Belle II observation prospects for axionlike particle production from B meson annihilation decay

Yi Zhang*

Key Laboratory of Nuclear Physics and Ion-beam Application (MOE), Fudan University, 200433 Shanghai, China and Université Paris-Saclay, CNRS/IN2P3, IJCLab, 91405 Orsay, France

Akimasa Ishikawa

SOKENDAI (The Graduate University for Advanced Studies), Hayama 240-0193, Japan and High Energy Accelerator Research Organization (KEK), 305-0801 Tsukuba, Japan

Emi Kou†

Université Paris-Saclay, CNRS/IN2P3, IJCLab, 91405 Orsay, France

Daniel Thomas Marcantonio and Phillip Urquijo

School of Physics, The University of Melbourne, Victoria 3010, Australia

(Received 12 June 2023; accepted 11 December 2023; published 9 January 2024)

We investigate a new production mechanism of axionlike particle (ALP) from B meson annihilation decays and its observation potential at the Belle and Belle II experiments. This mechanism allows for the production of ALP from B meson decays in association with a large variety of mesons. In this article, we first estimate the branching ratios of such processes with a perturbative QCD method. Focusing on the most promising $B \to ha'$ ($h = K^{\pm}, \pi^{\pm}, D^0$ and D_s) channels, we perform sensitivity studies for a' decaying invisibly or into diphoton with Belle and Belle II experiments.

DOI: 10.1103/PhysRevD.109.016008

I. INTRODUCTION

It has been a long time since the axion particle has been introduced by Peccei and Quinn as a solution to the strong CP problem [1], which questions why the CP violation allowed in the Standard Model (SM) is extremely suppressed in nature as seen from the nonobservation of the neutron electric dipole moment. When we consider this suppression as a result of the underlying global $U(1)$ symmetry, the axion particle emerges as the Nambu-Goldstone boson of the spontaneous breaking of this $U(1)$ symmetry. In this article, we investigate the so-called axionlike particle (ALP), a generic name for pseudoscalar particles at a GeV scale, much heavier than the axion in the original model [2–17]. The purpose of this article is to investigate a novel ALP production mechanism, B meson annihilation decay, and its search potential at the Belle (II) experiment.

*zhangy17@fudan.edu.cn
†emi.kou@ijclab.in2p3.fr

In Refs. [9,18–23], it has already been shown that the Belle (II) experiment can provide a unique probe for ALP searches. ALP searches are being intensively pursued, especially in the early physics program of Belle II, the specialized study before the SuperKEKB accelerator reaches luminosities high enough to perform its flavor physics programs, e.g., with B, D mesons or τ leptons. For example, the $e^+e^- \to 3\gamma$ process where ALP could be produced through a $a'\gamma\gamma$ coupling, a specialized trigger was set up [24]. Taking advantage of its clean environment of the e^+e^- collider, Belle (II) can perform very unique searches for ALP which decay into $\gamma\gamma$ or nothing, i.e., invisible decay, even with existing datasets. Thus, we focus on these decay channels in this article.

In the following, we investigate the novel production mechanism of ALP from the B meson annihilation process. The annihilation process occurs as the constituent quarks of B mesons (antibottom and up quark for B^+ meson and antibottom and down quark for B^0 meson) annihilate via the four-fermion interaction and produce hadrons in the final state. On one hand, the observations of pure annihilation processes, $B^0 \to K^+K^-$ (5.2σ significance) [25,26] or $B_s \to \pi^+\pi^-$ (7.0σ significance) [27], prove the existence of such process. On the other hand, annihilation processes

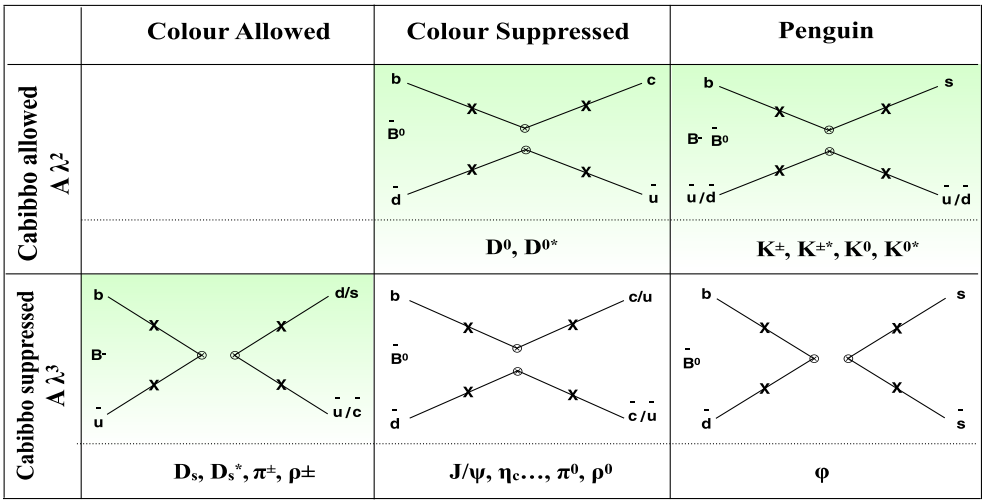

FIG. 1. ALP production from B meson annihilation decays: ALP can be produced at any of the quark line (cross mark), depending on the ALP coupling to different type of quarks. In this figure, we categorise the annihilation diagrams by the Cabbibo factors as well as color allowed/suppressed tree or penguin topologies.

have played important roles in B physics, such as the isospin violation of the $B \to \rho\gamma$ process [28–31] or the strong CP phase generated by the annihilation being the important source of the direct CP violation in the charmless B decays [32,33]. In this paper, we investigate the phenomenology of the annihilation process in ALP search at Belle (II) experiment. The mechanism of producing ALP we consider here is that the ALP is emitted from one of the quark lines of the four-fermion interaction, as shown in Fig. 1. ALP production from B mesons has been discussed in many articles but they mainly consider ALP production from the quarks in the loop (typically, top quark) or from the W boson. Thus, the final states are limited to $B \to K^{(*)}a'$ or $B \to \pi(\rho)a'$. The advantage of the annihilation mechanism is that the ALP can be produced associated with many other hadrons, such as $D, D^*, D_s\cdot$ as well as charmonium, $J/\psi, \eta_c \cdots$, which opens more channels to explore at Belle (II). Furthermore, it can occur from tree level process, whose branching ratio, in general, is larger than the loop level penguin process.

The annihilation diagram is calculable using either the so-called pQCD method [34] or the QCD factorization method [35]. The computation is very close to the annihilation contributions to the radiative $B \to \rho\gamma$ decay [30,31]. In these approaches, the initial and the final state meson distribution amplitudes are convoluted with a hard kernel, which, in this case, results in the a' emission. The hard kernel includes the propagators of the quarks, which are between the a' emission and the four-fermion interaction. Although the annihilation diagram is a $1/m_b$ contribution and is generally suppressed, when there is a light quark propagator (i.e., ALP emission from light quark), it is enhanced by a factor of $1/\Lambda_{\rm QCD}$. In addition, the ALP-quark coupling is proportional to the quark mass in the models where the ALP-fermion coupling is induced by derivative and this effect could compensate the $1/m_b$ suppression factor for the heavy quark cases. Thus, we will consider emission of ALP from all possible quarks in B decay, i.e., u, d, s, c, b and perform the sensitivity study for the best observation channels.

The article is organized as follows. In Sec. II, we describe our theoretical framework for ALP interactions, their couplings and the mass range we are interested in, and we review the computation of ALP production from the B meson annihilation diagram using the pQCD method. In Sec. III, we present our sensitivity study at Belle (II) experiment and we conclude in Sec. IV.

