

Using $B^+ \rightarrow K^+ \nu \bar{\nu}$ to search for the QCD axion

Merna Abumusabh^{1,*}, Giulio Dujany^{1,†}, Diego Guadagnoli^{2,‡},
Axel Iohner^{2,§} and Claudio Toni^{2,||}

¹*IPHC, Université de Strasbourg et CNRS, 67200 Strasbourg, France*

²*LAPTh, Université Savoie Mont-Blanc et CNRS, 74941 Annecy, France*



(Received 14 November 2025; accepted 6 April 2026; published 7 May 2026)

We introduce a model-independent framework to reinterpret Belle II results using only public data, analytically reconstructing the mapping between true and reconstructed kinematic variables within the statistically dominant Inclusive Tagging Analysis. This enables rare-decay measurements to probe light invisible particles—such as the QCD axion or axionlike particles, collectively denoted by a —without relying on internal simulations. Applying the method to $B^+ \rightarrow K^+ \nu \bar{\nu}$ yields the strongest bound on the branching fraction for $B^+ \rightarrow K^+ a$, improving existing limits by about a factor of 9 and constraining the axion’s fundamental flavor-changing coupling to b and s quarks. The approach establishes $B^+ \rightarrow K^+ \nu \bar{\nu}$ as a dual probe—simultaneously testing short-distance new physics and light invisible states, the two probes working independently to an excellent approximation—and provides a general strategy for model-independent reinterpretation of collider data.

DOI: [10.1103/nw6q-6nhz](https://doi.org/10.1103/nw6q-6nhz)

I. INTRODUCTION

Axions and axionlike particles (ALPs) are hypothetical light, spinless, pseudoscalar states actively sought through laboratory experiments [1–3], astrophysical observations [4–6], cosmological probes [7], and theoretical studies [8–10]. A central target is the “invisible” [11–14] QCD axion [15–18], henceforth *the* axion, which elegantly solves the strong- CP problem—the unexplained suppression of a CP -violating QCD term. Given the extreme smallness of this term, $O(10^{-10})$, neither coincidence nor anthropic reasoning offers a satisfactory explanation [19,20]. The axion is also a viable dark-matter candidate [21–23]. Throughout this paper, we denote the QCD axion or a generic ALP as a . Unlike generic ALPs, the QCD axion’s mass m_a and decay constant f_a are linked: A lighter QCD axion necessarily couples more weakly, its interactions scaling as $1/f_a$.

Although this relation places the QCD axion well below the pion mass [9], it simultaneously opens the door to its production at high-intensity colliders, where large event yields can compensate for its small couplings. In this regime, rare decays of heavy mesons such as K , D , and B —or their baryonic counterparts—provide a unique window: The axion would manifest as missing energy and momentum, while the

hadronic system, for suitably chosen channels, remains fully reconstructible. This defines a clean, general, and conservative search strategy capable of probing both QCD axions and generic ALPs.

A key experimental requirement for such searches is that the visible decay products be measured with high precision and the kinematics of the parent hadron be well constrained, enabling the identification of missing energy with minimal model dependence. The Belle II experiment [24] is particularly well suited in this regard: It collides e^- and e^+ beams with momenta of approximately 7 and 4 GeV [25], producing $b\bar{b}$ pairs almost exclusively through the $\Upsilon(4S)$ resonance. The latter decays with near-unit probability into charged or neutral $B\bar{B}$ meson pairs. Signal decays are then studied in one of the two B mesons, while the other is used as a tag, allowing the kinematics of the signal decay to be reconstructed and the flavor of the signal B to be identified.

In this work, we consider the decay $B^+ \rightarrow K^+ a$ [26]. The axion-hadron interaction relevant for this decay is obtained by matching the axion-quark interaction onto the low-energy chiral effective theory. As a result, the axion-hadron couplings are unambiguously determined—up to the standard hadronic low-energy constants at the chosen chiral order—by the underlying axion-quark couplings. Just below the electroweak scale, the latter can be written in the most general form as [27]

$$\mathcal{L}_a \supset \frac{\partial_\mu a}{2f_a} \sum_{i,j=1}^3 \bar{d}_i \gamma^\mu [(k_V)_{d_i d_j} + \gamma_5 (k_A)_{d_i d_j}] d_j + \{d \rightarrow u\} - \frac{\alpha_s}{8\pi} \frac{a}{f_a} G\tilde{G}, \quad (1)$$

where u_i , d_i denote up- and down-type quarks of generation i in the mass eigenbasis. The $G\tilde{G}$ operator [28] can be traded for an axion-field dependent quark mass matrix by redefining the quark fields through $q \rightarrow \exp[i(Q_V + Q_A \gamma_5) a / (2f_a)] q$,

*Contact author: merna.abumusabh@iphc.cnrs.fr

†Contact author: giulio.dujany@iphc.cnrs.fr

‡Contact author: diego.guadagnoli@lapth.cnrs.fr

§Contact author: axel.iohner@lapth.cnrs.fr

||Contact author: claudio.toni@lapth.cnrs.fr

with $\text{Tr}Q_A = 1$ [27]. Although the Hermitian matrices Q_V, Q_A can be chosen as diagonal, they introduce an unphysical reparametrization ambiguity [27].

The process $B^+ \rightarrow K^+ a$, on which this paper is focused, involves the complex coupling $(k_V)_{sb}$. The calculation of this process requires the following hadronic matrix element:

$$\langle K^-(p-q) | \bar{s}b | B^-(p) \rangle = \frac{m_B^2 - m_K^2}{m_b - m_s} f_0(q^2). \quad (2)$$

Here, form-factor values and conventions are taken from Ref. [29]. The relation in Eq. (2) implies the following branching ratio (BR) [30,31]:

$$\mathcal{B}(B^- \rightarrow K^- a) = \frac{\tau_B |(k_V)_{sb}|^2 |\vec{k}| m_B^2}{32\pi f_a^2} \left(1 - \frac{m_K^2}{m_B^2}\right)^2 f_0^2(m_a^2), \quad (3)$$

where $|\vec{k}| = \lambda^{1/2}(m_B^2, m_{K^{(*)}}^2, m_a^2)/(2m_B)$ and $\lambda(a, b, c) = a^2 + b^2 + c^2 - 2ab - 2ac - 2bc$. While the theoretically well-motivated QCD axion requires $m_a \simeq 0$, we will also discuss the general case of an ALP. Here and henceforth, the axion-quark couplings are understood at a scale close to m_B .

We search for the signal decay $B^+ \rightarrow K^+ a$ in data collected by Belle II for the $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurements. Belle II provides two nearly *uncorrelated* datasets for the $B^+ \rightarrow K^+ \nu \bar{\nu}$ analysis: the “classic” hadronic tagging analysis (HTA) and the more recent inclusive tagging analysis (ITA) [32–34]. In the HTA, the full tag- B decay chain is reconstructed, allowing the dineutrino invariant mass q^2 to be measured directly. In contrast, the ITA provides far larger statistics—and therefore dominates the sensitivity—but does not reconstruct q^2 ; instead, it uses a proxy variable q_{rec}^2 obtained as discussed below. Recasting the ITA sample for $B^+ \rightarrow K^+ a$ searches therefore requires knowledge of the q^2 – q_{rec}^2 mapping, which is normally accessible only through internal Belle II simulation.

