# Fusing photons into diphoton resonances at Belle II and beyond

**Francesca Acanfora** [a,b,c] **Roberto Franceschini** [c] **Alessio Mastroddi** [c] **and Diego Redigolo** [d]

[a]*Institute for Theoretical Particle Physics (TTP), Karlsruhe Institute of Technology (KIT),*
*D-76131 Karlsruhe, Germany*
[b]*Galileo Galilei Institute for Theoretical Physics,*
*Largo Enrico Fermi 2, I-50125 Firenze, Italy*
[c]*Università degli Studi and INFN Roma Tre,*
*Via della Vasca Navale 84, I-00146, Rome, Italy*
[d]*INFN, Sezione di Firenze,*
*Via G. Sansone 1, 50019 Sesto Fiorentino, Italy*

*E-mail:* francesca.acanfora@uniroma3.it, roberto.franceschini@uniroma3.it,
Alessio.Mastroddi@uniroma3.it, d.redigolo@gmail.com

ABSTRACT: We propose a new search for a diphoton resonance in the $e^+e^- + \gamma\gamma$ final state at Belle II where the full final state is reconstructed in the detector. This new channel improves the expected reach compared to the $\gamma + \gamma\gamma$ channel in most of the available mass range. For simplicity we show our results in the simple parameter space of an ALP coupled solely to Standard Model photons. In addition, we show how an extension of the forward coverage of Belle II, or another similar experiment at the high intensity frontier, could improve the reach in our channel. We show that such a forward extension can be advantageous even with a loss of a factor 100 in luminosity compared to Belle II.

# Contents

## 1 Introduction

Diphoton resonances have been an extremely important final state in the history of particle physics, which culminated with the Standard Model Higgs boson discovery at ATLAS and CMS [1, 2]. Without any doubt, finding a new peak in the diphoton invariant mass distribution will give a decisive hint of new physics. For this reason, independently on the theory motivations, it is of utmost importance to think about new ways to enhance the sensitivity of diphoton resonance searches at existing colliders. This program has been pursued in the recent years in different directions and lead to a significant extension of the reach on diphoton resonances for different production channels, with a special emphasis in the relatively unexplored region of low invariant mass for the diphoton pair [3–13].

In this paper, we propose a new search for a diphoton resonance at Belle II taking advantage of the "photon fusion" channel

$$e^+ e^- \to e^+ e^- + (\gamma\gamma)_{\rm res}, \tag{1.1}$$

where $(\gamma\gamma)_{\rm res}$ indicates the diphoton resonant pair. In concrete scenarios where a new scalar (or pseudoscalar) is coupled to the SM photons, we show how this channel can improve the reach with respect to the more commonly considered "ALP-strahlung" channel [14]

$$e^+ e^- \to \gamma + (\gamma\gamma)_{\rm res}. \tag{1.2}$$

Despite the smaller production cross-section with respect to $e^+ e^- \to \gamma + (\gamma\gamma)_{\rm res}$, the $e^+ e^- \to e^+ e^- + (\gamma\gamma)_{\rm res}$ process has characteristic features that make this signal highly distinguishable from the background, which is dominated by radiative corrections to Bhabha
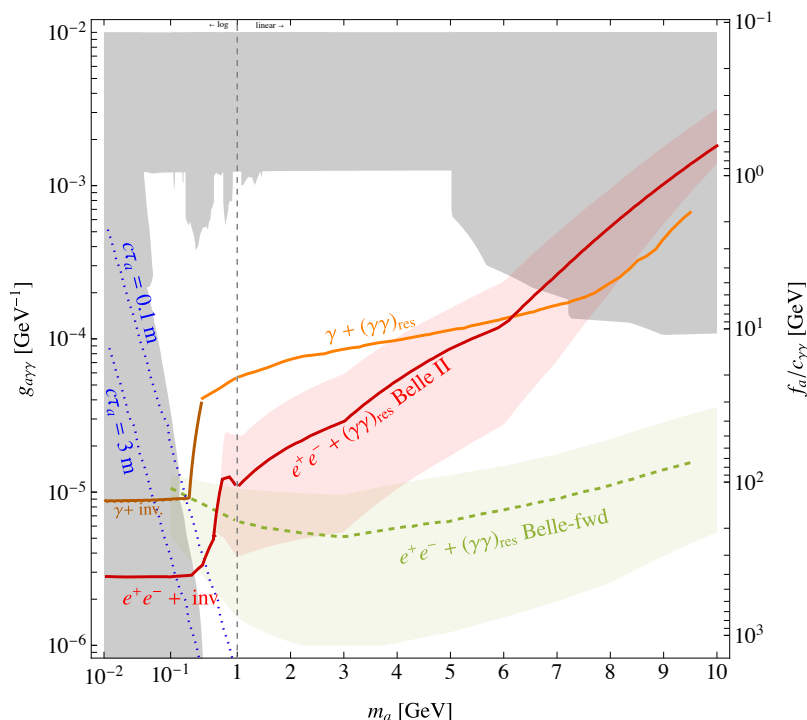
**Figure 1.** Expected sensitivity of Belle II at 95% C.L. to the ALP coupling to photons $g_{a\gamma\gamma}$ as defined in (1.3). In **red** we show the expected reach of $e^+e^- + (\gamma\gamma)_{\rm res}$ derived here and in **orange** the one of $\gamma + (\gamma\gamma)_{\rm res}$ derived in ref. [14] both assuming the present Belle II detector coverage (see eq. (1.5)) and $50~{\rm ab}^{-1}$ of integrated luminosity. The details of our analysis are given in section 2. The **green dashed** line corresponds to the reach of a hypothetical experiment with increased forward acceptance (see eq. (1.6)) and identical luminosity. The details are given in section 3. The **red error band** and the **green error band** on both expected reaches correspond to the different statistical treatment of the empty bins in the four-dimensional binned likelihood, as detailed in appendix A and appendix B. The **dotted blue** lines are isolines for the sample averaged ALP decay length measured in the lab-frame. When the ALP decays become long enough, the expected sensitivity of $e^+e^- +$ invisible [15] and of $\gamma +$ invisible [14] are also shown. The **gray shaded** region shows existing constraints, assuming $g_{a\gamma\gamma}$ is generated above the electroweak scale by an ALP coupling to the $U(1)_Y$ gauge boson (see refs. [14, 16] for the case where the ALP couples to $SU(2)$ gauge bosons). These include limits from beam dump experiments [17, 18], LEP and LHC limits recast in ref. [3], constraints from PrimEX data derived in ref. [7], the recent limits from the Belle II search on $\gamma + (\gamma\gamma)_{\rm res}$ [19], from BES III data [20] and from ATLAS [4] and CMS [12] searches on ultra-peripheral heavy ions collisions.

scattering. Ultimately, the features of the signal stem from the higher dimensional observable phase-space with respect to the "ALP-strahlung" process.

For instance, in the SM process the photons are preferably emitted at small angle with respect to the $e^\pm$ which radiated them and tend to carry little energy, while the signal allows for a large angular separation and energetic photons. Moreover, at low invariant masses the photons in the SM process tend to carry very little energy compared to the signal ones. In the following, we will elaborate further on the discrimination that can be attained depending on the invariant mass of the diphoton resonance.

For concreteness, we consider the simple model of an axion-like particle (ALP) coupled to SM photons, but we expect that our findings can be applied to a much broader class of diphoton resonances. The ALP Lagrangian that we consider is

$$\mathcal{L} = \frac{1}{2}(\partial_\mu a)^2 - \frac{m_a^2}{2}a^2 - \frac{g_{a\gamma\gamma}}{4}aF_{\mu\nu}\tilde{F}^{\mu\nu} \ ,  \tag{1.3}$$

and in figure 1 we summarize the expected reach on $g_{a\gamma\gamma}$ at Belle II with 50 ab$^{-1}$ of luminosity. The red line shows the projected sensitivity of our proposed $e^+e^- + (\gamma\gamma)_{\rm res}$ channel in the ALP parameter space compared with the one of the $e^+e^- \to \gamma + (\gamma\gamma)_{\rm res}$ channel derived in ref. [14]. The improvement of the reach in $g_{a\gamma\gamma}$ is substantial for an ALP lighter than 6 GeV promptly decaying into di-photons. For higher masses, the ALP searches at Belle II quickly become less sensitive than existing constraint from ultraperipheral heavy ion collisions at ATLAS [4] and CMS [12]. In section 2 we present the details of the analysis that lead to the sensitivity curve at Belle II of figure 1.

