/c^4$ [11], where results were consistent with the SM. The following study is an independent angular analysis of the very low q^2 region in Belle, and the first of its kind at an e^+e^- B -factory. To offset the smaller available dataset in Belle, this study includes the implementation of machine learning techniques to improve the performance of electron identification.

A. Differential decay rate

The $B \rightarrow K^*e^+e^-$ differential decay rate is expressed in terms of q^2 and three angular observables, θ_ℓ , θ_K , and ϕ [12]. The angle θ_ℓ is defined as the angle between the e^+ (e^-) candidate direction and the direction opposite to B ($\bar{B}$), in the dielectron rest frame. The angle θ_K is defined similarly between the kaon direction and the opposite direction of B in the K^* rest frame. Finally, ϕ is the angle between the plane containing the dielectron candidates and the plane containing the kaon and pion candidates in the B rest frame. In the low- q^2 region, the full differential decay rate [3] can be simplified by folding the ϕ distribution, which is done by adding π to values of ϕ below 0. The $K\pi$ S-wave contribution can be neglected in this q^2 region due to the polarization of the hadronic system [13], and electrons can be taken to be massless [11]. The differential decay rate, assuming there are no scalar or tensor contributions, is averaged over CP -conjugate modes throughout this paper and is given by

$$\begin{aligned} \frac{1}{d(\Gamma + \bar{\Gamma})/dq^2} \frac{d^4(\Gamma + \bar{\Gamma})}{dq^2 d\cos\theta_\ell d\cos\theta_K d\phi} = & \frac{9}{16\pi} \left(\frac{3}{4}(1 - F_L)\sin^2\theta_K + F_L\cos^2\theta_K + \left(\frac{1}{4}(1 - F_L)\sin^2\theta_K - F_L\cos^2\theta_K \right) \cos 2\theta_\ell \right. \\ & + \frac{1}{2}(1 - F_L)A_T^{(2)}\sin^2\theta_K\sin^2\theta_\ell\cos 2\phi + (1 - F_L)A_T^{\text{Re}}\sin^2\theta_K\cos\theta_\ell \\ & \left. + \frac{1}{2}(1 - F_L)A_T^{\text{Im}}\sin^2\theta_K\sin^2\theta_\ell\sin 2\phi \right), \end{aligned} \quad (1)$$

with four free parameters as functions of q^2 : the longitudinal polarization of K^* , F_L , the transverse asymmetry of K^* , $A_T^{(2)}$, and the real and imaginary components of the transversality amplitudes, A_T^{Re} and A_T^{Im} , respectively. The parameters A_T^{Im} and $A_T^{(2)}$ are related to $C_7^{(\prime)}$ at $q^2 = 0$ through the following equations [14]:

$$A_T^{\text{Im}}(q^2 = 0) = \frac{2\text{Im}(C_7^{\text{eff}} C_7^{\prime\text{eff}*})}{|C_7^{\text{eff}}|^2 + |C_7^{\prime\text{eff}}|^2}, \quad (2)$$

$$A_T^{(2)}(q^2 = 0) = \frac{2\text{Re}(C_7^{\text{eff}} C_7^{\prime\text{eff}*})}{|C_7^{\text{eff}}|^2 + |C_7^{\prime\text{eff}}|^2}, \quad (3)$$

where $C_7^{(\prime)\text{eff}}$ is proportional to $C_7^{(\prime)}$. In the SM, C_7' is helicity-suppressed by a factor of the ratio of the strange and bottom quark masses, $C_7' = (m_s/m_b)C_7$ [3]; hence, these parameters are expected to be near zero at $q^2 = 0$. Therefore, their measurement is sensitive to new physics scenarios. The other parameters, while not sensitive to new physics themselves, can be used to discern different beyond-SM scenarios.

B. The Belle detector and data sample

This study uses the full $\Upsilon(4S)$ data sample containing $(772 \pm 11) \times 10^6$ $B\bar{B}$ meson pairs recorded with the Belle detector [15,16] at the KEKB asymmetric-energy e^+e^- collider [17,18]. The Belle detector is a large-solid-angle magnetic spectrometer that consists of a silicon vertex detector, a 50-layer central drift chamber (CDC), an array of aerogel threshold Cherenkov counters, a barrel-like arrangement of time-of-flight scintillation counters, 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 instrumented to detect K_L^0 mesons and to identify muons (KLM). The detector is described in detail elsewhere [15]. The coordinate system is defined such that the positive z -axis aligns with the direction of the electron beam and is centered on the interaction point (IP).

Monte Carlo (MC) simulation studies are used to determine the analysis techniques. The MC samples include on-resonance $\Upsilon(4S) \rightarrow B\bar{B}$ events and continuum $e^+e^- \rightarrow q\bar{q}$ events with $q \in \{u, d, s, c\}$, which are generated using the EVTGEN [19], PYTHIA [20], and PHOTOS [21] packages with interference effects due to final-state

radiation being switched on. The B2BII package [22] is used to convert reconstructed events into a format compatible with the Belle II analysis software framework [23]. Samples of $B \rightarrow K^* e^+ e^-$ signal events are generated using the EVTGEN generators BTOSLLBALL [24], with form factors from Ref. [25], and BTOSLLNP [26], to study detector effects and understand any possible model dependence in the analysis procedure.

II. ELECTRON IDENTIFICATION

Electron identification in Belle is typically performed using a likelihood ratio (LHR), $\mathcal{L}_e/(\mathcal{L}_e + \mathcal{L}_\pi)$, where $\mathcal{L}_i$ is the likelihood of a charged particle hypothesis, i , determined using a combination of detector outputs [27]. We have implemented a new method for electron identification in Belle using machine learning algorithms. A boosted decision tree (BDT) algorithm [28] is used to classify an electron signal against all other long-lived charged particle hypotheses, exploiting the properties assigned to candidates by the detectors. BDTs are trained using 500 trees, a maximum depth of 3 and a shrinkage rate of 0.1, in different bins of momentum, angle and charge. For electron identification the BDT input variables are as follows: the binary LHR for hadron identification in the inner detectors [29] for all combinations of long-lived charged particles, the ratio of ECL cluster energy and measured momentum, and the ratio of the energy in the 3×3 and 5×5 ECL crystal grids around the center of an electromagnetic shower. Also included are the ECL cluster energy, the number of crystal hits in an ECL cluster and the lateral shower shape, as defined in Ref. [30]. These variables are assigned to long-lived charged particle candidates that have ECL clusters associated with hits in the CDC.