A first central result of this work is to show that this mapping can be derived analytically from kinematics alone, without relying on nonpublic experimental inputs. This insight enables a model-independent reinterpretation of the ITA data for light invisible states such as the axion. Using this mapping and publicly available efficiencies, we demonstrate that this search yields the strongest existing bounds on the magnitudes of the fundamental QCD-axion coupling $(k_V)_{sb}$. This represents our second key result.

Both the ITA and HTA measurements show central values above the standard-model (SM) prediction for $B^+ \rightarrow K^+ \nu \bar{\nu}$ [35–38]. This motivates introducing a signal-strength parameter $\mu_{\nu \bar{\nu}}$, defined through $d\mathcal{B}_{\nu \bar{\nu}}(\mu)/dq^2 \equiv \mu_{\nu \bar{\nu}} d\mathcal{B}_{\nu \bar{\nu}, \text{SM}}/dq^2$. Allowing $\mu_{\nu \bar{\nu}}$ to float ensures that possible short-distance new-physics effects in $B \rightarrow K \nu \bar{\nu}$ are not misinterpreted as, or artificially absorbed by, the two-body decay $B \rightarrow Ka$. This separation is essential: It renders the analysis sensitive to *both* types of phenomena and establishes $B^+ \rightarrow K^+ \nu \bar{\nu}$ as a dual probe—simultaneously testing short-distance modifications of the $b \rightarrow s \nu \bar{\nu}$ amplitude and the production of light invisible states a . As we will demonstrate, the two effects operate independently to an excellent approximation and this represents our third key result.

To implement this reconstruction explicitly, we recall that within the ITA, Belle II cannot directly reconstruct $q^2 \equiv$

$m_{\nu \bar{\nu}}^2 = (p-k)^2$ and thus constructs a proxy variable, q_{rec}^2 [32], in terms of the energy of the final-state kaon in the $B\bar{B}$ rest frame (denoted by a star) and of the known $s \equiv (p_{e^-} + p_{e^+})^2 = 4E_B^{*2}$. In this frame, the signal B -meson momentum has a fixed magnitude, $|\vec{p}_B^*| = 0.3298$ GeV, and is distributed isotropically. In the B rest frame (double star), the kaon from the two-body decay $B \rightarrow Ka$ would likewise be monochromatic for fixed $q^2 = m_a^2$; however, its observed momentum depends on the relative orientation between the two frames, as discussed next. One can write

$$\begin{aligned} q_{\text{rec}}^2 &\equiv \frac{s}{4} + m_K^2 - \sqrt{s} E_K^* \\ &= q^2 + E_B^{*2} - m_B^2 - 2\vec{p}_B^* \cdot \vec{p}_K^*(q^2) \\ &= q^2 + E_B^{*2} - m_B^2 - 2|\vec{p}_B^*| \gamma^{**} \\ &\quad \times (\beta^{**} E_K^{**}(q^2) + |\vec{p}_K^{**}(q^2)| \cos \theta^{**}), \end{aligned} \quad (4)$$

where $|\vec{p}_K^{**}| = \lambda^{1/2}(m_B^2, m_K^2, q^2)/(2m_B)$ and $\beta^{**} = |\vec{p}_B^*|/E_B^* = 0.06234$ is the boost parameter in the Lorentz transformation from the B rest frame to the $B\bar{B}$ rest frame. Finally, θ^{**} is the polar angle of the kaon momentum in the B -meson rest frame relative to the B -meson momentum in the $B\bar{B}$ rest frame. Importantly, θ^{**} constitutes the only unmeasured quantity in Eq. (4). The distribution of q_{rec}^2 for a given q^2 can therefore be derived from Eq. (4) with $c^{**} \equiv \cos \theta^{**} \in [-1, 1]$ uniformly distributed.

Equation (4) makes explicit the aforementioned dependence on the relative orientation between the $*$ and $**$ frames, showing that its lack of knowledge is structural in the ITA: While the kaon produced in $B \rightarrow Ka$ is monochromatic in the B -meson rest frame, reconstructing this energy (E_K^* in the q_{rec}^2 defining equation) requires knowledge of the full decay kinematics. This includes the angle θ^{**} , which in the ITA is not accessible. The monochromatic kaon energy is thus mapped onto a broadened distribution in q_{rec}^2 , whose shape is fixed almost entirely by kinematics: The sub-percent-level uncertainties [39] smearing the energies and momenta that appear in Eq. (4) have a completely negligible impact on the $q^2(q_{\text{rec}}^2)$ distribution.

The physical branching ratios and the observed yields are then related in the ITA as

$$\begin{aligned} \frac{dN_X}{dq_{\text{rec}}^2} &= N_B \int dq^2 \int dc^{**} \frac{1}{2} \delta(q_{\text{rec}}^2 - q_{\text{rec}}^2(q^2, c^{**})) \\ &\quad \times \epsilon_{\text{ITA}}(q^2) \frac{d\mathcal{B}_X}{dq^2} \\ &\equiv N_B \int dq^2 f_{q_{\text{rec}}^2}(q^2) \epsilon_{\text{ITA}}(q^2) \frac{d\mathcal{B}_X}{dq^2}, \end{aligned} \quad (5)$$

where $N_B = 387(6) \times 10^6$ [32,40] is the number of B candidates in the given decay channel X , $\mathcal{B}$ the corresponding BR, and $\epsilon_{\text{ITA}}(q^2)$ the ITA-specific selection efficiency as a function of q^2 , which can be obtained from Ref. [32]. The function $q_{\text{rec}}^2(q^2, c^{**})$ appearing in the delta function is the right-hand side (r.h.s.) of Eq. (4).

In the last equality of Eq. (5), we make contact with the smearing function $f_{q_{\text{rec}}^2}$ [31,41,42], which is typically assumed to require Belle II simulation. As discussed, this function is instead simply fixed by Eq. (4), shown in Fig. 1. This correlation

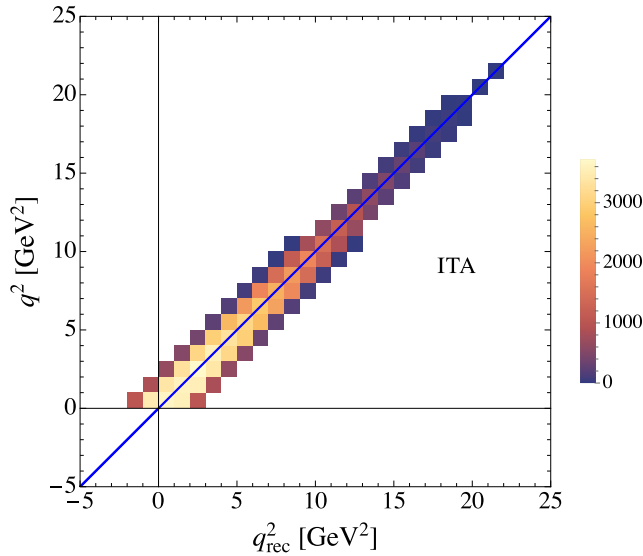


FIG. 1. Correlation between the true (q^2) and the reconstructed (q^2_{rec}) dineutrino invariant mass squared. The plot is normalized to 10^5 observed events.

plot may be compared with Fig. 5.1 of Ref. [43], which is the output of internal Belle II simulations. The two figures display good agreement, aside from an inconsequential low- q^2 tail visible in the figure of Ref. [43]. This tail originates from radiative decays with undetected photons. Strictly speaking, $B \rightarrow K \nu \bar{\nu} (+n\gamma)$ constitutes a distinct physical process from the nonradiative mode and would need to be subtracted if its contribution were sizable, as is the case for $B_s \rightarrow \mu^+ \mu^-$ [44]. In the present case, however, this effect is negligible—well below the percent level in Refs. [43,45]. Inclusion of the radiative tail in the referenced figure is therefore immaterial to the above comparison.