Given the ALP life-time

$$c\tau_a \simeq 1.5 \ \text{mm} \left[\frac{1 \ \text{GeV}}{m_a}\right]^3 \left[\frac{5\times 10^{-6} \ \text{GeV}^{-1}}{g_{a\gamma\gamma}}\right]^2 \ ,  \tag{1.4}$$

for sufficiently light ALP masses the decay length becomes macroscopic, ultimately exceeding the length of the Belle II decay volume. For this reason, the expected sensitivity in figure 1 accounts also for the possibility that the ALP decay happens so far out from the interaction point its production leaves no trace in the detectors, resulting in a $e^+e^-$ + invisible final state. We studied this signal in our earlier ref. [15]. The details of how this effect is taken into account in our simulation are given in appendix D. In the intermediate region, a dedicated search for displaced diphoton resonance could possibly further enhance the sensitivity. Here we do not distinguish prompt or displaced diphoton resonance and ignore any possible quality cut on the reconstruction of the diphoton vertex so that any displaced diphoton event is considered as prompt.[1]

As already noted in ref. [15], the Belle II geometry does not allow to fully take advantage of the photon-fusion production because of the limited acceptance to forward electron and positrons. The current angular coverage of Belle II can be written in terms of the electron (positron) pseudorapidity in the center of mass frame as

$$\text{Belle II}: \quad |\eta^*_{e\pm}| = \left|\log\left(\tan\left(\frac{\theta^*_{e\pm}}{2}\right)\right)\right| \leq 1.64 \, ,  \tag{1.5}$$

and corresponds to electron and positron with a *minimal* angle of 22° with respect to the beam axis.

Since the limited forward acceptance of Belle II cuts away a large fraction of the fusion production cross-section, it is interesting to explore how the reach of diphoton resonances could be improved by a hypothetical experiment with a wider angular coverage at an $e^+e^-$ facility. The result for this hypothetical detector, which we call Belle-fwd, are presented in green in

---

[1]In absence of such dedicated analysis, our sensitivity estimate suffers a small, but appreciable, drop in the mass range $m_a \simeq 0.1 - 1$ GeV. For masses around these values and for the coupling strength at which we can put a bound the decay length becomes comparable to the $\mathcal{O}(m)$ size of the detector, resulting in a loss of rate for both the search presented in this work and for that of our earlier ref. [15].

figure 1 assuming the same luminosity of Belle II. In order to fix a benchmark, we imagine a pseudo-rapidity coverage in the center of mass frame for both positrons and electrons of

$$\text{Belle-fwd}: \quad |\eta^*_{e\pm}| = \left|\log\left(\tan\left(\frac{\theta^*_{e\pm}}{2}\right)\right)\right| \leq 5\,, \tag{1.6}$$

which corresponds to a minimal angle with respect to the beam axis of $0.76°$. For reference, this roughly corresponds to the forward acceptance of the ATLAS calorimeter.

The possibility of extending the forward coverage at Belle II is limited by the presence of collimating magnets around the beam line which do not allow this region to be instrumented with detectors. The magnets are required to reach the expected integrated luminosity of the Belle II experiment [21]. Thanks to the impressive boost in the signal rate for an extended forward acceptance, even with reduced luminosity it is still possible to get an improvement in the reach compared to the Belle II experiment. This of course will also reduce the challenge of squeezing the beam to reach the luminosity targeted at Belle II. A more detailed illustration of how the reach in this channel depends on the forward acceptance and the luminosity is shown in figure 4 and discussed in section 3.

We conclude in section 4 leaving the technical details of our study for a series of appendices. In appendix A we discuss how we overcome the challenges of generating a smooth SM background prediction in our binned 4D likelihood analysis. In particular, we discuss the different possible statistical treatments of the empty bins in the MC, which result in the uncertainty in the expected reach in figure 1. In appendix B we further detail the features of our likelihood. In appendix C we quantify the range of applicability of the well-known effective photon approximation (EPA) [22–25] that is often advocated to carry out simplified cross-section computations and to develop intuition on the signal kinematics. Appendix D discusses the event by event reweighting we performed to account for the ALP lifetime.

## 2 Expected sensitivity at Belle II

In our signal we deal with $e^\pm$ and $\gamma$ detected in Belle II. According to Belle II detector performance [26] these objects are well reconstructed if they satisfy the following energy and angular requirements:

$$E^* > 0.25\,\text{GeV}, \quad \theta^* \in [22, 158]^\circ\,, \tag{2.1}$$

where the starred quantities are measured in the collision center of mass frame. In addition, as we want to discuss a resonance in the spectrum of resolved photons, we require

$$\Delta\theta_{\gamma\gamma} > 0.048 \quad \text{or} \quad \Delta\phi_{\gamma\gamma} > 0.048\,, \tag{2.2}$$

where $\Delta\theta_{\gamma\gamma}$ and $\Delta\phi_{\gamma\gamma}$ indicate the polar and azimuthal angular differences between the two resonant photons. This requirement on laboratory frame angles ensures that the two photons in the signal final state are reconstructed as separated photons at Belle II [26].

The photons need also to be separated from the electron to give rise to a clean final state, thus we require for all $e^\pm$-photon pairs

$$\Delta\theta_{\gamma e\pm} > 0.048 \quad \text{or} \quad \Delta\phi_{\gamma e\pm} > 0.048\,. \tag{2.3}$$
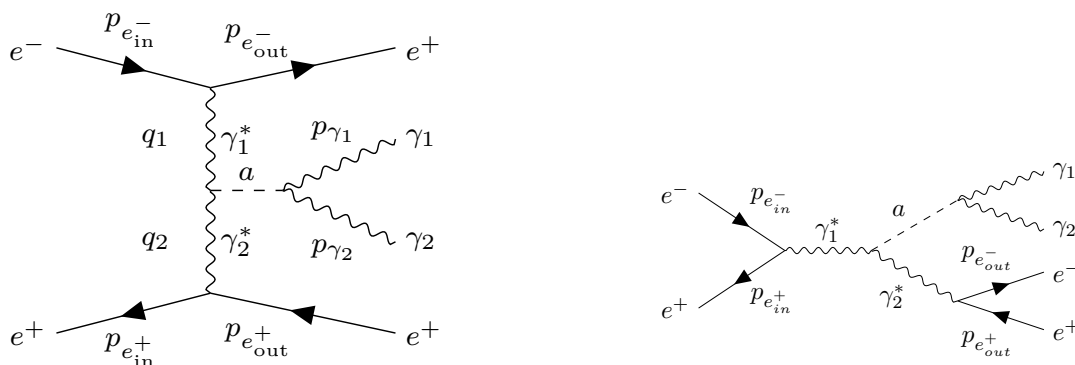
**Figure 2.** Feynman diagram for the $e^+e^- \to e^+e^- a$ production for photon-fusion (left) and Dalitz (right).

The angular separation and energy requirements of eqs. (2.1), (2.2), (2.3) cut-off soft and collinear divergences guaranteeing the convergence of the perturbative expansion so that both our signal and background samples are generated with MADGRAPH5_AMC@NLO v2.7.3 [27] at leading order in perturbation theory.

## 2.1 Signal vs background

Two possible ALP production mechanisms at a lepton collider are

$$e^+e^- \to \gamma_{\text{vis}} a \,, \tag{2.4}$$

$$e^+e^- \to e^+_{\text{vis}} e^-_{\text{vis}} a \,, \tag{2.5}$$

with $a \to \gamma\gamma$. We have denoted by the subscript $vis$ the final states that are in the acceptance of the Belle II detector [26] given in (2.1). Depending on the lifetime of the ALP, the production mechanisms above can give rise to either a visible resonant diphoton pair $a \to (\gamma\gamma)_{\text{res}}$, the processes eq. (2.4) and eq. (2.5), or to missing energy and momentum $a \to$ invisible.

The signal in eq. (2.4) is the well studied "ALP-strahlung" production (see ref. [14]), while the process in eq. (2.5) gives rise to two different processes: $i$) the "ALP-Dalitz" process, given by the ALP-strahlung eq. (2.4) with a photon conversion into $e^+e^-$; $ii$) the "photon-fusion" into ALP, given by a photon line exchanged in the $t$-channel that radiates the ALP depicted in figure 2. Similarly to what we found in [15] for the invisible ALP decay, the sensitivity for an ALP decaying into a diphoton pair at Belle II is dominated by the photon-fusion production. This dominance would be further enhanced in an experiment with extended forward detector coverage, as discussed in section 3.