A. Tag and probe procedure

The performance and agreement between data and MC for the BDT classifier is verified using a tag and probe method with $J/\psi \rightarrow e^+ e^-$ events for efficiency measurements and $K_S^0 \rightarrow \pi^+ \pi^-$ events to determine π - e misidentification rates in each phase-space bin. The J/ψ and K_S^0 candidates are taken inclusively from all $B\bar{B}$ and continuum events.

J/ψ candidates are reconstructed in a window around the known J/ψ invariant mass [31] from two electron candidate tracks that originate from the IP, where the distance of the tracks from the IP in the z -direction must be $|dz| < 5$ cm, and their radius in the r - ϕ plane is $|dr| < 2$ cm. The momentum is required to be $p_{\text{lab}} > 0.1$ GeV/ c , and there must be a match between a track in the CDC and a cluster in the ECL. A match is determined by extrapolating a charged track into the ECL and checking if any crystals it passes through are associated to an ECL cluster. A correction for bremsstrahlung energy loss is applied by adding the four-momenta of photon

candidates with $E_\gamma < 1$ GeV within an angular cone of 6° around the electron candidate's track direction from the IP. Low multiplicity $e^+ e^- \rightarrow (e^+ e^-) \ell^+ \ell^-$ and continuum events are reduced through selection criteria based on the event topology. The tagging criterion is a requirement that one $e^\pm$ candidate in the event, the tag, has an e vs π LHR [27], above 0.95.

K_S^0 candidates are reconstructed in a window around their known invariant mass using the same track, momentum, and topology requirements as J/ψ candidates. The cosine of the angle between the K_S^0 momentum vector and the decay vertex position vector, $\cos(\theta(\vec{p}_{K_S^0}, \vec{V}_{K_S^0}))$, is required to be above 0.998. No pion-tagging LHR criterion is required.

The number of J/ψ and K_S^0 candidates is determined with a binned maximum log-likelihood fit to the dielectron or dipion invariant mass distribution, respectively. For $J/\psi \rightarrow e^+ e^-$, the signal probability distribution function (PDF) is modeled by a Gaussian function added to a bifurcated Gaussian and a Crystal Ball [32] function, while for $K_S^0 \rightarrow \pi^+ \pi^-$, the signal component is modeled by a sum of three Gaussian functions. In both cases, the background is modeled by a second-order Chebychev polynomial. The shapes of these PDFs are defined based on MC, where the signal and background yields, the means, and a scale factor for the PDF widths remain floating in fits to data with all other parameters fixed.

The efficiency, ϵ , of applying a selection criterion on the probe is defined as follows:

$$\epsilon = \frac{N_{\text{pass}}^{\text{sig}}}{N_{\text{pass}}^{\text{sig}} + N_{\text{fail}}^{\text{sig}}}, \quad (4)$$

where $N_{\text{pass}}^{\text{sig}}$ is the number of signal candidates that pass the criterion, and $N_{\text{fail}}^{\text{sig}}$ is the number that do not. A simultaneous fit is performed over the mutually exclusive pass and fail datasets to avoid double counting the statistical uncertainty, where signal shape parameters are common for the two samples while those for the backgrounds are fit independently.

The statistical and systematic uncertainties are determined through the generation of pseudo-experiments. The statistical uncertainty is calculated in each phase-space bin and is determined by the following procedure. The population in each invariant mass bin is resampled according to a Poissonian probability, and the invariant mass fit is performed on each new pseudo-dataset to obtain a set of efficiencies. The width of the central 68% of a Gaussian PDF fit to this distribution of efficiencies is taken as the statistical uncertainty. For the systematic uncertainty, which is only assigned to data, an analogous pseudo-experiment method is used, where the fixed signal PDF parameters are varied according to their uncertainty. The efficiency ratios between data and MC in different bins, and the

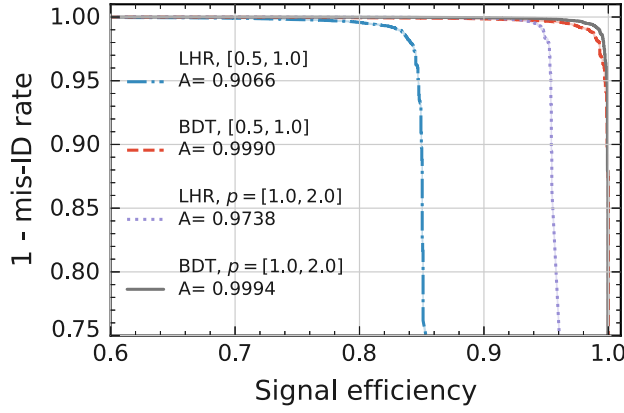


FIG. 1. ROC curves comparing the simulated electron identification performance of the BDT and LHR in the ECL barrel region are shown for two different momentum regions, in GeV/ c . The area under the curve, A , is also provided for comparison.

corresponding statistical and systematic uncertainties, are then calculated using these pseudo-datasets, and are used for corrections in the $B \rightarrow K^* e^+ e^-$ study. An additional study of track isolation effects was performed, finding that although performance degrades when tracks are in close proximity to each other, it is well described by simulation and no additional systematic uncertainty needs to be assigned.

B. Performance comparison

A comparison between the performance of the LHR and BDT methods is shown using receiver operating characteristic (ROC) curves for simulated electrons in Fig. 1, with the misidentification, or mis-ID, rate plotted against the signal efficiency. The BDT has a larger area under the curve and therefore outperforms the standard LHR. Similar performance is found for positrons.

