

DISSERTATION

submitted to the

*Combined Faculty of Mathematics, Engineering and
Natural Sciences of Heidelberg University, Germany*

for the degree of

Doctor of Natural Sciences

Put forward by:

Ruth Schäfer

born in:

Hamburg, Germany

Oral examination: October 19, 2022

NEW PHYSICS SEARCHES IN FLAVOUR OBSERVABLES

Referees: Prof. Dr. Susanne Westhoff

Prof. Dr. Tilman Plehn

ABSTRACT (IN GERMAN)

In dieser Doktorarbeit verfolgen wir zwei verschiedene Herangehensweisen um mögliche Theorien der Teilchenphysik jenseits des Standardmodells zu untersuchen, Suchen nach langlebigen Teilchen und effektive Feldtheorien. Dabei folgen wir meinen Veröffentlichungen während meiner Zeit als Doktorandin. Wir untersuchen das Potential von Suchen nach langlebigen Teilchen in den e^+e^- Beschleunigern Belle II und ILC und vergleichen ihre Sensitivitäten mit denen anderer Suchen und Detektoren. Wir stellen fest, dass diese Detektoren sehr sensitiv auf langlebige Teilchenzerfälle sind. Wir untersuchen die Möglichkeit, diese Detektoren mit weit entfernten Detektoren zu erweitern, um ihre Sensitivität auf Suchen nach langlebigen Teilchen weiter zu verbessern, stellen aber fest, dass diese in realistischen Ausmaßen keine großen Verbesserungen erzielen. Wir vergleichen außerdem die Suche nach langlebigen Teilchen in Suchen nach verschobenen Zerfallsvertices mit Suchen nach fehlender Energie und stellen fest, dass beide Suchen neue Regionen des Parameterraums erkunden. Während Suchen nach fehlender Energie ein breiteres Massenspektrum erkunden können, haben Suchen nach verschobenen Zerfallsvertices eine höhere Sensitivität bei hohen Massen. Mit effektiven Feldtheorien untersuchen wir die Flavourstruktur der neuen Physik an hohen Energieskalen, indem wir Wilsonkoeffizienten der SMEFT an Topquark- und Bottomquarkobservablen fiten. In beiden Herangehensweisen legen wir den Schwerpunkt auf Observablen aus der Flavourphysik, und insbesondere auf die Zerfälle von B -Mesonen.

ABSTRACT (IN ENGLISH)

In this thesis, we follow two approaches to explore the space of possible models of physics beyond the Standard Model, long-lived particle searches and effective field theories, following the papers I published during my Ph.D. studies. We explore the potential of long-lived particle searches at the e^+e^- colliders Belle II and ILC, comparing their sensitivities to different other detectors and search methods. We find that these detectors have a strong sensitivity to these searches. We also explore the possibilities of adding far detectors to improve their sensitivities and find that realistic far detectors do not offer a great improvement in sensitivity to the Belle II and ILC detectors. We also compare the different search strategies of missing energy and displaced vertex searches for an axion-like particle model at Belle II, finding that both have the potential to explore new parameter space. While the missing energy search has a broader reach in mass, the displaced search can reach smaller couplings for heavy masses. We use the effective field theory approach to explore the flavour structure of UV physics by fitting SMEFT Wilson coefficients to top and bottom observables. In both approaches, we focus on observables from flavour physics, specifically B meson decays.

PUBLICATIONS

This thesis is based on my collected publications up to this date which have all been written during my time as a Ph. D. student:

- [1]: A. Filimonova, R. Schäfer, S. Westhoff, *Probing dark sectors with long-lived particles at Belle II*, Phys. Rev. D **101**, 095006 (2020), arXiv:1911.03490 [hep-ph]
- [2]: S. Bruggisser, R. Schäfer, D. van Dyk, S. Westhoff, *The Flavor of UV Physics*, JHEP **05**, 257 (2021), arXiv:2101.07273 [hep-ph]
- [3]: M. Borsato et al., *Unleashing the full power of LHCb to probe stealth new physics*, Rept. Prog. Phys. **85**, 024201 (2022), arXiv:2105.12668 [hep-ph]
- [4]: D. Dreyer et al., *Physics reach of a long-lived particle detector at Belle II*, Snowmass Contribution (2021), arXiv:2105.12962 [hep-ph]
- [5]: T. Ferber, A. Filimonova, R. Schäfer, S. Westhoff, *Displaced or Invisible? ALPs from B decays at Belle II*, Publication Pending (2022), arXiv:2201.06580 [hep-ph]
- [6]: R. Schäfer, F. Tillinger, S. Westhoff, *Near or Far Detectors? Optimizing Long-lived Particle Searches at Electron-Positron Colliders*, Publication Pending (2022), arXiv:2202.11714 [hep-ph]
- [7]: The ILC International Development Team and the ILC community, *The International Linear Collider: Report to Snowmass 2021*, Snowmass Contribution (2022), arXiv:2203.07622 [physics.acc-ph]

Section III A focuses on the results of [1]. There, I focus on our results about displaced searches as the section on missing energy searches is largely based on the results of my master thesis [8]. Section II A 5 discusses the work and results of [2], section III B that of [4] and [6], and section III C that of [5]. The paper [3] is a review paper. My contribution to it, section 4.3.1, is based on [1]. Similarly, my contribution, found in sections 10.6 and 11.3 therein, to the paper [7], a Snowmass report, is based on [6]. The papers [1, 2, 4, 5] and [6] are all strongly based on my own work and contributions.

CONTENTS

Publications	
I. Introduction	1
II. Theory	8
A. Effective Field Theories	8
1. Fermi's theory of β decay	8
2. EFT concepts and usage	10
3. EFT construction	13
4. Weak Effective Theory	14
5. Standard Model EFT	16
6. Matching and Running in EFTs	18
B. Extensions of the Standard Model	20
1. Gauge Invariance and Portal Models	20
2. Scalar Portal	23
3. Vector Portal	26
4. Neutrino Portal	26
5. Axions and Axion-Like Particles	27
C. Statistics	28
1. Probability Distributions	29
2. Hypotheses and setting limits	33
III. Long-lived Particles	36
A. Long-lived scalars at Belle II	36
1. Belle II	36
2. Scalar portal and meson decays	37
3. Displaced decay sensitivity	42
B. Far Detectors at electron-positron colliders	45
1. Detector Design	46
2. Models and Modelling	53
3. Long-Lived Particle Reach	56
4. Far Detectors at the ILC	61
C. Complementary LLP signatures	68
1. Long-Lived ALP Decays	69
2. Missing Energy Signatures	75
3. Comparing LLP Signatures	85
IV. SMEFT	88
A. SMEFT operators	88
B. WET operators	89
C. Flavour structure and Minimal Flavour Violation	91
D. MFV in SMEFT	94
E. Matching SMEFT to WET	96
F. Observables	99
G. Fit and fit results	101
V. Conclusions	104
References	107

I. INTRODUCTION

The Standard Model of Particle Physics (SM) describes all particles and their interactions as known to us to this day. It is made up of fermions, spin- $\frac{1}{2}$ particles that make up matter, and bosons, spin-1 particles that mediate the interactions between them. There are two groups of fermions, quarks and leptons, two types of quarks, up-type and down-type quarks, and two types of leptons, charged leptons and neutrinos. Of each of these fermion types there are three generations of particles with different masses. The particles in the different generations are each known by a different name which we refer to as flavour. The up-type quarks are the up quark, the charm quark, and the top quark. The down-type quarks are the down quark, the strange quark, and the bottom quark. The quark flavours are thus given by up, down, strange, charm, bottom, and top. The charged leptons are electrons, muons, and tau leptons. The neutrinos are the electron neutrino, the muon neutrino, and the tau neutrino. All fermions also have antiparticles.

There are three interactions in the Standard Model: the strong interaction, the weak interaction, and the electromagnetic interaction.

The strong interaction is described by the theory of quantum chromodynamics (QCD). It works only on quarks, not on leptons. Quarks are charged under the strong interaction as one of three colour states. Anti-quarks, the antiparticles of quarks, are charged as one of three anti-colours, the complementary charges to each of the colour charges. The strong interaction is mediated by massless gluons, eight bosons carrying different non-unity combinations of colour and anti-colour charges. The strong interaction is named such because at low energies it becomes very strong, gaining a non-perturbative coupling and confining the quarks and gluons. This means that at low energies, quarks and gluons only exist in the form of bound state particles called hadrons. There are two classes of hadrons, mesons and baryons¹. Mesons are made up of a quark and an antiquark, while baryons are made up of three quarks (and antibaryons of three antiquarks).

The weak interaction is mediated by the massive W and Z bosons. The Z boson is neutral, while the W bosons are charged under the electromagnetic interaction. The Z boson couples to uncharged, flavour-diagonal fermion currents. Flavour-diagonal means that both fermions in the current are of the same flavour. The W bosons couple only to left-handed fermion currents. As they are charged, they couple to charged fermion currents, like charged lepton neutrino currents or up-type down-type quark currents. In addition to changing the type of fermion through their coupling, they also change the flavour, for example coupling a bottom quark to a charm or an up quark. We call this

¹ Additional bound states of quarks and antiquarks have been predicted and discovered in the form of tetraquarks and pentaquarks which are made up of four and five quarks and antiquarks, respectively.

a flavour changing charged current (FCCC). In contrast, we call an interaction which changes flavour but is not charged, coupling for example a bottom quark to a strange quark, a flavour changing neutral current (FCNC). Since the Z boson couples flavour-diagonally, these do not exist at tree-level in the Standard Model, though they can be loop-induced.

The electromagnetic interaction is described by the theory of quantum electrodynamics (QED) and mediated by the massless photon. The charged leptons carry a negative charge of -1 , neutrinos are neutral, up-type quarks carry a positive charge of $\frac{2}{3}$, and down-type quarks carry a negative charge of $-\frac{1}{3}$. Antiparticles carry a charge given by the same magnitude of their corresponding particle but with a flipped sign. The charged W bosons carry a charge of $+1$ and -1 , respectively, being each others' antiparticles.

The last particle of the Standard Model is the Higgs boson. It is a scalar particle of spin 0 and couples to fermion currents with changing chirality, that is of one left-handed and one right-handed fermion of the same type and flavour, as well as to weak bosons. Through its coupling, the W and Z bosons, the quarks, and the charged leptons gain their masses. The Standard Model contains no right-handed neutrinos, so that the Higgs boson is not predicted to couple to them. The Higgs boson was the last particle of the Standard Model to be discovered, which happened in 2012 at the LHC [9, 10].

In the decade since the Higgs was discovered, we have not gained any grand new insights into the nature of the Standard Model. Searches for the famously favoured theory of supersymmetry [11] which many expected the LHC to find evidence of [12] have so far come up empty [13, 14]. And beyond it, we also have not discovered any other particles or forces beyond the Standard Model (BSM) or any direct particle physics evidence of new physics (NP) yet.

While the predictions of the Standard Model have mostly been very confirmed by measurements, there are hints of physics beyond it to be found, too. Many experiments register anomalies in the measurements of some B meson observables [15, 16]. The anomalous magnetic moment of the muon [17], $(g - 2)_\mu$ was only recently confirmed by the Muon $g-2$ Collaboration at Fermilab [18]. Through neutrino oscillations [19], we have determined that neutrinos carry a mass which is unexplained in the Standard Model due to the absence of right-handed neutrinos therein. The hierarchy of the fermion masses, and the absence of CP violation in the strong interaction are both unexplained by the Standard Model. Above all, the indisputable astrophysical evidence for dark matter [20] shows us that there are parts of particle physics that lie beyond the Standard Model and that we are yet to discover.

In this thesis, we will focus on two separate ways of exploring the space of physics beyond the Standard Model in interesting ways. On the one hand, we will consider searches for long-lived particle decay signatures at collider experiments. On the other, we will use precision data collected at colliders to illuminate the space of effective field theories to see what we can tell about possible new physics at currently unobtainably high energies through obtainable observables. Both cases will involve flavour observables.

With flavour physics and flavour observables we generally mean the decays of mesons through the weak interaction. The B^+ meson, a meson made up of an anti-bottom quark and an up quark, for example, decays through the weak interaction into a multitude of final states which are already and continue to be precisely measured by several experiments focussing on flavour observables, like LHCb and Belle II. Particularly interesting for new physics searches are observables that are predicted to be very small in the Standard Model as this means that new physics contributions of the same size would be easier to detect over the smaller Standard Model expectation. An example for this are observables involving FCNCs which are loop-suppressed in the Standard Model but could possibly exist at tree-level in new physics models. The decay $B^+ \rightarrow K^+ X$ is such an FCNC decay, with the B^+ meson's bottom quark decaying via a W boson loop to the kaon K^+ 's strange quark and other final state particles X .

When looking for light new physics, using a specific well-motivated model to predict a decay like $B^+ \rightarrow K^+ X$ with X being a new particle is often a good strategy. Historically, these searches have usually been either prompt searches or missing energy searches. Prompt searches are searches for particles that are produced at the interaction point (IP) of the collider, in this case, directly from a B meson decay, and decay immediately, or at least at a distance so close to it that we cannot resolve its decay vertex as different from its production point in the detector. Missing energy searches on the other hand are searches for neutral particles produced within the detector that do not decay within it at all, either decaying far outside of it or being stable themselves. Between these two cases there are many orders of magnitude in lifetime that are not probed by either approach. These lifetimes can be probed by long-lived particle searches which have recently become more popular.

As lifetimes generally depend on coupling anti-proportionally,

$$\tau \propto \frac{1}{|c|^2}, \quad (1)$$

we expect that searches for particles with longer lifetimes let us probe smaller couplings. Missing energy searches also allow us to look for particles with longer lifetimes than prompt searches, and even stable particles, which do not decay within the detector

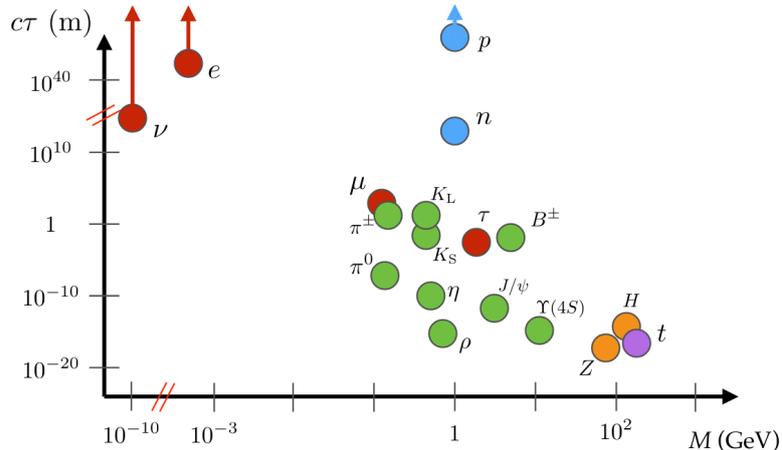


FIGURE 1: Selection of Standard Model particles, shown in the mass-lifetime plane. Leptons are marked in red, baryons in blue, mesons in green, the Z and Higgs boson in orange, and the top quark t in purple. Taken from [21] with credit going to Brian Shuve [22].

volume. On the other hand, it is difficult to determine much about a model from a missing energy search as we do not get to measure the particle as directly as we can when we measure its decay.

We refer to particles as long-lived particles (LLP) if they decay at a measurable distance from the interaction point of a collider. Figure 1 shows selected Standard Model particles in the plane of their mass and lifetime. As we can see, the Standard Model contains, next to the stable electron, proton and neutrinos, several particles with long lifetimes. Notable among them are the kaons K^\pm , K_S^0 , K_L^0 , pions π^\pm and muons μ^\pm . Since we know the masses and lifetimes of the Standard Model LLPs well, we can expect there to be little background in other regions of the mass-lifetime parameter space. This is in contrast to both prompt and missing energy searches which tend to have high Standard Model backgrounds.

While we are reasonably certain that there are no charged light BSM LLPs that we have missed, whose tracks would be visible in our tracking detectors and which we would therefore expect to have already found if they existed, neutral light BSM LLPs are harder to exclude, as they do not leave tracks in the detector. Instead, we have to search for the signatures of their decay to find them. In figure 2, we show a number of example signatures for the decays of long-lived particles. Shown are several notable classes of signatures.

First, there are signatures where the neutral long-lived particle (shown as a dotted black line) leaves the detector, leaving behind missing energy. We see this in the first three decays starting from the top and going clockwise. These are a disappearing track, where

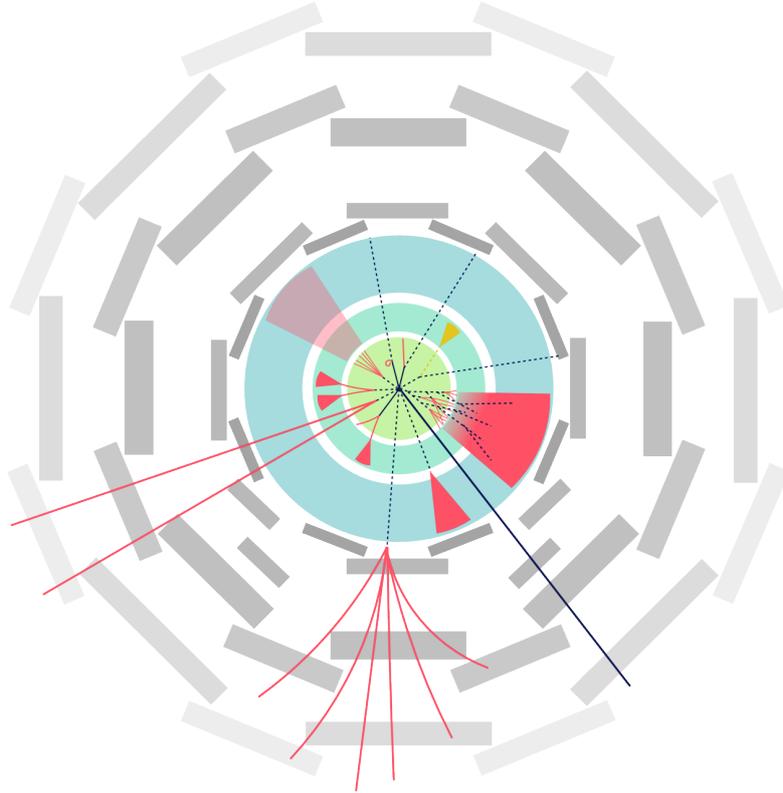


FIGURE 2: Different signatures of long-lived particle decays within a detector. The decay products shown in red are charged particles, those in yellow are photons, while dotted black lines signify a neutral long-lived particle. Taken from [21] with credit going to Heather Russell [23, 24].

the LLP is produced from a charged particles track by emitting a very soft charged particle that is not seen as a track due to its softness; a kinked track, where the charged particle produced in conjunction with the LLP is hard enough to leave a track; and a single photon which is produced outside of the interaction point.

Second, there are signatures where the LLP decays into pairs of charged particles, like muons or light mesons that are registered in the tracking detector and the calorimeter. In addition to pairs of charged particles, an LLP can also decay into a jet or a pair of jets, or into a shower of charged particles or photons. These decays are generally called displaced decays and their detection tends to rely on finding the vertex that the decay products point to, as it is at a distance from the interaction point.

A last signature, the emergent jet shown in the bottom right of figure 2, occurs when the LLP decays to further particles which are also neutral and invisible to the detector before all of these particles decay back to measurable Standard Model particles. This results in a jet or shower of Standard Model particles which may point towards a common vertex, but whose vertex is not shown by the produced tracks, as the shape of the cone

is determined by decays within a dark sector.

We will focus in this thesis on using displaced LLP decays into charged particle pairs to explore the parameter spaces of several LLP models, specifically in section III.

The other method we will use for exploring new physics relies on effective field theories. Effective field theories (EFTs) are a model-independent way of encapsulating the interactions between the relevant degrees of freedom of a theory without needing to define or know the underlying theory that these interactions follow from. This only works when none of the degrees of freedom of the new physics are relevant at the scale which we describe. Thus, this approach works well for very heavy new physics.

Since we do not impose any underlying theory, there is usually a large number of parameters. To fit these, we also need to use data from many different observables at many different experiments, and their predictions expressed in the parameters of the effective theory. From the fit results and the patterns therein, we then try to reconstruct any properties of the underlying theory that may become apparent.

In particular, in section IV of this thesis, we will try to explore the flavour structure of high-energy BSM physics through a fit of the Standard Model Effective Field Theory (SMEFT) to both high-energy top-quark observables and B meson decay flavour observables. To include the low-energy observables of flavour physics, we use match the SMEFT onto the Weak Effective Theory (WET). We express the flavour structure of the SMEFT operators in terms of Minimal Flavour Violation (MFV).

The two approaches of LLP searches and global fits to EFTs are complementary. LLP searches look for direct effects of long-lived particles light enough to be directly produced within the detector. The EFT approach, on the other hand, deals with new physics so heavy that it only has indirect effects on observables accessible to us. With the LLP searches we are proposing new search strategies, whereas in the EFT approach we are exploring the parameter space of the EFT with already existing data. Still, both approaches have the potential for uncovering more hints of BSM and possibly discovering it.

In section II of this thesis, we will lay the groundwork for the theory that is necessary in the following sections. Therein, we will first explain effective field theories in section II A, introducing them with the help of Fermi's theory of β decay before explaining their concepts and the specific EFTs we will work with. Then, we will introduce extensions of the Standard Model in section II B, first going over the concept of gauge invariance and then through several common portal models. Lastly, we will give a short introduction to the statistics necessary for this thesis in section II C, first introducing some basic

concepts of probability and then discussing the setting of limits on parameter spaces. In section III, we will go through the work and results of the papers I have written on long-lived particle searches, with section III A covering [1], section III B covering [4] and [6], and section III C covering [5]. Section IV will then cover my paper on EFTs, [2], before we conclude with section V.

II. THEORY

A. Effective Field Theories

The idea behind effective theories is that we do not need to know the full detail of a theory to be able to make predictions from it. For example, an architect constructing a skyscraper will need to know about the properties of steel and concrete, but not about the microscopic forces that keep the materials stable or how their properties come to be. For the task of constructing a skyscraper, all that is relevant are the macroscopic properties of the building materials. This is often the case in daily life: if we consider processes at a specific size scale, we can ignore effects of much smaller sizes.

Similarly, in particle physics we may also want to use effective theories, either because observables are easier to calculate than in the full theory or because we may not know the full theory to begin with. To use it in particle physics, our effective theory has to be a quantum field theory (QFT). It needs a Lagrangian that we can extract Feynman rules from to calculate observables, and needs to be regularisable and renormalisable [25]. We call these theories effective field theories (EFTs). In particle physics, the most important scale in our processes tends to be an energy or momentum, often the centre-of-mass energy or the transmitted momentum of a process.

1. Fermi's theory of β decay

The first prominent example of an EFT in particle physics is Fermi's theory of β -decay. When Fermi proposed his "quantitative" theory in 1933² [26, 27], Wolfgang Pauli had only just proposed the necessity of the neutrino a few years prior [29]. A pressing question about β decay then was where the electron and neutrino come from, as they could not otherwise be observed in the nuclei. Fermi boiled the process down to a simple equation of particles before and after the decay,

$$n \rightarrow p^+ e^- \bar{\nu}_e, \quad (2)$$

a neutron decaying to a proton, an electron and an (electron anti-)neutrino. This equation implies that the electron and neutrino are produced rather than just procured from the nucleon, and that thus their numbers are not constant in time, which Fermi proposed in analogy to his theory of spectroscopic absorption and emission of photons [28]. Also in analogy to the emission of photons, Fermi proposed a structure of vector currents for both the nucleon and the lepton side [30]. As there was no indication to a substructure

² He proposed it in 1933 at a conference in Brussels and submitted it to *Nature*, but was rejected. He instead published two versions in 1934 in an Italian [26] and a German [27] journal [28]

yet, Fermi proposed a simple contact interaction between the nucleons and the fermions. Fermi used the language of second quantisation, but in today's Lagrangian notation his theory can be expressed as [28]

$$\mathcal{L}_\beta = g (\bar{\psi}_p \gamma^\mu \psi_n) (\bar{\psi}_e \gamma_\mu \psi_\nu) \quad (3)$$

where ψ describe fermion states and the coefficient g was experimentally determined.

Fermi's interaction turned out to be much more general than just describing β -decay. With different fermions in the two currents, not only other nuclear phenomena like muon capture could be explained, but also processes like pion or muon decay [31]. For this reason, the term Universal Fermi Interaction (UFI) was coined in the 1940s [32–34]. In the 1950s, upon closer measurements of many of the relevant observables, UFI was in a crisis until the tension between theory and experiment was resolved with the introduction of the now famous V-A structure [35, 36]. With these changes, the Lagrangian for UFI now looks like [37]

$$\mathcal{L}_\beta = G_F \left(\bar{u} \gamma^\mu \frac{1 - \gamma^5}{2} d \right) \left(\bar{e} \gamma_\mu \frac{1 - \gamma^5}{2} \nu \right). \quad (4)$$

The coefficient G_F is today known as the Fermi constant, $G_F = 1.166 \cdot 10^{-5} \text{ GeV}^2$ [38].

Already in 1949, Lee, Yang, and Rosenbluth proposed that there might be an intermediate vector boson (IVB) mediating these UFI processes [39]. They labelled this boson W for its weak interaction. In the 1960s, Glashow [40], Weinberg [41], and Salam [42] proposed the electroweak theory that contained such a W boson. The W boson was finally discovered in 1983 at the $Spp\bar{p}S$ at CERN [43–48]. The weak interaction regarding the interaction between the W boson and fermions is described by the Lagrangian [38]

$$\mathcal{L}_{\text{weak}} = - \frac{g}{\sqrt{2}} \bar{f}_L \gamma^\mu (W_\mu^+ T^+ + W_\mu^- T^-) f_L \quad (5)$$

with g the weak coupling, W^\pm the W bosons, T^\pm the weak isospin raising and lowering operators, respectively, and the fermions f as flavour eigenstates. Hereby, the four-vertex of Fermi's interaction was resolved into the heavy mediator particle, the W boson, that couples via two three-vertices to pairs of fermions, as illustrated in figure 3.

Here and in the following, we will be treating all quarks in up-alignment, which means that the up-quark flavour eigenstates are defined to coincide with their mass eigenstates, while the down-quark flavour eigenstates are related to their mass eigenstates via the CKM matrix V_{CKM} , $d_{Lp}^{(\text{mass})} = V_{pr} d_{Lr}^{(\text{flavour})}$. We also only consider the flavour eigenstates of neutrinos, ν_e , ν_μ , and ν_τ .

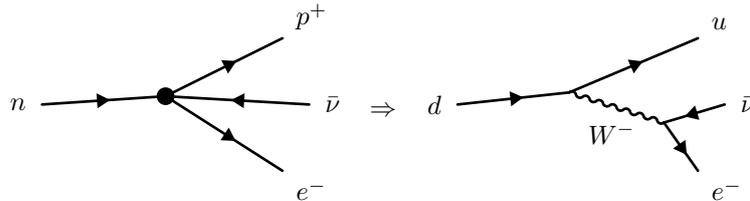


FIGURE 3: The resolution of the W -boson mediation of the weak interaction from the four-point interaction in Fermi's theory, cmp. [49].