II. ANNIHILATION DIAGRAM COMPUTATION IN PQCD

A. ALP production from annihilation diagrams

In this subsection, we show all the Feynman diagrams that can produce ALP from annihilation B decays. We first categorise the final states by the corresponding CKM matrix elements as well as the color-allowed tree, color-suppressed tree, or penguin topologies. For the purpose of this section, it is enough to consider only the dominant contributions. That is, $B \to Ka'$ and $B \to \pi a'$ have both penguin and tree contributions but they are, respectively, considered as the penguin and the tree processes. Similarly, the charmonium final state is considered as a tree process. The result, taking into account diagrams up to λ^3 order, is listed on the Table I. Naively we expect larger branching ratios for the tree diagrams. Among them, we investigate those with Cabbibo allowed color suppressed tree processes, and Cabbibo suppressed color allowed tree processes in the following, i.e., $D^0, D^{*0}, D_s, D_s^*, \pi^+, \rho^+$ final states. In addition, we also include Cabbibo allowed penguin processes, i.e., K^+, K^{*+}, K^0, K^{*0}, as their experimental sensitivity is known to be quite high. The charmonium production is suppressed with respect to the two tree

TABLE I. The pQCD computation result of $\mathrm{Br}(B \to ha')g_q^2$ of the $B \to ha'$ processes (where $h = D^0, D^{*0}, D_s, D_s^*, K^\pm, K^{*\pm}, K^0, K^{*0}, \pi^\pm, \rho^\pm$) in the unit of $\times 10^{-6}$.

	up	down	strange	charm	bottom
D^0	0.270	3.700	0.000	0.062	0.004
D^{*0}	0.155	3.462	0.000	0.011	0.000
D_s	3.573	0.000	0.305	0.129	0.002
D_s^*	3.403	0.000	0.066	0.025	0.000
$K^\pm$	2.860	0.000	1.659	0.000	0.001
$K^{*\pm}$	0.765	0.000	1.370	0.000	0.000
K^0	0.000	2.220	1.510	0.000	0.0012
K^{*0}	0.000	1.090	1.590	0.000	0.000
$\pi^\pm$	2.080	0.320	0.000	0.000	0.000025
$\rho^\pm$	2.810	0.175	0.000	0.000	0.000062

processes mentioned above but it can be interesting for future study. It should be mentioned that the pQCD method is not the most suitable for charmonium production and we would need to perform an independent theoretical investigation for these processes.

B. Computation of the ALPs production from the annihilation diagram in pQCD method

In this section, using the $\bar{B}^0 \to D^0 a'$ process as an example, we demonstrate how to compute ALP emission from the annihilation diagram in the pQCD method.

Let us start with the weak Hamiltonian:

$$H^{\mathrm{eff}} = \frac{G_F}{\sqrt{2}} \sum_{i=1,2} V_{cb}^* V_{ud} C_i O_i^c \tag{1}$$

where G_F is the Fermi constant and the $V_{qq'}$ is the CKM matrix element. C_i is Wilson coefficient. The operators O_i^c are defined as

$$O_1^c = (\bar{u}_i d_j)_{V-A} (\bar{c}_j b_i)_{V-A}, \tag{2}$$

$$O_2^c = (\bar{u}_i d_i)_{V-A} (\bar{c}_j b_j)_{V-A}. \tag{3}$$

where $(\bar{\psi}_i \psi'_j)_{V-A} = \bar{\psi}_i \gamma^\mu (1-\gamma_5)\psi'_j$ and (i, j) are the color indices. For annihilation decay, the initial b quark and d quark must be in the same current. Thus, we use the Fiertz transformed operators

$$O_1^{c\mathrm{F}} = (\bar{c}_j u_j)_{V-A} (\bar{d}_i b_i)_{V-A}, \tag{4}$$

$$O_2^{c\mathrm{F}} = (\bar{c}_j u_i)_{V-A} (\bar{d}_i b_j)_{V-A}. \tag{5}$$

Then, the amplitude is given as

$$A^{1(a,b)} = \frac{G_F}{\sqrt{2}} V_{cb}^* V_{ud} a_1 \langle D^0 a' | O_1^{c\mathrm{F}} | \bar{B}^0 \rangle^{(a,b)}, \tag{6}$$

where $a_1 = C_1 + C_2/N_C$. The indices a and b indicate ALP emission in Fig. 1 from the initial $b, \bar{d}$ and the final $c, \bar{u}$ quarks, respectively.

Using the definitions of decay constants, distribution functions and form factors given in Appendix A, as well as the axion coupling defined as

$$\mathcal{L} = g_q a' \bar{q} \gamma_5 q; \; q = u, d, s, c, b, \tag{7}$$

we obtain the amplitudes for $\bar{B}^0 \to D^0 a'$:

$$A^{1a}(\bar{B}^0 \to D^0 a') = \frac{G_F}{\sqrt{2}} V_{cb}^* V_{ud} (i f_D) a_1 [f_1^{D^0}(p_D \cdot p_D) + f_2^{D^0}(p_D \cdot p_{a'})], \tag{8}$$

$$A^{1b}(\bar{B}^0 \to D^0 a') = \frac{G_F}{\sqrt{2}} V_{cb}^* V_{ud} (-i f_B) a_1 [f_3^{D^0}(p_B \cdot p_D) + f_4^{D^0}(p_B \cdot p_{a'})], \tag{9}$$

where the Ba' form factors ($f_{1,2}$) and Da' form factors ($f_{3,4}$) are computed as a convolution of the distribution function of the B and D mesons and the hard kernel thst represents the a' emission [36]. The results yield

$$a_1 f_1^{D^0} = \int_0^1 dx \int_0^\infty db \left[(i g_d) N_c a_1(t) \left(\frac{-i}{\sqrt{2N_c}} \Phi_B(t) \right) \times (-1) 4 m_B (-x_1) |b_1| K_0(m_B \sqrt{x_1}, |b_1|) \right] + \int_0^1 dx \int_0^\infty db \left[(i g_b) N_c a_1(t) \left(\frac{-i}{\sqrt{2N_c}} \Phi_B(t) \right) \times (-1) 4 m_B (1-x_1) |b_1| K_0(m_B \sqrt{1+x_1}, |b_1|) \right] \tag{10}$$

$$a_1 f_2^{D^0} = \int_0^1 dx \int_0^\infty db \left[(i g_d) N_c a_1(t) \left(\frac{-i}{\sqrt{2N_c}} \Phi_B(t) \right) \times (-1) 4 m_B |b_1| K_0(m_B \sqrt{x_1}, |b_1|) \right] \tag{11}$$

$$a_1 f_3^{D^0} = \int_0^1 dx \int_0^\infty db \left[(i g_u) N_c a_1(t) \left(\frac{i}{\sqrt{2N_c}} \Phi_D(t) \right) \times (-1) 4 m_D (-x_2) |b_2| \frac{i\pi}{2} H_0(m_B \sqrt{x_2}, |b_2|) \right] + \int_0^1 dx \int_0^\infty db \left[(i g_c) N_c a_1(t) \left(\frac{i}{\sqrt{2N_c}} \Phi_D(t) \right) \times (-1) 4 m_D (1-x_2) |b_2| \frac{i\pi}{2} H_0(m_B \sqrt{1-x_2}, |b_2|) \right] \tag{12}$$

$$a_1 f_4^{D^0} = \int_0^1 dx \int_0^\infty db \left[(i g_u) N_c a_1(t) \left(\frac{i}{\sqrt{2N_c}} \Phi_D(t) \right) \times (-1) 4 m_D (-) |b_2| \frac{i\pi}{2} H_0(m_B \sqrt{x_2}, |b_2|) \right] + \int_0^1 dx \int_0^\infty d|b| \left[(i g_c) N_c a_1(t) \left(\frac{i}{\sqrt{2N_c}} \Phi_D(t) \right) \times (-1) 4 m_D |b_2| \frac{i\pi}{2} H_0(m_B \sqrt{1-x_2}, |b_2|) \right] \tag{13}$$

where x_i and b_i are the parameters describing the momentum of the light quarks insides of the B and D mesons ($i = 1$ for B and $i = 2$ for D, find notations, e.g., in [31,32]). The integration of x and $|b|$ must be taken in the range $\frac{\sqrt{2}}{m_B}\Lambda_{\rm QCD} < x < 1, 0 < |b| < \frac{1}{\Lambda_{\rm QCD}}$ after multiplying though the Wilson coefficient,

$$a_1(t) \quad \text{with} \quad t = \max[\sqrt{X}m_B, 1/b] \tag{14}$$

where X represents the arguments of the Bessel functions, K_0, H_0, in Eqs. (10)–(13), i.e., $x_1, 1+x_1, x_2, 1-x_2$. As is well known, the a_1 value depends strongly on the renormalization scale, t, e.g., from ~ -0.32 to $\sim +0.06$ for $t = 1$–5 GeV. In pQCD, t varies according the momentum fraction carried by the light quark in the mesons.

Equations (10)–(13) show that this process is sensitive to the $g_{u,d,c,b}$ couplings. As mentioned in the Introduction, the dominant contribution comes from the a' emission from the spectator, namely the $f_2^{D^0}$ term. This is because the distribution function of the B meson peaks at smaller value of x_1 while the Bessel function K_0 is suppressed at a higher value of x_1.

The amplitudes for the Cabibbo allowed tree processes ($D^{(*)-}, D_s^{(*)}$) are computed in a similar way, with Wilson coefficient $a_1(t)$, instead of $a_2(t)$, which varies more moderately from $\sim$1.2 to $\sim$1.0 for $t = 1$–5 GeV. These processes are sensitive to the $g_{u,d,c,b}$ and $g_{u,s,c,b}$ couplings, respectively. The ($\pi^{\pm}/\rho^{\pm}, \bar{K}^{(*)0}, K^{(*)-}$) processes come from both penguin and tree diagrams and receive contributions from $a_2(t), a_4(t), a_6(t)$. The computation for $B \to Ka'$ is given in Appendix B as an example. They are sensitive to the $g_{u,d,b}$, $g_{u,s,b}$, and $g_{d,s,b}$ couplings, respectively.