For a comparison in data, probe selection criteria thresholds are chosen for each phase-space bin such that the simulated electron signal efficiency is 95%, which is in agreement between fits to $J/\psi \rightarrow e^+ e^-$ events for data

and MC. The π - e misidentification rates in $K_S^0 \rightarrow \pi^+ \pi^-$ are then measured in data using the same criteria. The results for data, integrated over momentum, are presented in Table I, where a significant reduction in the misidentification rate is found for the BDT in all angular regions except for the backward endcap of the ECL, $2.23 < \theta < 2.71$ rad.

It should be noted that this performance is dependent on the presence of a match between a particle's track in the CDC and a shower in the ECL, and by requiring this match, the signal efficiency when using the BDT is reduced. Low-momentum electrons with $p < 1.0$ GeV/ c make up 20% of the kinematic phase-space of the $B \rightarrow K^* e^+ e^-$ decay, where this track-cluster matching has a reduced efficiency. Therefore for optimal electron identification performance, a combination of the BDT and LHR is used in the following analysis, using the BDT when there is a track and cluster match, and using LHR otherwise.

III. $B \rightarrow K^* e^+ e^-$ SELECTION AND RECONSTRUCTION

We reconstruct $B \rightarrow K^* e^+ e^-$ events in two channels: the neutral mode with $K^{*0} \rightarrow \pi^+ \pi^-$, and the charged mode with $K^{*+} \rightarrow K_S^0 \pi^+$. Pairs of electron and positron candidates are combined with pion and kaon candidates to form B meson candidates. These $e^\pm$ candidates are selected using a combination of the new electron identification BDT discussed earlier (BDT_e) and the standard LHR. The thresholds for each charge and identification method, inclusive of momentum and angle, are determined simultaneously from MC by maximizing a figure of merit (FOM), given by

$$\text{FOM} = \frac{N_S}{\sqrt{N_S + N_B}}, \quad (5)$$

where N_S is the number of $e^+ e^-$ candidate pairs with correctly assigned particle hypotheses and N_B is the number of candidate pairs with any incorrectly assigned hypothesis, after applying the BDT_e and LHR criteria.

TABLE I. The misidentification rates in data are compared between the BDT and LHR methods. The first errors are statistical and the second are systematic. The results use pions from $K_S^0 \rightarrow \pi^+ \pi^-$ candidates in data integrated over the full momentum space, measured using a probe selection threshold for the corresponding identification method that results in a 95% signal MC efficiency in $J/\psi \rightarrow e^+ e^-$ candidates.

Charge	θ region and range (rad)	BDT mis-ID in data $\times 10^3$	LHR mis-ID in data $\times 10^3$
+	Forward (0.22, 0.56)	$0.80 \pm 0.05 \pm 0.01$	$5.72 \pm 0.11 \pm 0.04$
-	Forward (0.22, 0.56)	$0.52 \pm 0.04 \pm 0.01$	$4.81 \pm 0.12 \pm 0.03$
+	Barrel (0.56, 2.23)	$0.53 \pm 0.02 \pm 0.01$	$3.44 \pm 0.03 \pm 0.01$
-	Barrel (0.56, 2.23)	$0.31 \pm 0.02 \pm 0.01$	$3.24 \pm 0.03 \pm 0.01$
+	Backward (2.23, 2.71)	$26.05 \pm 0.28 \pm 0.35$	$25.30 \pm 0.27 \pm 0.33$
-	Backward (2.23, 2.71)	$16.73 \pm 0.23 \pm 0.14$	$18.42 \pm 0.24 \pm 0.23$
+	Integrated (0.22, 2.71)	$0.78 \pm 0.02 \pm 0.01$	$4.45 \pm 0.03 \pm 0.01$
-	Integrated (0.22, 2.71)	$0.54 \pm 0.01 \pm 0.01$	$4.07 \pm 0.03 \pm 0.01$

The thresholds correspond to a signal efficiency of 94% in simulated $J/\psi \rightarrow e^+ e^-$ candidates, and the data/MC correction factor is measured to be $1.00 \pm 0.02 \pm 0.01$, where the first uncertainty is statistical and the second is systematic. Therefore the results in Table I are a good representation of the lepton misidentification rates in this analysis.

Charged pion candidates are required to have a binary π/K LHR [29], $\mathcal{L}_{\pi/K}$, above 0.6 which retains 95% of true pions and removes 55% of the incorrectly assigned candidates, while charged kaon candidates have $\mathcal{L}_{\pi/K} < 0.9$, which retains 95% of true kaons and removes 80% of the incorrectly assigned candidates. These thresholds are chosen by maximizing the FOM from Eq. (5). Both charged hadrons are required to originate from near the IP, where the distance from the IP in the z -direction must be $|dz| < 5$ cm, while the requirement of their radius in the r - ϕ plane is $|dr| < 2$ cm. K_S^0 candidates are selected from a pair of oppositely charged tracks that form a detached vertex, applying selection criteria for their invariant mass and reconstructed vertex, depending on their assigned momentum [33].

The $K\pi$ invariant mass is required to be near the mass of the $K^*(892)$ meson, $0.7 < M_{K\pi} < 1.0$ GeV/ c^2 . The upper threshold for the dielectron invariant mass is chosen to be $q^2 < 1.12$ GeV $^2/c^4$ to match the 2015 LHCb study [10], as this maximizes the available data sample while remaining near the upper q^2 threshold region in which the simplifications of the differential decay rate that give Eq. (1) are valid [5].

The beam-energy constrained mass is required to be $M_{bc} = \sqrt{E_{\text{beam}}^{*2}/c^4 - p_B^{*2}/c^2} > 5.23$ GeV/ c^2 , where p_B^* is the momentum of the B meson and E_{beam}^* is the beam energy, both in the center-of-mass frame (denoted by the symbol $*$). The energy-difference variable, $\Delta E = E_B^* - E_{\text{beam}}^*$, where E_B^* is the energy of the B meson, is required to be $|\Delta E| < 0.3$ GeV. In each event, only the candidate with the lowest value of $|\Delta E|$ is retained.