With the discovery of the full theory, we are now able to express G_F , the parameter in our effective or quantitative Lagrangian, in terms of the parameters from the full theory, [50]

$$G_F = \frac{\sqrt{2}g^2}{8m_W^2} \quad (6)$$

where g is the weak coupling and m_W is the mass of the W -boson. Creating such a relation between the EFT and the full theory is what we call matching. We will describe this process more closely in section II A 6.

2. EFT concepts and usage

With the analogy of Fermi's theory of β decay from the last section, we can more easily define some useful terms for EFTs. Scale separation is the difference between the scale of a process μ and the scale of its underlying fundamental physics Λ . For our example, the scale of β decay is at nuclear energies while the scale of its fundamental underlying physics, the W boson exchange, is at energies around the W boson mass. If there is no scale separation, that is, the scale of the process μ is close to Λ , the EFT is generally not valid anymore. In this case, Fermi's interaction works well for low-energy observables, but once we get close in energy to the W boson mass, we can see its resonance in our observables, which is not explained by Fermi's theory. We call the scale Λ near which the EFT becomes invalid its cut-off scale. For Fermi's theory, this cut-off is the W boson mass, $\Lambda = m_W$. The finite range of validity of EFTs is a feature that not all QFTs share. We call QFTs that do not have a cut-off scale UV-complete or full theories. We often call the theory that an EFT is a low-energy limit of its UV theory, even when it is not UV-complete.

We generally assume that, as in the case of Fermi's theory and weak interactions, all the physics inducing effective operators comes, approximately³, from the same scale Λ .

³ In addition to effects from the W^\pm boson at the weak scale, we also have effects from Z^0 bosons, top quarks t , and the Higgs boson h^0 . While these all have different masses, they still become relevant at very similar scales and we can generally say that all their effects on low-energy physics observables are well-described by a common cut-off scale $\Lambda \sim m_W$. If there are significant differences in scale between different particles in the UV theory, we may find it useful to construct an intermediate EFT of the effective effects of the heavier particles, to deal with the separation of scales within the UV theory.

Then, since the cut-off scale is the same for all terms in the Lagrangian of an EFT, it makes sense for us to separate terms in the Lagrangian into three components: the operators, which are comprised of all fields and the group structure that connects them, a power of the cut-off scale Λ and the Wilson coefficient (WC), which is a dimensionless scalar that contains the numerical factors that are left over. In the case of Fermi's theory, the operator is $(\bar{\psi}_L \gamma^\mu \psi_L) (\bar{\psi}_L \gamma_\mu \psi_L)$, while G_F contains both the Wilson-coefficient and the cut-off scale. Resolving G_F as seen in equation (6), we find the scale coefficient to be $\Lambda^2 = m_W^2$ and the Wilson coefficient $\frac{\sqrt{2}g^2}{8}$.

In addition to experimentally resolving the vertex or discovering the resonance of the mediating particle, another way to see at which scale the theory will break down is through partial wave unitarity. Partial wave unitarity puts a bound on the total growth possible for a cross section as derived from the optical theorem. This usually means that at a certain energy near or above the scale Λ , the calculated cross section is larger than allowed (by the optical theorem), which simply tells us that a UV completion is necessary. The derivation is beyond the scope of this work, but may be found in [51].

The success of an effective theory as a stepping stone towards a full theory is shown well by Fermi's theory. While the full theory was unknown and kinematically unprobeable at the time, the effective theory of a four-point interaction was able to describe the data. It also helped us classify that all weak decays fell under the same kind of interaction, without knowing what exactly this interaction was. The value for G_F , while at first experimentally determined, can now, with the knowledge of the full theory, be expressed by its parameters. This is a great example for why we still use EFTs today. We have good reasons to believe that physics beyond the Standard Model exists, but since we have not found supersymmetry (nor any other new particles or strong hints of a specific UV theory) there has not been an obvious contender for what theory should describe all new physics. We can (and do) still construct models and explore observables in their parameter space, and we will go into this in section IIB, but it is just as valid an approach to searching for new physics to use an EFT to encapsulate what we observe today and to look for hints of the full theory from there, as was done with the UFI in the past.

Another reason to work with EFTs is in cases where the full theory may be very difficult to calculate [52]. This is the case, for example, with quantum chromodynamics (QCD), the theory of quarks and gluons, and chiral perturbation theory (χ PT), which deals with light mesons [25]. Since QCD becomes non-perturbative at low energies, calculations on the basis of quarks and gluons become very difficult if not impossible. Since they form bound states in the form of mesons, it makes sense to calculate observables at these low energies directly in terms of the mesons instead of with non-perturbative QCD. In

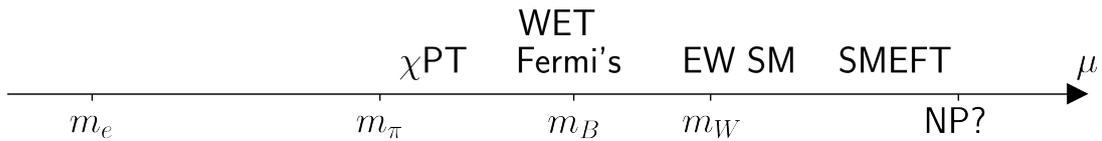


FIGURE 4: A range of scales μ with the masses of Standard Model particles at those scales (beneath the axis) and the theories that can be used to describe the interactions at those scales well (above the axis). More information on the theories shown is given in the text.

this case, matching the theories to each other is a bit more difficult, since the states of one theory cannot be neatly described by the other. We thus often use experimentally determined parameters for the EFT, until it is possible to calculate a matching between the EFT and its full theory. In χ PT specifically, matching is difficult but being improved continuously [53, 54].

One should keep in mind, that the UV theory of an EFT does not necessarily have to be a UV-complete theory [52]. Fermi's theory is the EFT to the theory of the weak interaction, but the weak interaction is only part of the electroweak theory and the Standard Model and the Standard Model itself is not complete and may only be the low-energy EFT of some as-of-yet unknown UV theory itself. Between each of these theories, there is a separation of scales: at β -decay energies, Fermi's interaction is a good description, while at energies of the mass of the W -boson, we need the full description of the weak interaction to describe processes, and at energies much higher than we can currently access with our experiments, we may resolve even higher-scale physics.

Figure 4 shows different theories at the scales that they are valid at. Fermi's theory of the weak interaction is shown there as being valid for scales at around the B -meson mass, which are well below the mass of the W -boson. If we get closer to it, Fermi's theory becomes less valid as the vertex is resolved into a W boson exchange. Then, we instead use the Standard Model (SM) with its electroweak theory (EW) to make predictions. Above the scales of the SM, there may be new physics (NP), which may make the SM calculations invalid from some scale on. To encapsulate well the behaviour at the higher scales, we can use the Standard Model Effective Field Theory (SMEFT), which we will go into detail on in section II A 5. LEFT and WET, which are listed together with Fermi's theory at the scale of meson masses, are the Low Energy Effective Theory and the Weak Effective Theory, respectively, and are a modern approach to Fermi's theory using different coefficients for every interaction term to encapsulate possible inconsistencies in the weak interaction that indicates new physics. We will describe it more closely in section II A 4. Also shown is chiral perturbation theory (χ PT) which describes the

interactions of light mesons below the scale where the non-perturbativity of the strong interaction makes the direct calculations from quarks and gluons accessible. For readability's sake, its UV theory of QCD has not been added. It becomes relevant at scales of a few GeV.

3. EFT construction

When constructing an EFT, it is important to first determine which the relevant fields are. In the case of Fermi's theory, these are fermions — electrons, neutrinos, neutrons and protons — while for χ PT these are light mesons. Often the relevant particles are obvious from the scale of observations as the particles which can be produced as free, physical particles at this scale. Then, we need to know the symmetries that are valid in our EFT. We may know these from observation or if we know the full theory, we can extract them from it. If we do not know the symmetries of our EFT, we do not impose any and may find that we see correlations between our coefficients later and learn about new symmetries that way. Based on the particles and the symmetries, we construct all possible terms for the EFT's Lagrangian, making sure to adhere to Lorentz- and gauge-invariance, leading to Lagrangian like [52]

$$\begin{aligned}\mathcal{L}_{\text{EFT}} &= \mathcal{L}_{\leq 4} + \sum_i \frac{\mathcal{C}_i^{(5)} \mathcal{O}_i^{(5)}}{\Lambda} + \sum_i \frac{\mathcal{C}_i^{(6)} \mathcal{O}_i^{(6)}}{\Lambda^2} + \dots \\ &= \mathcal{L}_{\leq 4} + \sum_d \sum_i \frac{\mathcal{C}_i^{(d)} \mathcal{O}_i^{(d)}}{\Lambda^{d-4}}\end{aligned}\tag{7}$$

with the operators $\mathcal{O}_i^{(d)}$ being Lorentz- and gauge-invariant combinations of the fields we have determined to be relevant for the EFT. We construct these terms mass-dimension by mass-dimension, and to each one we give a coefficient, the Wilson coefficient, and a power of a scale Λ that describes the scale near where the effective theory breaks down and the full theory is expected to lie. For Fermi's theory, the scale is $\Lambda \sim m_W$ and the Wilson coefficient is $G_F \cdot m_W^2$. The scale is used to normalise every term of the Lagrangian to mass-dimension 4.

The EFT thus works as an expansion in the small parameter of the quotient of scales, μ/Λ , that is, the ratio of the scale of a process and the cut-off scale of the theory, which describes how much the full theory influences processes that we calculate in the EFT. This forms an additional dimension of perturbation theory, where EFT observables are calculated up to a finite order in the operator's mass dimensions as well as a finite order in the fundamental couplings.

4. Weak Effective Theory

The Weak Effective Theory⁴ [55, 56] (WET) is an EFT that operates at scales far below the mass of the W boson, similar to Fermi's interaction. Throughout the history of Fermi's interaction, it became clear that, to the precision available to experiments at the time, all processes described by it were found to share the same coupling constant, G_F . As we know, this culminated in the weak theory that confirmed and explained the value of G_F . Nowadays, we have reasons to expect that there is physics beyond the Standard Model (BSM) and experiments with much higher precision checking the validity of the Standard Model's predictions. For low-energy observables, the WET encodes all allowed operators between quarks, leptons, photons and gluons, excepting the top-quark which is heavier than the W boson and thus not a resonant degree of freedom at low energies. In contrast to the UFI, which has a common coefficient⁵ for all operators, G_F , the WET introduces separate Wilson coefficients for every one of its operators. Since the WET can be seen as the low-energy EFT of the SM, we have predictions for all of these WCs through matching with the SM. However, we want to use the WET to see possible effects of new physics beyond the SM, which is why we use free parameters for our WCs.

In addition to the photon field, A^μ , and the gluon fields, $G^{a\mu}$, we then have $n_u = 2$ up-type quarks, u_{Lr}, u_{Rr} , $n_d = 3$ down-type quarks, d_{Lr}, d_{Rr} , $n_\ell = 3$ charged leptons, ℓ_{Lr}, ℓ_{Rr} , and $n_\nu = 3$ neutrinos, ν_{Lr} , where L and R denote the particle's chirality and r is its flavour. We consider neutrinos as left-handed only and give them no SM mass-term.

The WET Lagrangian is made up of all $SU(3)_c \times U(1)_Q$ invariant operators, sorted by their mass dimension. The operators with dimensions smaller or equal 4 are the QED and QCD Lagrangian [56]

$$\begin{aligned} \mathcal{L}_{QED+QCD} = & -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{4}G^{a\mu\nu}G_{\mu\nu}^a + \theta_{QCD}\frac{g_s^2}{32\pi^2}G^{a\mu\nu}\tilde{G}_{\mu\nu}^a \\ & + \sum_{\psi=u,d,\ell,\nu_L} \bar{\psi}_r i\gamma_\mu (\partial^\mu + ieQ_\psi A^\mu + ig_s T^a G^{a\mu}) \psi_r \\ & - \sum_{\psi=u,d,\ell} \bar{\psi}_{Rr} (M_\psi)_{rs} \psi_{Lr} + h.c. \end{aligned} \quad (8)$$

with the photon and gluon field strength tensors $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ and $G^{a\mu\nu} = \partial^\mu G^{a\nu} - \partial^\nu G^{a\mu} + g_s f^{abc} G^{b\mu} G^{c\nu}$, the QCD CP phases θ_{QCD} , the dual field strength tensors of the gluon $\tilde{G}^{a\mu\nu} = \varepsilon^{\mu\nu\rho\sigma} G_{\rho\sigma}^a$, the QED coupling e , the QED charge of ψ as Q_ψ , the strong coupling g_s , the generators T^a and f^{abc} of $SU(3)_c$ in the fundamental and

⁴ This theory is also often referred to as Low-energy Effective Field Theory (LEFT) in the literature.

⁵ While the coefficient G_F is the same for all currents, flavour-changing quark currents bring CKM-matrix elements with them that change the overall coupling strength between operators with different fermions.

adjoint representation, and the mass matrix M_ψ for ψ in flavour space. Both here and in the following, we use Greek letters μ, ν, ρ, σ , etc. for space-time indices, a, b, c , etc. as the colour indices of $SU(3)_c$, i, j, k, ℓ , etc. as the weak indices of $SU(2)_L$, and p, r, s, t , etc. as flavour indices.

Beyond QED and QCD, we again construct all operators that are invariant under $SU(3)_C \times U(1)_Q$, dimension by dimension. We consider here operators up to dimension six. At mass dimension three, we have the Majorana mass term for neutrinos,

$$\mathcal{O}_{\nu}^{pr} = \nu_{Lp}^T C \nu_{Lr} \quad (9)$$

where C is the charge conjugation operator and the first neutrino ν^T is transposed. At dimension five, we find dipole operators for the quarks and leptons, e. g.

$$\mathcal{O}_{d\gamma}^{pr} = (\bar{d}_{Lp} \sigma^{\mu\nu} d_{Rr}) F_{\mu\nu}, \quad (10)$$

At dimension six, we find gluon self-interaction operators, e. g.

$$\mathcal{O}_G = f^{abc} G_\mu^{a\nu} G_\nu^{b\rho} G_\rho^{c\mu}, \quad (11)$$

and four-fermion operators, e. g.

$$\mathcal{O}_{de}^{V,LR}{}_{prst} = (\bar{d}_{Lp} \gamma^\mu d_{Lr}) (\bar{e}_{Rs} \gamma_\mu e_{Rt}). \quad (12)$$

The four-fermion operators can further be separated by the fermions being coupled (denoted by a subscript of two to four out of u, d, e, ν), their type of currents (denoted by superscripts of S for scalar currents, V for vector currents, and T for tensor currents), their chiralities (denoted by superscripts of two from L and R after the type of current), and the gauge-representation of the currents (denoted by a superscript of 1 or 8 between the current and chirality labels for operators that can have currents that are colour-octets). For example,

$$\mathcal{O}_{uu}^{V8,LR}{}_{prst} = (\bar{u}_{Lp} \gamma^\mu T^a u_{Lr}) (\bar{u}_{Rs} \gamma_\mu T^a u_{Rt}) \quad (13)$$

is labelled uu for its two currents of down quarks, V because they are vector currents, LR because the first current is left-handed and the second current is right-handed, 8 because the currents carry colour-charge in the adjoint representation and $prst$ for its quarks' flavour indices. In addition to these labels, we often differentiate between operators that conserve the accidental SM symmetries of baryon and lepton number and ones that violate them.

For the preceding examples and all others following in section II, we use the operator basis from [56, 57] which is complete, intuitive, and commonly used in the field. In later chapters of this thesis, we will instead be working with the EOS basis of operators [58, 59] to ease the use of the EOS software [60]. This basis is constructed to be particularly convenient for interpreting the measurements of flavour observables, for example by defining the operators to contain factors of coupling, CKM coefficients, and masses that commonly appear in the corresponding processes. This is done such that the SM contribution to these Wilson coefficients is of order one, $\mathcal{O}(1)$, and independent of the relevant quark masses and also dimensionless. We will explore the operators relevant to our analysis in detail later. To avoid any confusion between SMEFT and WET operators of the same or similar names, we will denote WET operators as \mathcal{O}_i and WET Wilson coefficients as C_i .

5. Standard Model EFT

The Standard Model Effective Field Theory [61, 62] (SMEFT) is an EFT of all possible interaction terms between Standard Model particles. In the absence of knowing the true UV theory for physics beyond the Standard Model, we can predict observables with it and fit their experimental values to it, to explore the space of new physics without imposing⁶ a model on it.

The SMEFT is constructed from all SM fields using the SM gauge symmetries. Explicitly, these are the left-handed quark q_r and lepton ℓ_r doublets, the right-handed up-quarks u_r , down-quarks d_r , and electrons e_r , the weak bosons W_μ^i of the symmetry group $SU(2)_L$, the hypercharge boson B_μ of $U(1)_Y$, and the gluons G_μ^a of $SU(3)_c$, as well as the scalar Higgs doublet H .

We use the SM operators up to dimension 6 here, as in the WET case. A commonly used basis for SMEFT is the Warsaw basis [64] which we will also be using.

Under the SM symmetry group, the only allowed dimension 5 operator is

$$O_{\nu\nu} = \varepsilon_{ij}\varepsilon_{kl} (\ell_{ip}^T C \ell_{kr}) H_j H_l \quad (14)$$

which, after electroweak symmetry breaking, gives a neutrino Majorana mass term similar to equation (9) and where $\varepsilon_{ij\dots}$ is the fully anti-symmetric Levi-Civita tensor.

The dimension five operators we saw in the WET, the dipole operators, are now forbidden

⁶ While the SMEFT strives to be model-independent, there are some caveats to this. The SMEFT treats the Higgs doublet as one of its fields, while the Higgs EFT (HEFT) treats the Higgs boson and its associated Goldstone bosons as independent fields, which is sometimes seen as the more model-independent approach [63]. In addition, low-energy new physics cannot be modelled through a high-scale EFT, so for searches for light new physics, a different mode of operations is needed.

due to the addition of the weak $SU(2)_L$ to the symmetry group. They instead show up at dimension six with an additional Higgs doublet, for example compare

$$O_{dW} = (\bar{q}_p \sigma^{\mu\nu} d_r) \tau^i H W_{\mu\nu}^i \quad (15)$$

with equation (10). Similar to equation (11) in the WET, we also find gluon self-interaction operators, e. g.

$$O_G = f^{abc} G_\mu^{a\nu} G_\nu^{b\rho} G_\rho^{c\mu} \quad (16)$$

though now we find the same type of operator also for the W bosons which are non-resonant degrees of freedoms at WET scales, e. g.

$$O_W = \varepsilon^{ijk} W_\mu^{i\nu} W_\nu^{j\rho} W_\rho^{k\mu}. \quad (17)$$

Of course, SMEFT also has many four-fermion operators, for example

$$O_{qe} = (\bar{q}_p \gamma^\mu q_r) (\bar{e}_s \gamma_\mu e_t) \quad (18)$$

where we notice that in comparison to equation (12) here we have the right-handed charged lepton current coupled to a left-handed quark doublet current, that is, both the left-handed up-quark current and the left-handed down-quark current instead of coupling to them independently like in the WET. This is a result of the weak $SU(2)_L$ being a part of the SM (and thus the SMEFT) gauge group. As in the WET, the four-fermion operators are further divided into categories based on the fermions they couple together (denoted as q, ℓ, u, d, e) and the gauge-representations of the currents ((1) in case of a unitary current, if non-unitary versions of the operator exist as otherwise this label is omitted, (3) for weak-triplet currents in the adjoint representation of $SU(2)_L$, and (8) for colour-octet currents in the adjoint representation of the strong $SU(3)_c$). While these operators also differ in the types of currents and the chiralities of the fermions, these do not need to be explicitly labelled as they are easily inferable from the fermion labels. As an example,

$$O_{qu}^{(8)} = (\bar{d}_p \gamma_\mu T^a q_r) (\bar{u}_s \gamma^\mu T^a u_t) \quad (19)$$

is labelled qu for having a left-handed quark current and a right-handed up-quark current, (8) for the fact that the two currents are both colour-octets, and $prst$ for its flavour indices. If the currents have mixed chiralities, all fermions in the corresponding current

are listed, e. g.

$$O_{\ell edq} = (\bar{\ell}_p^j e_r) (\bar{d}_s q_t^j). \quad (20)$$

Due to the addition of the Higgs boson and the W bosons, which are not resonant degrees of freedom in the WET, the SMEFT contains additional operators combining these fields, for example

$$O_H = (H^\dagger H)^3 \quad (21)$$

$$O_{dH} = (H^\dagger H) (\bar{q}_p d_r H) \quad (22)$$

$$O_{HWB} = H^\dagger \tau^i H W_{\mu\nu}^i B^{\mu\nu} \quad (23)$$

A full list of operators can be found in [64]. We will discuss the operators relevant for our analysis in detail later. To avoid any confusion between SMEFT and WET operators of the same or similar names, we will denote SMEFT operators as O_i and SMEFT Wilson coefficients as C_i .

6. Matching and Running in EFTs

Like all other QFTs, EFTs also need to be renormalised. For a review of renormalisation, the need for it, and renormalisation group equations, see [51, 65]. In consequence, the couplings of EFTs may run with scale, that is, an EFT's Wilson coefficients depend on the scale μ of the observable they are a part of, $C_i = C_i(\mu)$. This is in general not just a prefactor to the original Wilson coefficient at the UV scale Λ , but often involves the mixing of operators, that is, the Wilson coefficient of an operator at a scale μ is a function of all (or many) of the Wilson coefficients at the UV scale,

$$C_i(\mu) = \sum_j f_{ij}(\mu, \Lambda) C_j(\Lambda) \quad (24)$$

with f_{ij} being functions specific to the Wilson coefficients. This means that even if a coefficient is expected to be $C_i(\Lambda) = 0$ at the UV scale, it may still be present with non-zero values at other scales $\mu \neq \Lambda$, if it is not explicitly forbidden by a symmetry. For example, in section II B 5 we will introduce an EFT with only one coupling at its UV scale Λ , a coupling between an ALP a and either W bosons or flavour-universally to diagonal fermion currents. Later, we will use this model to predict observables at a lower scale, around the B meson mass, and find that at this scale, the running of the

WCs has generated a non-zero WC for the flavour changing neutral current $b \rightarrow s$,

$$\mathcal{C}_{abs}(m_B) \sim f(m_B)\mathcal{C}_{ff/WW}(\Lambda) \neq 0. \quad (25)$$

If an EFT is the low-energy limit of a known UV theory, we know that the UV theory and the EFT have to give the same result for observables that are within the scale of validity of the EFT. With this, we can determine the Wilson coefficients of the EFT, by predicting observables with both theories and equating them to solve for the Wilson coefficients of the low energy EFT as a function of (possibly Wilson) coefficients of the UV theory (which may also be an EFT). To do this, we of course have to run the UV theory down to the scale at which the EFT is valid and predict the observable with the run-down theory. For example, in the case of Fermi's theory of β decay, we can run the Standard Model down to the scale of β decay and compare the resulting formulas for β decay in both theories to find

$$G_F = \frac{\sqrt{2}g^2}{8m_W^2}. \quad (26)$$

In general, the running of WCs and the mixing of operators mean that when we match a low-energy EFT onto a UV theory, we often do not have one-to-one matchings. Often, this is quite intuitive. For example, the SMEFT treats its fields as being at energies above electroweak symmetry breaking, meaning that the electroweak interaction is mediated by W and B bosons, while in the WET, we have interactions with photons instead and the interactions with W bosons are completely integrated out. Thus, the matching condition for an operator of the WET to the SMEFT may look like [56]

$$\mathcal{C}_{d\gamma}^{pr}(\mu) = \frac{v}{\sqrt{2}} (\cos \theta_W \mathcal{C}_{dB}^{pr}(\mu) - \sin \theta_W \mathcal{C}_{dW}^{pr}(\mu)) \quad (27)$$

for

$$\mathcal{O}_{d\gamma}^{pr} = (\bar{d}_{Lp}\sigma^{\mu\nu}d_{Rr}) F_{\mu\nu} \quad (28)$$

$$\mathcal{O}_{dW}^{pr} = (\bar{q}_p\sigma^{\mu\nu}d_r) H\tau^i W_{\mu\nu}^i \quad (29)$$

$$\mathcal{O}_{dB}^{pr} = (\bar{q}_p\sigma^{\mu\nu}d_r) HB_{\mu\nu} \quad (30)$$

and where μ the matching scale, v the Higgs vacuum expectation value and θ_W the weak mixing angle, both of which get contributions from SMEFT Wilson coefficients themselves. In addition, throughout their running from the high cutoff-scale of SMEFT the SMEFT operators also get contributions from additional operators [66]. Similarly,

if we keep running the WET operator further down to other scales, we will again get admixtures of additional operators [55].

Due to the complex nature of the renormalisation group (RG) evolution of these and other EFTs, we will not be explicitly calculating any running of Wilson coefficients or matching between EFTs in the course of this thesis. Wherever we encounter observables and theories of different scales, we will use numerical computational tools to run couplings and match theories to each other. The tools we use are `wilson` [67] and [68]. They use matching relations and anomalous dimensions for running parameters to the highest theoretical precision known. We will still be discussing the consequences of the RG running and results of the matching wherever they are relevant and interesting, but will not show the details of the calculations.

B. Extensions of the Standard Model

1. Gauge Invariance and Portal Models

As we have established earlier, the Standard Model has the gauge group $SU(3)_c \times SU(2)_L \times U(1)_Y$. The colour gauge group $SU(3)_c$ corresponds to the strong interaction which all quarks are charged under (in the fundamental representation $\mathbf{3}$ of $SU(3)$) and which is transmitted by gluons G_μ^a (in the adjoint representation $\mathbf{8}$ of $SU(3)$). The $SU(2)_L \times U(1)_Y$ correspond to the electroweak interaction. The weak isospin gauge group $SU(2)_L$ is transmitted by the W bosons W_μ^i (in the adjoint representation $\mathbf{3}$ of the $SU(2)$) and acts on the left-handed quarks and leptons (in the fundamental representation $\mathbf{2}$). The hypercharge gauge group $U(1)_Y$ acts on all fermions and the Higgs doublet. We can write down the finite gauge transformations of the fields as [51]

$$B_\mu \rightarrow B_\mu + \frac{2}{g_1} \partial_\mu \theta \quad (31)$$

$$W_\mu \rightarrow UW_\mu U^\dagger + \frac{2i}{g_2} U \partial_\mu U^\dagger \quad (32)$$

$$G_\mu \rightarrow VG_\mu V^\dagger + \frac{i}{g_s} V \partial_\mu V^\dagger \quad (33)$$

$$H = \begin{pmatrix} h_+ \\ \frac{h_0 + v + i\eta_0}{\sqrt{2}} \end{pmatrix} \rightarrow e^{-i\theta} UH \quad (34)$$

$$\ell_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \rightarrow e^{i\theta} U \ell_L \quad (35)$$

$$q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix} \rightarrow e^{-i\frac{\theta}{3}} UV q_L \quad (36)$$

$$e_R \rightarrow e^{2i\theta} e_R \quad (37)$$

$$u_R \rightarrow e^{-\frac{4}{3}i\theta} V u_R \quad (38)$$

$$d_R \rightarrow e^{\frac{2}{3}i\theta} V d_R \quad (39)$$

with θ the $U(1)_Y$ gauge phase, U the unitary 2×2 $SU(2)_L$ gauge transformation matrix, and V the unitary 3×3 $SU(3)_c$ gauge transformation matrix. When we said we constructed gauge-invariant operators in the last section, what we meant is that we construct operators such that the operator is unchanged under the gauge transformations of the constituent fields,

$$\mathcal{O}_i \rightarrow \mathcal{O}_i. \quad (40)$$

Gauge invariance not only needs to hold for EFTs, but for all QFTs. When constructing new models of new physics beyond the Standard Model, we can first look at what gauge-invariant combinations of fields we can form to inform our model building.