Equation (4) and the ensuing discussion put to work an important distinction between the SM decay $B^+ \rightarrow K^+ \nu \bar{\nu}$ and the exotic two-body decay $B^+ \rightarrow K^+ a$: The latter corresponds to a fixed invariant mass $q^2 = m_a^2$, whereas the former exhibits a continuous q^2 spectrum. This difference is fully exploited in our analysis: It is encoded at the level of the differential branching fractions, which enter as a smooth distribution in q^2 for the dineutrino mode and as a Dirac delta function for the axion signal. The apparent loss of monochromatic information in the ITA does not reflect a limitation of principle, but arises solely from the fact that the experimentally accessible variable is q^2_{rec} rather than the true q^2 .

In contrast, within the HTA, to which we turn next, Belle II can reconstruct the full event kinematics, allowing direct access to q^2 . In this case, the axion signal appears as a narrow peak in the q^2 spectrum, equivalent to a monochromatic kaon energy in the B -meson rest frame, and the analysis reduces to a standard “bump search.” The physical BRs and observed yields are simply related as

$$\frac{dN_X}{dq^2} = N_B \epsilon_{\text{HTA}}(q^2) \frac{dB_X}{dq^2}, \quad (6)$$

where $\epsilon_{\text{HTA}}(q^2)$ is the HTA-specific selection efficiency as a function of q^2 , also available in Ref. [32].

We conclude by emphasizing that, within our approach, the two tagging strategies exploit the same underlying kinematic information. The observed differences arise solely from degrees of freedom that are not experimentally accessible in the ITA, rather than from any discarded physical information. In particular, the distinction between two-body (signal) and three-body ($B \rightarrow K \nu \bar{\nu}$) kinematics is fully incorporated in both tagging strategies.

II. ANALYSIS

To probe the $B^+ \rightarrow K^+ a$ decay through the $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurements at Belle II, we construct a likelihood ratio test statistic based solely on the public data of Fig. 17 (ITA) and Fig. 20 (HTA) of Ref. [32]. The likelihood function $\mathcal{L}$ is constructed as the product of two contributions, i.e., $\mathcal{L} = \mathcal{L}_1 \mathcal{L}_2$.

The first and main term in the likelihood function represents a Poisson distribution P for each bin, labeled by an index i , assuming the bins to be uncorrelated: $\mathcal{L}_1 = \prod_i^{n_{\text{bin}}} P(n_{\text{obs}}^i; \lambda^i)$. Here, n_{obs}^i is the number of observed events in bin i , defined by $q^2 \in [q_i^2, q_{i+1}^2)$, and λ^i is the corresponding expected number of events, which depends on $\mathcal{B}(B^+ \rightarrow K^+ a) \equiv \mathcal{B}_a$ and on the signal strength $\mu_{\nu \bar{\nu}}$ —the two key quantities to be determined:

$$\lambda^i(\mathcal{B}_a, \mu_{\nu \bar{\nu}}) = \mathcal{B}_a n_a^i + \mu_{\nu \bar{\nu}} n_{\nu \bar{\nu}}^i + n_{\text{bkg}}^i. \quad (7)$$

This expression applies to either the ITA or HTA dataset, where n_a^i , $n_{\nu \bar{\nu}}^i$, and n_{bkg}^i are the yields in the i th bin for the signal, the $B^+ \rightarrow K^+ \nu \bar{\nu}$ decay, and the background, respectively. Within the two methods, the first two components on the r.h.s. of Eq. (7) are defined as

$$n_{a,\text{ITA}} = N_B \int_{i\text{th bin}} dq_{\text{rec}}^2 f_{q_{\text{rec}}}^2(m_a^2) \epsilon_{\text{ITA}}(m_a^2), \quad (8)$$

$$n_{\nu \bar{\nu},\text{ITA}} = N_B \int_{i\text{th bin}} dq_{\text{rec}}^2 \int dq^2 f_{q_{\text{rec}}}^2(q^2) \epsilon_{\text{ITA}}(q^2) \frac{dB_{\nu \bar{\nu},\text{SM}}}{dq^2}, \quad (9)$$

$$n_{a,\text{HTA}} = N_B \int_{i\text{th bin}} dq^2 \epsilon_{\text{HTA}}(m_a^2) g_a(m_a^2), \quad (10)$$

$$n_{\nu \bar{\nu},\text{HTA}} = N_B \int_{i\text{th bin}} dq^2 \epsilon_{\text{HTA}}(q^2) \frac{dB_{\nu \bar{\nu},\text{SM}}}{dq^2}. \quad (11)$$

These expressions can be understood directly from Eqs. (5) and (6), bearing in mind that, for the signal, the differential branching ratio is proportional to a Dirac delta function with support at $q^2 = m_a^2$, and that the overall normalization $\mathcal{B}_a$ is factored out in Eq. (7).

In addition, for the HTA signal contribution one must account for the finite invariant-mass resolution of the experiment. We implement this by replacing $\delta(q^2 - m_a^2)$ with a normal distribution $g_a(q^2)$ centered at m_a^2 , with variance taken as $\sim (450 \text{ MeV})^2$, comparable to the beam-constrained mass resolution of B mesons reconstructed with the hadronic tag [46].

This prescription also allows us to comment more generally on detector-induced smearing effects. In principle, the same smearing procedure should be applied to the dineutrino contribution in the HTA [Eq. (11)]. While in the axion case this treatment is essential, in the dineutrino case the impact is much milder because the q^2 spectrum of $B \rightarrow K \nu \bar{\nu}$ is already smooth. Numerically, the smeared and unsmeared

TABLE I. Upper limits on the axion BR and lower limits on the coupling-rescaled Peccei-Quinn scale at 90% C.L. for the different cases considered. All limits assume $m_a = 0$. See text for more details on the significant digits quoted for the different tagging strategies.