A. Continuum and signal selection BDTs

Event shape variables including event kinematics, the ratios of Fox-Wolfram moments [34], and spherical harmonic moments of the momenta of particles are used to train a BDT to classify signal events originating from $\Upsilon(4S)$ decays against continuum background, BDT $_{q\bar{q}}$.

The BDT $_{q\bar{q}}$ output is then used as one of the inputs to a BDT used to classify signal events against all other background, BDT $_S$. The other inputs to BDT $_S$ are ΔE , the distribution of which peaks at 0 GeV for signal events, and vertex information for the dielectron system, as true $e^+ e^-$ pairs in a signal event will originate from the same point. The agreement between data and MC for the inputs and the output of BDT $_{q\bar{q}}$ and BDT $_S$ is verified using the control modes $B \rightarrow K^* c\bar{c}(\rightarrow e^+ e^-)$ and $B \rightarrow K^* \gamma(\rightarrow e^+ e^-)$ and the $M_{bc} < 5.27$ GeV/ c^2 sideband of $B \rightarrow K^* e^+ e^-$

candidates. Here, $c\bar{c}$ candidates are reconstructed from dielectron candidates with an invariant mass within a 1σ window of the mass of the J/ψ and $\psi(2S)$ resonances [31], while γ candidates, which are reconstructed from dielectron candidates created in photon conversions in the detector material, are required to have $q^2 < 0.1$ GeV $^2/c^4$ and have a dielectron vertex radius in the x - y plane of at least 0.52 cm. The output of BDT $_{q\bar{q}}$ and BDT $_S$ for signal-like $B \rightarrow K^* c\bar{c}$ events and the background-like M_{bc} sideband events are shown in Fig. 2.

Deviations between data and simulations are found in the sideband sample, and a study for any potential bias is conducted in Sec. V.

B. Selection criteria optimization

To reduce $B \rightarrow K^* \gamma$ background events that peak in the signal M_{bc} region, a q^2 lower threshold and a dielectron vertex radius upper threshold are simultaneously determined by maximizing the FOM from Eq. (5). In this case, the background is taken to be the remaining simulated $B \rightarrow K^* \gamma$ events. The q^2 lower threshold is determined to be 0.0008 GeV $^2/c^4$, coincidentally the same as that of LHCb [11], and the radius upper threshold is 0.52 cm.

After applying these criteria, the lower threshold for BDT $_S$ is determined to be -0.79 by maximizing the FOM against all remaining background. Taking all selection criteria into account, for the $K^{*0} \rightarrow K^+ \pi^-$ ($K^{*+} \rightarrow K_S^0 \pi^+$) channel, approximately 18 (5) simulated signal events and 16 (8) background events are expected to remain in the $M_{bc} > 5.27$ GeV/ c^2 region, corresponding to a signal reconstruction efficiency of 10.0% (5.8%). This is a 90% increase in signal yield with similar remaining background when compared to using the same selection criteria but replacing the electron identification criteria with the often used binary LHR threshold of 0.9. In this M_{bc} region, 80% of simulated background events have both electrons correctly identified. Continuum events are the largest background (24% of the remaining sample, or 11 events) in simulations for $M_{bc} > 5.27$ GeV/ c^2 , followed by lepton candidates that originated from different particles (18% or 9 events). Peaking background from $B \rightarrow K^* \gamma$ events contributes to about 2%, or 1 event, of the remaining simulated sample that will be used to develop the fitting procedure.

IV. ANALYSIS METHOD

A two-stage binned log-likelihood fit is performed, first for M_{bc} , and then simultaneously for projections of $\cos \theta_\ell$, $\cos \theta_K$ and ϕ to measure the free parameters in the differential decay rate given in Eq. (1).

A. Fit to M_{bc}

A one-dimensional fit to the M_{bc} distribution is performed over the range $5.23 < M_{bc} < 5.29$ GeV/ c^2 to