We will now consider some of the combinations with the lowest mass dimensions. The first of which is the field strength tensor for the B boson,

$$\begin{aligned} F_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu \\ &\rightarrow \partial_\mu B_\nu + \frac{2}{g_1} \partial_\mu \partial_\nu \theta - \partial_\nu B_\mu - \frac{2}{g_1} \partial_\nu \partial_\mu \theta \\ &= F_{\mu\nu} \end{aligned} \quad (41)$$

which is SM gauge- but not Lorentz-invariant. While the B boson has a gauge-invariant field strength tensor, the same is not true for the W boson and gluon field strength tensors, which transform as

$$W_{\mu\nu} \rightarrow U W_{\mu\nu} U^\dagger \neq W_{\mu\nu}, \quad (42)$$

$$G_{\mu\nu} \rightarrow V G_{\mu\nu} V^\dagger \neq G_{\mu\nu}. \quad (43)$$

After electroweak symmetry breaking, that is, under the gauge group $SU(3)_c \times U(1)_{em}$, the photon also follows equation (41), even though it is a mixture between the W and B bosons. This is because after electroweak symmetry breaking, the electroweak symmetry group $SU(2)_L \times U(1)_Y$ is reduced to $U(1)_{em}$, so the W 's gauge non-invariance under $SU(2)_L$ is not relevant anymore. We can also directly see this by pointing out that the photon is the gauge boson of $U(1)_{em}$.

Another gauge-invariant field combination is

$$H^\dagger H \rightarrow (e^{-i\theta} U H)^\dagger e^{-i\theta} U H \quad (44)$$

$$\begin{aligned}
&= e^{-i\theta} e^{i\theta} H^\dagger U^\dagger U H \\
&= H^\dagger H
\end{aligned}$$

which is additionally also Lorentz-invariant.

Other gauge-invariant combinations involve fermions, like

$$\begin{aligned}
\bar{q}_L \gamma^\mu q_L &\rightarrow e^{-i\frac{\theta}{3}} e^{i\frac{\theta}{3}} \bar{q}_L V^\dagger U^\dagger \gamma^\mu U V q_L \\
&= \bar{q}_L \gamma^\mu q_L
\end{aligned} \tag{45}$$

and similarly

$$\begin{aligned}
\bar{\ell}_L \gamma^\mu \ell_L &\rightarrow \bar{\ell}_L \gamma^\mu \ell_L \\
\bar{u}_R \gamma^\mu u_R &\rightarrow \bar{u}_R \gamma^\mu u_R \\
\bar{d}_R \gamma^\mu d_R &\rightarrow \bar{d}_R \gamma^\mu d_R \\
\bar{e}_R \gamma^\mu e_R &\rightarrow \bar{e}_R \gamma^\mu e_R.
\end{aligned} \tag{46}$$

and also

$$\begin{aligned}
\bar{\ell}_L i\sigma_2 H^* &\rightarrow e^{i\theta} e^{-i\theta} \bar{\ell}_L U^\dagger i\sigma_2 U^* H^* \\
&= \bar{\ell}_L i\sigma_2 H^*.
\end{aligned} \tag{47}$$

You may notice that terms of the form $\bar{f}_{L/R} f_{R/L}$ or $\bar{f}_{L/R} \sigma^{\mu\nu} f_{R/L}$ are not gauge-invariant. This is due to the incompatibility of the left- and right-handed fermion fields' transformations under $SU(2)_L$. These combination are gauge-invariant only after electroweak symmetry breaking where $SU(2)_L \times U(1)_Y$ breaks down to the electromagnetic interaction $U(1)_{em}$, which is why we saw dipole operators only in the operators of the WET and not for the SMEFT and we otherwise saw scalar currents only in pairs to balance each other out.

Having built these gauge-invariant field combinations, we can now think about how physics beyond the Standard Model could look. The small deviations we have so far detected from Standard Model physics as well as our evidence for dark matter indicate that new physics is probably at most weakly coupled to the Standard Model — or alternatively that it is very heavy, so that it has very little effective consequences on our comparatively low-energy observables, in which case we may have a good chance of exploring it with the EFT approach covered in the last section. Whether the new physics is weakly coupled or not, however, we usually want to couple any new particles to the Standard Model gauge-invariantly and with operators of low mass dimension. Often, we

assume that there may be a number of new particles, but that only one of them interacts with the SM. We call these models portal models. In the following, I will introduce a few commonly used portals.

2. Scalar Portal

A common portal is based on equation (44) and couples a new real SM gauge singlet scalar ϕ to the Standard Model Higgs via terms

$$\mathcal{L} \supset -\frac{\delta_1}{2} H^\dagger H \phi - \frac{\delta_2}{2} H^\dagger H \phi \phi \quad (48)$$

where the cubic coupling δ_1 has mass-dimension one and the quartic coupling δ_2 is dimensionless. Due to its coupling to the scalar (or Higgs) sector of the Standard Model, this portal is usually called the scalar portal or Higgs portal. We now follow [69] in introducing the mixing between the SM Higgs and the new scalar, using unitary gauge,

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ h_0 + v \end{pmatrix}. \quad (49)$$

With electroweak symmetry breaking, the Higgs vacuum expectation value (vev) v is introduced and the Lagrangian turns into

$$\mathcal{L} \supset -\frac{1}{2} \delta_1 v^2 \phi - \frac{1}{2} \delta_2 v^2 \phi \phi - \delta_1 v h_0 \phi - \frac{1}{2} \delta_1 h_0 h_0 \phi - \delta_2 v h_0 \phi \phi - \frac{1}{2} \delta_2 h_0 h_0 \phi \phi. \quad (50)$$

In addition to the cubic and quartic coupling terms for the field combinations $h_0 - h_0 - \phi$, $h_0 - \phi - \phi$, and $h_0 - h_0 - \phi - \phi$, there are also terms with two fields only. These are mass terms. An additional mass term, that has not shown up yet, comes from the Higgs-only part of the Lagrangian and gives the Higgs boson h_0 its mass. With it, the complete mass Lagrangian is given by

$$\mathcal{L} \supset -\frac{m_{h_0}^2}{2} h_0 h_0 - \delta_1 v h_0 \phi - \frac{\delta_2 v^2}{2} \phi \phi. \quad (51)$$

The first and third term here are mass terms for the fields h_0 and ϕ . The second term is a mixing term between h_0 and ϕ , signalling that these terms are not stable under propagation and are not the mass eigenstates of the Lagrangian. To find these instead, and with them states that propagate stably, we can first write equation (51) as

$$\mathcal{L} \supset -\frac{1}{2} \begin{pmatrix} h_0 & \phi \end{pmatrix} \begin{pmatrix} m_{h_0}^2 & \delta_1 v \\ \delta_1 v & \delta_2 v^2 \end{pmatrix} \begin{pmatrix} h_0 \\ \phi \end{pmatrix}. \quad (52)$$

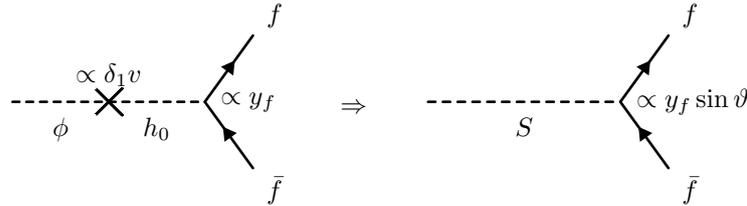


FIGURE 5: Diagrams showing the basis change from the Lagrangian basis of h_0 and ϕ coupling the new scalar ϕ to fermions via propagator mixing, to the mass basis of h and S where the new scalar S couples to fermions directly through its h_0 component.

From this we can find the mass eigenstates of the systems as the field mixtures that are stable under propagation by identifying the eigenvalues and eigenvectors of the mass matrix. The two physical, non-mixing, states are given by

$$\begin{pmatrix} h \\ S \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} h_0 \\ \phi \end{pmatrix} \quad (53)$$

with the masses

$$m_{h/S}^2 = \frac{m_{h_0}^2 + \delta_2 v^2}{2} \pm \left(\frac{m_{h_0}^2 - \delta_2 v^2}{2} \right) \sqrt{1 + x^2}, \quad (54)$$

the mixing angle

$$\tan \theta = \frac{x}{1 + \sqrt{1 + x^2}}, \quad (55)$$

and

$$x = \frac{\delta_1 v}{m_{h_0}^2 - \delta_2 v^2}. \quad (56)$$

Equation (53) rotates the fields from the basis of fields in which the Lagrangian is defined to the basis of mass eigenstates. There, h is the physical Standard Model Higgs boson in contrast to the Lagrangian field h_0 and S is the new physical scalar particle. Neither of them now mix into each other during propagation. In equation (54), we have assumed that S will be lighter than h because this is true for the model in section III A. Except for $\pm \rightarrow \mp$ in equation (54), nothing else changes when S is the heavier particle.

Through its mixing with the Higgs boson, the new scalar gains couplings to the other Standard Model fields with the suppression⁷ factor of the mixing angle, for example to

⁷ We can generally assume that portal couplings are small as their effects have not yet been detected, which is why we can speak of suppression here.

fermions

$$\mathcal{L} \supset y_f \sin \theta \bar{f} f S \quad (57)$$

with y_f the Yukawa coupling between fermions and the Higgs boson. We show this diagrammatically in figure 5. If we couple our new scalar ϕ to anything else outside the Standard Model, we of course also induce a coupling to this in the Higgs boson. These couplings are suppressed with a factor of $\sin \theta$, just like the coupling S gains to SM fermions. If, for example, ϕ is a mediator to a dark sector which includes dark matter, searching for invisible Higgs decays may be a way of probing this model. If such a connection to dark matter exists, the scalar is often called the dark scalar.

There are many ways to construct a scalar portal in different ways or to add onto the simple model we just introduced. We could introduce self-coupling terms between the scalars,

$$\mathcal{L} \supset -\delta_3 \phi \phi \phi - \delta_4 \phi \phi \phi \phi, \quad (58)$$

an explicit mass for the scalar independent of the Higgs vacuum expectation value,

$$\mathcal{L} \supset -\frac{1}{2} m_\phi^2 \phi \phi, \quad (59)$$

give the new scalar a vacuum expectation value as well

$$\phi \rightarrow \phi' + v_\phi, \quad (60)$$

or impose a \mathbb{Z}_2 symmetry on it,

$$\phi \rightarrow -\phi, \quad (61)$$

such that the cubic term in equation (48) is forbidden. There is also the possibility to introduce more particles into the model, for example dark sector particles that only the dark scalar starts out as coupling to. Instead of a real singlet scalar, we could also introduce a more complicated scalar, like a complex scalar or a scalar doublet. The latter are used in two Higgs doublet models, which are utilised for example in supersymmetry. All of these changes of course change the explicit dependence of the mixing angle on the Lagrangian parameters, but the general formalism of the new scalar mixing with the Higgs still follows this calculation.

3. Vector Portal

Another common portal is based on the gauge-invariance of the photon's field strength tensor, equation (41). We introduce an additional $U(1)$ symmetry, often called $U(1)_X$, with its own gauge boson X_μ . The symmetry group of our new vector mediator is often based on one of the accidental global symmetries of the Standard model, e. g. $B - L$ [70]. The new boson is often called the dark or hidden photon A' , especially when it has a very low mass [70], or the Z' , especially when it has a mass closer to the weak scale [71, 72]. The boson has the field strength tensor $X_{\mu\nu}$. With this, we can construct a gauge- and Lorentz-invariant operator, [70, 73]

$$\mathcal{L} \sim -\frac{\varepsilon'}{2} F^{\mu\nu} X_{\mu\nu} \quad (62)$$

the so-called kinetic mixing term. Depending on whether the SM fermions are charged under the new $U(1)$, we may have additional terms

$$\mathcal{L} \sim \bar{f} \gamma^\mu f X_\mu \quad (63)$$

as well. These terms show up even when we do not put them into the initial Lagrangian explicitly, because the kinetic mixing term lets the dark photon couple to the SM fermions via the insertion of a SM photon. We can compare this to the mixing of the Higgs boson with another scalar via mass mixing, which has a similar effect of giving the new particle a coupling to SM fermions. Like in the scalar portal case, we can also redefine our vector fields with a mixing angle to get a description with no kinetic mixing but instead explicit couplings between the sectors [73]. While in the scalar case, we need to first break the electroweak symmetry before mass mixing happens, here we preserve both $U(1)$ s. For a more detailed review of vector portals, which we will not be using further in this thesis, see [70–73].

4. Neutrino Portal

Another common portal makes use of equation (47) to couple the SM gauge-invariant term (that is missing a second fermion to be Lorentz-invariant) to a new SM singlet fermion N , [74]

$$\mathcal{L} \sim y_n \bar{\ell}_L i \sigma_2 H^* N. \quad (64)$$

The new fermion is often called a sterile neutrino, as it does not interact directly via any of the Standard Model interactions, or a heavy neutral lepton (HNL), since these particles are often constructed to be much heavier than neutrinos.

In the low-energy limit after electroweak symmetry breaking, equation (64) translates (with the introduction of the Higgs vacuum expectation value) to a coupling between neutrinos and N directly, as a mass mixing term [74]

$$\mathcal{L} \supset \frac{y_n v}{\sqrt{2}} \nu_L N. \quad (65)$$

Additionally, we have couplings of the new fermion N to the Higgs boson and the neutrino, and to the neutrino or the charged lepton together with a neutral or charged Goldstone boson, respectively.

From the mass mixing term in equation (65), we gain a mixing between the SM and the sterile neutrinos, similar to the mixing in the scalar portal earlier in this section. Through this mixing, the sterile neutrino gains access to the SM interactions proportional to the neutrino's interactions. We will not be working with HNLs in this thesis, but refer to [74] for a deeper look at their theory and current status.

5. Axions and Axion-Like Particles

Axion-like particles (ALPs) [75] differ from our portal discussions so far in the fact that they do not generally come from Lagrangian terms with a mass dimension of four or less. While there are models that UV-complete ALPs, the interaction terms of the ALPs with Standard Model particles are usually of mass dimension 5. Axion-like particles a are pseudo-scalars that are pseudo-Nambu-Goldstone bosons of approximate global symmetries beyond the Standard Model broken at a scale f_a much higher than the electroweak scale. They have a shift symmetry $a \rightarrow a + c$ that is softly broken by their mass m_a . The mass can be explained from UV models as well [76]. The name axion-like particle refers to the QCD axion [77–79], a specific ALP constructed to solve the strong CP problem.

The general effective Lagrangian for ALPs looks like [76, 80]

$$\begin{aligned} \mathcal{L} \supset & \frac{\partial_\mu a}{f_a} \sum_F \bar{\psi}_F \gamma^\mu \mathbf{c}_F \psi_F + c_{GG} \frac{g_s^2}{16\pi^2} \frac{a}{f_a} G_{\mu\nu}^a \tilde{G}^{a\mu\nu} \\ & + c_{WW} \frac{g^2}{16\pi^2} \frac{a}{f_a} W_{\mu\nu}^i \tilde{W}^{i\mu\nu} + c_{BB} \frac{g'^2}{16\pi^2} \frac{a}{f_a} B_{\mu\nu} \tilde{B}^{\mu\nu} \end{aligned} \quad (66)$$

where $F \in \{q_L, \ell_L, u_R, d_R, e_R\}$ are the fermion fields, the Wilson coefficients/couplings \mathbf{c}_F are hermitian matrices in flavour/generation space, and the dual tensor is defined as $\tilde{X}^{\mu\nu} = \frac{1}{2} \varepsilon^{\mu\nu\rho\sigma} X_{\rho\sigma}$. We can see that all the terms in this Lagrangian are invariant under the ALP's shift symmetry. Terms of higher mass dimension can be neglected since they would be further suppressed by higher orders of the symmetry breaking scale f_a .

Specific UV models of the ALP may impose relations between the ALP mass m_a and its symmetry breaking scale f_a or any of the couplings in equation (67), but we will treat them as free, independent parameters for now.

The couplings in equation (67) run with the scale of the process they are used in. If we only introduce few couplings in the UV, we will still induce other couplings through renormalisation group running, as described in section II A 6. The renormalisation group evolution of the couplings in equation (67) is described in [76] and we will use the code [68] for its calculation.

Later, we will use two ALP models in which either a universal diagonal fermion coupling c_{ff} or a coupling to the W bosons c_{WW} is the only nonzero coupling in the UV. The Lagrangians in the UV are then

$$\mathcal{L} = \frac{\partial_\mu a}{f_a} \sum_F c_{ff} \bar{\psi}_F \gamma^\mu \psi_F \quad (67)$$

and

$$\mathcal{L} = c_{WW} \frac{g^2}{16\pi^2} \frac{a}{f_a} W_{\mu\nu}^i \tilde{W}^{i\mu\nu}. \quad (68)$$

We will also use $\Lambda = 4\pi f_a$ as a cut-off scale in analogy to the usage in the EFT section. The running of the Wilson coefficients in the ALP Lagrangian generates effective couplings that are not initially put in, giving terms of the form

$$\mathcal{L}(\mu \ll \Lambda) \supset \frac{\partial_\mu a}{f_a} c_{bs} \bar{b} \gamma^\mu s \quad (69)$$

scale-dependently coupling the ALP to flavour-changing quark transition. The flavour changing coupling c_{bs} is a function of the scale μ of the process that it is used in, the cut-off scale Λ and the UV coupling(s) it is generated from, in this case $c_{ff}(\Lambda)$ or $c_{WW}(\Lambda)$.

C. Statistics

In particle physics, we are often concerned with using experiments to determine which theories are more likely to be true than others. To handle and understand the data from these experiments properly, we need an understanding of probabilities and statistics. In the first part of this section, section II C 1, we will introduce some of the basics of statistics in the form of probabilities and probability distributions. Then, in section II C 2, we will explain how to set limits on specific hypothesis. We will follow the pedagogical treatment found in [38, 81, 82].

1. Probability Distributions

In statistics, a random variable X describes the outcome of a single experiment. This can for example be the number rolled on a die, the card pulled from a well-shuffled deck of cards, or the set of particles a specific particle decays into. These events are described by probability distributions that tell us how likely each outcome is. For example, the probability for rolling a specific number, say 3, on a six-sided die is

$$\mathbb{P}(X = 3) = \frac{1}{6}. \quad (70)$$

We can also define probabilities for the case of more than one outcome, for example the probability that we draw a card of a specific suit, say heart, from a standard deck of cards is given by the sum of probabilities of drawing each individual heart card,

$$\mathbb{P}(X = \heartsuit) = \sum_{\heartsuit\text{-cards } i} \mathbb{P}(X = i) = \sum_{\heartsuit\text{-cards } i} \frac{1}{52} = \frac{1}{4}. \quad (71)$$

In addition to discrete variables like cards or integers, we can also have continuous distributions of probability defined with a probability density function (PDF) p . Its left-sided integral is the cumulative distribution function (CDF). The probability for any specific range of values is also given by an integral over the PDF in this range or equivalently by the difference between values of the CDF.

As an example, let us compute the probability that a radioactive decay of a single atom happens between the times t_0 and t_1 . We begin with the observation that the rate of spontaneous decay per unit time is a constant λ . In other words, this process is one in which the instantaneous probability for decay is independent of how much time has elapsed. The rate of decays per unit time is given by this constant probability times the number of radioactive atoms N ,

$$\frac{dN}{dt} = -\lambda N, \quad (72)$$

where the minus sign signifies depletion of the population. The number of atoms N that are still undecayed after a time t when starting out with N_0 atoms is then given by

$$N(t) = N_0 e^{-\lambda t}. \quad (73)$$

From this, we can extract the probability that a single atom is undecayed after a time t as

$$\tilde{P}(t) = \frac{N(t)}{N_0} = e^{-\lambda t}. \quad (74)$$

We are instead interested in the probability that a single atom has decayed,

$$P(t) = \frac{N_0 - N(t)}{N_0} = 1 - e^{-\lambda t}. \quad (75)$$

This is the CDF. It describes the probability that a single atom decays at any point before the time t after observing it at $t = 0$. The PDF instead describes the probability that a single atom decays at any specific point in time. It is given by the normalised derivative of the CDF, that is,

$$P(t) = \int_0^t dt' p(t'), \quad (76)$$

such that,

$$p(t) = \frac{e^{-\lambda t}}{\lambda}. \quad (77)$$

Therefore, the probability that a single atom decays in a range of time between t_0 and t_1 is given by,

$$\begin{aligned} \mathbb{P}(t_0 \leq t \leq t_1) &= \int_{t_0}^{t_1} dt' p(t') \\ &= P(t_1) - P(t_0) \\ &= \frac{e^{-\lambda t_0} - e^{-\lambda t_1}}{\lambda}. \end{aligned} \quad (78)$$

In the first line, we express the probability as an integral over the PDF, in the second line as a difference of values of the CDF, both of which give us the resulting closed form of the probability given in the third line.

Until now, we have described the probability that a single atom decays within a range of time. We have found that this decay process is spontaneous and independent of any of its surrounding atoms. We can now consider the probability of the distribution of the number of particle decays within a range of time instead. Processes like radioactive decay, in which events are spontaneous and independent of history, are part of a class of processes known as Poisson processes. The Poisson distribution describes the probability of a discrete random variable X which is usually interpreted as a point-count. In the case of radioactive decay, X is the number of atoms that decay in a specific range of time. The probability that the distribution yields a point-count of exactly k , which in our example corresponds to exactly k atoms decaying, is given by [38]

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}. \quad (79)$$

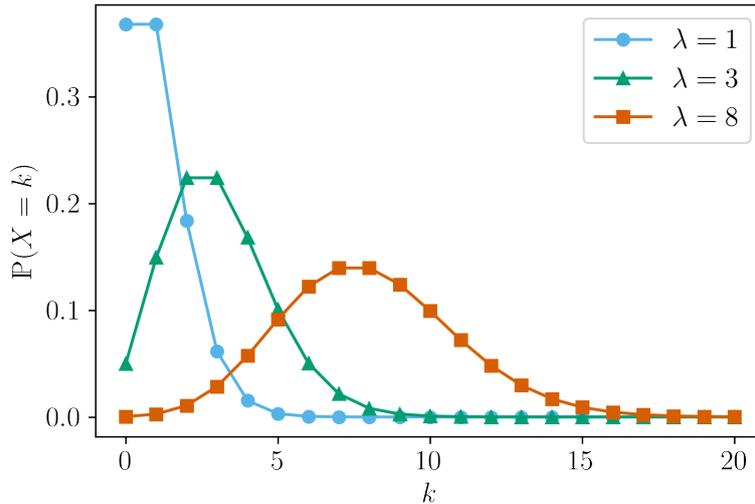


FIGURE 6: Poisson distributions with different mean values λ .

This is the probability mass function (PMF) of the Poisson distribution. It is the discrete analogue of the PDF. As in the previous example, λ can be interpreted as the constant independent count rate or more generally as the average or expected count. Its CDF corresponds to the probability that the point-count is k or lower, which in our example means that up to k atoms decay. It is given, due to the Poisson distribution's discrete nature, by a sum over the PMF,

$$\mathbb{P}(X \leq k) = \sum_{i \leq k} \mathbb{P}(X = i) = \sum_{i \leq k} \frac{\lambda^i}{i!} e^{-\lambda}. \quad (80)$$

Later in this thesis, we will look at particle collisions with specific final state particles and placing them in a parameter space of several kinematic variables. We will place cuts on this parameter space, and the placement of an event inside or outside of that region can also be described by the Poisson distribution. A sample of several Poisson distributions is shown in figure 6. Notice that the Poisson distribution describes a discrete count. However, when λ becomes very large, the distribution may be treated as approximately continuous. We will come back to this case later in the section.

We now turn to the normal distribution. The normal, or Gaussian, distribution describes the symmetric probability of a random variable X about a mean value μ with a characteristic spread given by its standard deviation σ which is related to the variance V as $V = \sigma^2$. This is the distribution often assumed for symmetric experimental errors. The PDF for the normal distribution is given by [38]

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \quad (81)$$

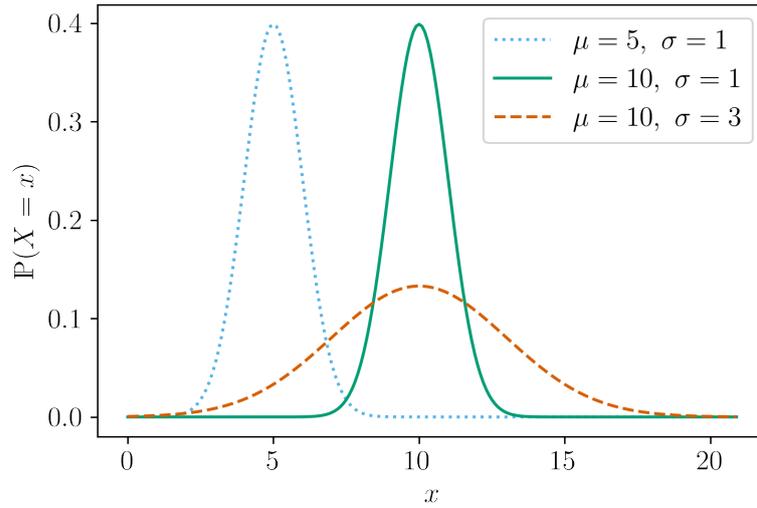


FIGURE 7: Gaussian distributions with different means μ and standard deviations σ .

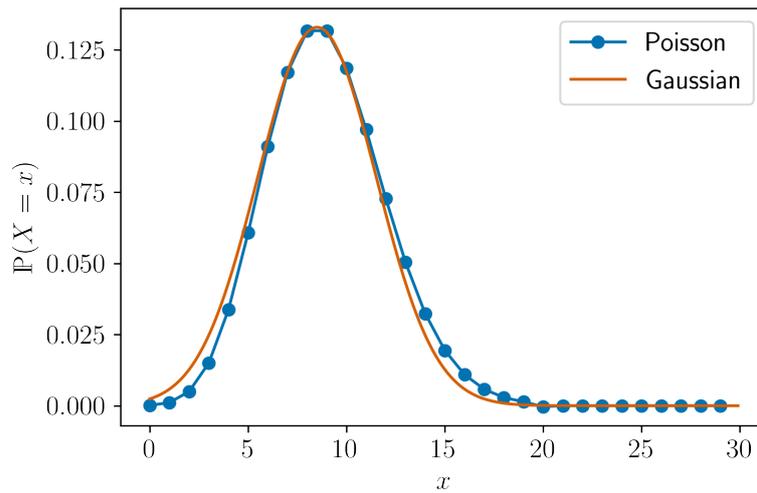


FIGURE 8: Comparison between a Poisson distribution of $\lambda = 8$ and its Gaussian approximation. While the two functions look very similar, it is clearly visible that the Poisson distribution is not fully symmetric and curves differently in its slopes than the Gaussian does.