	ITA	HTA	Combined
$\mu_{\nu\bar{\nu}} = 1$	$n_{\text{dof}} = 2$ $\mathcal{B}_a < 1.1 \times 10^{-6}$ $ (F_V)_{sb} > 8.4 \times 10^8 \text{ GeV}$	$n_{\text{dof}} = 2$ $\mathcal{B}_a < 5 \times 10^{-6}$ $ (F_V)_{sb} > 4 \times 10^8 \text{ GeV}$	$n_{\text{dof}} = 3$ $\mathcal{B}_a < 1.4 \times 10^{-6}$ $ (F_V)_{sb} > 7.6 \times 10^8 \text{ GeV}$
$\mu_{\nu\bar{\nu}} \neq 1$	$n_{\text{dof}} = 3$ $\mathcal{B}_a < 1.0 \times 10^{-6}$ $ (F_V)_{sb} > 8.9 \times 10^8 \text{ GeV}$	$n_{\text{dof}} = 3$ $\mathcal{B}_a < 6 \times 10^{-6}$ $ (F_V)_{sb} > 4 \times 10^8 \text{ GeV}$	$n_{\text{dof}} = 4$ $\mathcal{B}_a < 1.2 \times 10^{-6}$ $ (F_V)_{sb} > 8.0 \times 10^8 \text{ GeV}$

distributions differ only marginally [47] and the resulting likelihood and axion limits are unchanged within the two-digit precision of our results (see Table I).

In the ITA, the situation is fundamentally different. The axion signal is already broadened at the kinematic level: The monochromatic kaon energy in the B rest frame (or, equivalently, the Dirac-delta structure of the axion q^2 distribution) is mapped onto a smooth distribution in the experimentally accessible variable q_{rec}^2 . The finite experimental resolution on E_K^* therefore introduces only a subleading additional smearing, whose effect we have verified to be numerically negligible in practice.

We next address the role of uncertainties on the selection-efficiency factors ϵ appearing in Eqs. (8)–(11). The statistical uncertainties on the efficiencies are at the percent or subpercent level for both channels, as reported in Ref. [32]. The systematic components are, in turn, approximately 3% in ITA and 16% in HTA (see Tables I and II of Ref. [32]). Since the expected signal yields depend linearly on the efficiencies, these uncertainties propagate directly to the extracted BR limits quoted in Table I. Because the bounds are largely driven by the ITA sample, the corresponding variation in the ITA and combined results would affect only the third significant digit, whereas the HTA-only limits would be impacted already at the second digit. This is reflected in the precision with which we report the results for the two tagging strategies. In practice, varying the efficiencies within their quoted uncertainties does not modify the bounds at the level of precision displayed in Table I.

We now pause on the treatment of systematic uncertainties affecting the last component of Eq. (7), namely, the background normalization. In the Belle II analysis, the impact of a given systematic uncertainty is quantified by the reduction in the total uncertainty when that source is removed from the fit, as shown in Tables I and II of Ref. [32]. While the individual components cannot be summed in quadrature to reconstruct the total uncertainty, their relative magnitudes provide a useful estimate of their importance. From this comparison, the combined effect of the systematic uncertainties other than the background normalization is at the level of $O(50\%)$ or less relative to the background normalization contribution.

To model this component, we reason as follows. Although the uncertainties on the Monte Carlo generated background events are not explicitly provided, we infer that the relative uncertainty on the total normalization is, respectively, $\sigma_{\text{ITA}}^{\text{bkg}} \approx 1\%$ and $\sigma_{\text{HTA}}^{\text{bkg}} \approx 10\%$. These estimates have been obtained from the q^2 distribution and its corresponding pull shown

in Figs. 17 and 20 of Ref. [32]. One can identify the pull as $|n_{\text{obs}}^i - n_{\text{hst}}^i| / \sqrt{n_{\text{obs}}^i + (\sigma_{\text{hst}}^i)^2}$, where the subscripts “obs” and “hst” denote, respectively, the data and the value of the stacked histograms, and σ_{obs}^i the uncertainty in the data point. Inverting the pull distribution in terms of σ_{hst}^i , we find that the relative uncertainties $\sigma_{\text{hst}}^i / n_{\text{hst}}^i$ are approximately constant across bins. We therefore treat them as a single relative uncertainty σ^{bkg} on the overall background normalization, modeled as a nuisance parameter θ in $n_{\text{bkg}}^i = (1 + \sigma^{\text{bkg}} \theta) \langle n_{\text{bkg}}^i \rangle$. This can be constrained, independently in the ITA and HTA analyses, by the Gaussian likelihood terms

$$\mathcal{L}_2 = \frac{1}{\sqrt{2\pi}} e^{-\theta_{\text{ITA}}^2/2} \times \frac{1}{\sqrt{2\pi}} e^{-\theta_{\text{HTA}}^2/2}, \quad (12)$$

where the expected value $\langle n_{\text{bkg}}^i \rangle$ of the number of background events for each bin, n_{bkg}^i , can be read out from Ref. [32].

A strong test that the above procedure captures the bulk of the systematic uncertainties can be performed by studying our profile likelihood $\mathcal{L}$ as a function of the signal strength $\mu_{\nu\bar{\nu}}$ in the case where the axion is absent, i.e., where $\mathcal{B}_a = 0$, thus making contact with the counterpart in Ref. [32]. The result is shown in Fig. 4 in the Appendix, whose two panels show the case of neglecting versus including the systematic uncertainty on the background normalization. The figure shows that we obtain results well consistent with the collaboration’s counterparts only when the background normalization uncertainty is included. This indicates that our procedure captures the leading systematic effect. As a further test that the remaining systematics is negligible, we also calculated the same profile likelihood, with uncertainty inflated by a factor of 1.5, as if we were missing a residual $O(50\%)$ component. The resulting profile is essentially identical to the case with noninflated uncertainty.

III. RESULTS

For either the ITA or the HTA method, we then minimize $\chi^2(\mathcal{B}_a, \mu_{\nu\bar{\nu}}, \theta) \equiv -2 \ln \mathcal{L}$ and require that $\Delta\chi^2 \equiv \chi^2 - \chi_{\text{min}}^2 \leq \chi^2(90\%, n_{\text{dof}})$ [30]. This yields a limit on $\mathcal{B}_a$ as a function of m_a , with the number of degrees of freedom n_{dof} determined according to two alternative hypotheses. Specifically, we may either fix the dineutrino contribution to its SM value, $\mu_{\nu\bar{\nu}} = 1$, or allow it to float, which increases n_{dof} by 1. We then evaluate bounds on $\mathcal{B}_a$ using ITA or HTA individually, or their combination. Since the ITA and HTA datasets are uncorrelated to a very good approximation, the combined