For each putative $m_a$ we require the two final state photons to have an invariant mass in a range

$$m_{\gamma\gamma} - m_a \in [-3, 1.5] \cdot \sigma_{m_a} \,, \tag{2.6}$$

where $\sigma_{m_a}$ denotes the expected resolution on a di-photon resonance [26], which is dominated by the uncertainty on the photon energy depositions in the ECAL. This requirement removes large fractions of the backgrounds for all $m_a$ of interest.
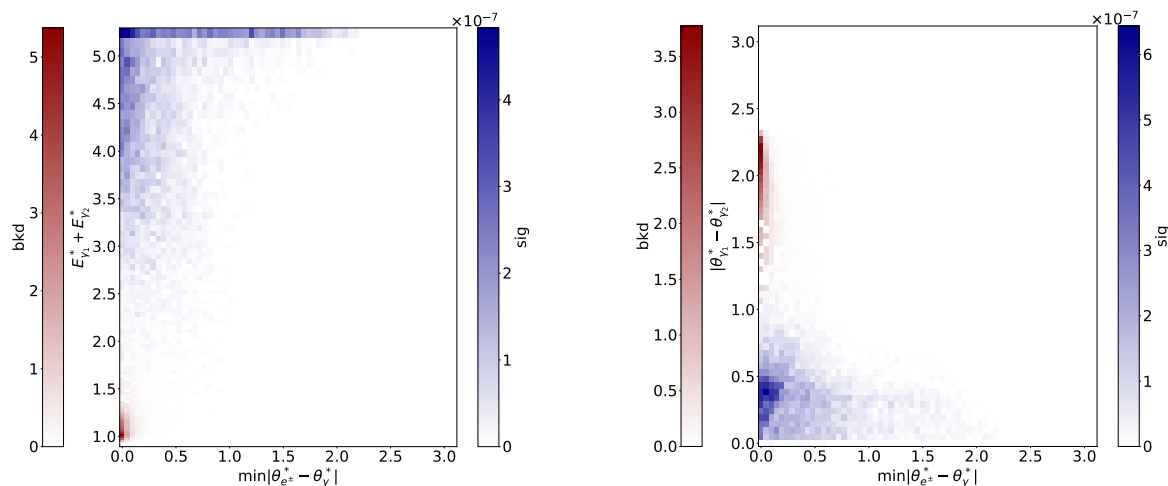
**Figure 3.** Distributions of signal events for the ALP production in photon fusion in eq. (2.5) (in **blue**) and background events for the radiative Bhabha in eq. (2.7) (in **red**). A fixed diphoton invariant mass of $m_{\gamma\gamma} = 1$ GeV within the range of eq. (2.6) is imposed, as well as the acceptance and isolation cuts at Belle II given in eqs. (2.1), (2.2), (2.3). *Left:* distributions as a function of the minimal angle between the positron/electron direction and the photon and the total energy of the photon pair, *Right:* distributions as a function of the minimal angle between the positron/electron direction and the photon and the angular separation between the two photons.

The most relevant background to the photon fusion production in eq. (2.5) is given by radiative corrections to Bhabha scattering

$$e^+ e^- \to e^+ e^- \gamma\gamma \,. \tag{2.7}$$

The LO rate for this process in the phase-space identified by eqs. (2.1)–(2.3) and eq. (2.6) in the mass window [0.95, 1.02] GeV is 13.40 pb, which is much larger than the signal rate $5.30 \cdot 10^{-5}$ pb $\cdot \left(\frac{g_{a\gamma\gamma}}{10^{-4}}\right)^2$. Despite the large imbalance of signal versus background, we will find sensitivity to the signal thanks to a highly differential likelihood. A large fraction of the cross-section for this process can be understood qualitatively as $e^+ e^- \to e^+ e^-$ dressed by two $e^\pm \to e^\pm \gamma$ splittings. As a consequence of the soft and collinear enhancement of the splittings, a distinguishing feature of this background process is the tendency of each photon to fly collinear to the fermion that has emitted it, as well as being preferred to be soft compared to the parent $e^\pm$. As a further consequence of the collinear preference of the photons emerging from Bhabha scattering, we observe that a large fraction of the background events tends to have photons flying in opposite directions. This is expected when the two photons are arising from different fermionic lines in the hard scattering.

In figure 3 the distributions of signal and background at fixed invariant mass of 1 GeV are shown. As discussed above, the background collinear enhancement will favor a small angle between the photon and the corresponding fermionic line (either the electron or the positron). As a consequence, in both panels in figure 3 we can see how the background event distribution favors a small *minimal* angle between the photon and electron (or positron) in the final state. In both panels we see that the signal does not have such a strong preference

for small angular separation between electrons and photons, thus the $\gamma$-$e^{\pm}$ angular separation can be used to reject background as necessary.

Moreover, for a small enough diphoton invariant mass, the signal events passing the Belle II acceptance cut typically feature a boosted ALP resonance (as already shown in ref. [15] but see also figure 8 in appendix C), which results in an energetic photon pair with relatively small opening angle between the photons. Conversely, the enhancement of the splitting function for soft radiation prefers the photons energies to be small and the angle between them could easily be large to make the required invariant mass if for instance they are split in opposite directions. The separation between signal and background becomes less sharp at higher invariant masses, when the boost of the signal resonance diminishes. This feature qualitatively explains the degradation of the sensitivity at Belle II apparent in figure 1.

Before leaving this section, we comment on further backgrounds beyond the one in eq. (2.7). For instance one could consider processes in which the final state $e^{\pm}$ originate from a photon conversion, e.g. $e^+ e^- \to 3\gamma$ followed by photon conversion $\gamma^* \to e^+ e^-$. These are similar to the Dalitz-like contribution to the signal itself. These backgrounds are populating a region of phase-space where our signal is basically absent, thus they do not contribute to our likelihood introduced below in eq. (2.11).

As a further source of background one can consider light-by-light scattering, which has been considered as a background to gamma fusion ALP production. For instance in refs. [28, 29] it has been studied for FCC and ILC colliders, thus at a much larger center of mass energy than Belle II. This process can in principle give rise to a potentially harmful background also at Belle II as it populates a region of phase-space similar to that of our signal. In any case, being this process a loop process, and being a non-resonant $2 \to 4$ process, we expect it to start to be relevant only for $g_{a\gamma\gamma} \lll 10^{-4}\,\mathrm{GeV}^{-1}$, which fall outside of the reach we presented in figure 1. In any case a proper evaluation of light-by-light scattering with a full $2 \to 4$ matrix elements calculation will be necessary to confirm these estimates. One further possibility to obtain background is the rare mis-reconstruction of hadrons as electrons and fake photons, e.g. from mis-identification of neutral mesons, e.g. a $\pi^0$ detected as a single photon. The efficiency-vs-purity performance of the Belle II experiment on particle-ID is a subject of active development (see e.g. ref. [30]). A realistic estimate of these backgrounds is challenging, especially when multiple hadrons arise from a copious two-body hard scatterings such as $e^+ e^- \to q\bar{q}$. The final background rate depends mainly on the detector operation, e.g. isolation requirements, and should be performed by the experimental collaboration. In addition we mention that SM diphoton resonances, e.g. $\pi^0$, $\eta$, ..., should be accounted for in the relevant $m_a$ ranges.

## 2.2 Sensitivity

For an optimal signal vs background separation we pursue a full four dimensional likelihood, which can take full advantage of their different kinematic structure. In order to efficiently tessellate the four dimensional phase space we define, in analogy with the usual Mandelstam

variables, the Lorentz invariants $t_\pm$ and $s_\pm$ as

$$
\begin{aligned}
t_+ &= \left( p_{e_{\text{in}}^+} - p_{e_{\text{out}}^+} \right)^2 , \\
t_- &= \left( p_{e_{\text{in}}^-} - p_{e_{\text{out}}^-} \right)^2 , \\
s_- &= \left( p_{e_{\text{in}}^+} + p_{e_{\text{in}}^-} - p_{e_{\text{out}}^+} \right)^2 , \\
s_+ &= \left( p_{e_{\text{in}}^+} + p_{e_{\text{in}}^-} - p_{e_{\text{out}}^-} \right)^2 ,
\end{aligned}
\tag{2.8}
$$

where the positron, electron, and photon momenta are defined in figure 2. A conversion between the Lorentz invariant variables and the measured energy and angles of the positron, electron, and photons can be easily derived. Neglecting the electron mass, we can write

$$
t_\pm \simeq -\sqrt{s}\, E^*_{e_{\text{out}}^\pm} \left( 1 \pm \cos\theta^*_{e_{\text{out}}^\pm} \right) \quad , \quad s_\pm \simeq \sqrt{s}\left( \sqrt{s} - 2 E^*_{e_{\text{out}}^\pm} \right) ,
\tag{2.9}
$$

where $\theta^*_{e_{\text{out}}^\pm}$ is the polar angle of the positron/electron with respect to the beam axis in the CoM frame. The above expressions show that $t_\pm \to 0$ controls the collinear singularity of the photon fusion production at $\theta^*_{e_{\text{out}}^\pm} \to 0$. Using energy conservation, the sum of the other two Mandelstam variables can be written in terms of the energy of the resonance $s_+ + s_- \simeq 2E_a^*\sqrt{s}$, while their difference gives the asymmetry between the positron and the electron energy.