C. Which final states are to be searched at Belle (II)?

Now let us see which processes are more sensitive to each g_i coupling. The obtained results are presented in Table I.

Let us consider different $g_q \neq 0$ scenarios and look into each cases to decide which channels are most promising for the Belle II experiment.

(i) $g_u \neq 0$: In this case, the $D_s, D_s^*, K^{\pm}, \rho^{\pm}, \pi^{\pm}$ channels obtain equally large contributions. The channel with a D_s^* is more challenging to reconstruct as it decays via $D_s^* \to D_s\gamma$ with a relatively low energy photon (over 90% branching ratio). The $\rho^{\pm}$ decays into two pions, with one of them neutral and its experimental measurement suffers from much more background due to the broad width. Thus, the best channels would be $B \to (D_s, K^{\pm}, \pi^{\pm})a'$.

(ii) $g_d \neq 0$: In this case, D^0 and D^{*0} obtain large contributions. However, the channel with a D^{*0} decaying to $D^0\pi^0$ or $D^0\gamma$ makes it more challenging than D^0 to be observed. So the best channel would be the $B \to D^0a'$ for this case.

(iii) $g_s \neq 0$: In this case, K^+, K^{*+}, K^0, K^{*0} obtain the largest contributions. Experimentally the best option would be the $B \to K^{\pm}a'$ channel.

(iv) $g_c \neq 0$: In this case, the best channel seems to be $B \to D_sa'$.

(v) $g_b \neq 0$: In this case, the best channel is $B \to D^0a'$, where D decays to $K^-\pi^+$, $K^-\pi^+\pi^0$ or $K^-\pi^+\pi^+\pi^-$.

III. BELLE (II) SENSITIVITY STUDY

In this section, we perform a Monte Carlo (MC) study to estimate the sensitivity of the ALP search with the Belle and Belle II experiments at the KEKB/superKEKb e^+e^- colliders [37–40]. e^+e^- colliders have an advantage over hadron machines, for missing energy and photon final states. Thus, we concentrate on the ALP decays, $a' \to$ invisible and $a' \to \gamma\gamma$ in this section. In the following, we perform a sensitivity study based on Belle MC, which can be extrapolated to Belle II.

The Belle detector [37] located at the interaction point (IP) of the KEKB collider, which operated at a centre-of-mass energy of 10.58 GeV. From 1999 to 2010, the Belle experiment collected an integrated luminosity of about 711 fb^{-1}. The detector [37] consisted of concentric cylindrical subdetectors; a silicon vertex detector, a central drift chamber, particle ID detectors, an electromagnetic calorimeter (ECL), and a K_L^0- muon detector. In this work, the Belle detector geometry and response is simulated with `Geant3` [41].

The decay channel of $B \to ha'$, where h is π, K, D^0, D_s, are studied. The MC samples of $B^{\pm} \to \pi^{\pm}a', K^{\pm}a', D_s^{\pm}a'$ and $\bar{B}^0 \to D^0a'$ are generated using `EvtGen` [42]. The a' is forced to promptly decay to neutrinos (to mimic a generic final state), or to a di-photon final state. The signal samples contain 1 million events for 33 different ALP mass hypotheses for $B \to D, D_s$ and 45 mass hypotheses for $B \to \pi, K$ processes due to the larger allowable mass range.

In this study, we assume that the branching ratios for these two decay channels are unity. This may come from merely the quark coupling given in Eq. (7) or it may require additional axion coupling. We will leave the detailed model building for the future work.

The background samples used in this work are MC samples produced by the Belle collaboration. They contain $e^+e^- \to B^+B^-, B^0\bar{B}^0$ (both at 10× the Belle integrated luminosity), and $q\bar{q}$ (at 6× the Belle integrated luminosity) where $q = u, d, s,$ or c. There is also a sample with "rare" B decays produced with 25 times of the Belle integrated luminosity. This sample contains B decays that despite having small or poorly measured branching fractions, could be important background to consider, e.g., $B^+ \to \pi^+K^0$.

A. Invisible decay of ALP

Events are reconstructed by first converting Belle MC to a Belle II software (`basf2`) [43] readable format using

b2bii [44]. The signal-side B-meson (B_{sig}) is reconstructed by reconstructing the signal-side hadron (π, K, D_s or D^0). Charged particles (pions and kaons) are required to have their point of closest approach to the interaction point (IP) less than 2 cm and 3 cm in the directions transverse and parallel to the beam direction, respectively, as well as meet some particle identification requirements and have lab-frame momentum greater than 200 MeV/c. The $B \to D_s$ and D channels use further selection on the invariant masses of combined tracks to find D meson candidates, and a vertex fit is done on the D candidate. The other B meson (B_{tag}) in the event is reconstructed with the full event interpretation algorithm (FEI) [45], which uses over 10, 000 channels We refer to the remaining tracks and calorimeter clusters not used in the reconstruction as the rest of event (ROE).

The B_{tag} is required to have a beam-energy-constrained mass of more than 5.27 GeV/c^2 and FEI must return a signal probability of greater than 0.005. The total unused ECL energy must be less than 0.4 GeV. Two multivariate classifiers are trained to separate the signal process from $q\bar{q}$ and nonsignal B events respectively. These are trained on variables that describe the shape of the events.

The search method looks for resonances in the signal-side hadron's B_{sig}-frame momentum, which should peak due to two-body kinematics. The signal-side B frame momentum is calculated by subtracting the four vectors of tag-side and ROE from that of the e^+e^- beam.

The signal peak is fitted for each generated mass hypothesis using a convolution of two Gaussians. The free parameters of this probability density function (PDF) are then parametrized as a function of the mean of the Gaussians in order to have a well-defined signal PDF anywhere in the search range. The background distribution is parametrized using a kernel density estimator [46].

A combined fit is performed over background-only MC, with only the yield of the signal and background PDFs allowed to float, and a 90% confidence interval on the number of signal events yielded by the fit is calculated. We use this upper limit to estimate the branching fraction sensitivity of the analysis. Our result is shown in Fig. 2. The largest discrepancy in the sensitivity between MC and real data is driven by the difference in the FEI performance between MC and data. The efficiency of the FEI can be 30% less on data compared to MC. The sensitivity limits shown here include a correction factor for the FEI efficiency derived from an analysis of $B^\pm \to K^\pm J/\psi$, $J/\psi \to \ell^+\ell^-$ decays where $\ell \in e, \mu$.

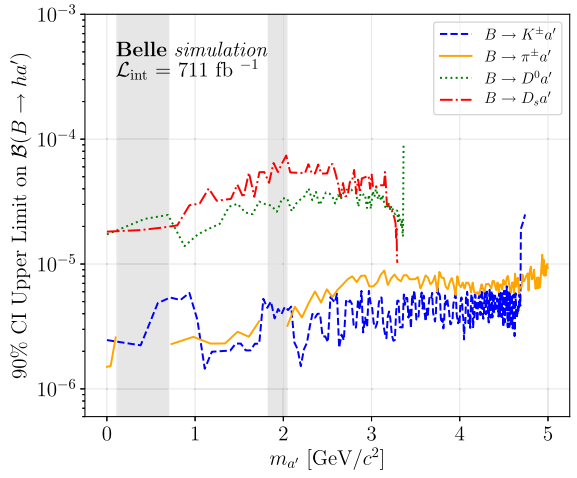

FIG. 2. Expected upper limit for $B \to h(a' \to \text{invisible})$ at Belle experiment, where h is π, K, D^0, D_s. The filled gray bands are the veto ranges of $B \to K^0\pi$ and $B \to D^{(*)0}\pi$, respectively.