Since this is a continuous distribution, its CDF is given by the left-sided integral over its PDF. It is related to the error function erf as

$$\begin{aligned} P(x \leq b) &= \int_{-\infty}^b dx p(x) \\ &= \frac{1}{2} \left(1 + \text{erf} \left(\frac{b - \mu}{\sigma} \right) \right). \end{aligned} \quad (82)$$

The probability that the value of x lies in a specific range $a \leq x \leq b$ is given by

$$\mathbb{P}(a \leq x \leq b) = \int_a^b dx \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \quad (83)$$

A few sample Gaussian distributions are shown in figure 7. As shown in figure 8, for large λ the Poisson distribution is also well-approximated by a normal distribution with $V = \mu \approx \lambda$.

The peak probability happens for the mean μ and its width is described by σ . Specifically, $\mu \pm \sigma$ are the positions of the inflection points of the symmetric function and the half width at half measure is given by $\sqrt{2 \ln 2} \sigma$. The standard deviation has another useful property. Within one standard deviation (1σ) of the mean value, one finds approximately 68% of the probability. Similarly, 95% is enclosed within 2σ and 99.7% within 3σ . In the following, we will use the number of ‘‘sigmas’’ and their corresponding enclosed probability interchangeably.

It should be noted that we need to differentiate between one-sided and two-sided intervals. As shown in the left frame of figure 9, a two-sided interval is one in which the PDF has been integrated symmetrically about the mean such that the quoted value of $x_0 \pm \Delta x$ encloses the given probability. Therefore, two-sided intervals are interpreted such that the value of x lies between $x_0 - \Delta x$ and $x_0 + \Delta x$ with a given probability. As shown in the right frame of figure 9, a one-sided probability is one in which the PDF has been integrated from negative infinity as in the CDF. Therefore, one-side intervals are interpreted such that the value of x lies below $x < x_0$ with a given probability.

2. Hypotheses and setting limits

A statistical hypothesis is a theory about the shape of a probability distribution for a specific problem. For example, if we look at a specific particle decay, we may have the hypothesis that the Standard Model is true. We would calculate a decay rate based on this hypothesis and base our probability distribution on a Poisson distribution depending on this decay rate.

Often, we have several competing hypotheses. The currently accepted theory H_0 is usually called the null or default hypothesis. In our case in particle physics, this is usually the Standard Model. Another hypothesis may be, for example, that in addition to the Standard Model we have a Higgs portal with an additional scalar particle, as discussed in section II B 2. This will give us another prediction for the decay rate and, following from that, a different probability distribution. We may call this hypothesis H_m , where m is a parameter (or a set of parameters) of our hypothesis. In our example

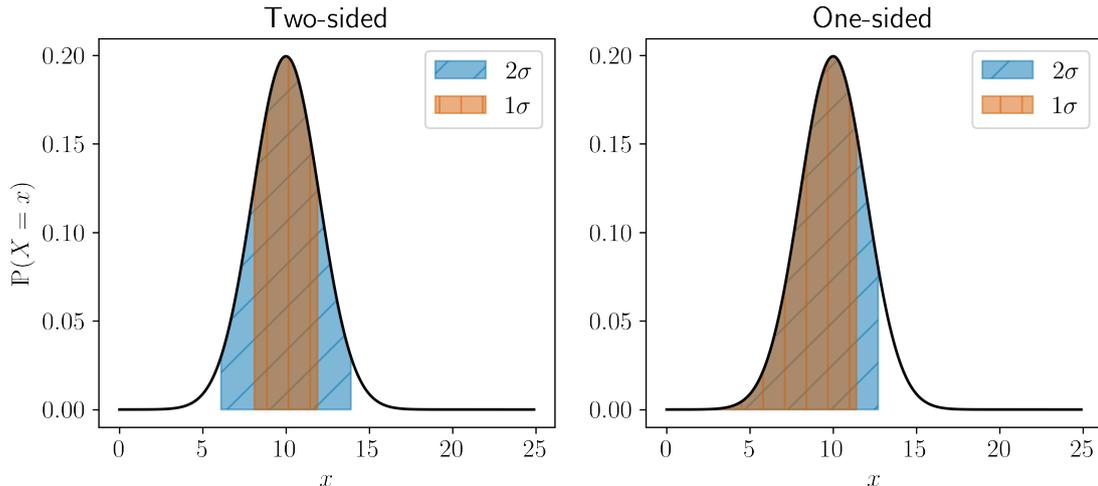


FIGURE 9: Comparison between the one-sided and two-sided intervals spanning one and two σ worth of events, both for the Gaussian probability density function. The two-sided interval is symmetric around the mean while the one-sided interval starts at $-\infty$ and reaches only part of the way down the second slope.

of the Higgs portal, this could for example be the mass of the new scalar. A hypothesis test is the process by which we try to distinguish which of the hypotheses is true given some data, and the determination of the certainty with which we can say this.

Say we have a counting experiment, where we have some decays happening in a parameter space and we try to determine if they are explainable fully by Standard Model processes. This is our null hypothesis. Otherwise, they may be explained by some physics beyond the Standard Model (BSM). If we have a specific BSM model, we may call this additional hypothesis H_m . Say we do an experiment that we expect could show hints of BSM physics. We can now do two separate things: look for a discovery, or set exclusion bounds.

When we look for a discovery, we are talking about the probability that a specific BSM model is true. For exclusion bounds, we instead mean to determine the probability that the Standard Model is wrong. That is, that there are signs of BSM in the data at all, not necessarily of a specific model. For discovery, we need more proof than for exclusions, since we need data that fits with the specific model rather than just not fitting with the Standard Model. In addition, we tend to ask for a higher level of confidence for discoveries than for exclusions. Discovery is usually claimed when the significance of a measurement interpreted in the context of H_m instead of H_0 is 5σ , that is, with a probability greater than $1 - 6 \cdot 10^{-5}$. On the other hand, for exclusion regions, we use 90% or 95% (2σ).

To set a limit of a minimal number $k > n_{min}$ of events necessary for our null hypothesis

to be rejected with a confidence level CL, we require

$$\mathbb{P}(X > n_{min}) = \text{CDF}(\text{Poisson})(X > n_{min}) > \text{CL}. \quad (84)$$

The inversion of the CDF of the Poisson distribution is not a trivial matter, but can be performed numerically and is accessible through tables and can be computed iteratively relatively easily. A common example of this minimal value n_{min} is that for zero predicted background ($\lambda = 0$), one needs to find 3 or more events to be confident to a CL of 95 % that the result is not in agreement with the prediction of zero background. That is, to be 95 % confident that the null hypothesis is wrong or incomplete the minimum measured value must be at least 3. In particle physics, this finding would usually mean that there is physics beyond the Standard Model that is relevant for predicting this process. For a confidence level of 3σ at zero background, we need six or more events.

In a collider counting experiment such as described above, the number of expected background events and thus the minimal number of signal events needed to bound or discover a theory, depends on the kinematic cuts we place on the parameter space. This is because we count the number of events that land within the critical region. We choose our kinematic variables such that signal and background are situated in parts of the parameter space that are as separable as possible. We know that a set of cuts is better when it contains fewer background events and more signal events.

A good formula to optimise when choosing cuts is the one introduced by Punzi in [81]. He uses a Gaussian approximation for the Poisson distribution and defines a as the number of sigmas corresponding to the desired confidence level, $\varepsilon(t)$ as the signal efficiency and $B(t)$ as the number of background events based on the cuts t . The signal efficiency is defined as the fraction of signal events that lie within the region whose boundary is delineated by the cuts, that is, what remains after cuts. The function to be optimised is then [81]

$$\frac{\varepsilon(t)}{\frac{a}{2} + \sqrt{B(t)}}. \quad (85)$$

Since this uses the signal efficiency and not the number of signal events in the region within the cuts S , we do not need to make assumptions on the total number of signal events produced and can leave this open as a parameter of our hypothesis. This is not the case for another common optimisation function, $\frac{S}{\sqrt{B+S}}$ which requires this knowledge. Another advantage of Punzi's formula is that it works even for very low background counts whereas $\frac{S}{\sqrt{B}}$, which is also often used for optimisation, is biased towards regions of low background at the cost of over-depleting signal.

The acceleration of the beams is asymmetric such that the centre-of-mass system is boosted in the forward direction (for the electron beam) [83]. There is also a slight angle between the two beams of 83 mrad [83]. The average boost of the B mesons is $\langle\gamma\beta\rangle = 0.3$ [83], which is low in comparison with B mesons produced at the LHC, which have an average boost of $\langle\gamma\beta\rangle \approx 15$ [84]. That is to say, while B mesons at the LHC move on average with a velocity of $\langle v\rangle = 0.97c$, B mesons at Belle II move at a much slower velocity of $\langle v\rangle = 0.3c$. Since they move much slower, they stay inside of the detector for longer. This is not only useful for the precise measurement of B and D meson lifetimes and mixing angles [83], but also for BSM long-lived particle searches.

In comparison to the LHC, it is also important to note that Belle II has a much lower background rate and much cleaner decay processes since the colliding particles are electrons and positrons and not hadrons. At a hadron collider like the LHC, the centre-of-mass energy of a specific process is also harder to determine because the particles colliding are the quarks and gluons inside of the hadrons which carry only part of the full collision energy. At an electron positron collider, we do not have this additional complication. A schematic view of the detector layout is shown in figure 10.

The subdetectors we will focus on now are the silicon pixel detector (PXD), silicon vertex detector (SVD) and the central drift chamber (CDC). While they all have different specifications and precisions to their measurements, they all are able to detect charged particle tracks and thus detect displaced particle pairs in this analysis. Thus, we will refer to all three of them together as the tracker or tracking detector. Here, we will simplify the tracker to be in the region between $\vartheta_{min} = 17^\circ$ and $\vartheta_{max} = 150^\circ$ in polar angle, unrestricted in φ , and have a radius between $\rho_{min} = 0.05$ cm and $\rho_{max} = 113$ cm. This simplification of the tracking detector at Belle II is shown in figure 11. For a more complete introduction to the Belle II detector and its research goals, see [83].

2. Scalar portal and meson decays

Now that we have introduced the detector whose sensitivity we will explore here, let us specify what model we use. We work with a scalar portal model close to how we introduced it in section II B 2. Its Lagrangian is

$$\mathcal{L} \supset -\frac{1}{2}m_\phi^2\phi^2 - \delta_1 H^\dagger H \phi \quad (86)$$

which means that in comparison to our earlier model we neglect the quartic term, but add an explicit mass term. As described before, the scalars mix with the mixing angle

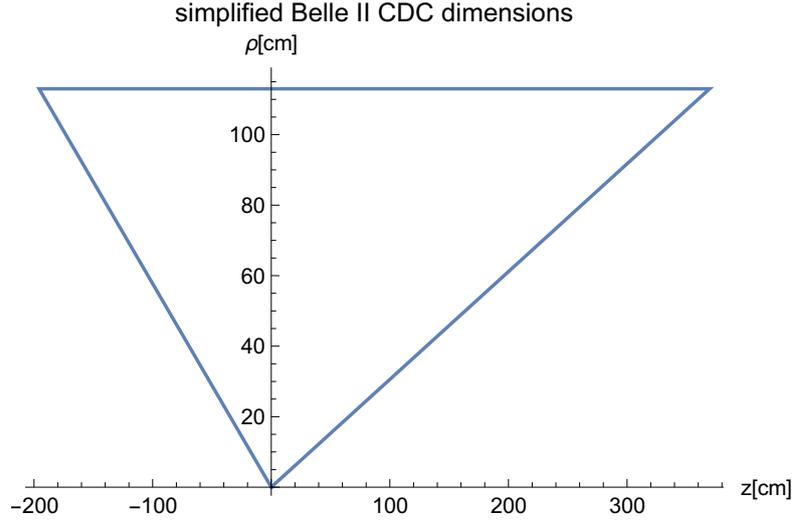


FIGURE 11: Simplification of the tracking detector of Belle II (CDC) which we use for the analysis in this section.

θ into the physical states h and S , as

$$m_{h/S}^2 = \frac{1}{2} (m_\phi^2 - m_{h_0}^2 \pm \Delta m^2) \quad (87)$$

$$\sin^2 \theta = \frac{1}{2} \left(1 + \frac{m_\phi^2 - m_{h_0}^2}{\Delta m^2} \right) \quad (88)$$

$$(\Delta m^2)^2 = 4v^2 \delta_1^2 + (m_\phi^2 - m_{h_0}^2)^2. \quad (89)$$

This leads to the following couplings of the scalars to SM fermions

$$\mathcal{L}_f = - \sum_f \frac{m_f}{v} (\cos \theta \bar{f} f h + \sin \theta \bar{f} f S). \quad (90)$$

As we can see, not only does the mixing give our new scalar S a suppressed coupling to the Standard Model fermions, it also slightly diminishes the coupling of the Higgs boson to them. While this also opens avenue for searches to probe this model's parameter space, we will here focus exclusively on the interactions between S and fermions.

The decay rate of S into leptons can be easily determined to be

$$\Gamma_{S \rightarrow \ell\ell} = m_S \frac{m_\ell^2}{8\pi v^2} \sin^2 \theta \left(1 - \frac{4m_\ell^2}{m_S^2} \right)^{\frac{3}{2}} \quad (91)$$

which is simply the rate of the SM Higgs boson decaying to fermions scaled by the mixing angle $\sin^2 \theta$. Due to confinement, the rate of S decaying to quarks is more difficult to calculate, such that we use the hadronic rates calculated in [85].

While the scalar S does couple to electrons and could thus be produced directly from the

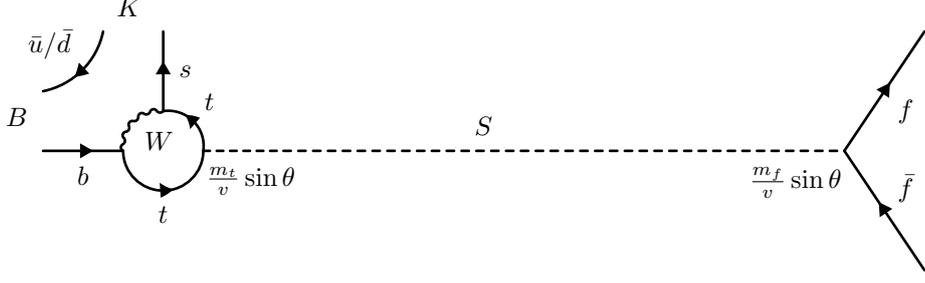


FIGURE 12: The dominant loop diagram that induces the effective b - s - S coupling as part of the full process of S production and displaced decay to a muon pair.

collision, it would be suppressed by the small electron mass in the coupling, $\frac{m_e}{v} \sin \theta \approx 10^{-6} \sin \theta$. Since there are a lot of B mesons produced at Belle II, it is a reasonable next step to check the production of scalars S in B meson decays. This proceeds via a loop diagram, since the scalar has no flavour-changing coupling. The dominant diagram is the one shown in figure 12. The branching ratio of B to a kaon K and a scalar S is then given by

$$Br(B^\pm \rightarrow K^\pm S) = \frac{\sqrt{2}G_F |c_{bs}|^2}{64\pi\Gamma_{B^\pm} m_B^3} \left(\frac{m_b - m_s}{m_b + m_s}\right)^2 f_0^2(m_S^2) (m_B^2 - m_K^2)^2 \lambda^{\frac{1}{2}}(m_B^2, m_K^2, m_S^2) \quad (92)$$

with the effective loop-induced flavour-changing coupling c_{bs} at lowest order in $\frac{m_S}{m_W}$ being

$$c_{bs} = \frac{3\sqrt{2}G_F m_t^2}{16\pi^2} V_{tb} V_{ts}^* \sin \theta, \quad (93)$$

the form factor $f_0(q^2)$ from [86], the full B^\pm meson decay rate Γ_{B^\pm} , and

$$\lambda(a, b, c) = a^2 + b^2 + c^2 - 2ab - 2bc - 2ca. \quad (94)$$

The branching ratio turns out to be quite sizeable, $Br(B^\pm \rightarrow K^\pm S) \approx 0.5 \sin^2 \theta$, so B decays are a very good option with which to produce the new scalar at Belle II.

Seeing as all decays of S come from the Standard Model Higgs boson through mixing, its decay will also always be suppressed by a factor of $\sin^2 \theta$. Thus, the full process of production and decay of the scalar is proportional to the fourth power of the mixing angle $\sin \theta$. Since a scalar S produced in a B meson decay will always decay to particles lighter than $2m_{\text{light}} \leq m_B - m_K$, we get a suppression of the decay of S in comparison to its production by a factor of $\frac{m_{\text{light}}}{m_t}$ from the mass hierarchical coupling of the Higgs boson, see figure 12. That means that the lifetime of S is longer than in a similar model with non-hierarchical couplings, and that we can still retain a good production rate of scalars S even while they have a long lifetime and thus decay displaced.

While prompt decays have been well-explored in these models before [87], a better limit still comes from LHCb's displaced muon searches, $B \rightarrow KS$, $S \rightarrow \ell^+\ell^-$ [88, 89]. Since particles at Belle II are boosted less than at LHCb, particles of longer lifetimes would have a higher likelihood of still decaying within the tracker, allowing us to probe smaller couplings/mixing angles. Belle II also has a much larger solid angle than LHCb which is situated in forward direction with an opening angle of 250 mrad [90].

To compare with the LHCb results, let us first consider a displaced decay into a muon pair. The number of expected displaced muon pairs from S decays produced in $B^\pm \rightarrow K^\pm$ decays within the detector can be expressed as

$$N_{\mu^+\mu^-} = 2N_{B^+B^-} Br(B^\pm \rightarrow K^\pm S) Br(S \rightarrow \mu^+\mu^-) \int_{\text{Belle II}} dr d\vartheta_0 \frac{1}{2} \frac{\sin \vartheta_0}{d_S} e^{-\frac{r}{d_S}} \quad (95)$$

with $N_{B^+B^-}$ the number of B^\pm pairs produced at Belle II. Since each pair is made up of two mesons, either of which can decay into KS , we multiply this number by a factor of two. The integral over the decay rate with the lifetime $d_S = \gamma_S \beta_S c \tau_S$ determines how many of the decays of S happen within the volume of the detector where they can be detected. ϑ_0 is the polar angle of the momentum direction of S in the rest frame of the B meson that produced it. The full derivation and calculation of this integral can be found in [91].

Now, seeing as only about half of the B mesons produced at Belle II are charged, we can adapt this formula to add the contribution of the neutral process, $B^0 \rightarrow K^0 S$, $S \rightarrow \mu^+\mu^-$ as well. Due to the decay rates of the $\Upsilon(4S)$ from which the B mesons are produced, 51.4% of the produced pairs are charged while 48.6% are neutral [38]. While the decay rate of $B \rightarrow KS$ is the same for both charged and neutral B mesons, since the only difference are in the spectator quark that is not directly involved in the decay and the masses of the charged and neutral B meson differ very little, the two particles do have different lifetimes. This means that their branching ratios differ accordingly. We can summarise this as

$$\begin{aligned} N_{\mu^+\mu^-} &= 2N_{B\bar{B}} \left(0.486 \frac{\tau_{B^+}}{\tau_{B^0}} + 0.514 \right) Br(B^\pm \rightarrow K^\pm S) Br(S \rightarrow \mu^+\mu^-) \quad (96) \\ &\quad \times \int_{\text{Belle II}} dr d\vartheta_0 \frac{1}{2} \frac{\sin \vartheta_0}{d_S} e^{-\frac{r}{d_S}} \\ &= 1.93N_{B\bar{B}} Br(B^\pm \rightarrow K^\pm S) Br(S \rightarrow \mu^+\mu^-) \int_{\text{Belle II}} dr d\vartheta_0 \frac{1}{2} \frac{\sin \vartheta_0}{d_S} e^{-\frac{r}{d_S}}. \end{aligned}$$

In addition to using both neutral and charged B mesons, we can also add additional kaon resonances to our search. While K , the lowest kaon resonance, is a pseudoscalar particle, the other relevant resonances are not. Thus, we use other production branching ratios

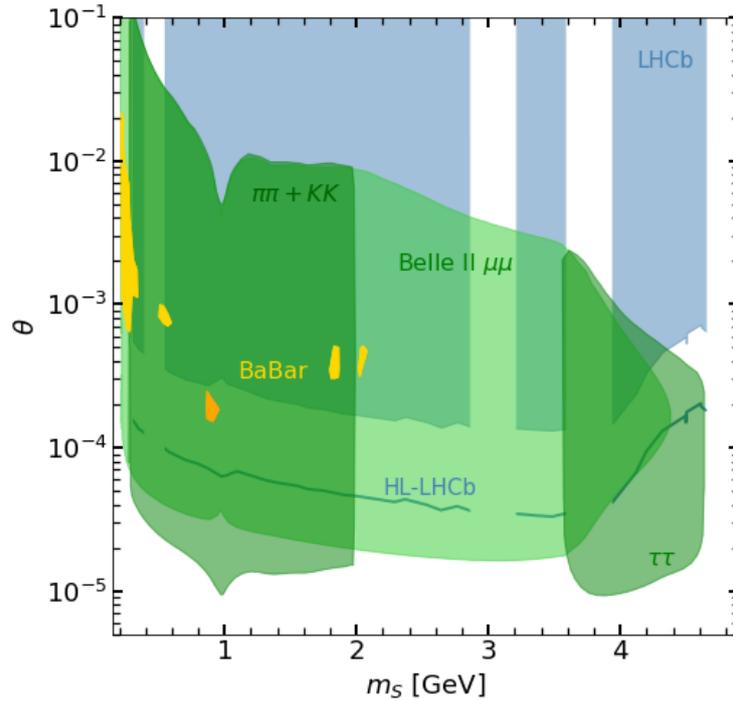


FIGURE 13: Shown are exclusion bounds in the parameter space of the mass m_S and the mixing angle θ of our new scalar. In blue, we see the 95 % CL exclusion bounds from the search for $B^+ \rightarrow K^+ S$, $S \rightarrow \mu^+ \mu^-$ at LHCb [88, 89]. In yellow and orange, we have the 90 % CL exclusions from the inclusive displaced BaBar searches for pion and muon pairs, $B \rightarrow X_s S$, $S \rightarrow \pi^+ \pi^-, \mu^+ \mu^-$ [95]. In green we see the sensitivity region of Belle II with the final states of a muon pair, a pair of pions or kaons, or a tau pair, which is defined as the region where 3 or more events are expected, which corresponds to a CL of 95 % at zero background. As a comparison, we also show a scaled up version of the LHCb search for the High Luminosity LHC (HL-LHC) as a blue line. Taken from [1].

as calculated in [92, 93]. The kaon resonances we consider are the pseudoscalar kaon K , the scalar kaons $K_0^*(700)$ and $K_0^*(1430)$, the vector kaons $K_1(1270)$ and $K_1(1400)$, and the axial vector kaons $K^*(892)$ and $K^*(1410)$ [38]. We neglect any other kaon resonances as they are either too little known to predict or reconstruct them well, or do not give a big enough contribution to include. This is particularly true for heavier kaon resonances whose sensitivity range lies at small scalar masses m_S where, as we will see in the following section, we already have a very strong sensitivity. In addition to differing in their production branching ratios, they also have different form factors. For the pseudoscalar kaon we use the form factors from [94], for the vector kaons we use [86], and for the other kaons we follow [93].

3. Displaced decay sensitivity

In figure 13, we show the sensitivity region of this search. It is defined as the region where at least 3 signal events are predicted, which corresponds to a rejection of the background-only hypothesis at 95 % CL, that is, if 3 or more events are found, which we would expect from this model in the shown area, then we are confident at 95 % that this cannot be explained by purely the Standard Model. This is valid under the assumption that all events are measured with 100 % efficiency and that there is no background in this channel. Due to this, the full realistic experimental reach is likely a little smaller. An efficiency ε smaller than 100 % has an effect of changing the sensitivity by a factor of $\sqrt{\varepsilon}$, while a small background of $N_{bg} \leq 3$ only changes the bound strength by a factor of around 2. A dedicated study of the detector response and the background predictions would have to be done to fully work out a realistic sensitivity region.

In addition to the displaced decay into muon pairs, we also consider the channels into displaced tau pairs and displaced pairs of pions or kaons. The pion and kaon channel ends abruptly at a scalar mass of 1 GeV because this is where our model of calculating the branching ratio into them breaks down [85]. The tau channel starts late because the mass m_τ is much higher than that of the muons, pions, and kaons. A $D\bar{D}$ channel would likely have a similar reach to the $\tau^+\tau^-$ channel, but its predictions have a very high uncertainty due to the charmonium resonances in the region [85]. This uncertainty also affects the sensitivity regions for the taus and muons through the lifetime, so the predicted sensitivity region near charmonium resonances should be treated with a little less certainty than the rest of the mass range.

The sensitivity region spans a range of mixing angles from 10^{-2} to 10^{-5} . On the upper edge of the curve, the region ends because most of the scalars produced decay too far inward inside the detector, in the prompt region ($\rho < 0.05$ cm). Most of the area where this is the case is already excluded by prompt searches and the LHCb search. The sensitivity at the lower edge of the curve is doubly suppressed. On the one hand, at such small mixing angles only very few scalars are even produced in the first place. On the other, the few scalars that are produced also have a very long lifetime and tend to decay outside of the viable detector region.

The shape of the contours is explained by the lifetime of the scalar. At 1 GeV, there is a dip in the upper edge of the contour, a lessening of reach in the lower edge for muons, and a heightening for pions and kaons. This is due to the fact that there is a resonance in the decay of S . Near 1 GeV lies the scalar hadronic resonance $f_0(980)$ which decays to pion or kaon pairs and makes up a large amount of the decay width of S at this mass. Thus, the lifetime is lower and the upper bound thus shrinks. On the other end,

the muon branching ratio gets smaller and so does its bound, but the branching ratio to pions and kaons increases significantly and so they have an improved bound around this resonance, too. Similarly, the tau channel makes a big improvement over the muon channel once it is kinematically viable, since its branching ratio is much larger due to the mass-hierarchical couplings. We can see from the sensitivity regions for our three channels that while muons may seem the most promising overall, the meson and tau channels add regions of the parameter space that the muon cannot explore.

The BaBar search shown in figure 13 was the first search for long-lived particles at an e^+e^- collider. It is an inclusive search, meaning that instead of choosing specific mesons for the B meson to decay into like we do, they only require that the hadronic particles that are produced contain an s quark, $B \rightarrow X_s S$. They then look for displaced decays of the scalar into pion pairs or muon pairs, $S \rightarrow \pi^+\pi^-, \mu^+\mu^-$. Their results are quite promising, even though they look small in comparison to our predictions, as they had to cut out the mass regions around known Standard Model resonances due to their inclusive search strategy and had only a limited data set available that was not optimised for long-lived particle searches. In particular, the luminosity available for the BaBar search was 489 fb^{-1} [95], while our Belle II sensitivity region is projected for 50 ab^{-1} .

The LHCb bound shown in figure 13 is a reinterpretation of the studies [88, 89] which look for displaced muon pairs from long-lived scalars in the processes $B^+ \rightarrow K^+ S$, $S \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow K^* S$, $S \rightarrow \mu^+\mu^-$. The cut out regions for the SM resonances K_S , $\psi(2S)$ and $\psi(3770)$ are partly covered by the combination of the two searches in our reinterpretation. To predict the improvement of the bound through the high luminosity run of the LHC (HL-LHC), we have scaled up the current LHCb bounds with the additional luminosity of the HL-LHC following [96] under the assumption that the displaced region stays background free. Even with this scaled HL-LHCb bound, Belle II's sensitivity region still reaches lower mixing angles. This is due to the lower boost of Belle II, that allows its particles a longer effective lifetime to traverse the detector. While at LHCb only particles of lifetimes up to about $c\tau \approx 30 \text{ cm}$ decay inside the vertex detector in relevant amounts, at Belle II the scalars regularly decay within the tracker with lifetimes of up to about $c\tau \approx 10 \text{ m}$.