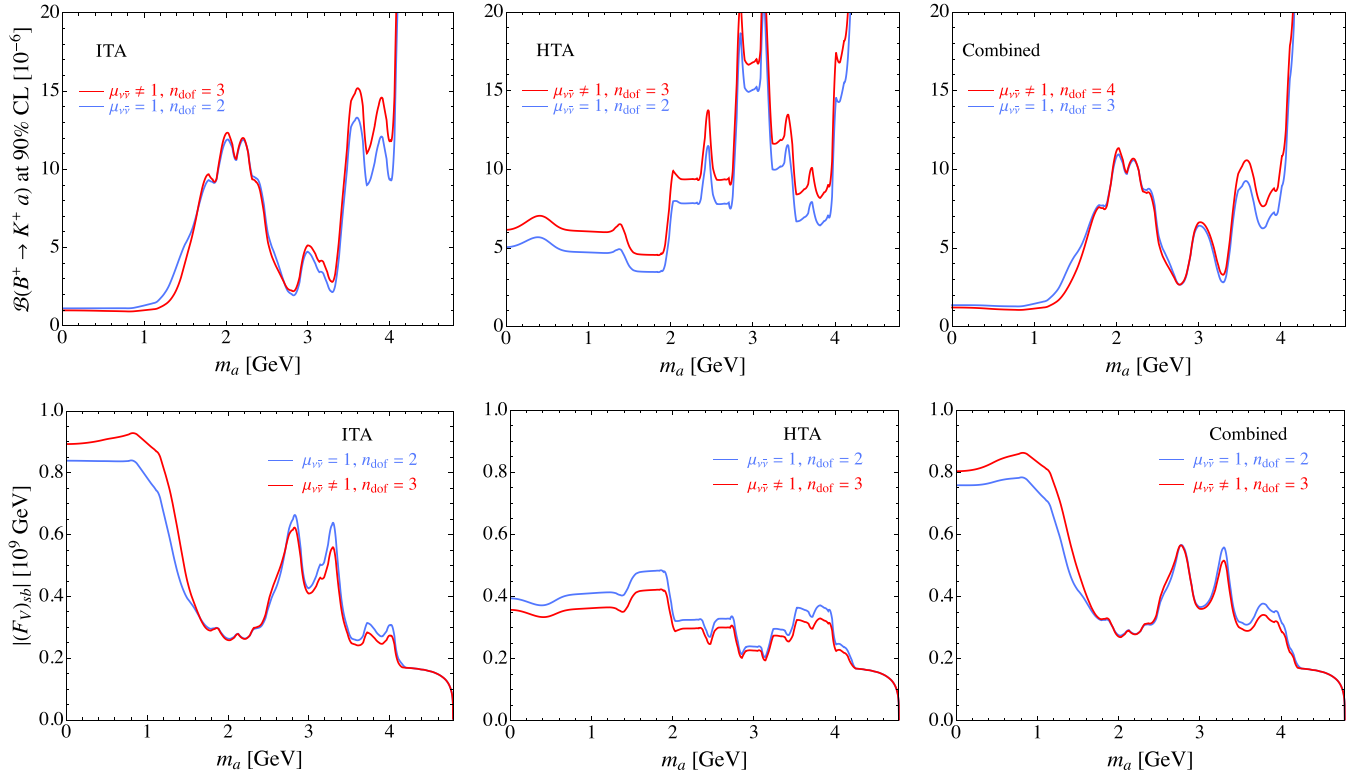


FIG. 2. Upper limit on $\mathcal{B}_a$ (upper row) and lower limit on the coupling-rescaled Peccei-Quinn scale $(F_V)_{sb}$ (lower row) at 90% C.L. as a function of the mass m_a , and with $\mu_{\nu\bar{\nu}}$ set to unity (blue line) or left floating (red).

analysis corresponds to multiplying the individual likelihoods. In this case, we minimize $\chi^2(\mathcal{B}_a, \mu_{\nu\bar{\nu}}, \theta_{\text{ITA}}, \theta_{\text{HTA}})$.

The different cases described and the resulting bounds are summarized in Table I, which assumes $m_a = 0$ [48]. Results for a general value of the axion mass are shown in Fig. 2. This figure quantitatively conveys two important pieces of information. First, as one would expect, the bound is weakened by an order of magnitude in the $m_{\nu\bar{\nu}}$ bins where data and the SM $B^+ \rightarrow K^+ \nu \bar{\nu}$ prediction are in tension. These discrepant bins can actually be used to set both an upper and a lower bound on $\mathcal{B}_a$ [31,42,49,50], as shown in Fig. 5 in the Appendix.

Second, and less obviously, the bound depends only marginally on the assumption of new physics in $\mathcal{B}_{\nu\bar{\nu}}$ —namely, on whether $\mu_{\nu\bar{\nu}}$ is set to unity or left floating. This is an important conclusion, making the $B^+ \rightarrow K^+ \nu \bar{\nu}$ decay a *double* probe: of new short-distance physics in the $B^+ \rightarrow K^+ \nu \bar{\nu}$ amplitude, testable by $\mu_{\nu\bar{\nu}} \neq 1$, and of new light, elusive particles, produced as $B \rightarrow Ka$, with a BR sensitivity proportional to the total $\mathcal{B}_{\nu\bar{\nu}}$ uncertainty [51]. This second point is also conveyed by the similarity between the bounds in the first and second rows of Table I. In the case where $\mathcal{B}_{\nu\bar{\nu}}$ is fixed to its SM value, we verify that varying it by as much as 10% has a negligible impact on these results. To corroborate this conclusion even further, we evaluate the bound on $\mathcal{B}_a$ assuming fixed values of $\mu_{\nu\bar{\nu}}$ but different from unity. As shown in Fig. 6 of the Appendix, the resulting bound has only a mild dependence on the specific value of the dineutrino signal strength—even for huge beyond-SM contributions to the $B^+ \rightarrow K^+ \nu \bar{\nu}$ amplitude. The BR limit in Fig. 2 (upper row) can be translated into a bound on the coupling-rescaled Peccei-Quinn scale $(F_V)_{sb} \equiv$

$2f_a/(k_V)_{sb}$ [52], to obtain a bound on this scale. This bound is shown in Fig. 2 (lower row), as a function of m_a for the two alternative $\mu_{\nu\bar{\nu}}$ possibilities. Limits for $m_a \simeq 0$ are collected in Table I. As the table shows, we obtain $|(F_V)_{sb}| \geq 0.9 \times 10^9$ GeV for the QCD axion, which improves by one order of magnitude the preexisting bound [52].

We now close the circle with the discussion in the Introduction concerning the distinction between the theoretically motivated but experimentally challenging QCD axion and generic ALPs, whose appeal lies mostly in their experimental accessibility. With our approach (based on a fully reproducible analysis strategy), we obtain bounds on the QCD axion—despite the fact that the QCD axion is generally much more challenging than ALPs at colliders, since its mass scale lies far below any collider “scales,” as clarified below.

More quantitatively, the axion mass is related to the $(F_V)_{sb}$ scale as

$$m_a = \frac{\sqrt{2}F_\pi m_\pi}{(F_V)_{sb}(k_V)_{sb}} \frac{\sqrt{m_u m_d}}{m_u + m_d}, \quad (13)$$

where $F_\pi = 93\sqrt{2}$ MeV. Our analysis yields lower bounds on $(F_V)_{sb}$, with the weakest at about 4×10^8 GeV and the strongest close to 9×10^8 GeV, depending on the treatment of the dineutrino signal strength and on the tagging strategy (hadronic or inclusive) (see Table I). One can therefore study the corresponding range of m_a values as a function of $|(k_V)_{sb}| \lesssim 10$, as imposed by perturbativity. The result is shown in Fig. 3.