B. Diphoton decay of ALP

In the $a' \to \gamma\gamma$ study, signal-side is fully reconstructed by combining a hadron $h = \pi, K, D^0, D_s$ and a combination of two photons. All charged particles are required to have distance from the e^+e^- IP smaller than 0.2 cm in the plane transverse to the beams and smaller than 1.0 cm along the beam direction. The binary ratio $\mathcal{L}(K/\pi) \equiv \frac{\mathcal{L}(K)}{\mathcal{L}(K)+\mathcal{L}(\pi)}$, is used to identify the species of charged particles, where $\mathcal{L}(h)$ is the likelihood for a kaon or pion to produce the signals observed in the detectors. Charged particles with $\mathcal{L}(K/\pi) > 0.6$ are identified as kaons and those with $\mathcal{L}(\pi/K) > 0.6$ as pions. The photons need to meet the criteria of energy $E > 0.05$ GeV. The π^0 and ALP candidates are formed from the combination of two photons, the former within the invariant mass range of $[0.125, 0.140]$ GeV/c^2, corresponding to $\pm 2.5\sigma$ away from the nominal π^0 mass. To suppress peaking backgrounds originating from π^0 and η in the mass spectrum of ALP candidates, we veto the diphoton mass ranges of $[0.120, 0.160]$ GeV/c^2 and $[0.520, 0.590]$ GeV/c^2. An additional requirement on the π^0 momentum to be larger than 0.5 GeV/c is implemented. The D mesons are reconstructed from the final states $K^\mp\pi^\pm$, $K^\mp\pi^\pm\pi^0$, and $K^\pm\pi^\mp\pi^\pm\pi^\mp$ within the invariant mass ranges of $[1.845, 1.885]$ GeV/c^2, $[1.833, 1.890]$ GeV/c^2, and $[1.850, 1.880]$ GeV/c^2, respectively. The $D_s^\pm$ mesons are reconstructed in the channel $K^\pm K^\mp \pi^\mp$ with an invariant mass requirement of $[1.950, 1.990]$ GeV/c^2. These ranges are approximately 3σ on either side of the nominal mass. The D and $D_s^\pm$ momenta are recalculated from a fit of the momenta of its decay products that constrains them to a common origin and their mass to the nominal mass of the D and $D_s^\pm$, respectively. The B mesons candidates are selected to have a beam-energy-constrained mass larger than 5.27 GeV/c^2 and energy difference in the range of $[-0.4, 0.4]$ GeV.

To suppress the background originating from $e^+e^- \to q\bar{q}$ events, a boosted decision tree (BDT) classifier is used to veto candidates from continuum events, training them on equal numbers of simulated signal and continuum. Differing from the invisible search, for $a' \to \gamma\gamma$, the nominal approach is to scan the resonances directly on the invariant mass spectrum of the two photons.

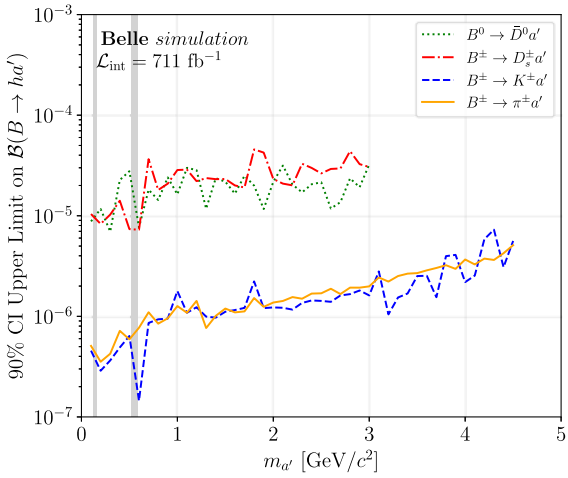

FIG. 3. Expected upper limit for $B \to h(a' \to \gamma\gamma)$ at Belle experiment, where $h = \pi, K, D^0, D_s$. The filled gray bands are the veto ranges of $\pi^0 \to 2\gamma$ and $\eta \to 2\gamma$, respectively.

The signal PDFs are determined by using simulated signal samples with different mass hypotheses. Generally, a combination of two gaussians and one asymmetric Gaussian is fitted to describe the signal shape, which is well-defined. The background is described by a second order polynomial function.

An unbinned maximum likelihood fit is performed on the background-only simulated sample. During the fit, the background yield is the only parameter allowed to float and a 90% confidence interval is calculated which are used to estimate the branching fraction upper limit. Figure 3 shows the corresponding branching fraction upper limit result from Belle simulation.

C. Observation potentials at Belle and Belle II

In this section, we discuss the observation potential of ALP produced from B meson annihilation decays at Belle and Belle II. In Figs. 4 and 5, we overlay the achievable

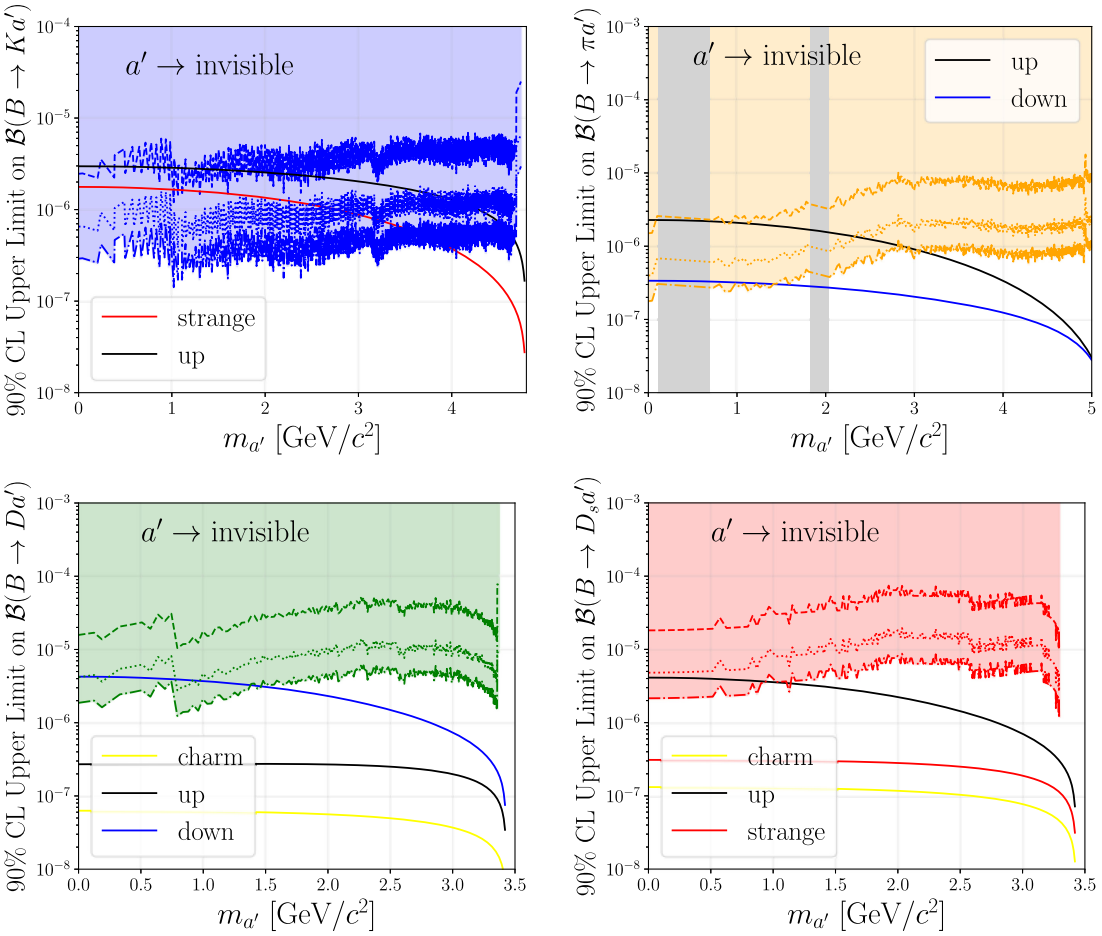

FIG. 4. Belle and Belle II sensitivity compared with the theoretical prediction for $B \to h(a' \to \text{invisible})$. The exclusion regions in blue, yellow, green and red, are those for $B \to ha'$ with $h = K^\pm, \pi^\pm, D^0, D_s$, respectively. The dashed line is simulated Belle sensitivity; The dotted and dash dotted lines are the extrapolation of 5 ab^{-1} and 10 ab^{-1}, respectively. The colored solid lines are theory predictions with $g_i = 1$ $(i = u, d, s, c)$. The filled gray bands are the veto ranges of $B \to K^0\pi$ and $B \to D^{(*)0}\pi$, respectively.