Starting from the energy and polar angle resolution of Belle II [26] we derive approximate resolutions for the measurement of $t_\pm$ and $s_\pm$. Each of these quantities contains energies and directions of photons and electrons. The energies are the least well measured, so we can take

$$
\frac{\delta\mu_i}{\mu_i} \simeq 2\frac{\delta E}{E} \simeq 4\% \quad \text{for} \quad \mu_i = \{s_+, s_-, t_+, t_-\} .
\tag{2.10}
$$

Based on the predicted differential rates for the SM background and for the signal, we construct a log-likelihood for 50 ab$^{-1}$ integrated luminosity at Belle II

$$
\Lambda = -2 \sum_{i,j} \ln \frac{L(S_{i,j}, B_{i,j})}{L(0, B_{i,j})} ,
\tag{2.11}
$$

where $i$ and $j$ run on the bins of the 4D space spanned by $s_+$, $s_-$, $t_+$, $t_-$ and $S$ and $B$ indicate respectively the expected number of signal or background events in each bin. In (2.11) $L(S_{i,j}, B_{i,j})$ is the Poisson factor computed in each bin defined as follows

$$
L(S, B) = \frac{(S+B)^B}{B!} e^{-(S+B)} .
\tag{2.12}
$$

The sensitivity shown in figure 1 corresponds to 95% C.L. and it is obtained by requiring $\Lambda < 4$.

Given the huge hierarchy between the signal rate and the background rate, our sensitivity benefits from regions of phase space where the background is strongly suppressed. To best exploit these features, we perform a full 4D likelihood analysis. This likelihood is challenging to construct from the viewpoint of the MC generation, but the improvement in the sensitivity is worth the effort.[2] For computational reasons we then tessellate the 4D phase space

---

[2]We note in passing that the expected sensitivity for a 3D likelihood obtained by collapsing our 4D one in any of the four directions is reduced by roughly one order of magnitude. This loss of sensitivity confirms that the full phase space information is needed to distinguish the photon fusion process from the background.

with different bin sizes depending on the ALP mass. Specifically for $m_a < 1$ GeV we take $\delta\mu_i/\mu_i = 15\%$, for $1$ GeV $\leq m_a < 3$ GeV we take $\delta\mu_i/\mu_i = 8\%$ and for $m_a \geq 3$ GeV we take the realistic Belle II resolution estimated in eq. (2.10). In this sense, the expected reach of figure 1 at low masses should be considered conservative and could be improved by reducing the bin size further.

Even with coarse grained bins, our expected reach has a residual dependence on the statistical treatment of the phase space cells with zero expected background within our MC uncertainty. Evaluating this uncertainty ultimately results in the uncertainty band in the expected reach shown in figure 1. A detailed description of our methods to generate a smooth background sample in the full 4D phase space and assigning a conservative uncertainty to it is given in appendix A.

We also checked the robustness of our result against possible systematic uncertainties. These are particularly dangerous as one approaches a high statistics regime and small couplings can be potentially probed. To compute the loss of sensitivity due to systematics we computed the log-likelihood keeping only bins for which the expected $S/B$ was larger than a certain threshold $\epsilon_{\rm sys}$. For $\epsilon_{\rm sys} < 10\%$ the expected sensitivity does not change appreciably, especially at low masses. This is because the likelihood is dominated by bins with very suppressed background yield, as discussed in detail in appendix B.

The variable $t_+, (t_-)$ in eq. (2.8) is related to the degree of collinearity of the initial and final $e^+, (e^-)$, which in turn can be seen as the collinearity of the virtual photons and the initial $e^+, (e^-)$. These quantities are typical markers of $t$-channel scattering, which dominates the photon-fusion production of the ALP. As it can be appreciated from figure 5 in appendix B, the Belle II acceptance cuts off this singularity reducing the signal acceptance. In the next section we show how a hypothetical experiment with extended forward acceptance would allow instead to explore kinematic configurations where both background and signal are closer to the collinear singularity and can be effectively distinguished by exploiting the full four dimensional phase space information.

## 3   Expected sensitivity of a forward flavor factory

In this section we entertain the possibility that a future experiment with extended forward coverage will be available at an $e^+e^-$ facility similar to SuperKEKB. Given that the expected luminosity of such a machine is likely to be correlated with its forward instrumentation, we want to study how the expected reach of the search discussed in section 2 depends on the forward acceptance, $\eta^*_{e\pm}$, and the luminosity, $\mathcal{L}$. We do not make a concrete technical proposal here on how to construct and operate this extended detector. We remark however that this is an active area of research for future collider detectors studies [31, 32].

Our results are shown in figure 4 for benchmark ALP mass of $3$ GeV and $\sqrt{s} = 10.58$ GeV, where the red and the green stars correspond to the Belle II reach and the Belle-fwd reach of figure 1. The reach in $g_{a\gamma\gamma}$ defined in eq. (1.3) decreases roughly like $\mathcal{L}^{1/4}$ as expected for a search that is not free from background. Conversely, increasing the forward acceptance of electron and positrons leads to a steep rise of the sensitivity which decreases only at high forward acceptances. We find that, thanks to the impressive gain in reach showed in figure 1, even an experiment collecting 100 times less luminosity than Belle II, profiting of
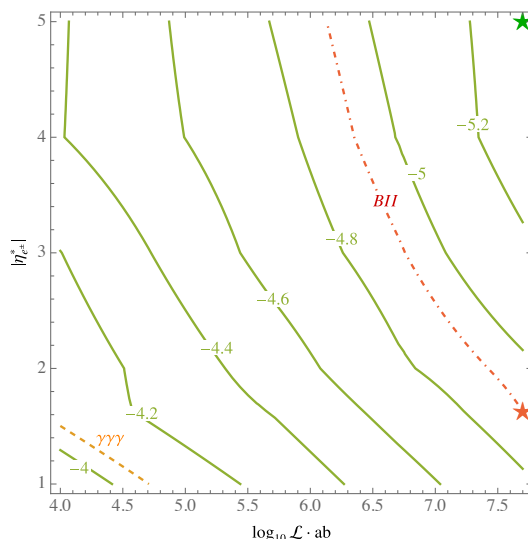
**Figure 4.** The **green contours** show the expected reach for $\log_{10}(g_{a\gamma\gamma})$ in the $e^+e^- \to e^+e^-(a \to \gamma\gamma)$ channel as a function of the luminosity and the final $e^\pm$ acceptance of a hypothetical lepton collider at $\sqrt{s} = 10.58$ GeV, given $m_a = 3$ GeV. The **red** and **green** stars are the benchmarks for the Belle II and Belle-fwd reach in figure (1). As a reference we show as a **orange** dashed line the Belle II reach for the $e^+e^- \to \gamma(\gamma\gamma)_{\rm res}$ channel. The red dashed line is the isoline of the value of the Belle II reach.

an extended forward acceptance down to $|\eta^*_{e^\pm}| < 5$, could lead to a substantial improvement in the reach compared to Belle II.

## 4 Conclusion

In this work we developed a new search for diphoton resonances at Belle II based on the photon fusion production in eq. (1.1). By exploiting a full tessellation of the four dimensional phase space of this $2 \to 3$ process we performed a detailed study of how this can be separated from the background given by the radiative Bhabha process in eq. (2.7). The result of this analysis is summarized in figure 1 which shows how this channel can lead to a substantial improvement in the reach compared to the "ALP-strahlung" production studied in ref. [14] at least for ALP masses below 6 GeV. Above 6 GeV the Belle II reach quickly becomes less sensitive than existing constraints from ultra-peripheral heavy ions collisions [4, 12]. Along the way, we illustrate in appendix A a method to associate a conservative uncertainty to the signal reach in analyses plagued by challenges in the MC sampling like ours.

In addition, we explored how the expected reach for a diphoton resonance in the $e^+e^- + \gamma\gamma$ channel can be further improved at a hypothetical experiment with extended forward coverage compared to Belle II. The results of figure 1 together with the ones in figure 4 show that with an extended forward coverage down to $|\eta^*_{e^\pm}| < 5$ the expected Belle II reach could be improved even with 100 time less luminosity.