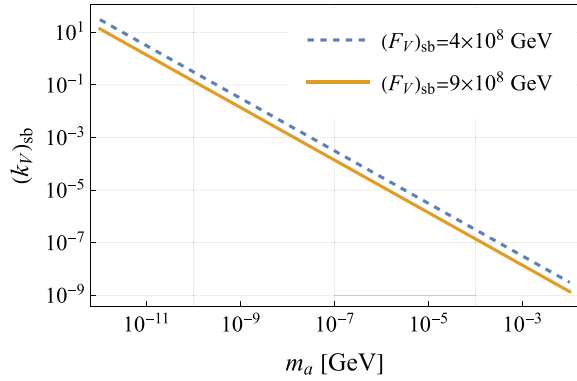


FIG. 3. QCD axion mass m_a as a function of the effective flavor-violating coupling $(k_V)_{sb}$, shown for two representative lower bounds on $(F_V)_{sb}$ derived in this work. The QCD axion lies effectively at $m_a \simeq 0$ on collider scales, so the resulting limits correspond to vertical constraints in axion mass vs coupling parameter space.

The figure shows that, over many orders of magnitude in the axion mass, the QCD axion is effectively located at $m_a \simeq 0$ on collider scales—meaning, collectively, the physical effective scales of the processes studied as well as the detection-induced resolution scales. Consequently, our limits act as essentially *vertical* constraints in axion parameter spaces of mass versus coupling, constraining the effective flavor-violating coupling $(k_V)_{sb}$ independently of m_a . This representation clarifies how our results should be interpreted within the broader axion-search landscape, which includes experiments that *are* sensitive to the axion mass range in our Fig. 3: While collider experiments are insensitive to the tiny QCD axion mass itself, they can nevertheless impose meaningful and constraints on its flavor structure.

In conclusion, we reinterpret the $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement as a probe of the two-body decay $B^+ \rightarrow K^+ a$, with a an axion ($m_a \simeq 0$) or a generic ALP ($m_a > 0$). A central ingredient of this work is the analytic reconstruction of the mapping between the measured q_{rec}^2 and the true q^2 , allowing the Inclusive Tag Analysis to be fully used without relying on internal simulations. Combined with publicly available efficiencies, this yields the strongest existing limit on the Peccei-Quinn scale rescaled by the fundamental vector flavor-

changing coupling relevant to our channel. A key outcome is that this bound is largely insensitive to assumptions about the dineutrino signal strength $\mu_{\nu\bar{\nu}}$. As a result, the decay $B^+ \rightarrow K^+ \nu \bar{\nu}$ acts as a dual probe: It simultaneously constrains short-distance new physics in the $b \rightarrow s \nu \bar{\nu}$ amplitude and the production of light invisible states such as the axion—with the two effects remaining statistically independent to excellent accuracy. The strategy developed here can be applied directly to $B^+ \rightarrow K^{*+} \nu \bar{\nu}$ and, more generally, to other Belle II channels as soon as the corresponding measurements become available, providing a model-independent framework for public-data reinterpretation in rare decays.

ACKNOWLEDGMENTS

We thank Joel C. Swallow for useful discussions. This work has received funding from the French ANR, under Contract No. ANR-19-CE31-0016 (“GammaRare”) and No. ANR-23-CE31-0018 (“InvISYble”), that we gratefully acknowledge.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request.

APPENDIX: SUPPLEMENTARY PLOTS

In this part, we collect additional plots that are not necessary to understand the results in the main text, but further corroborate them, or can be used to perform comparisons.

Figure 4 shows the log-likelihood profiles resulting from our analyses and their comparison with the Belle II counterparts. As the figure indicates, we obtain results very close to those of the collaboration, except in the region $\mu_{\nu\bar{\nu}} < 1$ for the HTA case. This small deviation is understood: It originates from our treatment of the background uncertainty. When the background is fixed to its central value with no associated error, the minimum coincides with the Belle II result; including the background error, however, broadens the minimum and slightly lowers the best-fit point.

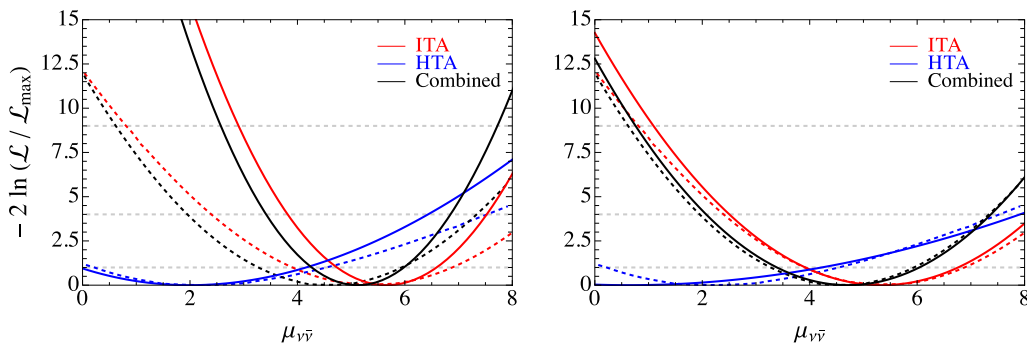


FIG. 4. Twice the negative profile log-likelihood ratio as a function of $\mu_{\nu\bar{\nu}}$ for the ITA (red), HTA (blue), and combined (black) analyses. Solid lines correspond to our results, while dashed lines reproduce the Belle II results shown in Fig. 16 of Ref. [32]. The horizontal gray dashed lines indicate the 1, 2, 3σ χ^2 thresholds for one degree of freedom. (Left) The number of background events in each bin is fixed to its expected value (n_{bkg}^i). (Right) The background normalization is floated as described in the text.