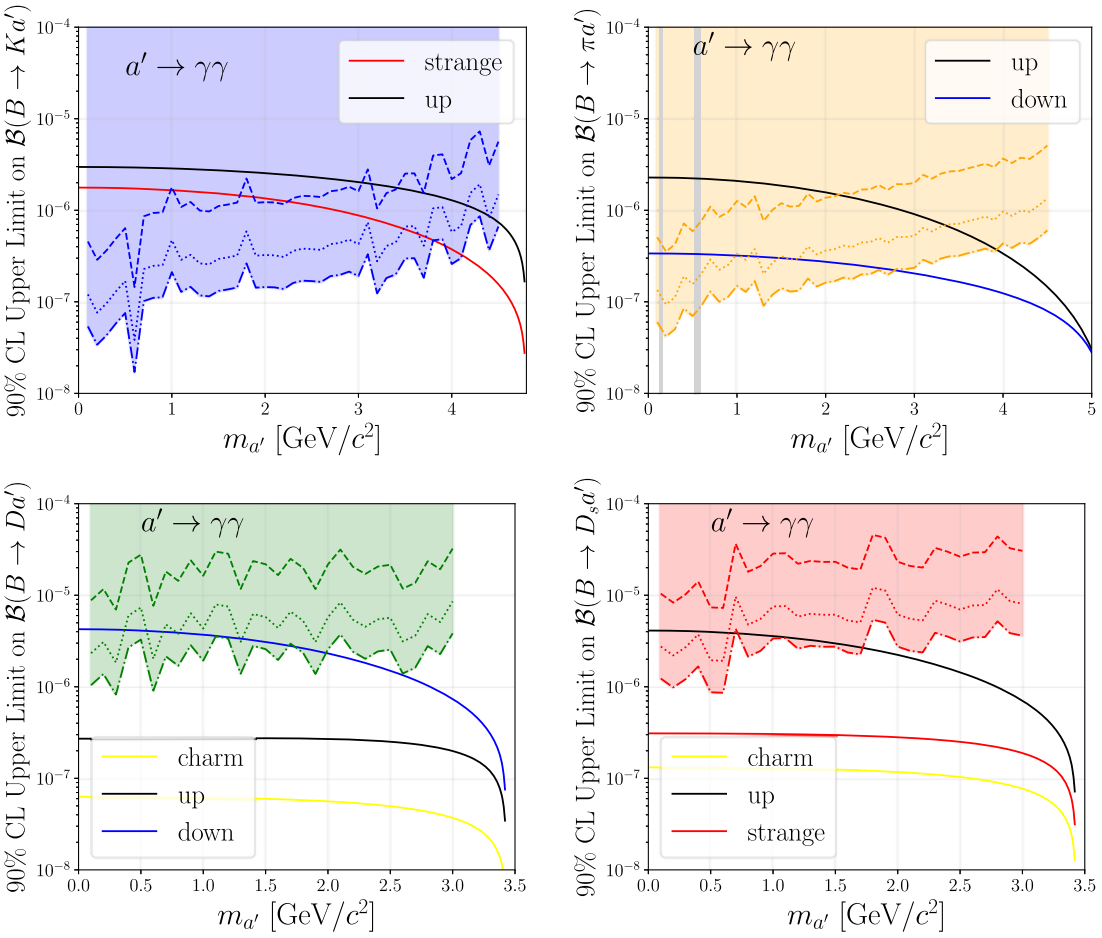

FIG. 5. Belle and Belle II sensitivity compares with the theoretical prediction for $a' \to \gamma\gamma$. The exclusion regions in blue, yellow, green and red, are those for $B \to ha'$ with $h = K^{\pm}, \pi^{\pm}, D^0, D_s$, respectively. The dash line is simulated Belle sensitivity; The dotted and dash dotted lines are the extrapolation of 5 ab^{-1} and 10 ab^{-1}, respectively. The colored solid lines are theory predictions with $g_i = 1$ $(i = u, d, s, c)$. The filled gray bands are the veto ranges of $\pi^0 \to 2\gamma$ and $\eta \to 2\gamma$, respectively.

upper limit of the branching ratios for each channel at Belle and Belle II (as obtained in the previous subsections) and pQCD predictions of the branching ratios (as obtained in Sec. II). The Belle II upper limits show in the Figs. 4 and 5 are the extrapolation results from Belle.

Figure 4 shows that for the invisible final state with the Belle dataset, the ALP that couples to up quarks with $g_u = 1$ are in the observable range for $B \to K^{\pm}a'$ and $B \to \pi^{\pm}a'$ channels. The strange type model, $g_s = 1$, can also be accessed with the $B \to K^{\pm}a'$ channel with Belle II in a few years time (i.e., 10 ab^{-1}). The models with nonzero down type coupling, $g_d = 1$, might be accessible with the full Belle II dataset via the $B \to D^0a'$ channel. The $B \to D^0a'$ and $B \to D_sa'$ channels are sensitive to the models where ALP couples to charm quark, though $g_c = 1$ is too small to be observed at Belle (II) unless there are some new physics model that allow g_c to be enlarged.

From Fig. 5, a similar conclusion can be drawn for the diphoton final state: the ALP that couples to up or strange quarks with $g_{u,s} = 1$ can already be observed in $B \to K^{\pm}a'$ channel at Belle and the down type model, $g_d = 1$, especially with lower mass of ALP, may be accessible in the $B \to D^0a'$ or $B \to \pi^{\pm}a'$ channels with Belle II. For the models where ALP couples to a charm quark, the best channel is $B \to D_sa'$ decay but again, a model that allows a large g_c is needed.

IV. CONCLUSIONS

In this article, we investigated ALP production from the weak annihilation transition in B meson decays. First, we have shown that a large variations of the final states is possible from this new mechanism, i.e., $B \to ha'$, where $h = D^{0(*)}, K^{\pm(*)}, K^{0,(*)}, D_s^{(*)}, \pi^{\pm}, \rho^{\pm}, \pi^0, \rho^0, \phi, J/\psi, \eta_c, \cdots$.

We performed a pQCD based theoretical computation to predict the branching ratios of these processes. It is found that assuming that the ALP-fermion coupling to be $g_i = 1$, the branching ratios are roughly 10^{-6} for the models where ALP couples to the light quarks (i.e., up, down, strange), 10^{-7} for the charm quark, and 10^{-9} for the bottom quark. The suppression for the bottom quark is the well known $1/m_b$ effect. Based on this computation, we conclude that the four channels with large branching ratios, $B \to ha'$ with

$h = D^0, K^\pm, D_s, \pi^\pm$, and we next performed sensitivity studies for these channels with Belle (II).

We focused on the ALP decays into invisible or diphoton final states, which are the strength of the Belle (II) experiment. In order to obtain the achievable upper limit on the branching ratios, we performed detailed sensitivity study using Belle MC. In both final states, the upper limit of the branching ratios for the $K^\pm$ and $\pi^\pm$ channels are an order of magnitude better than the ones for the D^0 and D_s channels. This is partially because the latter further decays weakly and only a part of their decay channels can be considered as the rest of the decay channels are higher multiplicity channels and difficult to analyse (i.e., much higher background). The $K^\pm$ and $\pi^\pm$ channels are sensitive to the models where ALP couples to up, down or strange quarks. Thus, we found that those models could be observed at the Belle (II) experiment. On the other hand, although our theoretical prediction shows that the D^0 and D_s final states are the most promising channels for the models with charm or bottom quarks, the predicted branching ratio is too small to be observed at Belle (II). However, these channels are still interesting to investigate for the models where the coupling of quarks and ALP are induced by the derivative coupling, for which the coupling is proportional to the quark masses and is enhanced for heavy quarks.

ACKNOWLEDGMENTS

We would like to acknowledge J. Zupan for his collaboration at the early stage of the project and his careful reading of the manuscript. The work of A. I. and E. K. is supported by Toshiko Yuasa Laboratory-France Japan Particle Physics Laboratory. A. I. is supported by JSPS KAKENHI Grant No. 22H00144. Y. Z. is supported by China Scholarship Council (CSC) under the State Scholarship Fund No. 201906100100. D. T. M. and P. U. are supported by the Australian Research Council with Grant No. DP210101900.