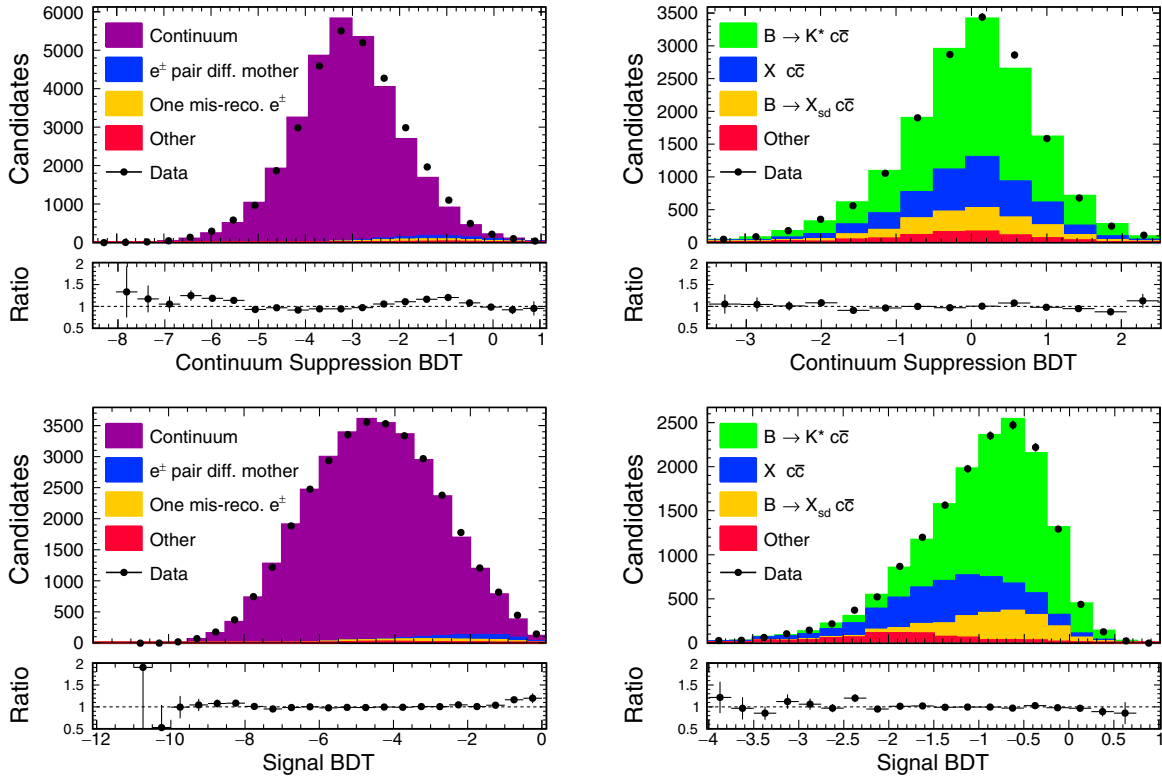


FIG. 2. The outputs of $\text{BDT}_{q\bar{q}}$ (top) and BDT_S (bottom) in control modes are shown for the $M_{bc} < 5.27 \text{ GeV}/c^2$ sideband (left) and $B \rightarrow K^* c\bar{c}$ events (right), comparing the full Belle dataset and simulations, scaled to the recorded data luminosity. The particle X_{sd} represents the S -wave $K\pi$ contribution. $Xc\bar{c}$ is any mis-reconstructed K^* with a correctly reconstructed dielectron pair from a $c\bar{c}$ resonance. $e^{\pm}$ pairs with different mothers are events where the $e^{\pm}$ candidates were reconstructed correctly but originate from different decays. One mis-reconstructed $e^{\pm}$ is for events with a single incorrectly assigned $e^{\pm}$ candidate. The ratios shown are for the number of candidates in data divided by the number of candidates in the scaled MC.

determine the relative contributions of signal events, f_S , peaking background events, f_P , and all other background, f_B . For signal and peaking background events, separate Gaussian PDFs are defined, where the means and widths are determined with simulations. For all other background, an ARGUS function [35] is used. In the M_{bc} fit, f_P is fixed to its expected value from simulations, 0.02, and all PDF shape parameters are fixed to the values determined in fits to each category separately. The value of f_B remains floating, with an expected value of 0.85, and is used after the M_{bc} fit to calculate the relative contributions of each category in the $M_{bc} > 5.27 \text{ GeV}/c^2$ signal region.

B. Fit to the angular distributions

The PDF for the projections of $\cos\theta_\ell$, $\cos\theta_K$ and ϕ for signal events in the $M_{bc} > 5.27 \text{ GeV}/c^2$ region is defined as a product of the differential decay rate from Eq. (1) and factorized acceptance functions for each angle [10]. The efficiency of signal events as a function of ϕ is expected to be uniform, whereas the acceptance functions for $\cos\theta_\ell$ and $\cos\theta_K$ are defined as fourth-order Legendre polynomials. The coefficients are determined by generating signal events with uniform angular distributions and then fitting to

the angular projections of events that remain after applying all selection criteria. The acceptance function for $\cos\theta_\ell$ is set to be symmetric.

Histogram PDFs are defined using simulations of the peaking background and all other background in the $M_{bc} > 5.27 \text{ GeV}/c^2$ region for each angular distribution. Cross-checks are performed for this choice of PDF definition, finding that results in MC are in agreement with fits using histogram PDFs defined using the $M_{bc} < 5.27 \text{ GeV}/c^2$ sideband region in data.

The angular PDFs are then summed using the coefficients that were determined from the M_{bc} fit. This combined PDF is then fit to the angular distributions, where $A_T^{(2)}$ and A_T^{Im} are floated, and F_L and A_T^{Re} are fixed to their value expected in SM, as determined by the FLAVIO package [36]. The SM values, while consistent with the LHCb result [11], are used to maintain independence from previous measurements. The decision to fix these parameters was due to the limited statistical power of the dataset, and the parameters were chosen because although they may differentiate between new physics scenarios, their impact on $\mathcal{C}_7^{(\prime)}$ constraints is minimal [14,37].

A pull test and a linearity test are performed to investigate potential bias from the M_{bc} and angular fitting procedures. Pseudo-datasets are generated at the expected Belle luminosity using $f_B = 0.85$, while for $A_T^{(2)}$ and A_T^{lm} , their SM values are used in the pull test, or over the range $[-1, 1]$ in the linearity test. The linearity test has a gradient consistent with unity and an intercept of zero, and the pull test finds a mean consistent with zero and a width consistent with one.

V. SYSTEMATIC UNCERTAINTIES

Various sources of potential systematic error are investigated, with uncertainties quantified where effects are not deemed negligible. They are described in turn below, and summarized in Table II.

Systematic uncertainties are assigned for the efficiency and misidentification of $e^\pm$ candidates and charged long-lived hadrons. We propagate uncertainties on corrections for particle identification performance in data based on measurements of D^* decays for $\mathcal{L}_{\pi/K}$ [29], while the analysis detailed in Sec. II is used to measure the LHR and BDT_e corrections used to propagate the uncertainty for electron identification.

Potential bias from BDT_S is investigated by measuring the ratio of the yield of $B \rightarrow K^* c \bar{c}$ from fits to data and simulations as a function of the BDT_S lower threshold. The ratio is found to be consistent in the region in which the threshold is applied to the $B \rightarrow K^* e^+ e^-$ analysis, and therefore, no systematic uncertainty is assigned for f_B , and effects on $A_T^{(2)}$ and A_T^{lm} are absorbed into the uncertainties for fit parameters, described below.

Although we find a fit bias to be consistent with zero within uncertainties, we assign the Gaussian mean value of the pull test multiplied by the expected statistical uncertainty as a systematic uncertainty. Potential biases that are folded into the detector response are also investigated by

using the BTOSLLNP event generator to generate signal events across a spectrum of $\mathcal{C}_7^{(i)}$ magnitudes and complex arguments followed by performing a pull and linearity test. The values for the four free parameters in Eq. (1) are calculated for each of these $\mathcal{C}_7^{(i)}$ values using FLAVIO [36] and a covariance matrix is determined. New pseudo-datasets are then created using randomly generated values of the four parameters while taking correlations into account [38], followed by a fit with F_L and A_T^{Re} fixed to their SM values. The width of the distributions of the A_T^{lm} and $A_T^{(2)}$ fit results quantify the uncertainty for fixing F_L and A_T^{Re} to their SM value.