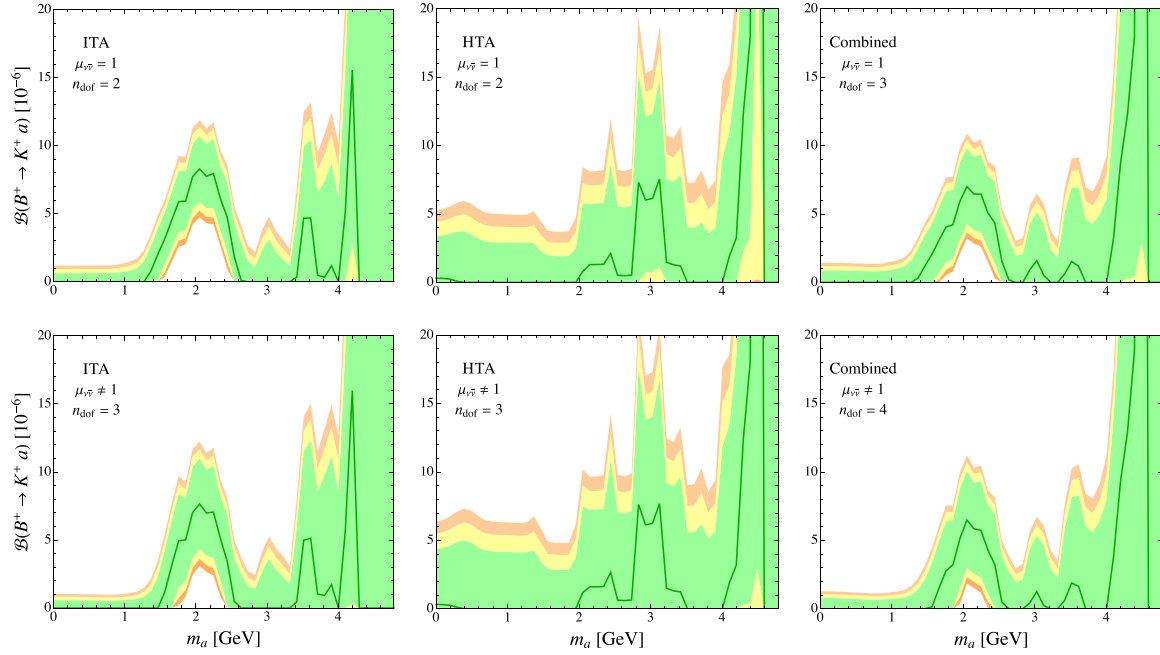


FIG. 5. Regions $\mathcal{B}_a$ at 68% (green), 95% (yellow), and 99% (orange) C.L. and best-fit value (solid green line) of $\mathcal{B}_a$ as a function of the mass with $\mu_{\nu\bar{\nu}}$ fixed to 1 (upper row) or left floating (lower row).

Figure 5 shows $\mathcal{B}_a$ as a signal, displaying the 1σ , 2σ , and 3σ regions for $\mathcal{B}_a$ as a function of the mass, with $\mu_{\nu\bar{\nu}}$ either fixed to unity or left floating.

In Fig. 6, we evaluate the bound on $\mathcal{B}_a$ for fixed values of $\mu_{\nu\bar{\nu}}$ different from unity, illustrating the mild dependence of the resulting bound on the assumed value of the dineutrino signal strength.

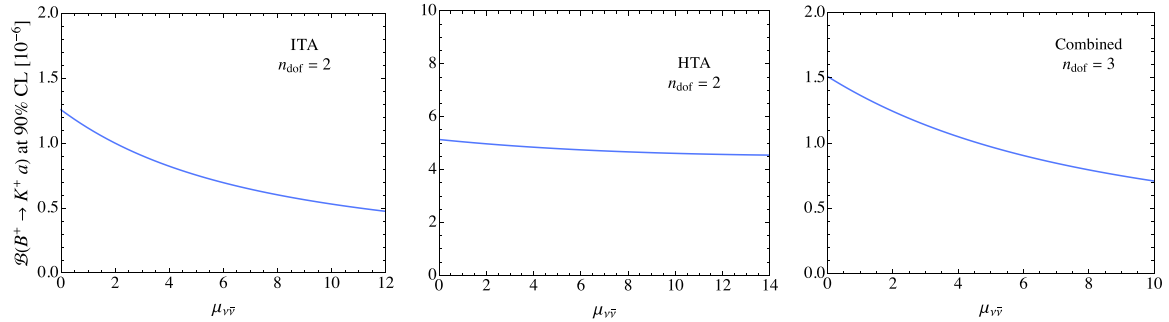


FIG. 6. Upper limit at 90% C.L. on $\mathcal{B}_a$ for fixed values of $\mu_{\nu\bar{\nu}}$.

- [1] I. G. Irastorza and J. Redondo, New experimental approaches in the search for axion-like particles, *Prog. Part. Nucl. Phys.* **102**, 89 (2018).
- [2] G. Lanfranchi, M. Pospelov, and P. Schuster, The search for feebly interacting particles, *Annu. Rev. Nucl. Part. Sci.* **71**, 279 (2021).
- [3] P. Sikivie, Invisible axion search methods, *Rev. Mod. Phys.* **93**, 015004 (2021).
- [4] G. G. Raffelt, Astrophysical methods to constrain axions and other novel particle phenomena, *Phys. Rep.* **198**, 1 (1990).
- [5] G. G. Raffelt, Astrophysical axion bounds, *Lect. Notes Phys.* **741**, 51 (2008).
- [6] A. Caputo and G. Raffelt, Astrophysical axion bounds: The 2024 edition, in 1st Training School of the COST Action COSMIC WISPerS (CA21106) (2024), *PoS (COSMICWISPerS)*, 041 (2024).
- [7] D. J. E. Marsh, Axion cosmology, *Phys. Rep.* **643**, 1 (2016).
- [8] J. E. Kim and G. Carosi, Axions and the strong CP problem, *Rev. Mod. Phys.* **82**, 557 (2010); **91**, 049902(E) (2019).
- [9] L. Di Luzio, M. Giannotti, E. Nardi, and L. Visinelli, The landscape of QCD axion models, *Phys. Rep.* **870**, 1 (2020).
- [10] K. Choi, S. H. Im, and C. Sub Shin, Recent progress in the physics of axions and axion-like particles, *Annu. Rev. Nucl. Part. Sci.* **71**, 225 (2021).
- [11] J. E. Kim, Weak interaction singlet and strong CP invariance, *Phys. Rev. Lett.* **43**, 103 (1979).
- [12] M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, Can confinement ensure natural CP invariance of strong interactions? *Nucl. Phys. B* **166**, 493 (1980).