APPENDIX A: FORM FACTOR COMPUTATIONS IN PQCD

In this section, we briefly describe the computation of the form factors, f_{1-7} in pQCD. We first define the kinematics in the light-cone frame [47]. First, we define the momentum of the initial and the final states. Here we ignore the term $m_{D^{(*)}}/m_B^2$ and $m_{a'}/m_B^2$:

$$p_B = \frac{m_B}{\sqrt{2}}(1,1,0), \quad p_{a'} = \frac{m_B}{\sqrt{2}}(0,1,0), \quad p_{D^{(*)}} = \frac{m_B}{\sqrt{2}}(1,0,0) \tag{A1}$$

Next, the momentum of the spectator quark of B meson in Fig. 1 (top middle), which we call it as k_1 and the spectator of $D^{(*)}$ meson in the same figure, which we call as k_2

$$k_1 = \left(\frac{m_B}{\sqrt{2}}x_1, 0, k_\perp^1\right), \qquad k_2 = \left(\frac{m_B}{\sqrt{2}}x_2, 0, k_\perp^2\right) \tag{A2}$$

The longitudinal and transverse polarizations of the D^* can be given, respectively, as

$$\epsilon^L = \frac{1}{\sqrt{2}}\frac{m_B}{m_{D^*}}(1,0,0), \qquad \epsilon^T = (0,0,\epsilon_\perp) \tag{A3}$$

We use the following spin projection and the wave function to describe the B and $D^{(*)}$ mesons:

$$\int \frac{d^4w}{(2\pi)^4} e^{ik\omega}\langle 0|\bar{d}_\alpha(0)b_\delta(w)|\bar{B}^0\rangle = \frac{-i}{\sqrt{2N_C}}[(\not{p}_B + m_B)\gamma_5]_{\delta\alpha}\phi_B(k)$$

$$\int \frac{d^4w}{(2\pi)^4} e^{ik\omega}\langle D^0|\bar{c}_\alpha(0)u_\delta(w)|0\rangle = \frac{i}{\sqrt{2N_C}}[\gamma_5(\not{p}_D + m_D)]_{\delta\alpha}\phi_D(k)$$

$$\int \frac{d^4w}{(2\pi)^4} e^{ik\omega}\langle D^{0*}|\bar{c}_\alpha(0)u_\delta(w)|0\rangle = \frac{1}{\sqrt{2N_C}}\{[\not{\epsilon}^*_{D^*}\not{p}_{D^*}]_{\delta\alpha}\phi_{D_T^*}(k) + [\not{\epsilon}^*_{D^*}m_{D^*}]_{\delta\alpha}\phi_{D_L^*}(k)\}$$

$$\int \frac{d^4w}{(2\pi)^4} e^{ik\omega}\langle K^-|\bar{s}_\alpha(0)u_\delta(w)|0\rangle = \frac{i}{\sqrt{2N_C}}[\gamma_5\not{p}_K\phi_K(k) + \gamma_5\mu_0\phi_K^p(k) + \gamma_5(\not{n}_+\not{n}_- - 1)\mu_0\phi_K^t(k)]_{\delta\alpha}$$

$$\int \frac{d^4w}{(2\pi)^4} e^{ik\omega}\langle \bar{K}^{*0}|\bar{s}_\alpha(0)u_\delta(w)|0\rangle = \frac{1}{\sqrt{2N_C}}[\not{\epsilon}_{K^*}\not{p}_{K^*}\phi_{K^*}^t(k) + \not{\epsilon}_{K^*}m_{K^*}\phi_{K^*}(k) + m_{K^*}\phi_{K^*}^s(k)]_{\delta\alpha}$$

where v is the unit vector in direction of p_K and n_- is the opposite direction. The meson distribution amplitudes are

$$\phi_B(k) = N_B x^2(1-x)^2 \exp\left[-\frac{1}{2}\left(\frac{xm_B}{\omega_B}\right)^2 - \frac{\omega_B^2 b^2}{2}\right]$$

$$\phi_{D^{(*)}}(k) = \frac{3}{\sqrt{2N_c}} f_{D^{(*)}} x(1-x)(1 + C_{D^{(*)}}(1-2x))$$

$$\phi_K(k) = \frac{3}{\sqrt{2N_c}} f_K x(1-x)(1 + 3a_1^K(1-2x) + \frac{3}{2}a_2^K(5(1-2x)^2 - 1))$$

$$\phi_K^p(k) = \frac{1}{2\sqrt{2N_c}} f_K \left(1 + \left(30\eta_3 - \frac{5}{2}\rho_K^2\right)\frac{1}{2}(3(1-2x)^2 - 1) + \left(-3\eta_3\omega_3 - \frac{27}{20}\rho_K^2 - \frac{81}{10}\rho_K^2 a_2^K\right)\frac{1}{8}(35(1-2x)^4 - 30(1-2x)^2 + 3)\right)$$

$$\phi_K^t(k) = \frac{3}{\sqrt{2N_c}} f_K (1-2x)\left(\frac{1}{6} + \left(5\eta_3 - \frac{1}{2}\eta_3\omega_3 - \frac{7}{20}\rho_K^2 - \frac{3}{5}\rho_K^2 a_2^K\right)(1 - 10x + 10x^2)\right)$$

$$\phi_{K^*}(k) = \frac{3}{\sqrt{2N_c}} f_{K^*} x(1-x)(1 + 3a_1^{\parallel}(1-2x) + \frac{3}{2}a_2^{\parallel}(5(1-2x)^2 - 1)$$

$$\phi_{K^*}^t(k) = \frac{1}{2\sqrt{2N_c}} f_{K^*}^T (3(1-2x)^2 + \frac{3}{2}a_1^{\perp}(1-2x)(3(1-2x)^2 - 1) + \frac{3}{2}a_2^{\perp}(1-2x)^2(5(1-2x)^2 - 3) + \frac{35}{4}\zeta_3^T(3 - 30(1-2x)^2 + 35(1-2x)^4))$$

$$\phi_{K^*}^s(k) = \frac{3}{\sqrt{2N_c}} f_{K^*}^T (a_1^{\perp}(1 - 6x + 6x^2) + (1-2x)\left(1 + \left(a_2^{\perp} + \frac{70}{3}\zeta_3^T\right)(1 - 10x + 10x^2)\right))$$

$$\phi_\pi(k) = \frac{3}{\sqrt{2N_c}} f_\pi x(1-x)\left(1 + 0.44\frac{3}{2}(5(1-2x)^2 - 1) + 0.25\frac{15}{8}(21(2x-1)^4 - 14(2x-1)^2 + 1\right)$$

$$\phi_\pi^p(k) = \frac{1}{2\sqrt{2N_c}} f_\pi \left(1 + 0.43\frac{1}{2}(3(2x-1)^2 - 1) + 0.09\frac{1}{8}(35(2x-1)^4 - 30(2x-1)^2 + 3\right)$$

$$\phi_\pi^t(k) = \frac{1}{2\sqrt{2N_c}} f_\pi (1-2x)(1 + 0.55(1 - 10x + 10x^2))$$

with $N_B = 92$ GeV, $\omega_B = 0.4$ GeV, $C_{D^{(*)}} = 0.7$. This wave function is normalized to match to the definitions of the decay constants and their derivatives introduced earlier:

$$\langle 0|\bar{d}\gamma_\mu\gamma_5 b|\bar{B}^0\rangle \equiv i f_B p_B^\mu \tag{A4}$$

$$\langle D^0|\bar{c}\gamma_\mu\gamma_5 u|0\rangle \equiv -i f_D p_D^\mu \tag{A5}$$

$$\langle D_L^{*0}|\bar{c}\gamma_\mu u|0\rangle \equiv f_{D^*} m_{D^*} \epsilon_{\mu D^*}^* \tag{A6}$$

$$\langle D_T^{*0}|\bar{c}\sigma_{\mu\nu} u|0\rangle \equiv -i f_{D^*}^T (\epsilon_{\mu D^*}^* k_\nu - \epsilon_{\nu D^*}^* k_\mu) \tag{A7}$$

$$\langle K^-|\bar{s}\gamma_\mu\gamma_5 u|0\rangle \equiv -i f_K p_K^\mu \tag{A8}$$

$$\langle K^-|\bar{s}\gamma_5 u)|0\rangle \equiv i\mu_0 f_K \tag{A9}$$

The distribution function, $\Phi_{B,D^{(*)}}(t) = \phi_{B,D^{(*)}}(x,b) S_t(x) \exp[-S_B(x,b,t)]$ includes the threshold factor, $S_t(x)$ and the Sukakov form factor, $\exp[-S_{B,D^{(*)}}(x,b,t)]$ which are

$$S_t(x) = \frac{2^{(1+2c)}\Gamma(3/2 + c)}{\sqrt{\pi}\Gamma(1+c)}(x(1-x))^c \tag{A10}$$

$$S_B(x_1, b_1, t) = \begin{cases} s\left(\frac{x_1 m_B}{\sqrt{2}}, b_1\right) + 2\int_{1/b_1}^{t} \frac{d\mu}{\mu}\gamma(\mu) & \text{for } S_B > 0 \\ 0 & \text{for } S_B < 0 \end{cases} \tag{A11}$$

$$S_{D^{(*)}}(x_2, b_2, t) = \begin{cases} s\left(\frac{x_2 m_B}{\sqrt{2}}, b_2\right) + 2\int_{1/b_2}^{t} \frac{d\mu}{\mu}\gamma(\mu) & \text{for } S_{D^{(*)}} > 0 \\ 0 & \text{for } S_{D^{(*)}} < 0 \end{cases} \tag{A12}$$

$$S_K(x_2,b_2,t)=\begin{cases} s\left(\frac{x_2 m_B}{\sqrt{2}},b_2\right)+s\left(\frac{(1-x_2)m_B}{\sqrt{2}},b_2\right)+2\int_{1/b_2}^{t}\frac{d\mu}{\mu}\gamma(\mu) & \text{for } S_K>0 \\ 0 & \text{for } S_K<0 \end{cases} \tag{A13}$$

$$s(Q,b)=\begin{cases} \int_{1/b}^{Q}\frac{d\mu}{\mu}\ln\left[\frac{Q}{\mu}\right]A(\mu)+B(\mu) & \text{for } Q>1/b \\ 0 & \text{for } Q<1/b \end{cases} \tag{A14}$$

where $c=0.4$ and $\gamma(\mu)=-\frac{\alpha_s(\mu)}{\pi}$.