The fixed values for the M_{bc} and acceptance PDF fit parameters are fluctuated according to their 1σ uncertainty determined from the individual MC fits. Pseudo-datasets of size 10^4 are then generated, and the fitting procedure is applied to determine an associated systematic uncertainty using parameter distribution widths. Any bias from factorizing the acceptance function and using a flat angular distribution is measured using a method similar to Ref. [39]. The dataset is reweighted with the reciprocal of the acceptance function of $\cos \theta_\ell$ to recover the flat angular distribution in $\cos \theta_\ell$, and it is then reweighted with the reciprocal of $\cos \theta_K$. A Legendre polynomial is fit to this distribution resulting in coefficients consistent with zero. The same effect is seen in events generated using BTOSLLBALL, and therefore, no systematic uncertainty is applied for any bias from factorization.

For f_P , the ratio of the yield of $B \rightarrow K^* \gamma (\rightarrow e^+ e^-)$ events between data and simulations is measured and found to be 0.96 ± 0.07 . Pseudo-datasets are then generated by correcting the value of f_P according to the ratio and fluctuating it within a 1σ uncertainty.

Beam energy and magnetic field mismodeling is investigated by fitting $B \rightarrow K^* c \bar{c}$ in data with the mean of the signal Gaussian and an overall width factor, w , allowed to remain floating. The results, $\mu = 5.27949 \pm 0.00003$ GeV/ c^2 and $w = 0.98 \pm 0.01$, are used to generate pseudo-datasets, for which the spread of parameter fit results is negligible and therefore no systematic uncertainty is assigned.

The uncertainty due to the limited number of simulated events is assigned by propagating the statistical error for each of the different event types, which were generated with different integrated luminosities between 5 and 10 times the size of the Belle dataset.

VI. RESULTS

The fitting procedure finds a yield of 21 ± 6 signal events and 29 ± 5 background events for $M_{bc} > 5.27$ GeV/ c^2 , corresponding to a value of $f_B = 0.86 \pm 0.04 \pm 0.01$, which is in agreement with simulations. Performing a profile likelihood test on this result, with the systematic uncertainties included, returns a significance of 3.3σ against the

TABLE II. A summary of all the systematic uncertainties in the $B \rightarrow K^* e^+ e^-$ angular analysis.

Uncertainty	$A_T^{(2)}$	A_T^{lm}	f_B
Electron ID efficiency	0.004	0.004	0.001
Electron mis-ID	0.001	0.002	0.001
Hadron ID efficiency	0.002	0.001	0.000
Hadron mis-ID	0.003	0.003	0.001
Fit bias	0.014	0.016	0.002
Signal generator	0.039	0.059	0.000
Fixed F_L and A_T^{Re}	0.009	0.033	0.000
Fixed M_{bc} params.	0.014	0.021	0.001
Fixed acceptance params.	0.022	0.015	0.000
Fixed f_P	0.008	0.012	0.000
MC statistics	0.097	0.095	0.011
Total	0.109	0.121	0.011

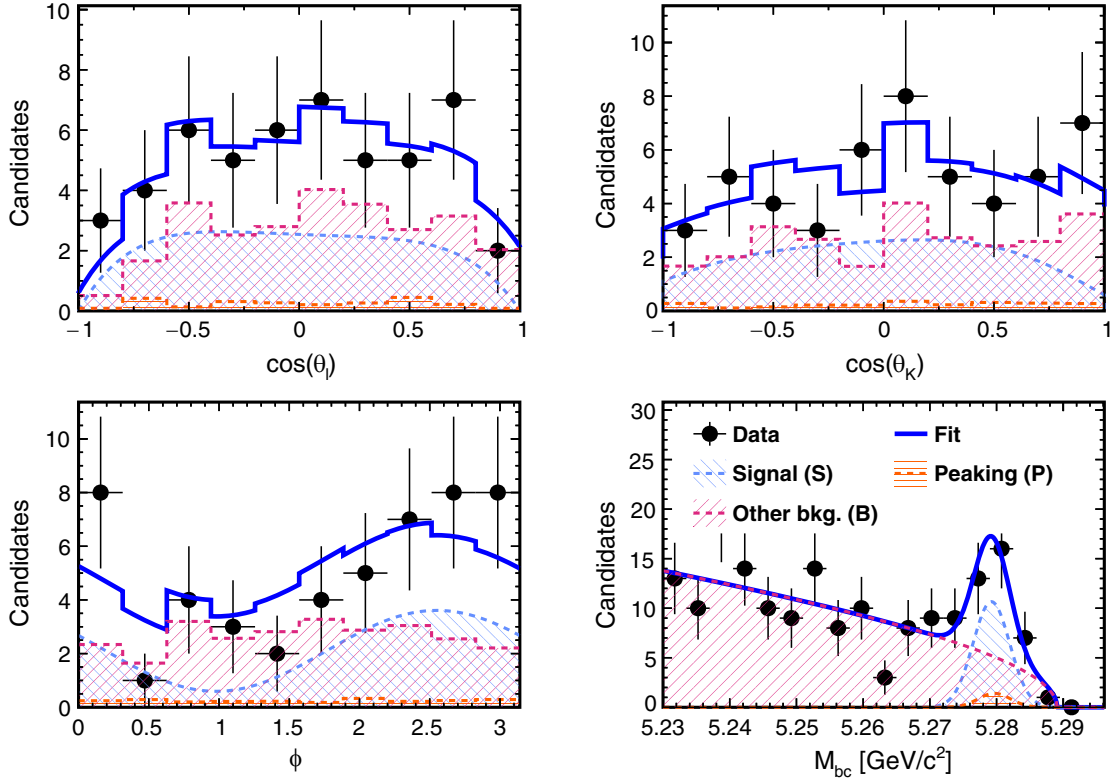


FIG. 3. Projections of $\cos \theta_\ell$, $\cos \theta_K$, ϕ , and M_{bc} are shown, determined from the fits to the full Belle dataset, noting that only $A_T^{(2)}$ and A_T^{Im} are floated in the angular PDF. The angular distributions are plotted for the $M_{bc} > 5.27 \text{ GeV}/c^2$ signal region.

background-only hypothesis. The fits to M_{bc} and the projections of the three angular observables are shown in Fig. 3.