## Acknowledgments

# A  Generation of the background sample for the 4D likelihood analysis

The generation of 4-fold differential cross-section requires intense computational resources. A relatively gross-grained sampling of each dimension with about 100 bins leads to a rather large set of $10^8$ bins in a single histogram. Obtaining the value of the cross-section with precision $1/p$ in each bin requires $p^2$ samplings of the differential cross-section, which, even for leading order matrix elements, takes some substantial time.

Significant speed-up of the generation of the information necessary to characterize our background in a 4D phase-space can be obtained i) parallelizing the generation of multiple MC samples, ii) using a clever re-weighting of the events, iii) applying suitable biases to the samples, and iv) combining the information from the several runs.

In our calculation we have applied a bias to the differential cross-section as to overpopulate tails at large $s_+$ and large $s_-$, and to increase the population of the area of phase-space around $t_+ \simeq t_- \simeq 0$. These regions are most difficult to describe precisely with MAD-GRAPH5_AMC@NLO both because they are tails and because they are rather narrow region with respect to the whole phase-space of the process. In addition, $t_+ \simeq t_- \simeq 0$ is a boundary of the physical phase-space, which poses extra difficulties. The accurate description of these regions of phase-space is of utmost importance for our results as they give a large discrimination of signal over background within our likelihood in eq. (2.11).

Particularly challenging is the generation of background events for $m_a \leq 1\,\mathrm{GeV}$ at Belle II, because the kinematics of the signal is such that the background suppression in specific regions of phase space is larger than for higher masses. In what follows we describe our MC strategy in this light mass range and mention the simplification at higher masses at the end.

For $m_a \leq 1\,\mathrm{GeV}$, we produce $\mathcal{O}(10)$ MC samples with MADGRAPH5_AMC@NLO with bias and cuts aimed towards several different combinations of small and large $s_+ + s_-$ and similarly for $t_+, t_-$. From each sample we obtain a weight, that corresponds to the cross-section for each given cell of the 4D phase-space. We denote this weight as $w_{c,s}$, where $c$ denotes the cell of the phase-space and $s$ the sample we are dealing with. Each such weight can be endowed of an *estimated* uncertainty, which originates from the finiteness of the MC samples that we generated. We estimate an uncertainty $\delta w_{c,s} = w_{c,s}/\sqrt{n_{c,s}}$, where $n_{c,s}$ is the number of weighted MC events in the cell $c$ from the sample $s$.

We then combine the weights of a given cell $c$ as if they were the results of different experiments into a combined value

$$w_{c,C} = \frac{\sum_s w_{c,s} \cdot \xi_{c,s}}{\sum_s \xi_{c,s}}, \tag{A.1}$$

where $\xi_{c,s} = \delta w_{c,s}^{-2}$, as in the usual combination of two or several measurements described by an underlying Gaussian probability density function. Linear uncertainty propagation is applied to this formula, finding the standard result for the uncertainty on the combined value

$$\delta w_{c,C} = \left( \sum_s \delta w_{c,s}^{-2} \right)^{-1/2}. \tag{A.2}$$

The samples $s$ included in the combinations are chosen so that only measurements that are *compatible* with each other are used. In particular, we combine only samples that have

non-zero weight on the cell $c$ and discard the samples that returned zero weight, either because $c$ was entirely out of the region of phase-space of the MC run that produced sample $s$, or because the bias in that run was put to highlight other tails of the phase-space. The runs with zero weight are discarded on the ground that they correspond to incompatible measurements, and therefore cannot be combined with the other group. In fact, the information from the samples returning zero weight would imply zero cross-section on that cell, which is clearly incompatible with the information from a dedicated sample that has found non-zero cross-section in that cell.

When a cell is not populated by any of our background samples simulations, we conservatively assume that the cross-section in that cell is non-zero and that each sample has given us an upper-bound on the cross-section in that cell. Each sample contributes an upper-bound on the weight $w_{c,s} < \Delta w_s$, where $\Delta w_s$ is the minimum weight in the sample $s$. The combined upper-bound then is

$$ w_{c,C} < \Delta w_{c,C} \quad \text{for} \quad \Delta w_{c,C} = \left( \sum_s \Delta w_s^{-2} \right)^{-1/2} . \tag{A.3} $$

From the above procedure we obtain a combined histogram in 4D that has either a measurement or an upper bound on the cross-section in each cell of phase-space. In the following we describe a method to assess the impact of the choice in the treatment of these cells for which we have obtained an upper bound.

In order to give an estimate of the error on the bounds on $g_{a\gamma\gamma}$ due to the finiteness of our MC samples, we produce replicas of our combined histogram. In each replica we fluctuate the number of expected events in the cell $c$ based on the underlying uncertainty on the combined weight $\delta w_{c,C}$. For the bins with non-zero background events the number of the expected events follows a truncated normal distribution where the gaussian mean and standard deviation are taken from eq. (A.1) and eq. (A.2).

For the empty bins we only have the upper bound in eq. (A.3), which leaves us the freedom of picking the probability distribution in the bin. Variations of the underlying probability distribution for the bins gives the uncertainty in the expected reach in figure 1. We identify three scenarios:

- The weakest constraints come from assuming that the distribution is a Dirac $\delta$ function centered at the upper bound in eq. (A.3). This gives the maximal number of expected background events compatible with having observed zero events in our Monte Carlo.

- The strongest constraints are obtained assuming a discrete Poisson distribution with mean set at the upper bound in eq. (A.3). The discrete nature of this distribution gives a sizable probability to have zero expected background, thus boosting the likelihood in eq. (2.11).

- Intermediate bounds between these two options are obtained by picking a continuous uniform probability distribution from zero to the upper bound in eq. (A.3). The central line in the bands in figure 1 corresponds to this choice, which we consider as the most reasonable estimate of the attainable constraints.
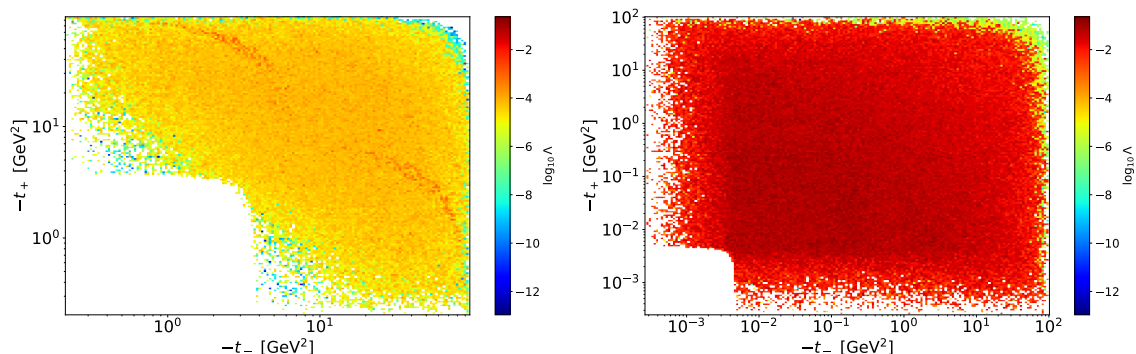
**Figure 5.** Distribution of log-likelihood defined in eq. (2.11) in the $(t_-, t_+)$ plane for $m_a = 3$ GeV. The 4D likelihood is maximized in the $(s_-, s_+)$ plane to be projected on a 2D plane. The signal strength is fixed to the expected reach of Belle-fwd with 50 ab$^{-1}$ of data which is $g_{a\gamma\gamma} = 5.5 \cdot 10^{-6}$ GeV$^{-1}$ as shown in figure 1. The white region is kinematically forbidden because of the detector acceptance. *Left* Belle II likelihood with acceptance defined in eq. (1.5) *Right* Belle-fwd likelihood with acceptance defined in eq. (1.6).

The background generation for $m_a \geq 1$ GeV is analogous to the one explained above, with the technical simplification that a single MC sample was sufficient to obtain a smooth background sample. Running multiple samples would have allowed us to reduce the bin size of the likelihood for 1 GeV $\leq m_a < 3$ GeV and in general reduce the uncertainty in our projected sensitivity for all the ALP masses. The projected sensitivity and its uncertainty shown in figure 1 should then be taken as a first conservative assessment of the discriminating power of the $e^+e^- \to e^+e^-(\gamma\gamma)_{\mathrm{res}}$ channel.

# B Features of the 4D likelihood

We now comment on the features of the 4D likelihood which result in the sensitivity at Belle II and at Belle-fwd in figure 1. Firstly, we show in figure 5 the distribution of our likelihood as defined in eq. (2.11) in the $(t_-, t_+)$ plane. For every $t_-, t_+$ pair, we pick the maximal likelihood value in the $(s_-, s_+)$ plane. In the left plot we see the distribution at Belle II while in the right one at a hypothetical experiment with extended forward coverage at an $e^+e^-$ facility and the same luminosity of Belle II.