The $V-A$ form factor for Ka' and $S\pm P$ form factors for Ka' can be obtained in a similar manner [48].

$$\begin{aligned}\langle K^- a'|(\bar{s}_i u_i)_{V-A}|0\rangle &= \int_0^1 dx\int_0^\infty d|b|(ig_u)N_c\left(\frac{i}{\sqrt{2N_c}}\phi_K^p(k)\right)\times(-1)\\ &\quad\times 4\mu_0[-x_2 p_K^\mu - p_{a'}^\mu]|b_2|\frac{i\pi}{2}H_0(m_B\sqrt{x_2},|b_2|)\end{aligned} \tag{A15}$$

$$\begin{aligned}\langle a'|(\bar{d}_j b_j)_{S-P}|\bar{B}^0\rangle &= \int_0^1 dx\int_0^\infty d|b|(ig_u)N_c\left(\frac{-i}{\sqrt{2N_c}}\phi_B(k)\right)\times(-1)\\ &\quad\times 4[-x_1(p_B\cdot p_K)+(p_B\cdot p_{a'})]|b_1|K_0(m_B\sqrt{x_1},|b_1|)\end{aligned} \tag{A16}$$

$$\begin{aligned}\langle K^- a'|(\bar{s}_i u_i)_{S+P}|0\rangle &= \int_0^1 dx\int_0^\infty d|b|(ig_u)N_c\left(\frac{i}{\sqrt{2N_c}}\phi_K(k)\right)\times(-1)\\ &\quad\times 4\mu_0[x_2 m_K^2+(p_K\cdot p_{a'})]|b_2|\frac{i\pi}{2}H_0(m_B\sqrt{x_2},|b_2|)\end{aligned} \tag{A17}$$

The integration must be taken in the range $\frac{\sqrt{2}}{m_B}\Lambda_{\rm QCD}<x<1, 0<b<\frac{1}{\Lambda_{\rm QCD}}$ after multiplying with the Wilson coefficient,

$$a_1(t)\quad\text{with}\quad t=\max[\sqrt{x}m_B,1/b] \tag{A18}$$

APPENDIX B: COMPUTATION OF THE PENGUIN PROCESS $B^-\to K^- a'$ IN PQCD

The weak Hamiltonian which gives the $B^-\to K^- a'$ process is

$$H^{\rm eff}=\frac{G_F}{\sqrt{2}}\left[V_{ub}^* V_{us}(C_1 O_1+C_2 O_2)-V_{tb}^* V_{ts}\sum_{i=1,10}C_i O_i\right] \tag{B1}$$

where the operators O_i are defines as

$$\begin{aligned} O_1&=(\bar{s}_i u_j)_{V-A}(\bar{u}_j b_i)_{V-A}, & O_2&=(\bar{s}_i u_i)_{V-A}(\bar{u}_j b_j)_{V-A}\\ O_3&=(\bar{s}_i b_i)_{V-A}(\bar{u}_j u_j)_{V-A}, & O_4&=(\bar{s}_i b_j)_{V-A}(\bar{u}_j u_i)_{V-A}\\ O_5&=(\bar{s}_i b_i)_{V-A}(\bar{u}_j u_j)_{V+A}, & O_6&=(\bar{s}_i b_j)_{V-A}(\bar{u}_j u_i)_{V+A}\\ O_7&=\frac{3}{2}e_u(\bar{s}_i b_i)_{V-A}(\bar{u}_j u_j)_{V+A}, & O_8&=\frac{3}{2}e_u(\bar{s}_i b_j)_{V-A}(\bar{u}_j u_i)_{V+A}\\ O_9&=\frac{3}{2}e_u(\bar{s}_i b_i)_{V-A}(\bar{u}_j u_j)_{V-A}, & O_{10}&=\frac{3}{2}e_u(\bar{s}_i b_j)_{V-A}(\bar{u}_j u_i)_{V-A}\end{aligned} \tag{B2}$$

In order to have the initial b quark and $\bar{u}$ quark be in the same current (see Fig. 1), $O_{3,10}$ must be Fiertz transformed in the Dirac space.

$$\begin{aligned} O_{3,9}^F&\propto(\bar{u}_j b_i)_{V-A}(\bar{s}_i u_j)_{V-A}, & O_{4,10}^F&\propto(\bar{u}_j b_j)_{V-A}(\bar{s}_i u_i)_{V-A}\\ O_{5,7}^F&\propto -2(\bar{u}_j b_i)_{S-P}(\bar{s}_i u_j)_{S+P}, & O_{6,8}^F&\propto -2(\bar{u}_j b_j)_{S-P}(\bar{s}_i u_i)_{S+P}\end{aligned} \tag{B3}$$

where $(\bar{\psi}\psi')_{S\pm P} = \bar{\psi}(1 \pm \gamma_5)\psi'$ After applying the Fiertz transformation to the color space, we find the combination of the Wilson coefficients $a_2 O_2$, $a_4 O_4^F$ and $a_6 O_6^F$ where $a_4 = C_4 + C_3/N_C + C_{10} + C_9/N_C, a_6 = C_6 + C_5/N_C + C_8 + C_7/N_C$.

The amplitude of $B^- \to K^- a'$ can be obtained by sandwiching this Hamiltonian by the initial and the final states:

$$A = \frac{G_F}{\sqrt{2}}\{V_{ub}^* V_{us} a_2 \langle K^- a'|O_2|B^-\rangle - V_{tb}^* V_{ts}[a_4\langle K^- a'|O_4^F|B^-\rangle + a_6\langle K^- a'|O_6^F|B^-\rangle]\} \tag{B4}$$

The amplitudes for the APLs emission from the initial b and $\bar{u}$ and the final s and $\bar{u}$ in Fig. 1 are given, respectively

$$A^{1a} = \frac{G_F}{\sqrt{2}}\{V_{ub}^* V_{us} a_2 \langle K^-|(\bar{s}_i u_i)_{V-A}|0\rangle\langle a'|(\bar{u}_j b_j)_{V-A}|B^-\rangle \tag{B5}$$

$$\begin{aligned} &- V_{tb}^* V_{ts} a_4 \langle K^-|(\bar{s}_i u_i)_{V-A}|0\rangle\langle a'|(\bar{u}_j b_j)_{V-A}|B^-\rangle \\ &- V_{tb}^* V_{ts}(-2a_6)\langle K^-|(\bar{s}_i u_i)_{S+P}|0\rangle\langle a'|(\bar{u}_j b_j)_{S-P}|B^-\rangle\} \end{aligned} \tag{B6}$$

$$A^{1b} = \frac{G_F}{\sqrt{2}}\{V_{ub}^* V_{us} a_2 \langle K^- a'|(\bar{s}_i u_i)_{V-A}|0\rangle\langle 0|(\bar{u}_j b_j)_{V-A}|B^-\rangle \tag{B7}$$

$$\begin{aligned} &- V_{tb}^* V_{ts} a_4 \langle K^- a'|(\bar{s}_i u_i)_{V-A}|0\rangle\langle 0|(\bar{u}_j b_j)_{V-A}|B^-\rangle \\ &- V_{tb}^* V_{ts}(-2a_6)\langle K^- a'|(\bar{s}_i u_i)_{S+P}|0\rangle\langle 0|(\bar{u}_j b_j)_{S-P}|B^-\rangle\} \end{aligned} \tag{B8}$$

Then, after defining the form factors as [49]

$$\langle a'|(\bar{u}_j b_j)_{V-A}|B^-\rangle \equiv p_K^\mu f_1^{K^-} + p_{a'}^\mu f_2^{K^-} \tag{B9}$$

$$\langle K^- a'|(\bar{s}_i u_i)_{V-A}|0\rangle \equiv p_K^\mu f_3^{K^-} + p_{a'}^\mu f_4^{K^-} \tag{B10}$$

$$\langle a'|(\bar{u}_j b_j)_{S-P}|B^-\rangle \equiv (p_K \cdot p_B) f_5^{K^-} + (p_{a'} \cdot p_B) f_6^{K^-} \tag{B11}$$

$$\langle K^- a'|(\bar{s}_i u_i)_{S+P}|0\rangle \equiv (p_K \cdot p_K) f_7^{K^-} + (p_{a'} \cdot p_K) f_8^{K^-} \tag{B12}$$