The results for the floated transversality amplitudes in the $0.0008 < q^2 < 1.12 \text{ GeV}^2/c^4$ range are

$$\begin{aligned} A_T^{(2)} &= 0.52 \pm 0.53 \pm 0.11, \\ A_T^{\text{Im}} &= -1.27 \pm 0.52 \pm 0.12, \end{aligned} \quad (6)$$

where the first uncertainty is statistical and the second is systematic. The correlation between these two values is found to be 0.10.

Constraints on the ratio of C_7'/C_7 are shown in Fig. 4, with the left-handed Wilson coefficient fixed to its SM value, $C_7 = -0.2915$ [40]. The figure includes constraints from the measurement of CP -violation parameters in $B_s^0 \rightarrow \phi\gamma$ decays at LHCb [41], $S = 0.43 \pm 0.30 \pm 0.11$ and $\mathcal{A}^\Delta = -0.67_{-0.41}^{+0.37} \pm 0.17$, and $B \rightarrow K^*\gamma$ or $B \rightarrow K^0\pi^0\gamma$ decays averaged by HFLAV [42], $A_{CP} = -0.006 \pm 0.011$ and $S = -0.16 \pm 0.22$. Also shown are the inclusive branching fraction $\mathcal{B}(B \rightarrow X_s\gamma) = (3.49 \pm 0.19) \times 10^{-4}$ [31], and the $B^0 \rightarrow K^{*0}e^+e^-$ angular analysis from LHCb [11], with $A_T^{(2)} = 0.11 \pm 0.10 \pm 0.02$ and $A_T^{\text{Im}} = 0.02 \pm 0.10 \pm 0.01$. The results with and without the LHCb measurement are combined into a global 1σ constraint.

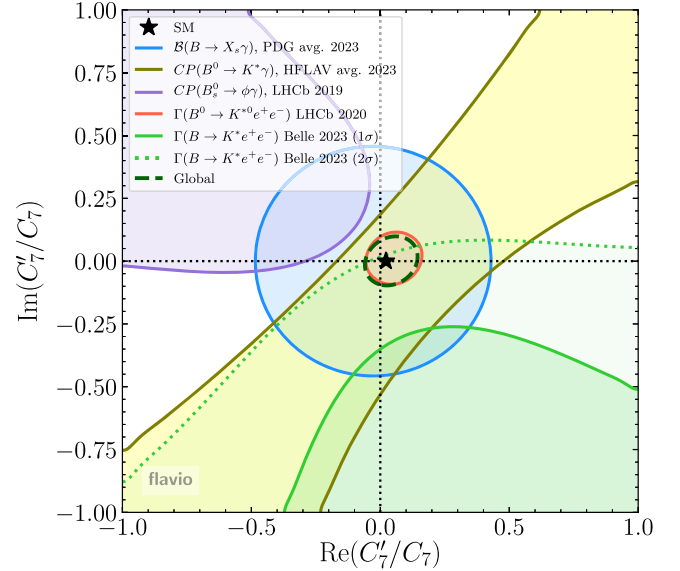


FIG. 4. Contours for confidence limits (at 1σ unless stated otherwise) on C_7'/C_7 are shown for the angular analyses in this paper, the LHCb 2020 study [11], and for several other branching fraction and CP -violation parameter measurements [31,41,42]. These constraints are determined in FLAVIO [36] with $C_7 = -0.2915$.

The results from this analysis alone constrain the value of C_7' to be in agreement with the SM within a 2σ confidence interval, and the combined results remain consistent with the SM, independent of the LHCb measurement.

VII. CONCLUSION

An angular analysis of $B \rightarrow K^* e^+ e^-$ decays for a dielectron invariant mass squared range of $0.0008 - 1.12 \text{ GeV}^2/c^4$ is performed using techniques in electron identification that are new to Belle. This new electron identification uses machine learning and reduces electron misidentification rates by up to a factor 5, expanding the capabilities of rare B decay searches at the experiment. The angular analysis finds $A_T^{(2)} = 0.52 \pm 0.53 \pm 0.11$ and $A_T^{\text{lm}} = -1.27 \pm 0.52 \pm 0.12$. This constrains non-SM right-handed contributions with the finding that C_7' is in agreement with the SM expectation within a 2σ confidence interval. Scaling these results to the design luminosity of Belle II will mean future measurements will have statistical sensitivity competitive with that of LHCb.

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), and the Tau-Lepton Physics Research Center of Nagoya University; the Australian Research Council including Grants No. DP210101900, No. DP210102831, No. DE220100462, No. LE210100098, No. LE230100085; Austrian Federal Ministry of Education, Science and Research (FWF) and FWF Austrian Science Fund No. P 31361-N36; National Key R&D Program of China under Contract No. 2022YFA1601903, National Natural Science Foundation of China and research Grants No. 11575017, No. 11761141009, No. 11705209, No. 11975076, No. 12135005, No. 12150004, No. 12161141008, and No. 12175041, and Shandong Provincial Natural Science Foundation Project No. ZR2022JQ02; 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 Forschungsgemeinschaft, 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. 2016R1D1A1B02012900, No. 2018R1A2B3003643, No. 2018R1A6A1A06024970, No. RS202200197659, No. 2019R1I1A3A01058933, No. 2021R1A6A1A03043957, No. 2021R1F1A1060423, No. 2021R1F1A1064008, No. 2022R1A2C1003993; 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. Aaij *et al.* (LHCb Collaboration), *Phys. Rev. Lett.* **125**, 011802 (2020).
 - [2] S. Wehle *et al.* (Belle Collaboration), *Phys. Rev. Lett.* **118**, 111801 (2017).
 - [3] W. Altmannshofer, P. Ball, A. Bharucha, A. J. Buras, D. M. Straub, and M. Wick, *J. High Energy Phys.* **01** (2009) 019.