The shape of the likelihood distribution can be explained by noticing that the absolute values of the variables $t_\pm$ as defined in eq. (2.9) are bounded from below by a quantity that depends on the detector acceptance which determines the minimal polar angle of the electron/positron that one can measure, $\theta^{\mathrm{min}}_{e^\pm_{\mathrm{out}}}$. In formulas we can write the bound for tip of the wedge with $t_+ \sim t_-$ as

$$|t_\pm| > \frac{s}{4}(1 - \cos\theta^{\mathrm{min}}_{e^\pm_{\mathrm{out}}}),\tag{B.1}$$

where $\theta^{\mathrm{min}}_{e^\pm_{\mathrm{out}}}$ can be extracted from Eq, (1.5) for Belle II and from eq. (1.6) for Belle-fwd. The comparison between the two panels shows clearly how the extended forward coverage opens up a new region of phase space where the likelihood is considerably larger, hence boosting the sensitivity to the ALP.
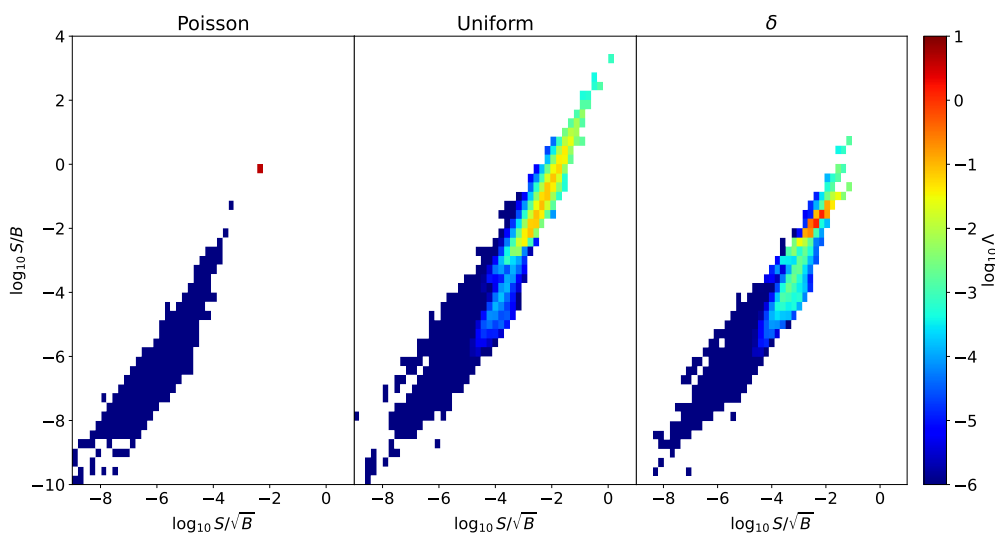
**Figure 6.** Likelihood distributions for $m_a = 3$ GeV in the plane $(S/\sqrt{B}, S/B)$ where $S$ is the expected number of signal events and $B$ the expected number of background events. The three panels correspond to the three different assumption on the distribution of the background events in the empty bins as illustrated in appendix A. *Left:* the events in the empty bins are distributed as a Poisson distribution with mean set to the upper bound of the cross-section derived in our MC. The likelihood correspond to the best expected sensitivity in figure 1 which is $g_{a\gamma\gamma} = 5.49 \cdot 10^{-6}$ GeV$^{-1}$ *Center:* the events on the empty bins are distributed as a uniform distribution with maximum given by the upper bound derived in our MC. This treatment correspond to the red line in figure 1 which gives $g_{a\gamma\gamma} = 2.88 \cdot 10^{-5}$ GeV$^{-1}$. *Right:* the events on the empty bins are distributed as a delta-function centered at the upper bound derived in our MC. This assumption gives the most conservative expected reach in figure 1 which is $g_{a\gamma\gamma} = 5.89 \cdot 10^{-5}$ GeV$^{-1}$.

We now illustrate the dependence of the likelihood on our treatment of the empty bins illustrated in appendix A. Calling $S$ the expected signal events and $B$ the expected background events, we show in figure 6 the distribution of the likelihood in the plane spanned by $S/\sqrt{B}$, which roughly controls the statistical weight, and $S/B$, which is a good proxy for the sensitivity of our result on the systematic uncertainties in the background rate. The three panel illustrate the three different treatments of the empty bins, and the coupling strength of the ALP in each panel is fixed at its corresponding expected sensitivity at Belle II.

In the left panel we show the likelihood assuming that the probability distribution of the empty bins is Poissonian and has its mean set at the upper bound of the background cross-section obtained from our MC exploration. The discrete nature of the likelihood makes most of the likelihood concentrated on the tail of the Poisson distribution with zero background events which is shown as a red square dot in the plot. The center panel shows the likelihood obtained under the assumption that the events in the empty bins are uniformly distributed with maximal value given by the upper bound derived in our MC. Here we see that the likelihood distribution becomes nicely continuous in the $(S/\sqrt{B}, S/B)$ plane. In the right panel we show the likelihood distribution where we assume the background rate in the empty bins to be fixed at the lower bound derived from the MC. This assumption gives the
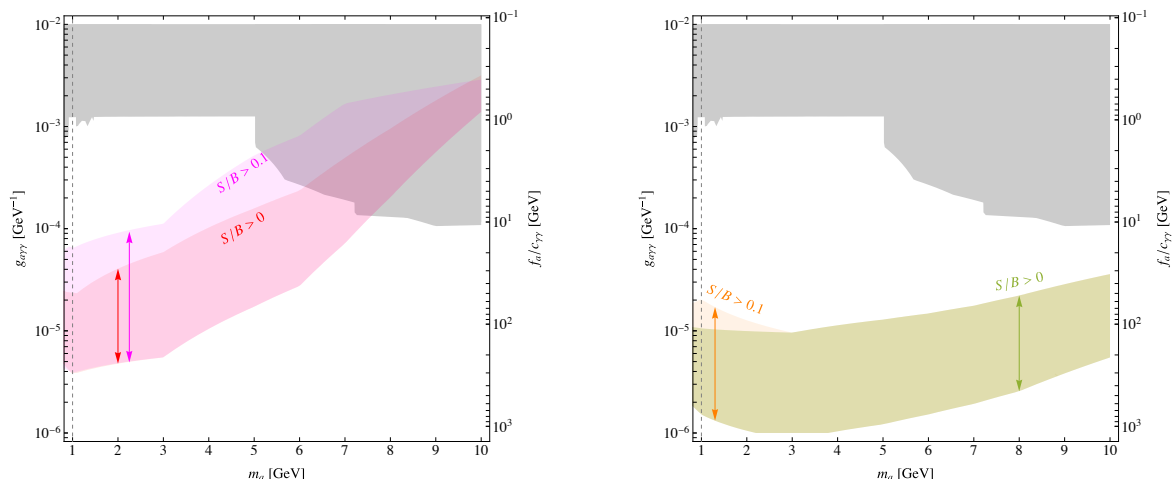
**Figure 7.** Dependence of the expected reach on the systematic uncertainties on the background rate. The gray shaded region shows existing constraints as in figure 1. *Left:* the **red** band shows the expected reach of $e^+e^- + (\gamma\gamma)_{\text{res}}$ as in figure 1, and in **magenta** the same reach with the requirement $S/B > 0.1$, as detailed in appendix B. For clarity we draw arrows of the same color of the shades to display the size each band. *Right:* the **green** band corresponds to the reach of Belle-fwd as defined in eq. (1.6) and detailed in section 3; the **orange** band is the same reach with the requirement $S/B > 0.1$.

most conservative expected sensitivity in figure 1. For both the center and right panel, we observe that a large fraction of the likelihood lives at large value of $S/B$, anticipating the low dependence of our expected reach on systematic uncertainties on the background.

The dependence of the expected reach on the systematic uncertainties is shown in figure 7. To quantify this effect, we restrict our calculation of eq. (2.11) to the bins with $S/B$ larger than a threshold value. Throwing away in our likelihood calculation all the regions of phase-space with $S/B$ smaller this threshold gives a conservative estimate of the bound one obtains in presence of systematic uncertainty on the background of the same size of the chosen threshold. For $S/B > 10\%$ we see that the uncertainty in our expected reach is increasing at Belle II (left panel figure 7), while it is left almost unchanged at Belle-fwd (right panel figure 7). In particular, at Belle II we see that systematic uncertainties at 10% level make the sensitivity worse by roughly a factor of five on the coupling. This is due to the fact that the upper part of the bands is obtained in the very conservative assumption that the upper bound on the cross-section derived in our MC is the number of expected background events. Especially at Belle II, being the signal-to-background worse than at Belle-fwd, this can result in larger losses of sensitivity if the backgrounds are not well understood. The lower parts of the bands are instead more robust to systematics. This is because they are obtained assuming the events in the empty bins to be distributed as a Poissonian. Then the likelihood is dominated by bins with $B = 0$ and therefore essentially insensitive to systematic uncertainties.