It is interesting to compare the results in figure 13 where we use all relevant kaon resonances with a scenario where we look for just the simplest case, only the charged process with the lightest pseudoscalar kaon, $B^\pm \rightarrow K^\pm S$. We show this comparison in figure 14. The inclusion of the neutral processes and the addition of the other kaon resonances leads to an improvement over the simple case of a factor of two or more for most of the lower edge of the sensitivity region. A full reconstruction of all kaon resonances, especially the heavier ones, is experimentally quite difficult, so that not

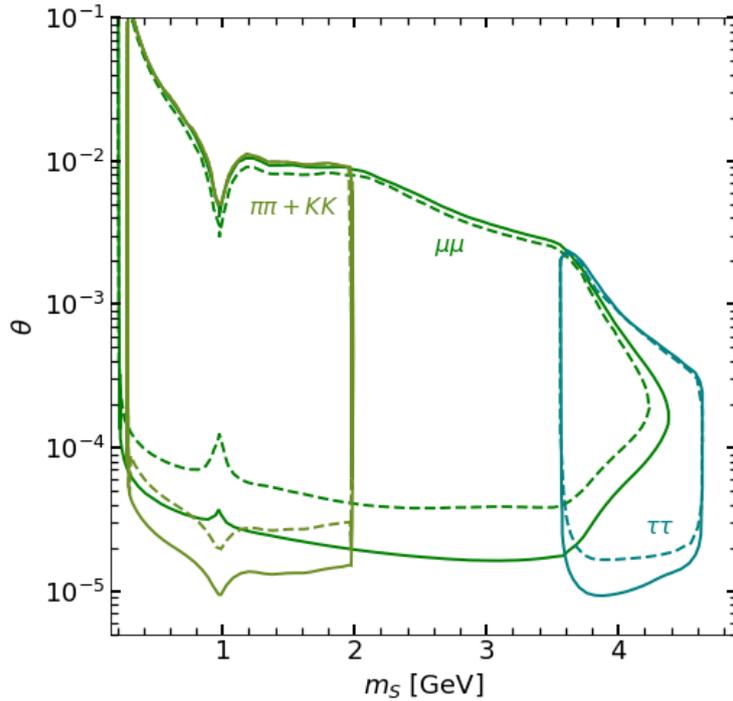


FIGURE 14: Comparison between the full sensitivity region for a search with all relevant kaon resonances and both charged and neutral processes, shown with a solid line, and the simpler case of only the charged process $B^\pm \rightarrow K^\pm S$ with the lowest-mass pseudoscalar kaon, shown with a dashed line. Both cases are shown for all three channels, as labelled in the plot. Taken from [1].

all resonances may be able to contribute as much as is shown here in a fully realistic scenario. The true reach of Belle II likely lies between the two scenarios shown here.

At the time of the writing of this paper, Belle II had not yet started taking data, which is why we found it important to compare it with other experiments whose run-time would also lie in the future. One of these is the high-luminosity run of the LHC, the LHCb prediction of which we have already shown in figure 13. In figure 15, we compile more such experimental predictions, comparing the full sensitivity region of all three channels with running, funded, and proposed future experiments. The bounds from the beam-dump mode of NA62 come from [85, 97], the ones for the funded long-baseline experiment FASER from [98] and its proposed future stage from [98, 99]. The bounds for the other proposed long-baseline experiments come from [100] for SHiP, [101] for CODEXb, and [102] for MATHUSLA.

It is impressive that Belle II has a similar sensitivity to many of these long-baseline experiments while using just the tracker with a total extension of 113 cm in transversal direction ρ . Still, the large proposed long-baseline experiments would be able to explore more than an additional order of magnitude in the mixing angle and thus greatly improve

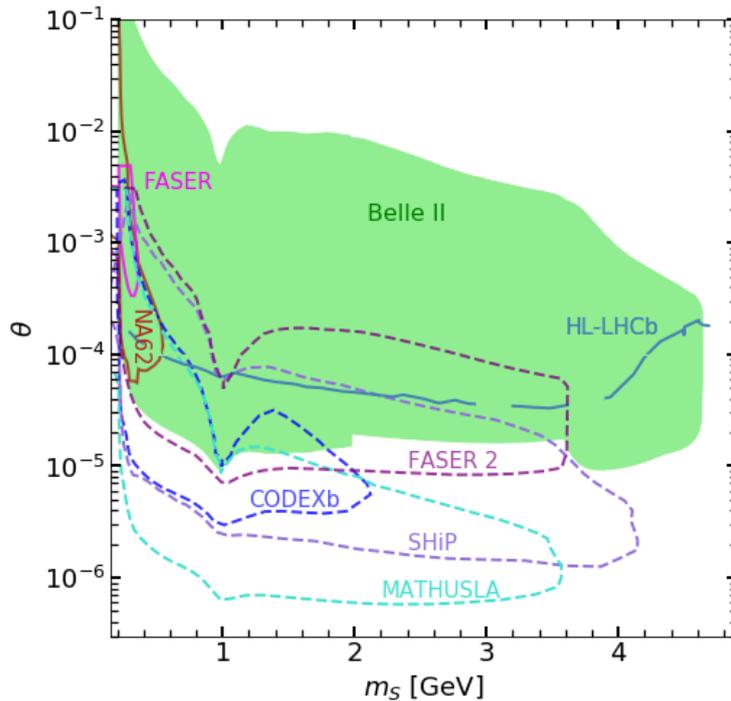


FIGURE 15: Comparison between our projections for the Belle II sensitivity and proposed future experiments (dashed lines) and running or future funded experiments (solid lines). More details in the text. Taken from [1].

on this result. As they are still awaiting funding, though, the promise of Belle II, which is at this point taking data and releasing first analyses [103–105], is a great step in the direction of uncovering new parameter space through displaced decays.

Overall, we have shown the suitability of Belle II as an experiment for long-lived particle searches. It benefits from its low boost and large solid angle, as well as from the clean background. Doing a search for displaced pairs of muons, pions, and kaons would follow through on the promise we saw in the BaBar search [95] and probe a large region of the parameter space. It could especially explore parameter space beyond what searches at the LHC seem to be able to achieve, and probe parameter regions that are usually seen as the purview of long-baseline detectors. This paper’s promising result of a large reach for long-lived particle searches at Belle II has greatly inspired the papers in the rest of this section III. It has also led to increased interest within the Belle II collaboration to work on long-lived particle searches in this and other models.

B. Far Detectors at electron-positron colliders

As we saw in the last section, Belle II outcompetes prompt and displaced measurements at the main detectors of the LHC, but far detectors at the LHC are expected to do

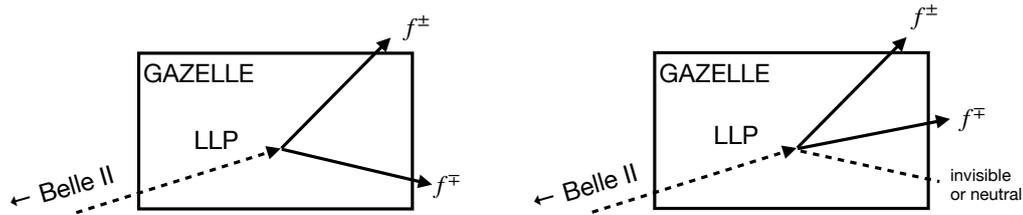


FIGURE 16: Signatures of a long-lived particle decaying into two charged particles in a far detector, with (right) or without (left) the addition of neutral particles. Taken from [4].

at least as well if not better than Belle II in looking for displaced decays, at least in the specific Higgs portal model we considered in the last section. This naturally begs the question whether a far detector placed at Belle II would also be able to greatly improve on its sensitivity. In this section, we will be going through [4] exploring this very question. We will first cover the design of the far detectors for Belle II in section III B 1, then moving on to the benchmark models and the detector modelling in section III B 2, before exploring the reach of the far detectors in comparison to the main detector in section III B 3. Lastly, in section III B 4 we will follow [6] in comparing far detectors and the main detector for the International Linear Collider (ILC). Anything not otherwise cited will come from these two references.

The signatures of long-lived particle decay we will be considering in this section are displaced decays into charged particle pairs, just like the decay of S to two muons or pions in the last section. We show a sketch of these signatures in figure 16. We know that Belle II can detect these kinds of signatures well. Now we want to construct far detectors that can do the same and are still cost-effective and realistic.

1. Detector Design

Let us start with the experimental requirements needed to detect charged particle tracks. For that, we first and foremost need a tracking system with a sufficiently high vertex resolution. As we intend to make far detectors, the exact precision of the vertices is not of a high importance, such that we settled on a vertex resolution of 10 cm. To make sure that we have, in fact, an incoming neutral particle decaying into two charged particles and not an incoming charged particle decaying into another charged particle and neutral particles, we want to be able to tell the flight direction of the tracks in the detector. For this, we need a timing resolution. This can also be used to determine the velocities of the particles leaving tracks. To measure the velocities of these particles well, we set this timing resolution requirement at 100 ps. As a last requirement, we need the latency of the detector response to be around $1 \mu\text{s}$ so that the far detector can be included in

the Belle II trigger. The trigger is the system of the detector that uses information about an event that is currently occurring in the detector to decide whether to save or discard it. Including the far detector in the trigger means that the far detector could trigger both itself and Belle II, and also that it would be triggered when an interesting event is detected in the main detector of Belle II. It might even open the possibility for defining more complicated states to trigger on where specific conditions are detected in both Belle II and the far detector.

While I will not go into specifics here on how such a detector can be achieved, these detector capabilities are possible and affordable and are discussed in [4]. Particularly the absence of a calorimeter and a magnetic field spanning the detector reduce the cost for such a far detector. A calorimeter would allow us to detect further final states including photons and neutral hadrons, which would broaden the physics possibilities a little but come at a large price increase. A magnetic field spanning the detector would allow us to fully reconstruct the momenta of particle tracks rather than just their velocities. While this would again broaden the physics case, allowing for additional and more specialised searches, it is simply not feasible for detectors of the size we propose. The cost of the potential far detector is especially crucial as it would be an addition to an already existing detector and as such has a much smaller budget.

A detector with these components can tell us a lot about the charged particle tracks we measure within them. From the charged particle tracks, we can first determine the vertex that they are produced at. From the direction of the two tracks, we can then also reconstruct the direction of our long-lived particle (LLP) and thus whether it came from the direction of the collision point at the centre of Belle II. From the opening angle between the tracks and the velocities of the particles that make them, we can reconstruct the mass of the LLP. We can also tell the absolute time, which we can use to check if the event seen happened coincidentally with a collision at Belle II and can thus be connected to it.

We can already see that these properties of the detector will greatly help in suppressing much of the background in the far detector. There are roughly three types of backgrounds that we need to take into account. The first is neutral hadrons coming from the main collision, shown on the upper left of figure 17. K_L^0 mesons are long-lived enough that they could be a background for a far detector, depending on its exact distance, and they dominantly decay into $\pi^\pm \ell^\mp \nu_\ell$ [38], making them an obvious background for our far detector. Most of the mesons get caught in the shielding around Belle II, though, so that they would not make up a big proportion of the background. A bigger problem are muons produced in the primary collision or as secondary decay products of the produced particles that interact with the shielding. There, they can produce new K_L^0

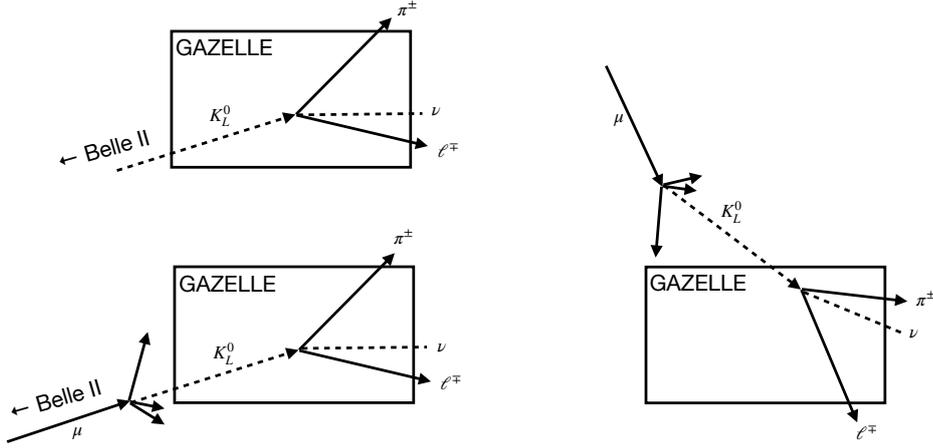


FIGURE 17: Signatures of backgrounds from kaon decays at Belle II’s far detector. On the upper left, we see a K_L^0 background from direct production in the main collision.

On the lower left, we see a K_L^0 background induced by a muon μ from the main collision. On the right, we see a K_L^0 background induced by a muon from cosmic rays. Taken from [4].

and other neutral hadrons like K_S^0 and Λ^0 that would not otherwise be long-lived enough to be backgrounds to the far detector on their own, were they produced in the primary collision. This is shown on the lower left of figure 17. Since this background originates from muons that are produced in the main detector, rejecting all events that look like an LLP but point in the direction of a coincident muon in the main detector can reduce this background significantly.

The second source of background events are cosmic rays. While the fraction of neutral hadronic Standard Model LLPs is negligible, there is a high amount of muons in cosmic rays that can induce neutral hadrons in the shielding and surroundings of the far detector as discussed in the last paragraph. This is shown on the right in figure 17. Since these muons do not come from a source that we measure, we have to find another way to reject such background events. Luckily, the cosmic ray muons come from the direction of the sky, meaning that the direction the tracks point to is very likely to not be the collision point of Belle II.

A third background of note is the misreconstruction of a muon decay as the decay of a neutral particle into two charged tracks, shown in figure 18. The time resolution of the far detector should usually get rid of this, and we have the additional bonus of being able to reject decays of muons produced in the main detector based on the wrong direction they point towards. This may not be the case for muons from cosmic rays, but there we find that the mass of the misreconstructed LLP is usually too high to have been produced at SuperKEKB. Signal events have additional properties that these background events do not have. They are related in time to the collisions at Belle II, though the exact time

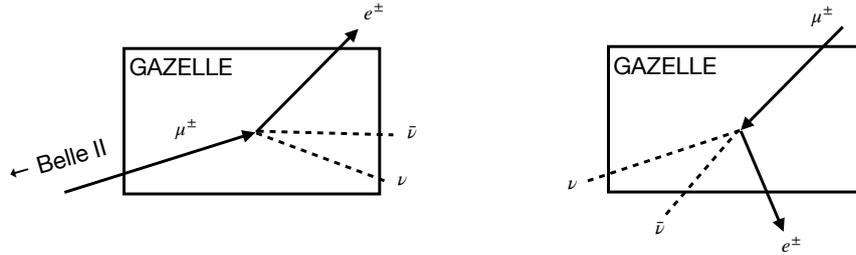


FIGURE 18: Signatures of backgrounds from muon decays at Belle II’s far detector, shown for muons from the primary collision on the left and for cosmic muons on the right. Taken from [4].

difference between the initial collision and the LLP decay in the far detector depends on the LLP mass and can thus not be restricted too closely. When the event in the far detector is coincident with a primary collision in the main detector, we can also check if the momentum of the LLP that we reconstruct matches the missing momentum found in the Belle II detector. For models in which the LLP decays only into two charged particle tracks and not into any additional neutral particles, the tracks and the line between the reconstructed decay vertex and the collision point at Belle II all lie on a plane, which is very unlikely to be the case for background events, where it can only happen by chance.

As we can see, the placement of the shielding in relation to the far detector is somewhat of an optimisation problem between placing enough of it to shield the detector from neutral particles and placing it far enough away that any muon-induced hadrons do not decay within the detector. The cosmic ray backgrounds could be reduced by adding an active veto on top of the detector, such that we can reject muon-induced backgrounds better. Overall, we are optimistic that the background is reducible to a very low level, though a full background study would be needed to produce quantitative results. For the remainder of this section, we will be assuming zero background events and 100% efficiency in detecting and reconstructing signal events.

Due to the low background expected and the long-lived particle search aim of it, we call the far detector we propose here GAZELLE. GAZELLE is the Approximately Zero-background Experiment for Long-Lived Exotics. Its name is a recursive acronym.

Let us now consider the specific shape and positioning of GAZELLE. To propose a realistic detector, we need to consider the space around Belle II and how much of it is unoccupied enough to place another detector into. Within the confines of Tsukuba hall, the hall in which Belle II resides, we propose two designs, with a third detector proposed close-by.

The first detector, which we will call BabyGAZELLE (labelled BG for short) on account

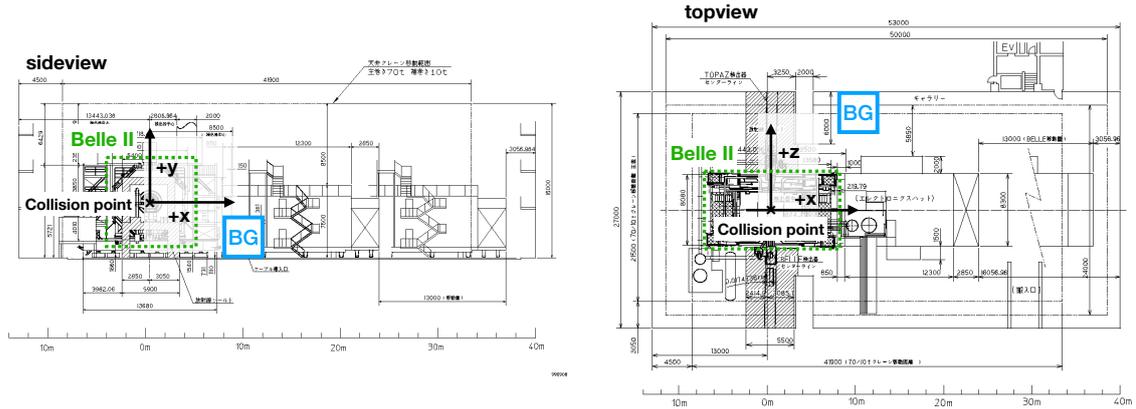


FIGURE 19: Schematic sideview (left) and topview (right) of the proposed far detector BabyGAZELLE at Belle II. Taken from [4].

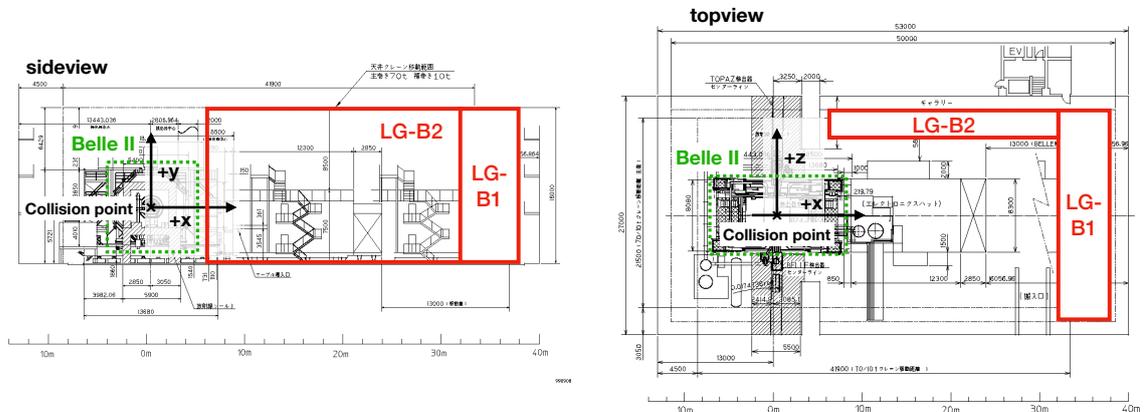


FIGURE 20: Schematic sideview (left) and topview (right) of the proposed far detector L-GAZELLE at Belle II. Taken from [4].

of it being the smallest, is a cube of 4 m side-length positioned on the floor of Tsukuba hall as close to the beam-line as possible and as forward as possible. This detector would also work well as a proof-of-concept first stage for a larger GAZELLE, like FASER is to FASER 2 [98] or CODEX- β to CODEX-b [106].

The second detector design we propose will be called L-GAZELLE (labelled LG for short) for its shape. It would be constructed along two of the walls in Tsukuba hall, the wall in forward beam direction and the one on the far side of the hall from the detector, parallel to the beam. We label these two parts of the detector B1 and B2 respectively. The detectors would span the whole height and length of the walls and have a depth of 6 m in the case of B1 and 3 m in the case of B2.

The last detector proposed is also the biggest, which is why it is named GODZILLA (labelled GZ for short). Instead of being inside of Tsukuba hall, it would occupy an empty plot close to it, taking up the space of a whole building.

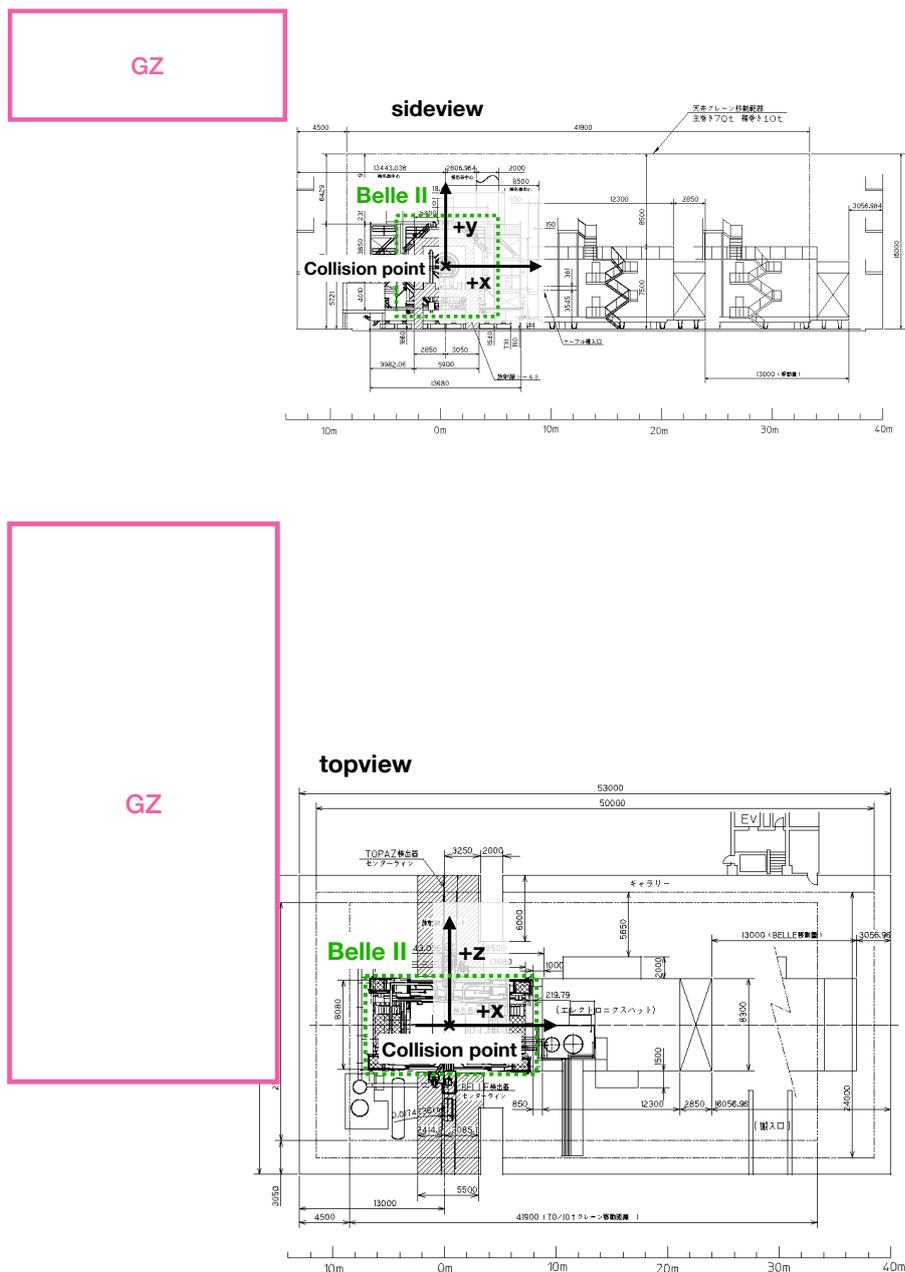


FIGURE 21: Schematic sideview (top) and topview (bottom) of the proposed far detector Godzilla at Belle II. Taken from [4].

The full measurements of all three detectors are shown in table I. Their coordinate system is built around the collision point within Belle II as its origin with the z axis following the forward beam direction, the y direction pointing straight up, and the x direction pointing towards the far end of Tsukuba hall, perpendicularly to the other two. The detectors are also shown along with their placement in or around Tsukuba hall and the Belle II detector in the figures 19 (BG), 20 (LG), and 21 (GZ).

Detector		Size [m]			Position [m]			Solid angle Ω	
Name	Abbr.	x	y	z	x	y	z	[sr]	[%]
BabyGAZELLE	BG	4	4	4	10	-3.7	10	0.12	0.95
L-GAZELLE B1	LG-B1	6	16	24	35	2.3	0	0.34	2.7
L-GAZELLE B2	LG-B2	26	16	3	19	2.3	10.5	0.76	6.0
GODZILLA	GZ	25	10	50	-27	18	20	0.74	5.9

TABLE I: The size and position of all three GAZELLEs, as well as their solid angles

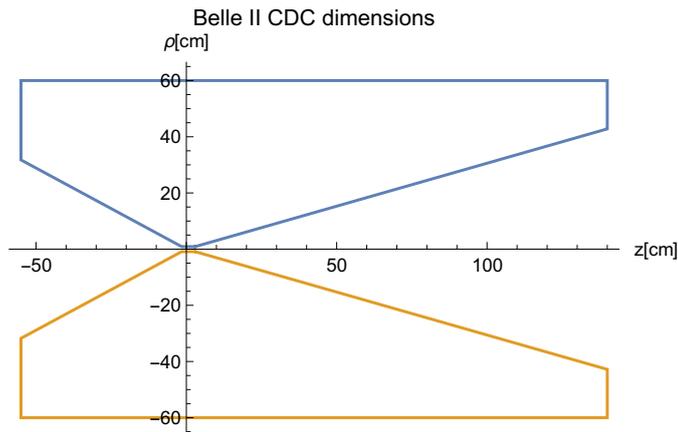


FIGURE 22: The representation of Belle II's tracker, the CDC, in the two-dimensional plane of ρ and z .

For our comparison to Belle II we use the measurements

$$\begin{aligned}
 17^\circ &< \vartheta < 150^\circ & (97) \\
 -55 \text{ cm} &< z < 140 \text{ cm} \\
 0.9 \text{ cm} &< \rho < 60 \text{ cm}.
 \end{aligned}$$

It is shown in figure 22. In comparison with the tracking system described in section III A 1 and shown in figure 11, the radial extension here is chosen to be much smaller. This is due to the fact that to detect tracks and to reconstruct vertices, a certain amount of hits in the detector are needed. So while particles decaying at a radial distance of, say, 80 cm from the central collision point still decay in the detector, their tracks may not be able to be reconstructed to find their decay vertex with a high certainty, which is why we limit the detector size here. The solid angle of the Belle II tracker is $\Omega = 11.5 \text{ sr}$ (90%). We will abbreviate the Belle II tracker in the following as CDC, since most of its volume is given by the CDC subdetector, as described in section III A 1.