We can obtain the amplitudes as:

$$A^{1a}(B^- \to K^- a') = \frac{G_F}{\sqrt{2}}\{V_{ub}^* V_{us}(if_K) a_2[f_1^{K^-}(p_K \cdot p_K) + f_2^{K^-}(p_K \cdot p_{a'})] \tag{B13}$$

$$\begin{aligned} &- V_{tb}^* V_{ts}(if_K) a_4[f_1^{K^-}(p_K \cdot p_K) + f_2^{K^-}(p_K \cdot p_{a'})] \\ &- V_{tb}^* V_{ts}(i\mu_0 f_K)(-2a_6)[f_5^{K^-}(p_K \cdot p_B) + f_6^{K^-}(p_{a'} \cdot p_B)]\} \end{aligned}$$

$$A^{1b}(B^- \to K^- a') = \frac{G_F}{\sqrt{2}}\{V_{ub}^* V_{us}(-if_B) a_2[f_3^{K^-}(p_B \cdot p_K) + f_4^{K^-}(p_B \cdot p_{a'})] \tag{B14}$$

$$\begin{aligned} &- V_{tb}^* V_{ts}(-if_B) a_4[f_3^{K^-}(p_B \cdot p_K) + f_4^{K^-}(p_B \cdot p_{a'})] \\ &- V_{tb}^* V_{ts}(im_B f_B)(-2a_6)[f_7^{K^-}(p_K \cdot p_K) + f_8^{K^-}(p_{a'} \cdot p_K)]\} \end{aligned} \tag{B15}$$

[1] R. D. Peccei and Helen R. Quinn, Phys. Rev. Lett. **38**, 1440 (1977).
[2] H. Georgi, D. B. Kaplan, and L. Randall, Phys. Lett. **169B**, 73 (1986).
[3] M. Bauer, M. Neubert, and A. Thamm, J. High Energy Phys. 12 (2017) 044.
[4] D. S. M. Alves and N. Weiner, J. High Energy Phys. 07 (2018) 092.
[5] I. Brivio, M. B. Gavela, L. Merio, K. Mimasu, J. M. No, R. del Rey, and V. Sanz, Eur. Phys. J. C **77**, 572 (2017).
[6] C. Cornella, P. Paradisi, and O. Sumensari, J. High Energy Phys. 01 (2020) 158.
[7] Jorge Martin Camalich, Maxim Pospelov, Pham Ngoc Hoa Vuong, Robert Ziegler, and Jure Zupan, Phys. Rev. D **102**, 015023 (2020).
[8] Martin Bauer, Matthias Neubert, Sophie Renner, Marvin Schnubel, and Andrea Thamm, J. High Energy Phys. 09 (2022) 056.
[9] L. Merlo, F. Pobbe, S. Rigolin, and O. Sumensari, J. High Energy Phys. 06 (2019) 091.
[10] Alexey Kivel, Julien Laux, and Felix Yu, J. High Energy Phys. 03 (2023) 078.
[11] A. Bharucha, F. Brümmer, N. Desai, and S. Mutzel, J. High Energy Phys. 02 (2023) 141.
[12] Xabier Cid Vidal, Alberto Mariotti, Diego Redigolo, Filippo Sala, and Kohsaku Tobioka, J. High Energy Phys. 01 (2019) 113.
[13] Babette Döbrich, Fatih Ertas, Felix Kahlhoefer, and Tommaso Spadaro, Phys. Lett. B **790**, 537 (2019).
[14] Daniel Aloni, Yotam Soreq, and Mike Williams, Phys. Rev. Lett. **123**, 031803 (2019).
[15] G. Alonso-Álvarez, M. B. Gavela, and P. Quilez, Eur. Phys. J. C **79**, 223 (2019).
[16] J. Bonilla, A. de Giorgi, B. Gavela, L. Merlo, and M. Ramos, J. High Energy Phys. 02 (2023) 138.
[17] M. B. Gavela, R. Houtz, P. Quilez, R. del Rey, and O. Sumensari, Eur. Phys. J. C **79**, 369 (2019).
[18] Torben Ferber, Anastasiia Filimonova, Ruth Schäfer, and Susanne Westhoff, J. High Energy Phys. 04 (2023) 131.
[19] Matthew J. Dolan, Torben Ferber, Christopher Hearty, Felix Kahlhoefer, and Kai Schmidt-Hoberg, J. High Energy Phys. 12 (2017) 094.
[20] Kingman Cheung, Abner Soffer, Zeren Simon Wang, and Yu-Heng Wu, J. High Energy Phys. 11 (2021) 218.
[21] Akimasa Ishikawa, Yasuhito Sakaki, and Yosuke Takubo, Prog. Theor. Exp. Phys. **2022**, 113B05 (2022).
[22] E. Kou *et al.* (Belle-II Collaboration), Prog. Theor. Exp. Phys. **2019**, 123C01 (2019); **2020**, 029201(E) (2020).
[23] Motoi Endo, Syuhei Iguro, and Teppei Kitahara, J. High Energy Phys. 06 (2020) 040.
[24] F. Abudinén *et al.* (Belle II Collaboration), Phys. Rev. Lett. **125**, 161806 (2020).
[25] R. Aaij *et al.* (LHCb Collaboration), J. High Energy Phys. 10 (2012) 037.
[26] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. Lett. **118**, 081801 (2017).
[27] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **108**, 211803 (2012).
[28] M. Beyer, D. Melikhov, N. Nikitin, and B. Stech, Phys. Rev. D **64**, 094006 (2001).
[29] Patricia Ball and Roman Zwicky, Phys. Rev. D **71**, 014029 (2005).
[30] Stefan W. Bosch and Gerhard Buchalla, Nucl. Phys. **B621**, 459 (2002).
[31] Cai-Dian Lü, Mao-Zhi Yang, M. Matsumori, and A. I. Sanda, Phys. Rev. D **72**, 094005 (2005).
[32] Yong-Yeon Keum, Hsiang-nan Li, and A. I. Sanda, Phys. Lett. B **504**, 6 (2001).
[33] Christian W. Bauer, Dan Pirjol, Ira Z. Rothstein, and Iain W. Stewart, Phys. Rev. D **70**, 054015 (2004).
[34] Y.-Y. Keum, H.-n. Li, and A. I. Sanda, Phys. Rev. D **63**, 054008 (2001).
[35] M. Beneke, G. Buchalla, M. Neubert, and C. T. Sachrajda, Nucl. Phys. **B591**, 313 (2000).
[36] The K_0 and H_0 terms represent the Bessel and the Hankel functions, which appear by the Fournier transformation of the $k_\perp$ of the spectator to the impact parameter b, which leads to $\int_{-\infty}^{\infty} dk_\perp^2 \frac{1}{k_\perp^2+a^2} = \int_0^{\infty} d|b||b|K_0(a|b|)$ and $\int_{-\infty}^{\infty} dk_\perp^2 \frac{1}{k_\perp^2-a^2} = \int_0^{\infty} d|b||b|\frac{i\pi}{2}H_0(a|b|)$.
[37] A. Abashian *et al.* (Belle Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 117 (2002).
[38] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods Phys. Res., Sect. A **499**, 1 (2003).
[39] K. Akai, K. Furukawa, and H. Koiso, Nucl. Instrum. Methods Phys. Res., Sect. A **907**, 188 (2018).
[40] T. Abe *et al.* (Belle II Collaboration), arXiv:1011.0352.
[41] R. Brun *et al.*, Report No. CERN-DD-EE-84-01 (1984), https://inspirehep.net/literature/252007.
[42] D. J. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 152 (2001).
[43] T. Kuhr, C. Pulvermacher, M. Ritter *et al.*, Comput. Software Big Sci. **3**, 1 (2019).
[44] M. Gelb, T. Keck, M. Prim *et al.*, Comput. Software Big Sci. **2**, 9 (2018).
[45] T. Keck *et al.*, Comput. Software Big Sci. **3**, 6 (2019).
[46] Kyle Cranmer, Comput. Phys. Commun. **136**, 198 (2001).
[47] The light-cone frame is defined by $p \equiv (p_+, p_-, p_\perp) = (\frac{p_0+p_3}{\sqrt{2}}, \frac{p_0-p_3}{\sqrt{2}}, (p_1, p_2))$. The product is given as $p \cdot p' = p_+p'_- + p_-p'_+ - p_\perp \cdot p'_\perp$.
[48] Note it turned out that the tensor term is small numerically. Thus, we do not write the result explicitly in the following.
[49] Note that the following expressions are correct up to the tensor wave function contribution. The tensor term makes the expression very messy while its contribution turned out to be very small numerically. Thus, we refrain from writing down the full expression.