## C  Effective photon approximation at flavor factories

The photon fusion process that we propose as a new channel to search for ALP production in Belle II is traditionally linked to the equivalent vector boson approximation [22–25]. In the
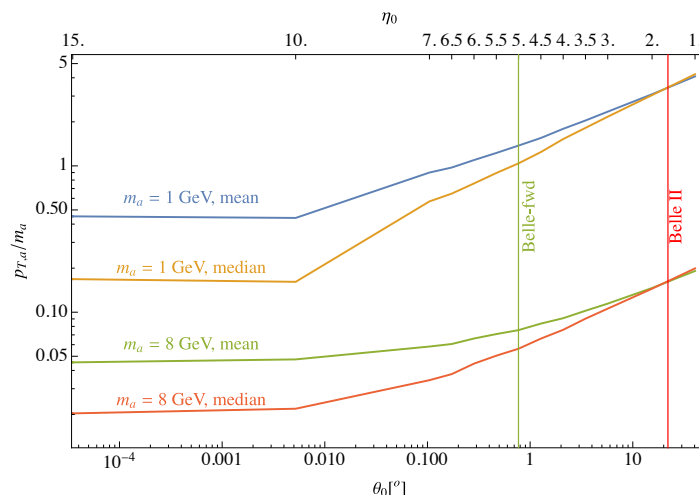
**Figure 8.** Mean and median $p_{T,a}/m_a$ for eq. (1.1) including phase-space up to a minimal angle $\theta_0$ (maximal pseudo-rapidity $|\eta_0|$).

context of LHC physics this approximation provides a simple computational tool to estimate the total cross-section of fusion processes, see e.g. refs. [33, 34] for early applications in the context of Higgs boson production in hadron colliders. While the computational aspect is no longer as pressing as it used to be, the equivalent boson picture can in principle be used to guide the isolation of the fusion signal from the backgrounds.

The basic idea of the approximation is to consider as on-shell the gauge boson field that enters in a scattering, as long as its virtuality can be neglected in comparison with the momentum exchanged in the scattering in which the boson participates. In our case the photon, in order to be approximately on-shell, must have negligible virtuality with respect to the mass of the $a$ that emerges from the photon-fusion. The physical foundation of this approximation relies on the fact that the process of photon fusion happens in a time of order $1/m_a$, which is too short to probe the virtuality of the photon, $q_i^2$ as defined in figure 2 if $q_i^2 \ll m_a^2$ [35]. The virtuality of the photon can only be "detected" in experiments sensitive to much longer time-scales.

As in our kinematics the photon virtualities are linked to the final state electron and positron transverse momenta

$$q_1^2 \simeq p_{T,e_{\text{out}}^-}^2, \quad q_2^2 \simeq p_{T,e_{\text{out}}^+}^2 \quad \text{and} \quad p_{T,e_{\text{out}}^+} + p_{T,e_{\text{out}}^-} = p_{T,a}, \tag{C.1}$$

it is instructive to look at the typical transverse momentum of the ALP divided by its mass. In figure 8 we show the median and the mean $p_{T,a}/m_a$ for $m_a = 1$ GeV $m_a = 8$ GeV including phase-space up to a minimal angle $\theta_0$ (maximal rapidity $|\eta_0|$) from the beam axis. We observe that for 1 GeV ALP mass, and for similar ALP masses compared to the center of mass energy, $p_{T,a}/m_a \sim 1$, which indicates that such light ALP is not accurately described by the equivalent photon approximation. This is especially true for $\theta_0 \simeq 1°$ and $22°$, the benchmarks for the Belle-fwd hypothetical detector and the actual Belle II detector respectively.

To corroborate this expectation, we show in the left panel of figure 9 the exact calculation of the fusion process eq. (1.1) and the EPA prediction as computed at its lowest order
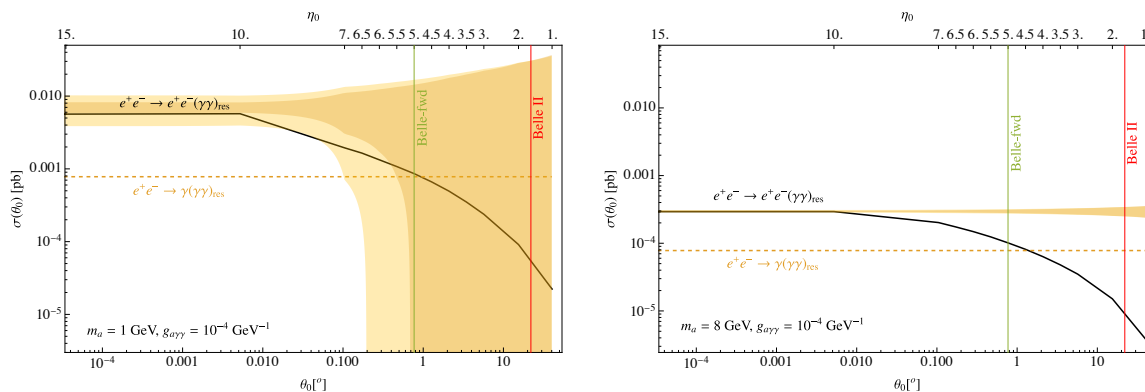
**Figure 9.** Comparison of the EPA prediction and the exact cross-section for the process eq. (1.1), in the case of $m_a = 1\,\text{GeV}$ (left) and $m_a = 8\,\text{GeV}$ (right). The *black* line is the cross-section for the exact computation of the process eq. (1.1) as a function of the lepton acceptance (see eq. (1.5)), expressed either as a polar angle measured in degrees $\theta_0$ or as a pseudo-rapidity $\eta_0$. The light yellow and the dark yellow error bands are the EPA with mean and median $p_{T,a}/m_a$ relative error, as detailed in the text. The vertical red and green lines correspond to the angular acceptance at Belle II and Belle-fwd respectively. For reference we also show the $\gamma + (\gamma\gamma)_{\text{res}}$ cross-section (which is invariant with respect to $\theta_0$) as an orange dashed line.

following the definition of ref. [24]. The EPA prediction is further equipped with an error band in light (dark) yellow, which corresponds to an uncertainty given by the mean (median) $p_{T,a}/m_a$ shown in figure 8. The agreement between the EPA and the exact computation degrades as $\theta_0 \gtrsim 0.01°$ ($|\eta_0| \lesssim 10$), with an order of magnitude disagreement expected for Belle-fwd and even worse for Belle II. In the right panel of figure 9 we show the same quantities for a heavy ALP mass of $8\,\text{GeV}$. In this case a much better agreement between EPA and the exact computation is observed, in full compliance with the smallness of the ratio $p_{T,a}/m_a \lesssim 0.1$ over the whole range of $\theta_0$.

From these analyses we conclude that the EPA works well in the cases in which the observed final states corresponds to small $t$-channel photon virtualities, but it is definitively possible to reach situations in which EPA does not provide a good estimate of the ALP total production cross-section. In these situations, also the kinematic picture of a signal dominated by small $p_{T,a}$ can be flawed. In fact, for ALPs mass $1\,\text{GeV}$ in figure 3 we have observed a strong preference for large $E_{\gamma_1} + E_{\gamma_2} \gg m_a$, in spite of the EPA expectation of an ALP produced nearly at rest.

Due to the difficulty to build forward detector in machines capable of reaching high instantaneous luminosity, such as flavor factories, the EPA should be used with great care to compute cross-sections at these machines. Similarly, the physics intuition that follows from the validity of the EPA can be inaccurate for experiments on these machines, due to the absence of coverage in the forward phase-space.

## D  Influence of ALP lifetime

For small enough ALP masses, the decay length becomes so large that most of the time the ALP decay happens far away from the interaction point, possibly even outside the detector.