2. Models and Modelling

Now we have to consider how to compare the different detectors with each other. We use three different models as benchmarks for different LLP scenarios. We generate samples for each benchmark model and several representative parameter points. Lastly, we analyse their behaviour regarding the different detectors with a software I wrote. In this section, we will first introduce the benchmark models and then move onto the analysis software.

The models we use to probe the LLP response are an HNL model [107, 108], an ALP model [109, 110] and a model with inelastic dark matter (iDM) [111], in which a dark scalar is the LLP. These models were chosen as a representative sample of LLP models as they all involve the use of portals and are commonly used in LLP analyses. They also nicely span the parameter space not only through their different model structures but also through the different production channels they are produced in and the different kinematics that follow from them. The HNL couples to the τ lepton through a mixing angle with the tau neutrino ν_τ and is thus produced in the decay of the τ leptons that are abundantly produced at Belle II. The ALP is produced in the decays of the equally abundant B mesons, similar to the dark scalar we discussed in section III A. The dark scalar of the iDM model, lastly, is produced in association with a dark photon directly from the e^+e^- collision. While we have considered all three models with the same care in [4], I will mostly limit this section to the ALP model with which I worked most closely.

The ALP model used here follows the effective Lagrangian

$$\mathcal{L} = -2g_{ij} \frac{\partial_\mu a}{\Lambda} \bar{d}_{Li} \gamma^\mu d_j + \frac{c_\ell}{2} \frac{\partial_\mu a}{\Lambda} \bar{\ell} \gamma^\mu \gamma^5 \ell, \quad (98)$$

giving the ALP an effective coupling to down-quark FCNCs and a diagonal coupling to leptons. As discussed in section II B 5, these couplings may either come directly from the UV theory of this ALP, or may instead only come about through its RG evolution. For simplicity, we will only set the couplings $g_{23} = g_{sb}$ and $c_\ell = c_\mu = c_e$ to nonzero values here. While other configurations will surely have a similar qualitative behaviour, their dependence of the parameters may be significantly different and more complicated. As this is not a study focussed on exploring the full parameter space of the ALP, but the space of possible far detectors, we will work only with this specific model configuration.

In this model, the production rate of the ALP in rare $B^\pm \rightarrow K^\pm$ decays is [112]

$$\Gamma_{B^\pm \rightarrow K^\pm a} = \frac{|g_{sb}|^2}{16\pi\Lambda^2} |f_0(m_a^2)|^2 \frac{(m_B^2 - m_K^2)^2}{m_B^3} \lambda^{\frac{1}{2}}(m_B, m_K, m_a) \quad (99)$$

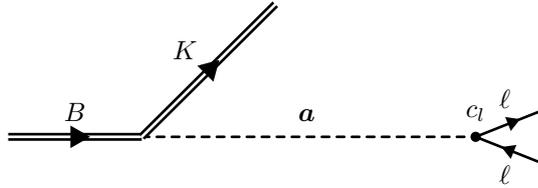


FIGURE 23: The production and decay of the ALP a defined by the Lagrangian in equation (98).

with the scalar hadronic form factor $f_0(q^2)$ [86] and $\lambda(a, b, c)$ defined in equation (94). The ALP's decay rate to leptons is [76, 113]

$$\Gamma_{a \rightarrow \ell^+ \ell^-} = \frac{m_\ell^2 |c_\ell|^2}{8\pi\Lambda^2} m_a \sqrt{1 - \frac{4m_\ell^2}{m_a^2}}. \quad (100)$$

In figure 23, we show the Feynman diagram of the ALP production and decay.

Now that we have our model in place, we generate 10,000 sample events for every choice of parameters in each model with the help of `EvtGen` [114] (or, in the cases of iDM and HNLs, `MadGraph` [115]). The resulting `.hepmc`-files are then run through a `MadAnalysis5` [116–123] script I wrote to read out the LLP's momentum and production vertex and write them to another file that is the input for the analysis program. I wrote the main program in `Mathematica` [124] and while we focus on the ALP case here it was used for the analyses of all three models.

After reading in the momenta and production vertices of the LLPs in every event, the program determines the probability that the LLP decays within a given detector. For this, the three GAZELLEs are coded into the program as cuboids and the code first checks if the ray given by an LLP's production vertex \vec{x}_0 and its momentum \vec{p} intersects with any of the cuboid's faces,

$$\vec{x}_0 + r \cdot \vec{p} = \vec{f}_0 + s \cdot \vec{f}_1 + t \cdot \vec{f}_2 \quad (101)$$

where \vec{f}_0 is one of the corners of the cuboid's face, and \vec{f}_1 and \vec{f}_2 are vectors pointing along opposite edges of the cuboid away from \vec{f}_0 , generating the surface. The variables r , s , and t are determined via `Mathematica`'s `Solve` function. If there is an intersection, we then check if it happens on the cuboid's surface rather than just in the plane of it, i. e.,

$$0 \leq s \leq 1 \quad 0 \leq t \leq 1, \quad (102)$$

and whether it happens in the flight direction of the particle,

$$r \geq 1. \quad (103)$$

If this is the case for any of the faces of the cuboid, then we know that the particle intersects with the detector as a whole. We then determine from the intersection points where the particle enters (\vec{x}_{in}) and exits (\vec{x}_{out}) the detector. From this, we determine the length of flight of the LLP at the moment of entering,

$$\ell_{in} = |\vec{x}_{in} - \vec{x}_0|, \quad (104)$$

and exiting,

$$\ell_{out} = |\vec{x}_{out} - \vec{x}_0|, \quad (105)$$

the detector. The probability that the LLP decays within a given detector is 0 if it does not intersect with in. Otherwise, the probability that the LLP decays within the detector is given by the difference in probabilities of it not decaying before entering the detector and it only decaying after the detector,

$$\mathbb{P}_i = \begin{cases} 0 & i \text{ misses detector} \\ e^{-\frac{\ell_{in}}{\gamma\beta_i c\tau}} - e^{-\frac{\ell_{out}}{\gamma\beta_i c\tau}} & i \text{ intersects between } \ell_{in} \leq \ell \leq \ell_{out}. \end{cases} \quad (106)$$

The probability is then averaged over all available samples,

$$\langle \mathbb{P} \rangle = \frac{1}{N} \sum_i \mathbb{P}_i. \quad (107)$$

This procedure works well for the cuboid GAZELLEs but is less applicable to the cylindrical Belle II itself, where the faces are less easy to work with in Cartesian coordinates. For Belle II, then, I wrote another version of the code following the same principals. There, Belle II is translated into cylindrical coordinates as already defined above, with *rho* describing the radius in transversal direction, *z* being the direction of the beam-line and thus the barrel direction, and the angle around the barrel being given by φ . Since Belle II is (approximately) invariant in φ , we can reduce the detector down to a shape in ρ - z space as shown in figure 22.

We then similarly reduce the particle ray's coordinates to the transversal and longitudinal directions as well. Then, for every edge of the irregular polygon that is the Belle II

tracker, we check if the particle ray intersects with it by solving

$$\begin{pmatrix} \rho \\ z \end{pmatrix} + r \cdot \begin{pmatrix} p_T \\ p_z \end{pmatrix} = \begin{pmatrix} e_0^T \\ e_0^z \end{pmatrix} + s \cdot \begin{pmatrix} e_1^T \\ e_1^z \end{pmatrix} \quad (108)$$

where \vec{e}_0 is one of the corners of the edge and \vec{e}_1 points in the direction of the other all along the edge. As before, there is an intersection if $r \geq 0$ and $0 \leq s \leq 1$. From the intersections with the edge we can again get ℓ_{in} and ℓ_{out} , from which the probability definitions follow as above.

Since we want to compare the reaches of the different GAZELLES to the reach at Belle II, we will later look at probability ratios instead of probabilities themselves, normalising $\langle \mathbb{P} \rangle_{BG/LG/GZ}$ by $\langle \mathbb{P} \rangle_{CDC}$ to discuss the different ways these probabilities evolve with the LLP couplings. In addition, we will also look at the absolute number of events that we expect to decay in each of the GAZELLES or in the Belle II tracker. This can be defined, here for the ALP case, as

$$N_{\text{events}} = N_{B\bar{B}} \times Br_{B^\pm \rightarrow K^\pm a} (m_a, g_{sb}) \times Br_{a \rightarrow \ell^+ \ell^-} (m_a, c_\ell) \times \langle \mathbb{P} \rangle (m_a, c_\ell). \quad (109)$$

This is valid with the assumptions of there being no background events and the efficiency in detecting any event being 100%. Comparing equation (109) with equation (95), we can see that we here switch out the integration over the volume of the detector with an average over a large sample size of particles probing just this detector volume. For the ALP, we use the maximal non-excluded value of g_{sb} in the calculation of N_{events} . These values are extracted from [125].

3. Long-Lived Particle Reach

Now that we have introduced our model and our analysis procedure, let us take a look at the results we receive. In figure 24, we see the probability ratio of GAZELLE over Belle II for each of the three GAZELLES and a range of ALP masses m_a . As we can clearly see, for small couplings which corresponds to long lifetimes GODZILLA (and for $m_a = 4 \text{ GeV}$ also L-GAZELLE) outperform Belle II, i. e., it is more likely for a very long-lived ALP to decay within GODZILLA than Belle II. In contrast, the probability for an ALP to decay within BabyGAZELLE reaches barely 10% of the decay probability within Belle II even for the highest mass.

The drop in the probability ratio corresponds to the couplings where the ALP becomes short-lived enough to decay dominantly within the Belle II tracker. There is a noticeable difference in probability between the light masses and the $m_a = 4 \text{ GeV}$ case in

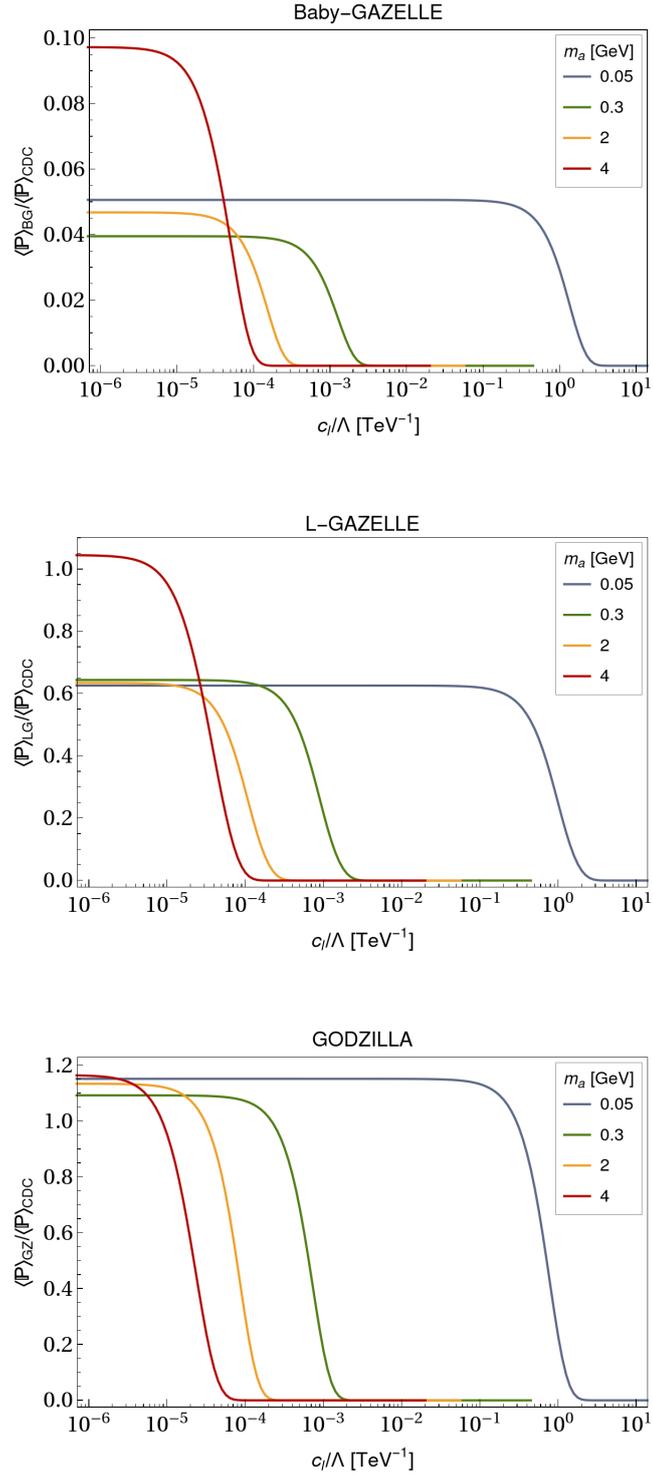


FIGURE 24: The ratio of the mean probability to decay within any of the three GAZELLEs normalised by that of decaying within the CDC as a function of the effective ALP coupling to leptons, $\frac{c_\ell}{\Lambda}$ and several ALP masses m_a . Taken from [4].

m_a [GeV]	g_{sb} [10^{-5}]	L-GAZELLE	Belle II	LG/CDC
0.3	3.9	$1.4 \cdot 10^{-5}$	$1.2 \cdot 10^{-5}$	1.2
2.0	3.8	$1.7 \cdot 10^{-6}$	$1.4 \cdot 10^{-6}$	1.3
4.0	3.5	$4.4 \cdot 10^{-7}$	$4.5 \cdot 10^{-7}$	1.0

TABLE II: The projected reach of L-GAZELLE and Belle II in the effective ALP coupling to leptons, $\frac{c_\ell}{\Lambda}$ for several ALP masses m_a . The second column shows the maximal allowed FCNC coupling from [125]. The last column shows the ratio between the reach of L-GAZELLE and Belle II. A number smaller than one means that L-GAZELLE has a stronger sensitivity than Belle II, while a number larger than one means that Belle II has the stronger sensitivity. Taken from [4].

L-GAZELLE and BabyGAZELLE. At $m_a = 4$ GeV, the ALP has to carry most of the momentum of the B meson that decays into it, and thus inherits most of its forward boost. Since both detectors are in the forward beam direction and get close to the beam, their probability to have the more forward ALP decay in them rises, in contrast to the lower, more isotropic masses, and GODZILLA, which is not placed in the forward direction.

Figure 25 shows the number of events expected for the GAZELLES and the different mass benchmarks. Also shown is the 3 events line which corresponds to a confidence level of 95 % that such an occurrence could not be explained by the null hypothesis, that is, without the ALP. We see that for higher masses, the probeable region of coupling becomes lower, as we also saw in figure 13 of section III A. While the lifetime gets smaller as the particle gets heavier and more decay channels open up, the velocity of the particle also gets much slower and it thus spends more time within the detector, giving us access to lower couplings. In table II, we show not only the maximal allowed value of g_{sb} from [125] for three masses, but also the lowest probeable coupling c_ℓ for both Belle II and L-GAZELLE at this mass, and the ratio of the two. This ratio tells us how much we can improve on the Belle II prediction with GAZELLES.

From both figure 24 and table II, we can see that the GAZELLES do not offer significant improvements over what we predict for Belle II. This turns out to be for geometric reasons. For particles that are very long-lived with regard to the detector we consider, $d = \gamma\beta c\tau \gg \ell_{in}, D$, the probability approximates to

$$\mathbb{P} = e^{-\frac{\ell_{in}}{d}} - e^{-\frac{\ell_{in}+D}{d}} \quad (110)$$

$$\approx \Omega \times \frac{D}{d} \quad (111)$$

in the limit of a fully isotropic angular distribution. The “goodness” of the detector, its fiducial acceptance, is given fully by its geometric properties, its solid angle Ω and its radial depth D . With the values from table I, we can now calculate the fiducial

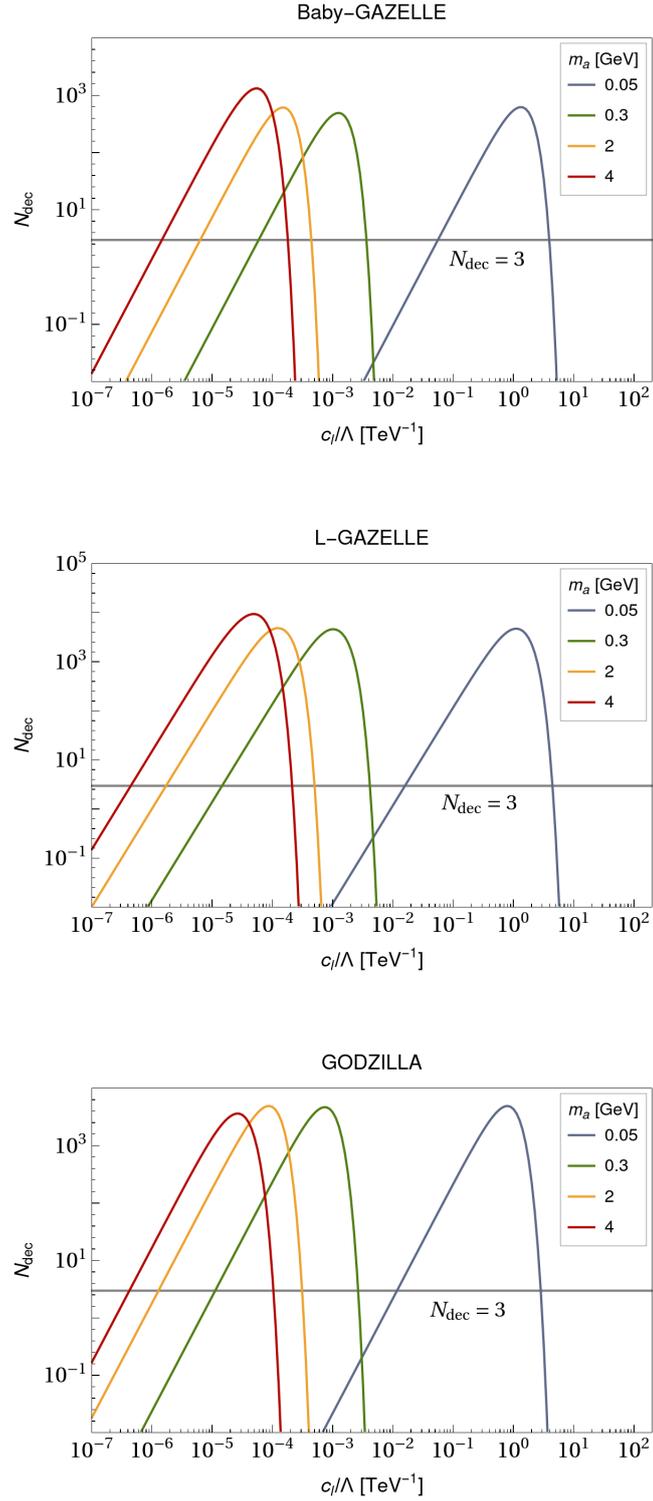


FIGURE 25: The expected number of events in the three GAZELLEs as a function of the effective ALP coupling to leptons, $\frac{c_\ell}{\Lambda}$ and several ALP masses m_a . Taken from [4].

acceptances for all three GAZELLES and for Belle II's tracker in comparison. They are

$$\begin{aligned}
 \text{Belle II} &: 11.5 \text{ sr} \times 0.6 \text{ m} &= 6.9 \text{ sr m}, & (112) \\
 \text{BabyGAZELLE} &: 0.12 \text{ sr} \times 2.5 \text{ m} &= 0.3 \text{ sr m}, \\
 \text{L-GAZELLE} &: 0.34 \text{ sr} \times 5.3 \text{ m} + 0.76 \text{ sr} \times 4.1 \text{ m} &= 4.9 \text{ sr m}, \\
 \text{GODZILLA} &: 0.76 \text{ sr} \times 4.1 \text{ m} &= 8.9 \text{ sr m}.
 \end{aligned}$$

As in the earlier plots, we see that BabyGAZELLE has a much smaller acceptance than the other detectors. It is both too thin and has too small of a solid angle as seen from the collision point to make much of an impact. On the other hand, L-GAZELLE and GODZILLA have fiducial acceptances of the order of Belle II's, explaining the closeness of their predicted number of events and lower bounds on the coupling, too.

Of course the fiducial acceptance does not always equal the probability quite as easily as in the very long-lived case, and not all models (under all choices of parameters) have isotropic distributions of LLPs, as we saw for the heavy ALP. Still, the fiducial acceptance is a good measure for comparing different detectors and their reach for long-lived particle searches. We can now see, that there is no possibility of an improvement of an order of magnitude or more without improving either the solid angle or the radial thickness of any of the detectors by an order of magnitude as well. Unfortunately, this is not possible within the constraints of Belle II's existing placement in Tsukuba hall. In the next section, we will thus discuss the usefulness of far detectors for a detector that is still in development, the International Linear collider (ILC).

Before that, however, let us quickly gather what we have concluded so far. Far detectors at Belle II, specifically L-GAZELLE and GODZILLA, are able to add $\mathcal{O}(1)$ contributions to the sensitivity to long-lived particle decays. While this is no overwhelming improvement, it is also not negligible. In addition to this small improvement, there are non-isotropic models and models with different signatures than the simple displaced decay vertex that one could discuss that would gain still more from a far detector, as shown in section 6 of our paper [4]. Even for the models shown here, there may be improvements our simple analysis could not encapsulate, like effects from the connection of the triggers of Belle II and GAZELLE which might lead to a higher chance of catching interesting missing energy signatures in the main detector and connecting them with the decay of the missing particles in GAZELLE. A third advantage to a far detector would be the measurement of LLP properties should one be detected, especially if it is first detected in missing energy searches that are not able to characterise their finds as well as a directly measured displaced decay. We will study the complementarity of displaced searches and missing energy searches for LLPs in more detail in section III C.

Detector		Size [m]			Position [m]			Fiducial acceptance		
Name	Abbr.	x	y	z	x	y	z	$\frac{\Omega}{4\pi}$ [%]	$\langle r \rangle$ [m]	$\Omega \cdot \langle r \rangle$ [sr m]
Shaft	S	18	30	18	0	45	0	2.6	16	5
Tunnel	T	140	10	10	0	-5	-35	4.6	11	6
Ground	G	1000	10	1000	0	75	0	44	23	126

TABLE III: The size and position of all three far detectors for the ILC, as well as their solid angles Ω , average radial depths $\langle r \rangle$, and fiducial acceptances $\Omega \cdot \langle r \rangle$.

4. Far Detectors at the ILC

Now let us take a quick look at the reach of far detectors at the International Linear Collider (ILC) [126], a proposed e^+e^- collider that would be run primarily as a Higgs boson factory in contrast to the B factory that Belle II is. We will work with the 250 GeV ILC with unpolarised beams and consider the ILD as its main detector. In this section, we will follow [6].

While the ILC has not been built yet, there is already a very complete plan for it and its surrounding structures. Thus, we have to take the surroundings into account in a similar way as at Belle II. After trying out several more designs, we finally settled on three. The first of these would be placed in the vertical shaft above the collision point, through which the main detector ILD will be lowered down. The second detector would be placed along the length of an access tunnel that runs in parallel to the cavern with the main detector. The third detector would be placed on the ground above the detector. While we construct the first two, which we call Shaft and Tunnel, respectively, with realistic dimensions, the third, named Ground, serves as an upper limit for large detectors and as such is itself larger than is realistic. The measurements of all three far detectors are shown in table III. Their size and positioning is shown in figure 26.

The ILD is made up of multiple subdetectors. The ones relevant for the tracking system and which we will consider here are the multi-layer vertex pixel detector (VTX) and the time projection chamber (TPC). Together they form a cylindrical decay volume which, in cylindrical coordinates, is given by

$$\begin{aligned}
 8^\circ < \vartheta < 172^\circ & \tag{113} \\
 -235 \text{ cm} < z < 235 \text{ cm} \\
 0.6 \text{ cm} < \rho < 180.8 \text{ cm}
 \end{aligned}$$

and shown in figure 27. The ILD's solid angle is $\Omega/4\pi = 99.9\%$, its average radial thickness $\langle r \rangle = 2.2$ m, and its fiducial acceptance $\Omega \cdot \langle r \rangle = 27$ sr m. From this and the values in table III, we would expect, to first order, that isotropic LLPs have a higher

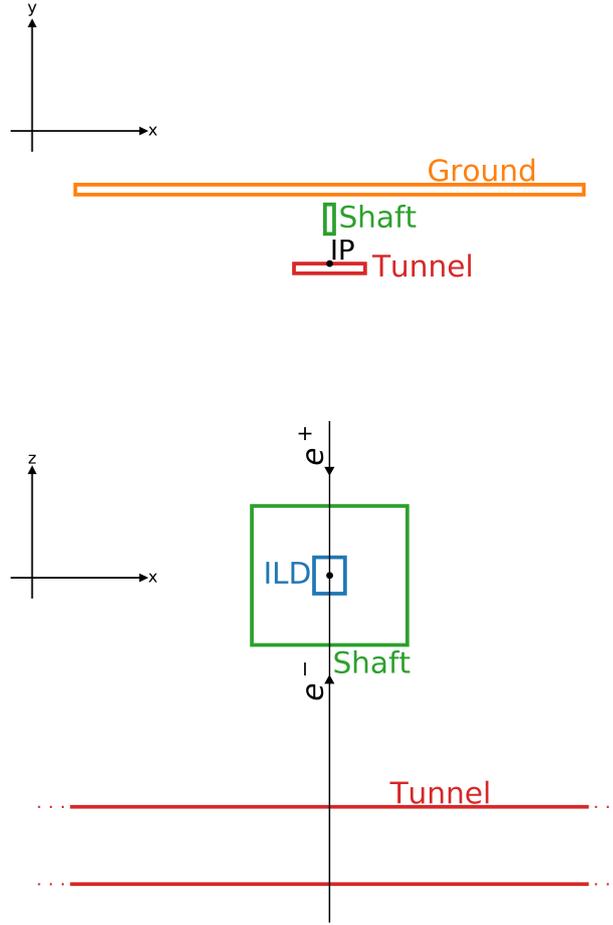


FIGURE 26: The three proposed far detectors for the ILC shown around the interaction point (IP) in side view (upper plot) and in top view (lower plot), with the Shaft detector in green, the Tunnel detector in red and the Ground detector in orange, as well as the ILD in blue. As the Ground detector is very large in the x - z plane, it is not shown in the top view. It is centred around $(x, z) = (0, 0)$. Taken from [6].

chance to decay within the ILD than in the Shaft or Tunnel detectors, but a higher chance still to decay in the Ground detector. To verify this, we now need to introduce a benchmark model to test our prediction with.