- [13] M. Dine, W. Fischler, and M. Srednicki, A simple solution to the strong CP problem with a harmless axion, *Phys. Lett. B* **104**, 199 (1981).
- [14] A. R. Zhitnitsky, On possible suppression of the axion hadron interactions (in Russian), *Sov. J. Nucl. Phys.* **31**, 260 (1980).
- [15] R. D. Peccei and H. R. Quinn, CP conservation in the presence of instantons, *Phys. Rev. Lett.* **38**, 1440 (1977).
- [16] R. D. Peccei and H. R. Quinn, Constraints imposed by CP conservation in the presence of instantons, *Phys. Rev. D* **16**, 1791 (1977).
- [17] S. Weinberg, A new light boson? *Phys. Rev. Lett.* **40**, 223 (1978).
- [18] F. Wilczek, Problem of strong P and T invariance in the presence of instantons, *Phys. Rev. Lett.* **40**, 279 (1978).
- [19] L. Ubaldi, Effects of theta on the deuteron binding energy and the triple-alpha process, *Phys. Rev. D* **81**, 025011 (2010).
- [20] M. Dine, L. Stephenson Haskins, L. Ubaldi, and D. Xu, Some remarks on anthropic approaches to the strong CP problem, *J. High Energy Phys.* **05** (2018) 171.
- [21] J. Preskill, M. B. Wise, and F. Wilczek, Cosmology of the invisible axion, *Phys. Lett. B* **120**, 127 (1983).
- [22] L. F. Abbott and P. Sikivie, A cosmological bound on the invisible axion, *Phys. Lett. B* **120**, 133 (1983).
- [23] M. Dine and W. Fischler, The not so harmless axion, *Phys. Lett. B* **120**, 137 (1983).
- [24] T. Abe *et al.* (Belle-II), Belle II technical design report, [arXiv:1011.0352](https://arxiv.org/abs/1011.0352).
- [25] The spread, or bite, around these momenta is tiny and completely negligible for the purposes of this paper.
- [26] The charge-conjugate mode is assumed to be included throughout.
- [27] H. Georgi, D. B. Kaplan, and L. Randall, Manifesting the invisible axion at low-energies, *Phys. Lett. B* **169**, 73 (1986).
- [28] The sign difference in the f_a -defining $aG\tilde{G}$ term in Eq. (1) with respect to Ref. [9] arises from our use of opposite ϵ -tensor conventions.
- [29] N. Gubernari, M. Reboud, D. van Dyk, and J. Virto, Dispersive analysis of $B \rightarrow K^{(*)}$ and $B_s \rightarrow \phi$ form factors, *J. High Energy Phys.* **12** (2023) 153; **01** (2025) 125.
- [30] S. Navas *et al.* (Particle Data Group), Review of particle physics, *Phys. Rev. D* **110**, 030001 (2024).
- [31] P. D. Bolton, S. Fajfer, J. F. Kamenik, and M. Novoa-Brunet, Signatures of light new particles in $B \rightarrow K^{(*)}E_{\text{miss}}$, *Phys. Rev. D* **110**, 055001 (2024); **111**, 039903(E) (2025).
- [32] I. Adachi *et al.* (Belle-II), Evidence for $B^+ \rightarrow K^+ \nu \bar{\nu}$ decays, *Phys. Rev. D* **109**, 112006 (2024).
- [33] M. T. Prim *et al.* (Belle), Search for $B^+ \rightarrow \mu^+ \nu_\mu$ and $B^+ \rightarrow \mu^+ N$ with inclusive tagging, *Phys. Rev. D* **101**, 032007 (2020).
- [34] F. Abudinén *et al.* (Belle-II), Search for $B^+ \rightarrow K^+ \nu \bar{\nu}$ decays using an inclusive tagging method at Belle II, *Phys. Rev. Lett.* **127**, 181802 (2021).
- [35] W. G. Parrott, C. Bouchard, and C. T. H. Davies (HPQCD), Standard Model predictions for $B \rightarrow K\ell + \ell^-$, $B \rightarrow K\ell 1 - \ell 2+$ and $B \rightarrow K\nu \bar{\nu}$ using form factors from $N_f = 2 + 1 + 1$ lattice QCD, *Phys. Rev. D* **107**, 014511 (2023); **107**, 119903(E) (2023).
- [36] D. Bečirević, G. Piazza, and O. Sumensari, Revisiting $B \rightarrow K^{(*)}\nu \bar{\nu}$ decays in the Standard Model and beyond, *Eur. Phys. J. C* **83**, 252 (2023).
- [37] A. J. Buras, J. Girrbach-Noe, C. Niehoff, and D. M. Straub, $B \rightarrow K^{(*)}\nu \bar{\nu}$ decays in the Standard Model and beyond, *J. High Energy Phys.* **02** (2015) 184.
- [38] J. Brod, M. Gorbahn, and E. Stamou, Two-loop electroweak corrections for the $K \rightarrow \pi \nu \bar{\nu}$ decays, *Phys. Rev. D* **83**, 034030 (2011).
- [39] V. Bertacchi *et al.* (Belle II Tracking Group), Track finding at Belle II, *Comput. Phys. Commun.* **259**, 107610 (2021).
- [40] The below-2% N_B error has a negligible impact on our analysis, because $n_{\nu \bar{\nu}, \text{bkg}}^i$ have an error approaching 10%.
- [41] K. Fridell, M. Ghosh, T. Okui, and K. Tobioka, Decoding the $B \rightarrow K\nu \nu$ excess at Belle II: Kinematics, operators, and masses, *Phys. Rev. D* **109**, 115006 (2024).
- [42] P. D. Bolton, S. Fajfer, J. F. Kamenik, and M. Novoa-Brunet, Impact of new invisible particles on $B \rightarrow K^{(*)}E_{\text{miss}}$ observables, *Phys. Rev. D* **112**, 035010 (2025).
- [43] C. Praz, Search for $B \rightarrow K\nu \nu$ decays with a machine learning method at the Belle II experiment, Ph.D. thesis, Department of Mathematics, University of Hamburg, Hamburg, 2022.
- [44] A. J. Buras, J. Girrbach, D. Guadagnoli, and G. Isidori, On the Standard Model prediction for $B(B_{s,d} \rightarrow \mu^+ \mu^-)$, *Eur. Phys. J. C* **72**, 2172 (2012).
- [45] The suppression arises because the lightest charged emitter is the K^+ , which makes the exponent b in
- $$\mathcal{B}(B \rightarrow K\nu \bar{\nu} (+n\gamma)) = \left(\frac{2\Delta E}{m_B} \right)^{\frac{a}{4\pi}b} \mathcal{B}(B \rightarrow K\nu \bar{\nu}),$$
- parametrically small. Here, ΔE denotes the minimal photon energy in the B rest frame required for detectability—a “theoretical” scale usually proxied by the $\sqrt{q^2}$ resolution.
- [46] T. Keck *et al.*, The full event interpretation: An exclusive tagging algorithm for the Belle II experiment, *Comput. Softw. Big Sci.* **3**, 6 (2019).
- [47] The dominant effect is that a few events with small positive q^2 migrate to slightly negative reconstructed values, thus shifting between adjacent bins of the Belle II spectrum.
- [48] Reference [53] obtains a bound slightly above 1×10^{-6} for the branching ratio in the low- m_a regime, based on a projected dataset of 0.5 ab^{-1} . Since Ref. [53] predates Ref. [32], a detailed numerical comparison of the limits is not straightforward. We note, however, that we achieve a very similar sensitivity—see results in Table I—using 0.362 ab^{-1} of actual data, dominated by the ITA selection.
- [49] W. Altmannshofer, A. Crivellin, H. Haigh, G. Inguglia, and J. Martin Camalich, Light new physics in $B \rightarrow K^{(*)}\nu \bar{\nu}$? *Phys. Rev. D* **109**, 075008 (2024).
- [50] J. Alda, M. Fuentes Zamoro, L. Merlo, X. Ponce Díaz, and S. Rigolin, Comprehensive ALP searches in meson decays, [arXiv:2507.19578](https://arxiv.org/abs/2507.19578).
- [51] M. Cavan-Piton, D. Guadagnoli, A. Iohner, D. Martinez Santos, and L. Vittorio, Probing QCD axions or axion-like particles in three-body K decays, *J. High Energy Phys.* **06** (2025) 039.
- [52] J. M. Camalich, M. Pospelov, P. N. H. Vuong, R. Ziegler, and J. Zupan, Quark flavor phenomenology of the QCD axion, *Phys. Rev. D* **102**, 015023 (2020).
- [53] T. Ferber, A. Filimonova, R. Schäfer, and S. Westhoff, Displaced or invisible? ALPs from B decays at Belle II, *J. High Energy Phys.* **04** (2023) 131.