- [4] A. Khodjamirian, T. Mannel, A. A. Pivovarov, and Y. M. Wang, *J. High Energy Phys.* **09** (2010) 089.
- [5] D. Becirevic, E. Kou, A. Le Yaouanc, and A. Tayduganov, *J. High Energy Phys.* **08** (2012) 090.
- [6] E. Kou, C.-D. Lü, and F.-S. Yu, *J. High Energy Phys.* **12** (2013) 102.

- [7] H. Eberl, K. Hidaka, E. Ginina, and A. Ishikawa, *Phys. Rev. D* **104**, 075025 (2021).
- [8] M. Blanke, B. Shakya, P. Tanedo, and Y. Tsai, *J. High Energy Phys.* **08** (2012) 038.
- [9] R. Malm, M. Neubert, and C. Schmell, *J. High Energy Phys.* **04** (2016) 042.
- [10] R. Aaij *et al.* (LHCb Collaboration), *J. High Energy Phys.* **04** (2015) 064.
- [11] R. Aaij *et al.* (LHCb Collaboration), *J. High Energy Phys.* **12** (2020) 081.
- [12] R. Aaij *et al.* (LHCb Collaboration), *J. High Energy Phys.* **08** (2013) 131.
- [13] R. Aaij *et al.* (LHCb Collaboration), *J. High Energy Phys.* **11** (2016) 047; **04** (2017) 142(E).
- [14] D. Becirevic and E. Schneider, *Nucl. Phys.* **B854**, 321 (2012).
- [15] A. Abashian *et al.* (Belle Collaboration), *Nucl. Instrum. Methods Phys. Res., Sect. A* **479**, 117 (2002).
- [16] J. Brodzicka *et al.* (Belle Collaboration), *Prog. Theor. Exp. Phys.* **2012**, 04D001 (2012).
- [17] S. Kurokawa and E. Kikutani, *Nucl. Instrum. Methods Phys. Res., Sect. A* **499**, 1 (2003).
- [18] T. Abe *et al.*, *Prog. Theor. Exp. Phys.* **2013**, 03A001 (2013).
- [19] D. J. Lange, *Nucl. Instrum. Methods Phys. Res., Sect. A* **462**, 152 (2001).
- [20] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, *Comput. Phys. Commun.* **191**, 159 (2015).
- [21] E. Barberio, B. van Eijk, and Z. Was, *Comput. Phys. Commun.* **66**, 115 (1991).
- [22] M. Gelb *et al.*, *Comput. Soft. Big Sci.* **2**, 9 (2018).
- [23] T. Kuhr, C. Pulvermacher, M. Ritter, T. Hauth, and N. Braun (Belle-II Framework Software Group Collaboration), *Comput. Soft. Big Sci.* **3**, 1 (2019).
- [24] A. Ali, P. Ball, L. T. Handoko, and G. Hiller, *Phys. Rev. D* **61**, 074024 (2000).
- [25] P. Ball and R. Zwicky, *Phys. Rev. D* **71**, 014029 (2005).
- [26] A. Sibidanov, T. E. Browder, S. Dubey, S. Kohani, R. Mandal, S. Sandilya, R. Sinha, and S. E. Vahsen, *J. High Energy Phys.* **08** (2024) 151.
- [27] K. Hanagaki, H. Kakuno, H. Ikeda, T. Iijima, and T. Tsukamoto, *Nucl. Instrum. Methods Phys. Res., Sect. A* **485**, 490 (2002).
- [28] T. Keck, *Comput. Soft. Big Sci.* **1**, 2 (2017).
- [29] E. Nakano, *Nucl. Instrum. Methods Phys. Res., Sect. A* **494**, 402 (2002).
- [30] A. Drescher, B. Gräwe, B. Hahn, B. Ingelbach, U. Matthiesen, H. Scheck, J. Spengler, and D. Wegener, *Nucl. Instrum. Methods Phys. Res., Sect. A* **237**, 464 (1985).
- [31] R. L. Workman *et al.* (Particle Data Group), *Prog. Theor. Exp. Phys.* **2022**, 083C01 (2022).
- [32] T. Skwarnicki, A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances, Ph.D. thesis, Institute for Nuclear Physics, Krakow, 1986.
- [33] F. Fang, Measurement of branching fractions and CP violation in $B \rightarrow \eta_c K$ and observation of $B^\pm \rightarrow p\bar{p}K^\pm$, Ph.D. thesis, University of Hawaii, 2003.
- [34] G. C. Fox and S. Wolfram, *Phys. Rev. Lett.* **41**, 1581 (1978).
- [35] H. Albrecht *et al.* (ARGUS Collaboration), *Phys. Lett. B* **241**, 278 (1990).
- [36] D. M. Straub, arXiv:1810.08132.
- [37] F. Kruger and J. Matias, *Phys. Rev. D* **71**, 094009 (2005).
- [38] H. F. Kaiser and K. Dickman, *Psychometrika* **27**, 179 (1962).
- [39] M. Borsato, Study of the $B^0 \rightarrow K^{*0} e^+ e^-$ decay with the LHCb detector and development of a novel concept of PID detector: The focusing DIRC, Ph.D. thesis, Paris U., IV, 2015.
- [40] T. Blake, G. Lanfranchi, and D. M. Straub, *Prog. Part. Nucl. Phys.* **92**, 50 (2017).
- [41] R. Aaij *et al.* (LHCb Collaboration), *Phys. Rev. Lett.* **123**, 081802 (2019).
- [42] Y. Amhis *et al.* (Heavy Flavor Averaging Group Collaboration), *Phys. Rev. D* **107**, 052008 (2023).