As for GAZELLE, we again use an ALP model. Since the ILC is not tuned to produce B mesons, we choose different production channels for our ALP, $e^+e^- \rightarrow a\gamma$ and $e^+e^- \rightarrow Z\gamma \rightarrow (a\gamma)\gamma$. The two production channels are shown in figure 28. We use the following Lagrangian (at high scales beyond the weak scale)

$$\mathcal{L}(\mu > \mu_w) = \frac{c_{\ell\ell}}{2} \frac{\partial_m u}{f_a} \bar{\ell} \gamma^\mu \gamma^5 \ell + c_{WW} \frac{\alpha_2}{4\pi} \frac{a}{f_a} W^{i\mu\nu} \tilde{W}_{\mu\nu}^i \quad (114)$$

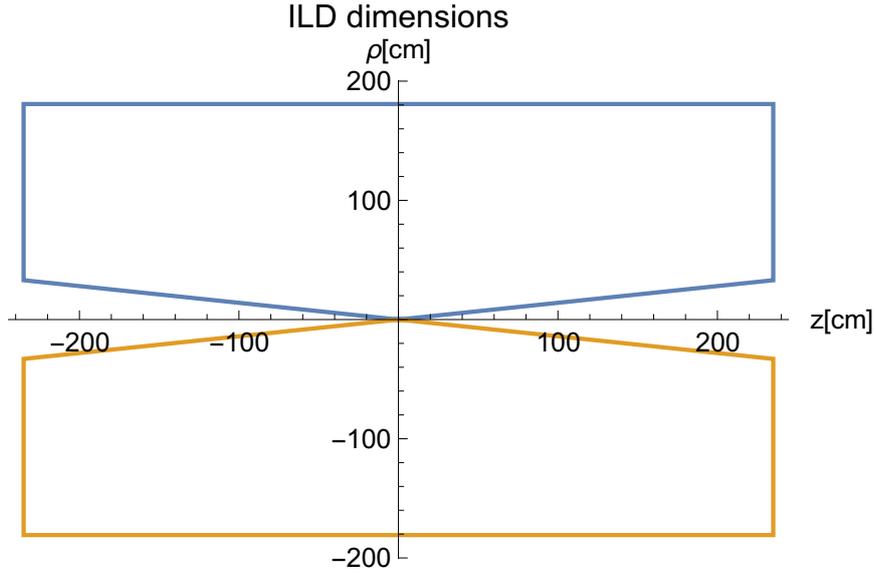


FIGURE 27: Schematic view of the ILD detector in the coordinates of the transverse/outward detector direction ρ and the beam direction z .

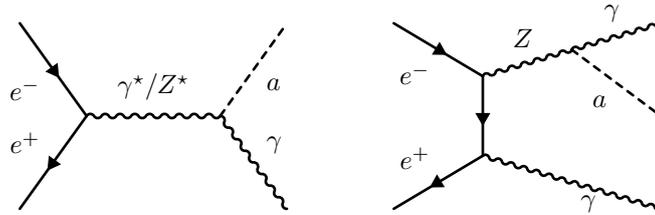


FIGURE 28: Feynman diagrams for the production channels considered for the ALP at the ILC.

which after electroweak symmetry breaking leads to terms of the form

$$\mathcal{L}(\mu < \mu_w) \supset \frac{\alpha}{4\pi} \frac{a}{f_a} \left(c_{\gamma\gamma} F_{\mu\nu} \tilde{F}^{\mu\nu} + 2 \frac{c_{\gamma Z}}{s_w c_w} F_{\mu\nu} \tilde{Z}^{\mu\nu} + \frac{c_{ZZ}}{s_w^2 c_w^2} Z_{\mu\nu} \tilde{Z}^{\mu\nu} \right) \quad (115)$$

with the couplings

$$c_{\gamma\gamma} = c_{WW} \quad (116)$$

$$c_{\gamma Z} = c_w^2 c_{WW} \quad (117)$$

$$c_{ZZ} = c_w^4 c_{WW} \quad (118)$$

and $s_w = \sin \theta_w$ and $c_w = \cos \theta_w$ the sine and cosine of the weak mixing angle θ_w , respectively.

The process $e^+e^- \rightarrow a\gamma$ couples the ALP through the photon coupling, $c_{\gamma\gamma}$ while $e^+e^- \rightarrow Z\gamma \rightarrow (a\gamma)\gamma$ with its physical intermediate Z boson couples it through the mixed photon-

Z boson coupling, $c_{\gamma Z}$. The cross sections to both processes are numerically given by

$$\sigma(e^+e^- \rightarrow a\gamma) \approx 298 \left(\frac{c_{WW}}{f_a [\text{TeV}]} \right)^2 \text{ fb} \quad (119)$$

$$\sigma(e^+e^- \rightarrow Z\gamma \rightarrow (a\gamma)\gamma) \approx 144 \left(\frac{c_{WW}}{f_a [\text{TeV}]} \right)^2 \text{ fb}. \quad (120)$$

We only allow the decay of the ALP to light leptons, so that

$$\Gamma_{a \rightarrow \ell^+ \ell^-} = \frac{m_a m_\ell^2}{8\pi} \left(\frac{c_{\ell\ell}}{f_a} \right)^2. \quad (121)$$

In the following, we will work with the mass benchmark $m_a = 0.3 \text{ GeV}$ and the coupling benchmark $c_{WW}/f_a = 1 \text{ TeV}^{-1}$. Then, the lifetime of the ALP is

$$c\tau_a = \frac{c}{\Gamma_{a \rightarrow \ell^+ \ell^-}} \quad (122)$$

$$\approx 50 \left(\frac{f_a [\text{TeV}]}{c_{\ell\ell}} \right)^2 \mu\text{m}. \quad (123)$$

While the production cross sections and the resulting kinematics do not change much with the mass of the ALP, the decay rate and lifetime do, so re-interpreting our following results for other masses should be done with care for the changing lifetime.

Kinematics are the reason that we consider two different production processes here. In the case of $e^+e^- \rightarrow a\gamma$, the ALP and photon are produced back to back and have a well-defined energy of exactly $E_a = E_\gamma = 125 \text{ GeV}$, as well as a maximal p_T of roughly the same value. The angular distribution of the ALP is biased toward the transverse direction near $\vartheta_a \approx \frac{\pi}{2}$. On the other hand, $e^+e^- \rightarrow Z\gamma \rightarrow (a\gamma)\gamma$ has a very different kinematic distribution. Since the ALP is here produced together with two photons, its energy is distributed over a wider range of values, with $E_a = 125 \text{ GeV}$ as the endpoint. The cross section for $e^+e^- \rightarrow Z\gamma$ is collinearly enhanced in beam direction [127], such that the ALP inherits the forward boost of the Z and is produced with bias in forward direction, in contrast to the other production channel. These kinematic distributions are shown in figure 29.

Now that we have detectors and a benchmark model, we can determine their reach for displaced decays. Using the same procedure as described above, though the program itself was ported to Python [128] and Jupyter [129], we generate 100,000 events with `MadGraph5.aMC@NLO` [115, 130] for each production channel and calculate first their mean decay probability within the detectors, $\langle \mathbb{P} \rangle$, and then their expected number of events,

$$N_a = \mathcal{L}\sigma(e^+e^- \rightarrow aX) \langle \mathbb{P} \rangle, \quad (124)$$

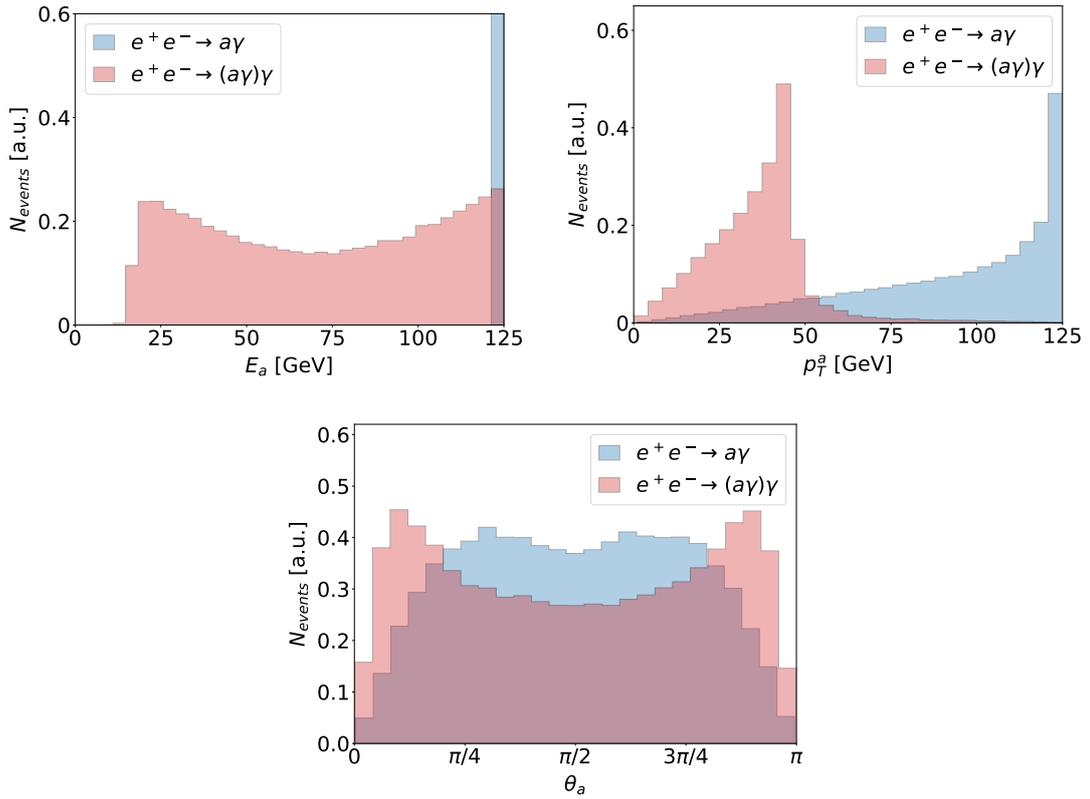


FIGURE 29: The kinematic distributions of energy E_a on the upper left, transverse momentum p_T^a on the upper right, and scattering angle θ_a in the lower plot of an ALP with $m_a = 300$ MeV, produced via $e^+e^- \rightarrow a\gamma$ (blue) or $e^+e^- \rightarrow (a\gamma)\gamma$ (red) at the ILC with $\sqrt{s} = 250$ GeV. Taken from [6].

where $\mathcal{L} = 250 \text{ fb}^{-1}$ is the luminosity of the ILC.

In figure 30, we see the expected number of events for the four detectors. Table IV shows the reach of the four detectors in the coupling, $c_{\ell\ell}/f_a$. The reach corresponds to the lowest coupling at which we expect at least 3 events, as this is the 95% confidence level of not being caused by the SM alone, assuming zero background and 100% efficiency. Figure 31 shows the reach in terms of the production cross section and the lifetime of the ALP.

As we expected from the fiducial acceptances, more ALPs decay within the ILD than at the smaller far detectors Shaft and Tunnel, and the huge Ground detector does have more events than the ILD for a significant range of couplings. We can also see that the Shaft detector has a very different response to the second production channel with the intermediate Z boson than the others, predicting much fewer events in comparison. This is due to the different kinematics in the two channels, with the Z channel having ALPs that are produced much more forward, which does not diminish the number of

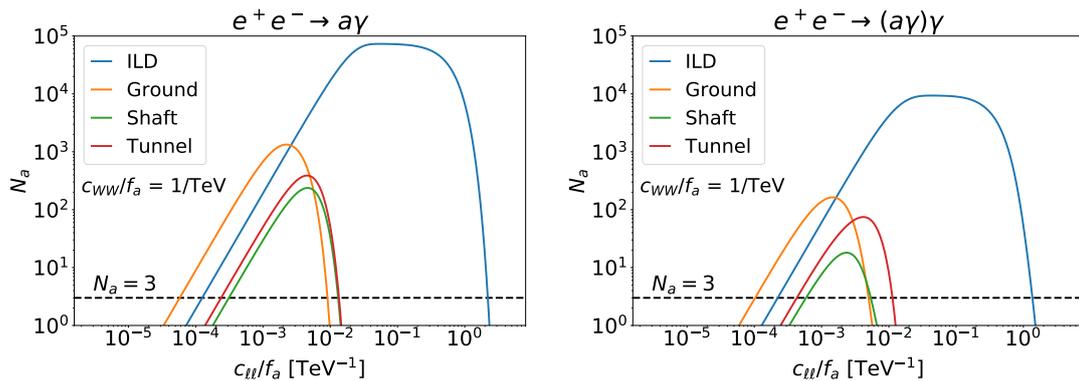


FIGURE 30: The expected number of signal events N_a in the three far detectors (Ground in orange, Shaft in green, Tunnel in red) and the main detector ILD (in blue) at the ILC as a function of the effective coupling to leptons, $c_{\ell\ell}/f_a$, with $c_{WW}/f_a = 1/\text{TeV}$ and $m_a = 300 \text{ MeV}$. On the left, we have the production channel $e^+e^- \rightarrow a\gamma$ and on the right $e^+e^- \rightarrow Z\gamma \rightarrow (a\gamma)\gamma$. Taken from [6].

$c_{\ell\ell}/f_a [10^{-4}/\text{TeV}]$	ILD	Ground	Shaft	Tunnel
$e^+e^- \rightarrow a\gamma$	1.3	0.58	3.1	2.4
$e^+e^- \rightarrow (a\gamma)\gamma$	1.2	0.54	3.0	2.2

TABLE IV: The expected reach of the four detectors in the two production channels in the ALP-lepton coupling $c_{\ell\ell}/f_a$ for $c_{WW}/f_a = 1/\text{TeV}$ and $m_a = 0.3 \text{ GeV}$. Taken from [6].

events in the Tunnel detector which is centred around the beam line, but does so for the Shaft detector, which is placed at a distance perpendicular to the beam. While the Ground detector is also placed in the transverse direction, there is very little change in its response due to its enormous size. A third feature of note is the plateau seen in the event distribution of the ILD. This is due to the fact that it is thick in comparison to its distance to the interaction point, $\frac{D}{\ell_{in}}$, spanning several orders of magnitude in lifetime for ALPs to decay within it.

In conclusion, while far detectors at the ILC can measure displaced decays, the realistic options do not add much sensitivity to what the main detector ILD can already achieve on its own, and even the massive Ground detector adds only a very small improvement. Like at Belle II before it, this is a sign that e^+e^- detectors are themselves already very well-suited to detecting long-lived particles due to their low background rates and, here even more strongly the case than for Belle II, the large solid angle of the main detectors. Of course, for a fully realistic comparison of far detectors and the main detector, a full background study would be needed. This, though, is beyond the scope of both [6] and this thesis.

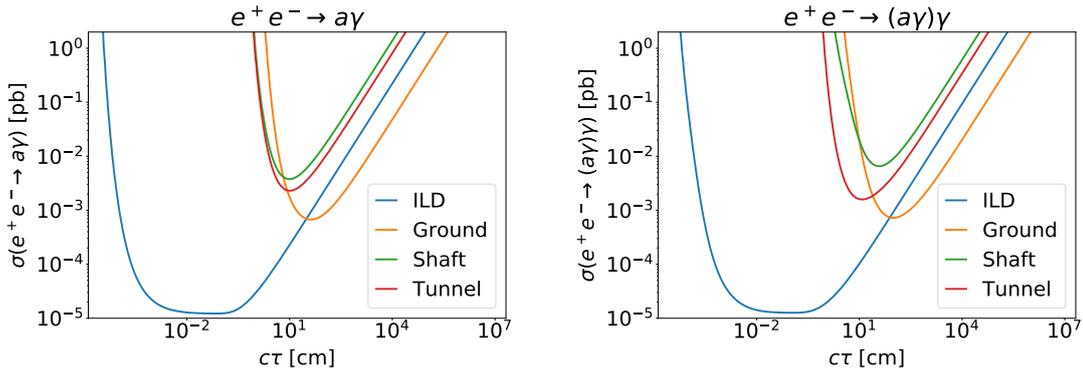


FIGURE 31: Contours of $N_a = 3$ for the two production channels of ALPs the three far detectors (Ground in orange, Shaft in green, Tunnel in red) and the main detector ILD (in blue) as a function of the ALP production cross section, σ , and its lifetime, $c\tau_a$. Taken from [6].

Instead, it makes sense to compare the two main detectors we have considered during this section, the ILD at the ILC and Belle II at SuperKEKB. To compare the above results with Belle II, we use the production of the ALP through a W -loop induced $B^\pm \rightarrow K^\pm$ transition, using

$$\mathcal{L}_{\text{eff}}(\mu < \mu_w) \supset c_{sb}(\mu) \frac{\partial_\mu a}{f_a} (\bar{s}_L \gamma^\mu b_L) \quad (125)$$

with the FCNC coupling [76]

$$c_{sb}(\mu_w) = -V_{ts}^* V_{tb} \frac{\alpha_t}{4\pi} \frac{3\alpha}{2\pi s_w^2} \frac{1 - x_t + x_t \ln x_t}{(1 - x_t)^2} c_{WW}(\mu_w), \quad (126)$$

where $\alpha_t = y_t^2/4\pi$, $x_t = m_t^2/m_W^2$, and further running below the electroweak scale can be neglected. With this, the ALP's production rate is

$$\Gamma(B \rightarrow Ka) = \frac{m_B}{16\pi} \left(\frac{c_{sb}(m_b)}{f_a} \right)^2 f_0^2(m_a^2) \left(1 - \frac{m_K^2}{m_B^2} \right)^2 \lambda^{1/2}(m_B^2, m_K^2, m_a^2), \quad (127)$$

with $f_0(q^2)$ from [86] and λ defined in equation (94).

We generate 10,000 events with `EvtGen` [114] and follow the same procedure as before. The comparison between the $N_a = 3$ contours of Belle II and the two production processes at the ILC are shown in figure 32. It should be noted that while the ILC bound is taken at a much higher centre-of-mass energy, $\sqrt{s}_{\text{ILC}} = 250 \text{ GeV}$, than Belle II, $\sqrt{s}_{\text{Belle II}} = 10.6 \text{ GeV}$, Belle II's prediction is for a luminosity of $\mathcal{L} = 50 \text{ ab}^{-1}$ compared to ILC's $\mathcal{L} = 250 \text{ fb}^{-1}$. The ILC not only gains an $\mathcal{O}(1)$ improvement over Belle II in both channels in $c_{\ell\ell}$, it also reaches around an order of magnitude further in c_{WW} than the ILC. The ILC has the advantage that due to its higher centre-of-mass energy it can

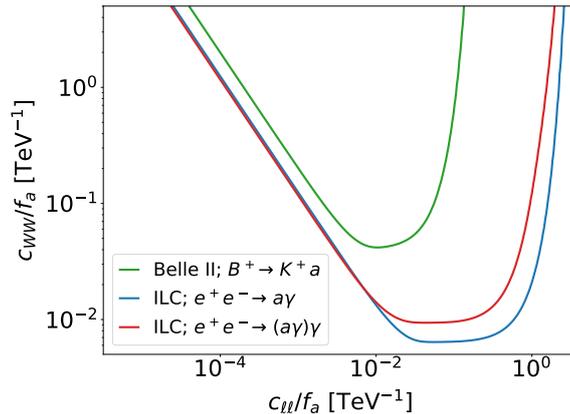


FIGURE 32: The sensitivity projections of displaced ALP decays at the ILC (in red and blue for its two production channels with and without an intermediate physical Z boson) and Belle II (in green, produced via the rare B decay $B^\pm \rightarrow K^\pm$) as a function of the lepton and W boson couplings, $c_{\ell\ell}/f_a$ and c_{WW}/f_a , respectively. Taken from [6].

probe ALPs of higher masses than Belle II can. A study of ALPs directly produced from the electron-positron collision at Belle II might help it gain additional sensitivity on c_{WW} as well.

In conclusion of this section, we can say that e^+e^- colliders are very good long-lived particle detectors, and that it is very difficult to construct realistic far detectors that improve significantly upon the sensitivity of the main detectors. In the next section, we will then explore more closely the different signatures of long-lived particle searches that we can explore at Belle II, comparing displaced searches and invisible searches.

C. Complementary LLP signatures

So far, we have focussed solely on one of the common signatures used to search for long-lived particles, displaced decays, while neglecting another that is at least equally as commonly used and that similarly able to illuminate the parameter space of LLP models, missing energy searches. We will explore this signature at the Belle II detector. In the last two sections we have seen that its large solid angle is very advantageous for displaced decay searches. Now we will see that it is also crucial for missing energy searches looking for long-lived particles.

In this section, we will first introduce the specific ALP model we work with in section III C 1, then introduce the background modelling and cut selection for our missing energy search in section III C 2, and then finally compare the sensitivity region of the missing energy search with a range of previous searches as well as to a displaced decay search at Belle II in section III C 3. Throughout the section, we will be following [5].

1. Long-Lived ALP Decays

As for the exploration of possible GAZELLE detectors, we will be working with an ALP model, compare sections IIB 5 and IIIB 2, here with the Lagrangian

$$\mathcal{L}_{eff}(\mu) = \sum_f \frac{c_{ff}(\mu)}{2} \frac{\partial_\mu a}{f_a} \bar{f} \gamma^\mu \gamma^5 f + c_{WW}(\mu) \frac{a}{f_a} \frac{\alpha_2}{4\pi} W_{\mu\nu}^i \tilde{W}^{i\mu\nu} \quad (128)$$

with the flavour universal and diagonal fermion coupling, $c_{ff}(\mu)$, and the coupling to W bosons, $c_{WW}(\mu)$, which are both scale dependent, and the ALP scale $f_a = \frac{\Lambda}{4\pi} = 1$ TeV. We will be considering two scenarios, the c_{ff} scenario, in which we set $c_{ff}(\Lambda) = 1$ and $c_{WW}(\Lambda) = 0$ and the c_{WW} scenario, where $c_{WW}(\Lambda) = 1$ and $c_{ff}(\Lambda) = 0$. In addition to this coupling, other couplings will be induced at lower scales $\mu < \Lambda$ via renormalisation group (RG) running [76, 131–133]. We will calculate the values of all relevant couplings at the relevant scales with the help of a numerical code [68]. In the rest of this section, we will give some approximate formulas for the RG-induced couplings at low-energy scales as well. Our results, however, will be based on the full numerical calculations.

The decay signatures we plan to compare are

$$B^\pm \rightarrow K^\pm a, a \rightarrow \text{inv.} \quad (129)$$

for the missing energy search and

$$B^\pm \rightarrow K^\pm a, a \rightarrow \ell^+ \ell^- \quad (130)$$

for the displaced decay. We use B decays, as in the last two sections, because Belle II is a B factory and as such both produces a very high rate of B mesons to produce the ALP this way. It is also constructed in such a way to be able to reconstruct B decays well.

As we have seen, both decay signatures rely on an ALPs produced in $B^\pm \rightarrow K^\pm$ decays. The decay rate of this process is

$$\Gamma_{B^\pm \rightarrow K^\pm a} = \frac{\pi}{4} \frac{|c_{sb}(\mu)|^2}{\Lambda^2} |f_0(m_a^2)|^2 m_B \left(1 - \frac{m_K^2}{m_B^2}\right)^2 \lambda^{\frac{1}{2}}(m_B^2, m_K^2, m_a^2) \quad (131)$$

with the scalar hadronic form factor $f_0(q^2)$ from [86] and the RG-induced flavour changing neutral current coupling $c_{sb}(\mu)$ approximately being described by [76]

$$c_{sb}(\mu_w) = V_{ts}^* V_{tb} \left[(1 - R_t(\mu_w, \Lambda)) c_{tt}(\Lambda) \frac{\alpha_t}{4\pi} \left(\frac{1}{2} \ln \frac{\mu_w^2}{m_t^2} - \frac{1}{4} - \frac{3}{2} \frac{1 - x_t + \ln x_t}{(1 - x_t)^2} \right) \right] \quad (132)$$

$$+\frac{1}{9}R_t(\mu_w, \Lambda) c_{tt}(\Lambda) - c_{WW} \frac{\alpha_t}{4\pi} \frac{3\alpha}{2\pi s_w^2} \frac{1-x_t+x_t \ln x_t}{(1-x_t)^2} \Big],$$

where $\alpha_t = y_t^2/4\pi$, $x_t = m_t^2/m_W^2$ and the function [76]

$$R_t(\mu_w, \Lambda) \approx \frac{9}{2} \frac{\alpha_t(\mu_w)}{\alpha_s(\mu_w)} \left(1 - \left(\frac{\alpha_s(\Lambda)}{\alpha_s(\mu_w)} \right)^{\frac{1}{7}} \right) \quad (133)$$

encapsulates the RG effects between the cutoff scale Λ and the weak scale μ_w on the top quark coupling c_{tt} where $c_{tt}(\Lambda) = c_{ff}(\Lambda)$. The running between the weak scale and the scale of B meson decay is small, $c_{sb}(\mu_b) \approx c_{sb}(\mu_w)$, but we will nonetheless include it in our numerical analysis.

With these formulas, the branching ratio for the ALP production is numerically

$$Br(B^\pm \rightarrow K^\pm a) = 0.25 (c_{ff}(\Lambda) + 0.0032 c_{WW}(\Lambda)) \left(\frac{f_0(m_a^2)}{f_0(0)} \right)^2 \frac{\lambda^{\frac{1}{2}}(m_B^2, m_K^2, m_a^2)}{m_B^2 - m_K^2} \quad (134)$$

in terms of the UV Lagrangian parameters $c_{ff}(\Lambda)$ and $c_{WW}(\Lambda)$ and the meson and ALP masses. It should be noted that for $m_a \rightarrow 0$, $\lambda^{\frac{1}{2}}(m_B^2, m_K^2, m_a^2) \rightarrow m_B^2 - m_K^2$, such that all terms are well normalised.

The ALP decay rate depends on its mass. Since we will be considering 12 different mass benchmark values,

$$m_a \in \{5 \text{ MeV}, 50 \text{ MeV}, 70 \text{ MeV}, 0.1 \text{ GeV}, 0.2 \text{ GeV}, \\ 0.25 \text{ GeV}, 0.3 \text{ GeV}, 0.5 \text{ GeV}, 1 \text{ GeV}, 2 \text{ GeV}, 3 \text{ GeV}, 4 \text{ GeV}\}, \quad (135)$$

we will now go over the range of different decay modes the ALPs can decay into. For light ALPs, $m_a < 2m_e$, the ALP decays only to photons via [76]

$$\Gamma_{a \rightarrow \gamma\gamma} = \frac{\alpha^2 m_a^3}{4\pi\Lambda^2} \left| c_{\gamma\gamma}^{\text{eff}}(m_a) \right|^2, \quad (136)$$

with the mass-dependent effective ALP-photon coupling [76]

$$c_{\gamma\gamma}^{\text{eff}}(m_a) = \begin{cases} c_{\gamma\gamma}(m_a) + \sum_{f \in \{\ell, Q\}} N_c^f Q_f^2 c_{ff}(m_a) B_1 \left(\frac{4m_f^2}{m_a^2} \right) & m_a > 1 \text{ GeV} \\ c_{\gamma\gamma}(m_a) + \sum_{f \in \{\ell, Q\}} N_c^f Q_f^2 c_{ff}(m_a) B_1 \left(\frac{4m_f^2}{m_a^2} \right) \\ - \frac{m_a^2}{m_\pi^2 - m_a^2} \frac{c_{uu}(m_a) - c_{dd}(m_a)}{2} \\ - \left(\frac{5}{3} + \frac{m_\pi^2}{m_\pi^2 - m_a^2} \frac{m_d - m_u}{m_d + m_u} \right) c_{GG}(m_a) & m_a < 1 \text{ GeV}, \end{cases} \quad (137)$$

which has contributions from the explicit photon coupling $c_{\gamma\gamma} = c_{WW}$, fermion loop contributions, which are suppressed for fermions with masses above the ALP mass, and contributions from chiral perturbation theory (seen in the third line), that is, contributions to the ALP-photon coupling from mixing between the ALP and the neutral pion. In the above equation and in the following, $\ell \in \{e, \mu, \tau\}$, $Q \in \{c, b, t\}$, and [76]

$$B_1(\tau) = 1 - \tau f^2(\tau) \quad (138)$$

$$f(\tau) = \begin{cases} \arcsin \frac{1}{\sqrt{\tau}} & \tau \geq 1 \\ \frac{\pi}{2} + \frac{i}{2} \ln \frac{1+\sqrt{1-\tau}}{1-\sqrt{1-\tau}} & \tau < 1 \end{cases}.$$

Since ALPs are pseudoscalar particles, the lowest number of pions they can decay to are three. Thus, the ALP decay rate between the electron threshold and the three pion threshold, $2m_e < m_a < 3m_\pi$, is determined by ALP decays to electrons and muons in addition to $a \rightarrow \gamma\gamma$,

$$\Gamma_{a \rightarrow \ell^+ \ell^-} = 2\pi m_a \frac{|c_{\ell\ell}^{\text{eff}}(m_a)|^2 m_f^2}{\Lambda^2} \sqrt{1 - \frac{4m_f^2}{m_a^2}} \quad (139)$$

where the effective ALP-lepton coupling is mostly $c_{\ell\ell}^{\text{eff}}(\mu) \approx c_{ff}(\mu)$ with a small additional loop contribution by c_{WW} .