In order to account for this effect, we reweigh every event by including the probability of the ALP to decay outside the detector length. The decay probability before some distance $L$ can be written as

$$P_{\text{decay}} = 1 - e^{-\frac{L}{\gamma_a \tau_a}} = 1 - e^{-\frac{L \Gamma_a m_a}{E_a}} \, , \tag{D.1}$$

where $m_a$ and $E_a$ are respectively the ALP mass and energy and

$$\Gamma_a = \frac{g_{a\gamma\gamma}^2 m_a^3}{64\pi} \tag{D.2}$$

its decay width. In order to make the most conservative estimate, we will choose $L$ to be the shortest distance between the interaction point and the ECL barrel which is $L = 1.25$ m at Belle II [26]. Each event will then have a weight which is a function of $g_{a\gamma\gamma}$ and also depends on the ALP mass and energy. We can then include this effect in the likelihood of eq. (2.11). The effects of the ALP lifetime, as expected, are visible only for $m_a \lesssim 0.5\,\text{GeV}$; for those masses there is effectively a reduction of the signal rate. As shown in figure 1, however, for ALP masses below $1\,\text{GeV}$, invisible searches become quickly more competitive, filling the gap of the visible ones.

Interestingly, the loss of promptly decaying events is less rapid at Belle-fwd (defined in eq. (1.6)) with respect to Belle II (see eq. (1.5)). This can be explained by computing the average boost of the particle as a function of the opening angle of the detector. For a more forward detector like Belle-fwd the average boost of the ALP will be smaller, making it easier for the ALP to decay inside the detector.

**Data Availability Statement.** This article has no associated data or the data will not be deposited.

**Code Availability Statement.** This article has no associated code or the code will not be deposited.

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution License (CC-BY4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

# References

[1] CMS collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30 [arXiv:1207.7235] [INSPIRE].

[2] ATLAS collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1 [arXiv:1207.7214] [INSPIRE].

[3] S. Knapen, T. Lin, H.K. Lou and T. Melia, *Searching for axionlike particles with ultraperipheral heavy-ion collisions*, *Phys. Rev. Lett.* **118** (2017) 171801 [arXiv:1607.06083] [INSPIRE].

[4] S. Knapen, T. Lin, H.K. Lou and T. Melia, *LHC limits on axion-like particles from heavy-ion collisions*, *CERN Proc.* **1** (2018) 65 [arXiv:1709.07110] [INSPIRE].

[5] A. Mariotti, D. Redigolo, F. Sala and K. Tobioka, *New LHC bound on low-mass diphoton resonances*, *Phys. Lett. B* **783** (2018) 13 [arXiv:1710.01743] [INSPIRE].

[6] X. Cid Vidal et al., *New axion searches at flavor factories*, *JHEP* **01** (2019) 113 [*Erratum ibid.* **06** (2020) 141] [arXiv:1810.09452] [INSPIRE].

[7] D. Aloni, C. Fanelli, Y. Soreq and M. Williams, *Photoproduction of axionlike particles*, *Phys. Rev. Lett.* **123** (2019) 071801 [arXiv:1903.03586] [INSPIRE].

[8] D. Wang, L. Wu, J.M. Yang and M. Zhang, *Photon-jet events as a probe of axionlike particles at the LHC*, *Phys. Rev. D* **104** (2021) 095016 [arXiv:2102.01532] [INSPIRE].

[9] S. Knapen, S. Kumar and D. Redigolo, *Searching for axionlike particles with data scouting at ATLAS and CMS*, *Phys. Rev. D* **105** (2022) 115012 [arXiv:2112.07720] [INSPIRE].

[10] ATLAS collaboration, *Search for boosted diphoton resonances in the 10 to 70 GeV mass range using 138 fb$^{-1}$ of 13 TeV pp collisions with the ATLAS detector*, *JHEP* **07** (2023) 155 [arXiv:2211.04172] [INSPIRE].

[11] G. Alonso-Álvarez, J. Jaeckel and D.D. Lopes, *Tracking axion-like particles at the LHC*, arXiv:2302.12262 [INSPIRE].

[12] CMS collaboration, *Evidence for light-by-light scattering and searches for axion-like particles in ultraperipheral PbPb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV*, *Phys. Lett. B* **797** (2019) 134826 [arXiv:1810.04602] [INSPIRE].

[13] R. Balkin et al., *Probing axion-like particles at the Electron-Ion Collider*, *JHEP* **02** (2024) 123 [arXiv:2310.08827] [INSPIRE].

[14] M.J. Dolan et al., *Revised constraints and Belle II sensitivity for visible and invisible axion-like particles*, *JHEP* **12** (2017) 094 [*Erratum ibid.* **03** (2021) 190] [arXiv:1709.00009] [INSPIRE].

[15] F. Acanfora, R. Franceschini, A. Mastroddi and D. Redigolo, *Fusing photons into nothing, a new search for invisible ALPs and dark matter at Belle II*, arXiv:2307.06369 [INSPIRE].

[16] T. Ferber, A. Filimonova, R. Schäfer and S. Westhoff, *Displaced or invisible? ALPs from B decays at Belle II*, *JHEP* **04** (2023) 131 [arXiv:2201.06580] [INSPIRE].

[17] E.M. Riordan et al., *A search for short lived axions in an electron beam dump experiment*, *Phys. Rev. Lett.* **59** (1987) 755 [INSPIRE].

[18] B. Döbrich et al., *ALPtraum: ALP production in proton beam dump experiments*, *JHEP* **02** (2016) 018 [arXiv:1512.03069] [INSPIRE].

[19] BELLE-II collaboration, *Search for axion-like particles produced in $e^+e^-$ collisions at Belle II*, *Phys. Rev. Lett.* **125** (2020) 161806 [arXiv:2007.13071] [INSPIRE].

[20] BESIII collaboration, *Search for an axion-like particle in radiative $J/\psi$ decays*, *Phys. Lett. B* **838** (2023) 137698 [arXiv:2211.12699] [INSPIRE].

[21] BELLE-II collaboration, *Belle II technical design report*, arXiv:1011.0352 [INSPIRE].

[22] E. Fermi, *On the theory of collisions between atoms and electrically charged particles*, *Nuovo Cim.* **2** (1925) 143 [hep-th/0205086] [INSPIRE].

[23] E.J. Williams, *Nature of the high-energy particles of penetrating radiation and status of ionization and radiation formulae*, *Phys. Rev.* **45** (1934) 729 [INSPIRE].

[24] V.M. Budnev, I.F. Ginzburg, G.V. Meledin and V.G. Serbo, *The two photon particle production mechanism. Physical problems. Applications. Equivalent photon approximation*, *Phys. Rept.* **15** (1975) 181 [INSPIRE].

[25] S. Frixione, M.L. Mangano, P. Nason and G. Ridolfi, *Improving the Weizsacker-Williams approximation in electron-proton collisions*, *Phys. Lett. B* **319** (1993) 339 [hep-ph/9310350] [INSPIRE].

[26] BELLE-II collaboration, *The Belle II physics book*, *PTEP* **2019** (2019) 123C01 [*Erratum ibid.* **2020** (2020) 029201] [arXiv:1808.10567] [INSPIRE].

[27] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079 [arXiv:1405.0301] [INSPIRE].

[28] N. Steinberg, *Discovering axion-like particles with photon fusion at the ILC*, arXiv:2108.11927 [INSPIRE].

[29] P. Rebello Teles, D. d'Enterria, V.P. Gonçalves and D.E. Martins, *Searches for axionlike particles via $\gamma\gamma$ fusion at future $e^+e^-$ colliders*, *Phys. Rev. D* **109** (2024) 055003 [arXiv:2310.17270] [INSPIRE].

[30] S. Stengel, C. Sfienti and T. Ferber, *Optimization of the $\pi^0$ reconstruction selections for the Belle II experiment*, Ph.D. thesis, Johannes Gutenberg University, Mainz, Germany (2019) [INSPIRE].

[31] C. Accettura et al., *Towards a muon collider*, *Eur. Phys. J. C* **83** (2023) 864 [*Erratum ibid.* **84** (2024) 36] [arXiv:2303.08533] [INSPIRE].

[32] M. Ruhdorfer, E. Salvioni and A. Wulzer, *Invisible Higgs boson decay from forward muons at a muon collider*, *Phys. Rev. D* **107** (2023) 095038 [arXiv:2303.14202] [INSPIRE].

[33] S. Dawson, *The effective W approximation*, *Nucl. Phys. B* **249** (1985) 42 [INSPIRE].

[34] R.N. Cahn and S. Dawson, *Production of very massive Higgs bosons*, *Phys. Lett. B* **136** (1984) 196 [*Erratum ibid.* **138** (1984) 464] [INSPIRE].

[35] P. Borel, R. Franceschini, R. Rattazzi and A. Wulzer, *Probing the scattering of equivalent electroweak bosons*, *JHEP* **06** (2012) 122 [arXiv:1202.1904] [INSPIRE].