The decay of the ALP to hadrons can be separated into two regions. Below 1 GeV, we will use the decay of the ALP to three pions, $a \rightarrow \pi^+ \pi^- \pi^0$ or $a \rightarrow \pi^0 \pi^0 \pi^0$, as predicted by chiral perturbation theory. Above 1 GeV, we treat the quarks and gluons perturbatively. The ALP decay rate to pions is given by [76]

$$\Gamma_{a \rightarrow \pi^0 \pi^i \pi^j} = \frac{m_a m_\pi^4}{384\pi f_\pi^2 \Lambda^2} \left(c_{uu}(m_a) - c_{dd}(m_a) + 2c_{GG}(m_a) \frac{m_d - m_u}{m_d + m_u} \right)^2 g_{ij} \left(\frac{m_\pi^2}{m_a^2} \right), \quad (140)$$

with [76]

$$g_{00}(r) = \frac{2}{(1-r)^2} \int_{4r}^{(1-\sqrt{r})^2} dz \sqrt{1 - 4\frac{r}{z}} \sqrt{\lambda(1, z, r)} \quad (141)$$

$$g_{+-}(r) = \frac{12}{(1-r)^2} \int_{4r}^{(1-\sqrt{r})^2} dz \sqrt{1 - 4\frac{r}{z}} \sqrt{\lambda(1, z, r)} (z-r)^2.$$

The perturbative decay to hadrons is given by [76]

$$\Gamma_{a \rightarrow \text{had.}} = \frac{2\alpha_s^2 m_a^3}{\pi} \frac{|c_{GG}^{\text{eff}}(m_a)|^2}{\Lambda^2} \left(1 + \left(\frac{97}{4} - \frac{7n_q}{6} \right) \frac{\alpha_s}{\pi} \right) + \sum_q \Gamma_{a \rightarrow q\bar{q}}, \quad (142)$$

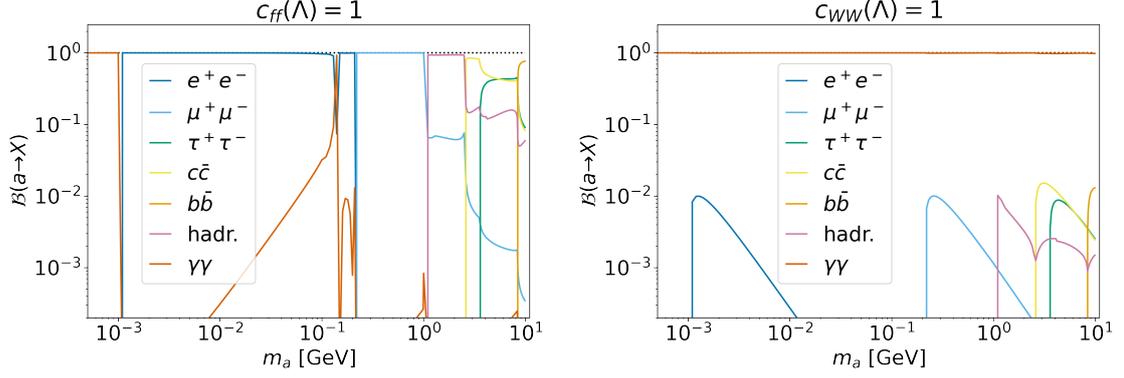


FIGURE 33: Branching ratios of ALPs decaying to different final states, $a \rightarrow X$ in the c_{ff} scenario on the left and the c_{WW} scenario on the right. Taken from [5].

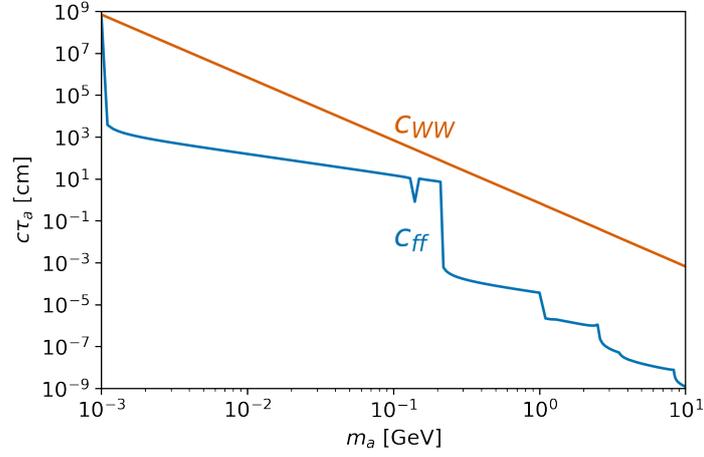


FIGURE 34: Proper decay length $c\tau_a$ of the ALP in the c_{ff} scenario in blue and the c_{WW} scenario in orange as a function of the ALP mass m_a . Taken from [5].

where $n_q = 3$ is the number of light quarks $q \in \{u, d, s\}$ and the effective gluon coupling is given by [76]

$$c_{GG}^{\text{eff}}(m_a) = c_{GG}(m_a) + \sum_{q'} \frac{c_{q'q'}(m_a)}{2} B_1 \left(\frac{4m_{q'}^2}{m_a^2} \right). \quad (143)$$

The last decay channel is that to heavy quarks, meaning charm and bottom quarks, $Q \in \{c, b\}$, for which the rate is

$$\Gamma_{a \rightarrow Q\bar{Q}} = 6\pi m_a \frac{|c_{QQ}^{\text{eff}}(m_a)|^2 m_f^2}{\Lambda^2} \sqrt{1 - \frac{4m_Q^2}{m_a^2}} \quad (144)$$

which equals the decay rate into leptons with the additional colour factor $N_c = 3$ and where again $c_{QQ}^{\text{eff}}(\mu) \approx c_{ff}(\mu)$.

The branching ratios of the ALP in the two coupling scenarios are shown in figure 33, and the corresponding lifetimes of the ALP are shown in figure 34. As we can clearly see, the c_{WW} scenario is dominated entirely by the decay rate into photons. This makes sense as its contributions to the lepton and quark couplings are loop-suppressed. Thus, it is unsurprising that the ALP generally has a longer lifetime in the c_{WW} scenario. The ALP of the c_{ff} scenario decays dominantly into lepton pairs from the electron threshold until $m_a = 1$ GeV and into hadrons and heavy quarks that decay into them for higher masses. Here, the ALP couples to fermions directly, but its coupling to photons is loop-suppressed which can be seen in its rapidly rising lifetime at masses below the electron threshold.

In the course of this analysis, we will only look at ALP decays to leptons or photons explicitly, though the ALP lifetime, which contains all decay rates,

$$\begin{aligned} c\tau_a &= \frac{c}{\Gamma_a} \\ &= \frac{c}{\sum_X \Gamma_{a \rightarrow X} \Theta(m_a - m_X)} \end{aligned} \quad (145)$$

will have the largest effect on our results. Here $\Theta(x)$ is the Heaviside function ensuring that only kinematically allowed processes are included.

As we introduced in the last section, the probability that an LLP of boost $\gamma\beta$ and lifetime τ has not decayed yet at the distance r from its production point is

$$\mathbb{P}(r) = e^{-\frac{r}{\gamma\beta c\tau}}. \quad (146)$$

If r_i is the distance an ALP i traverses between its production vertex and leaving Belle II, then we can average over a representative sample of generated ALP events to define the average probability of an ALP leaving the detector undecayed,

$$\langle \mathbb{P} \rangle = \frac{1}{N} \sum_i \mathbb{P}(r_i). \quad (147)$$

The distance between the particle's production and when it leaves the detector are calculated in the same way as in the last section. As explained in detail there, the particle's line of movement is defined by its production vertex \vec{x}_0 and its momentum \vec{p} . We check the intersection of this ray with every edge of the detector, working again in the two-dimensional plane of ρ and z . If there are no interaction points, the particle is always undecayed by the time it leaves the detector because it spends no time within the detector. If there are intersection points, the intersection point with the furthest distance to the production vertex defines the distance r , corresponding to ℓ_{out} in our

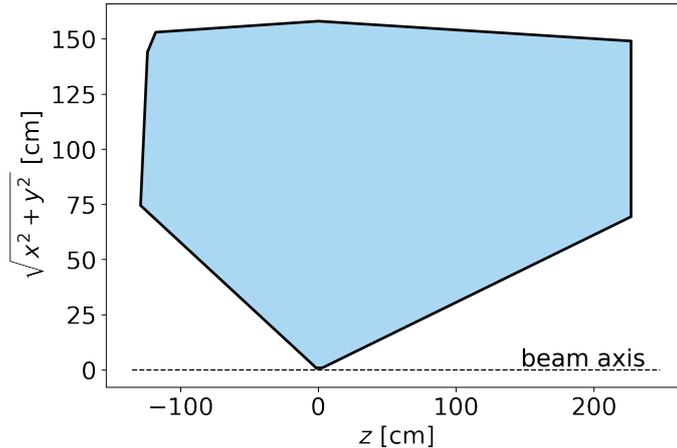


FIGURE 35: Schematic representation of Belle II as used for the selection of invisible ALP decays, including the calorimeter as well as the tracking system. Taken from [5].

earlier notation.

Now, we need to differentiate between displaced decays and missing energy searches. For prompt and displaced decays, the events that interest us are those where the ALP decays before the edge of the tracking detector. Thus, we use the same definition of Belle II as in the previous section, see figure 22 for our determination of $\langle \mathbb{P} \rangle$. The number of expected events is then

$$N_a (B^\pm \rightarrow K^\pm a, a \rightarrow X) = N_{B\bar{B}} Br (B^\pm \rightarrow K^\pm a) Br (a \rightarrow X) (1 - \langle \mathbb{P} \rangle) \quad (148)$$

where X is any visible final state.

For missing energy searches, which we will focus on in this section, our signal is defined by an ALP leaving the detector without decaying. As they are called missing *energy* searches, it is clear that in addition to the tracking detector, the ALP can also not decay within the calorimeter and still be considered a signal event. That is why in contrast to the displaced vertex search, we here use an extended model of Belle II which includes the volume of its electromagnetic calorimeter (ECL), as seen in figure 35. To differentiate between the probability regarding the tracker and the probability regarding the calorimeter in addition, we will label the latter as \mathbb{P}' while keeping the former as \mathbb{P} . With this probability, we can calculate the number of expected events as

$$N_a (B^\pm \rightarrow K^\pm + \text{inv.}) = N_{B\bar{B}} Br (B^\pm \rightarrow K^\pm a) (Br (a \rightarrow \text{inv.}) + Br (a \rightarrow X) \langle \mathbb{P}' \rangle) \quad (149)$$

if there is an invisible ALP decay rate. Since we do not assume this to exist here, the

number of events simplifies to

$$N_a(B^\pm \rightarrow K^\pm + \text{inv.}) = N_{B\bar{B}} Br(B^\pm \rightarrow K^\pm a) \langle P' \rangle \quad (150)$$

since without any invisible rate, the branching ratio of the ALP to visible final states is 1.

The calculation of the probability, as well as the rest of the analysis that follows in the next section, is done in Python [128]/Jupyter [129] in code based on my framework from section III B, but refactored and rewritten by me into Python to work with [68], the package for numerically calculating the ALP couplings at the scale of the ALP production and decay.

2. Missing Energy Signatures

While for displaced decays at Belle II we have only a negligible amount of background events, missing energy signatures have a variety of backgrounds. Neutrinos are an obvious background for invisible particles, but visible particles that are produced in the collision can be produced in the forward direction and leave the detector without detection. In addition, we have detector inefficiencies, the non-zero probability that a particle is just not detected by the detector, and particles that are too soft to be detected. Thus, to study the potential for probing the ALP parameter space in missing energy searches at Belle II, we perform an involved background study.

The signal we are looking for is the process $B \rightarrow Ka$ where a decays invisibly, that is, we are looking for a kaon and missing energy. Thus, we want to take into account any other background channels that can also produce an event with a kaon and missing energy. Kaons are a very common decay product of not only B decays, but also of hadronisation at the energies at Belle II, as well as being a common final state of τ decays. Missing energy, as mentioned above, comes not only in the form of neutrinos, which are frequently produced in all kinds of meson decays, but also in the form of missing the angular acceptance of the detector or particles simply being missed by the detector.

To estimate the frequency of events that look signal-like, we generate 10,000,000 events for each of the relevant background channels⁸, those being B meson pairs, $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\bar{B}$, both charged and neutral, continuum production of quark anti-quark pairs, $e^+e^- \rightarrow q\bar{q}$, and continuum production of τ lepton pairs, $e^+e^- \rightarrow \tau\bar{\tau}$. For each of

⁸ While beam-induced backgrounds are also expected to give a relevant contribution, their inclusion is beyond the scope of this work.

Channel	σ [nb]	Programs
B^+B^-	0.54	EvtGen [114]
$B^0\bar{B}^0$	0.51	EvtGen [114]
$u\bar{u}$	1.61	KKMC [136], PYTHIA8.2 [137], EvtGen [114]
$d\bar{d}$	0.4	KKMC [136], PYTHIA8.2 [137], EvtGen [114]
$s\bar{s}$	0.38	KKMC [136], PYTHIA8.2 [137], EvtGen [114]
$c\bar{c}$	1.3	KKMC [136], PYTHIA8.2 [137], EvtGen [114]
$\tau\bar{\tau}$	0.919	KKMC [136], TAUOLA [138]

TABLE V: The different background channels we use, their production cross section for normalisation [83], and the programs that were used to generate the samples.

these channels, we let the produced particles decay until only particles considered stable on the detector scale are left. These particles, we then base our analysis on. In table V, we quote the production cross sections for each of the background channels, as well as the programs used in their production.

For the signal, we produce 10,000 events for every mass benchmark using the Belle II analysis software framework [134, 135] and **EvtGen** [114]. The number of background samples we produce is much higher than the number of signal events, as many of the events produced will not actually be a background for our signal, while most of the signal events will work as such for us.

While we do not use the full Belle II detector simulation, we do attempt to handle the sample events in a realistic manner. The particles left over at the end of the decay chain are photons, electrons, muons, charged pions, charged kaons, protons, neutrons, long-lived neutral kaons K_L^0 , as well as neutrinos. We consider the last three neutral particles to be invisible to the detector.

For the photon, we impose a minimum energy of $E_\gamma \geq 50$ MeV to be considered visible, for the other charged particles a minimal transverse momentum $p_T \geq 0.2$ GeV. Of course, the particles also have to be within the angular acceptance of the detector, $17^\circ \geq \theta \geq 150^\circ$. For particles that clear these conditions, we assume an efficiency of 100% of detecting photons and of 99% to detect other charged particle tracks. We also smear the momenta of charged particles and the energy of photons with a normal distribution using $\frac{\Delta p}{p} = 0.5\%$ [139] for charged particles and $\frac{\Delta E}{E} = 5\%$ [83] to simulate the slight error in measuring these parameters.

Lastly, we apply misidentification rates to the pions and kaons. Following [140], we set the probability that a kaon is identified as a kaon at 80%, and the probability that a pion is misidentified as a kaon at 5%. All of these conditions are used for signal events just as for background events.

Now that we have an understanding of how we approximate Belle II, let us think about how we can choose our signal region to contain as much signal and as little background as possible. In a full experimental analysis, this would involve using many kinematic variables and boosted decision trees to construct complicated cuts to find the optimal regions. As this is a phenomenological study, we will go about it with fewer kinematic variables and cuts. This may hinder our sensitivity a little, but should still give us a good approximation of what is possible with a full experimental study.

Before we can set cuts, we need to think about the properties of our signal. As we mentioned above, the signal is a B meson decaying into a kaon and an ALP that stays invisible by decaying outside of the detector. The two important parameters of this event are the momenta of the two decay particles, p_K and p_a . Due to the invisible nature of the ALP, we need to reconstruct its momentum from the particles we do measure as best as we can. Assuming we could measure all other particles i , the ALP momentum would be given by the difference between the initial momentum of the $\Upsilon(4S)$ resonance that the B that decays into the ALP is produced from, and the sum of all the other particles. Given that we will not be able to fully and correctly measure the momenta of all particles, we instead use p_{miss} as an estimate for p_a , where

$$p_a \approx p_{miss} = p_{\Upsilon(4S)} - \sum_i p_i. \quad (151)$$

and the sum goes over all visible and detected particles i .

Another complication is that many events will contain more than one kaon, even in the signal case, as the second B meson in the event often decays to a number of particles that also contain a kaon. Studying the kinematic distributions shown in figures 36 and 37, we find that the kaon produced in $B \rightarrow Ka$ is usually the kaon with the highest p_T , at least for light ALPs. This makes sense as generic B decays often decay into more than two particles such that any momentum and energy has to also be divided among more particles, leaving the kaon less likely to gain as large a momentum. We will call the kaon with the highest p_T the leading kaon and assume it to be the kaon associated with the ALP.

Having reconstructed both our kaon and ALP momentum, we can now choose our kinematic variables. For this, let us consider the properties of our signal events. We have already mentioned the high likelihood of the kaon with the highest p_T to be the true signal kaon, the kaon that is produced together with the ALP in a B decay. It thus makes sense to choose the leading kaon's transverse momentum, $p_T(K)$, as one of our variables for making cuts. Another property of the signal is that the momenta of the ALP and the kaon reconstruct the B meson mass, $(p_a + p_K)^2 = m_B^2$. Since we will not be able

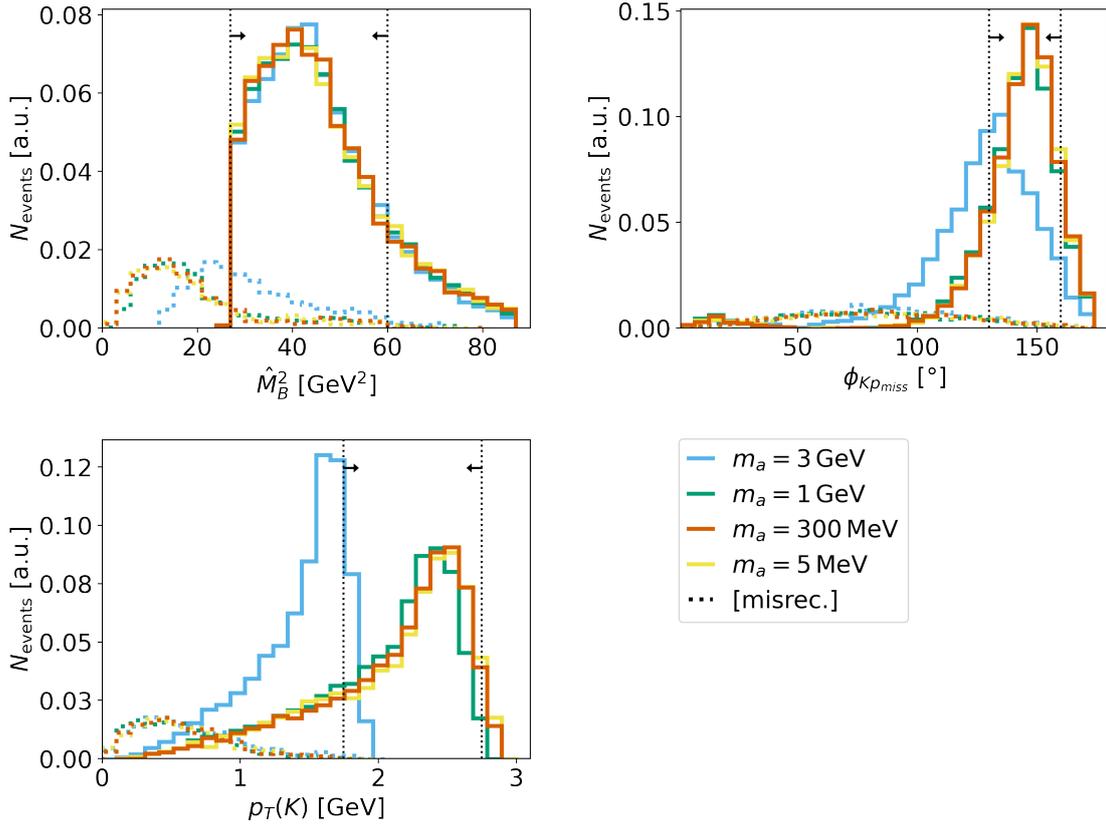


FIGURE 36: The distributions of properly reconstructed (solid lines) and misreconstructed (dotted lines) signal events in the three kinematic variables chosen for cutting for a representative selection of benchmark masses: the reconstructed B meson mass \hat{M}_B^2 in the top left, the opening angle between the leading kaon's momentum and the missing momentum, $\phi_{Kp_{miss}}$ in the top right, and the leading kaon's transverse momentum $p_T(K)$ on the lower left. The thin black dotted lines and arrows show the kinematic cuts to be made. Taken from [5].

to access the real p_a , we can instead use p_{miss} to estimate the B mass reconstruction, $\hat{M}_B^2 = (p_{miss} + p_K)^2$. We expect this to be another good indicator for whether an event is a signal event. As a third kinematic variable, we choose the angle between the leading kaon's momentum and the missing momentum, $\phi_{Kp_{miss}}$, as an estimate for the opening angle of the decay products of the B meson decay.

As we can gather from the definitions of our kinematic variables and their underlying momenta, as well as the detector response, not all signal events will be recognised as such. If the leading kaon is not the signal side kaon, then we refer to this event as a misreconstructed signal event. We will label these [misrec.] in the following. Another option is that no kaon appears to exist in the event. This could be the case because the signal kaon is outside of the angular acceptance of the detector, because it was misidentified as a pion (20% chance), or because its track was missed (1% chance). Table VI shows the number of signal events that are recognised as such and the number

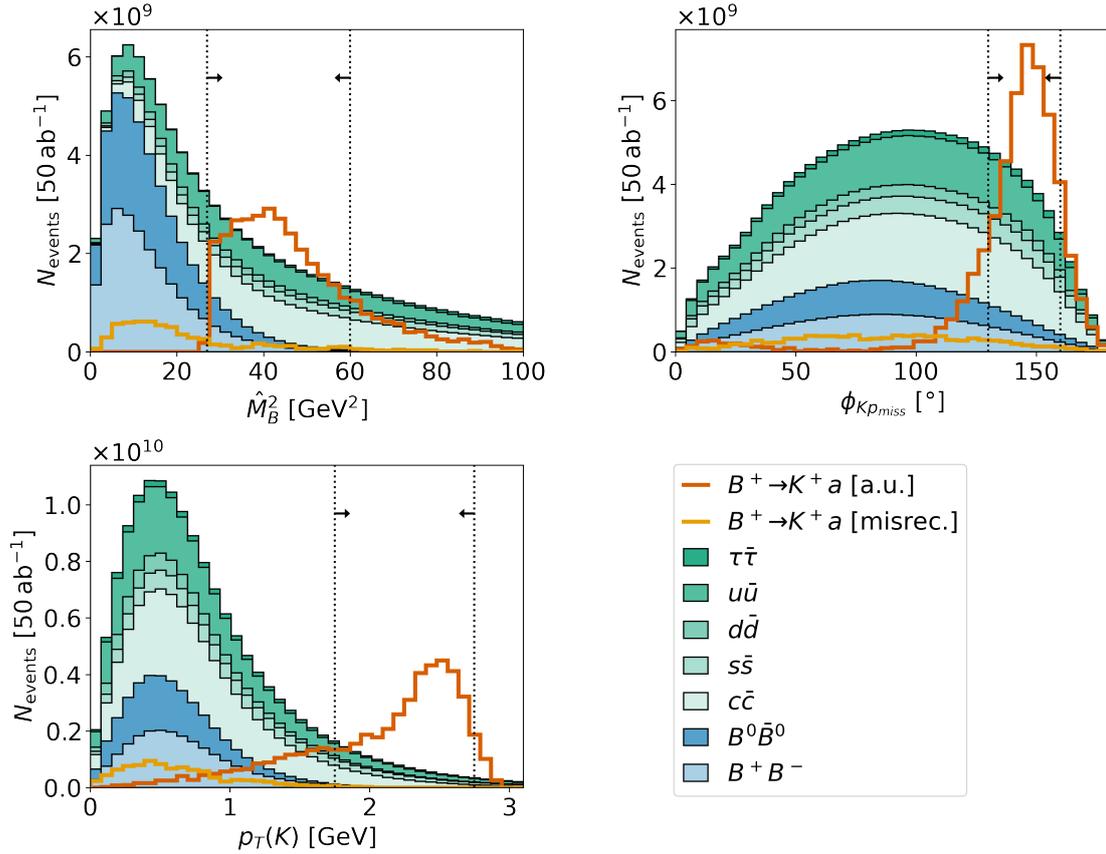


FIGURE 37: The distributions of properly reconstructed (orange-red line) and misreconstructed (yellow line) signal events for a mass of $m_a = 0.3 \text{ GeV}$ and for the different background channels (stacked, in solid colours) in the three kinematic variables chosen for cutting: the reconstructed B meson mass \hat{M}_B^2 in the top left, the opening angle between the leading kaon's momentum and the missing momentum, $\phi_{K p_{miss}}$ in the top right, and the leading kaon's transverse momentum $p_T(K)$ on the lower left. The background event numbers are normalised to the expected number for a luminosity of 50 ab^{-1} , while the signal events are normalised to be a tenth of the shown background events and the misreconstructed events scale with the signal events. The thin black dotted lines and arrows show the kinematic cuts to be made. Taken from [5].

of misreconstructed signal events.

Figure 36 shows the distribution of the kinematic variables of the signal event, $B \rightarrow Ka$ for four representative ALP masses. Figure 37 shows the kinematic distributions of the different background channels in comparison to the signal and misreconstructed signal for a mass $m_a = 0.3 \text{ GeV}$, which is the benchmark we will be using throughout these plots to compare signal and background distributions. Figure 38 shows scatter plots comparing the full background distribution not separated into channels in two-dimensional parameter space with the signal and misreconstructed signal distributions.

The two-body decay structure is well-reflected in the plots for the signal events. As we

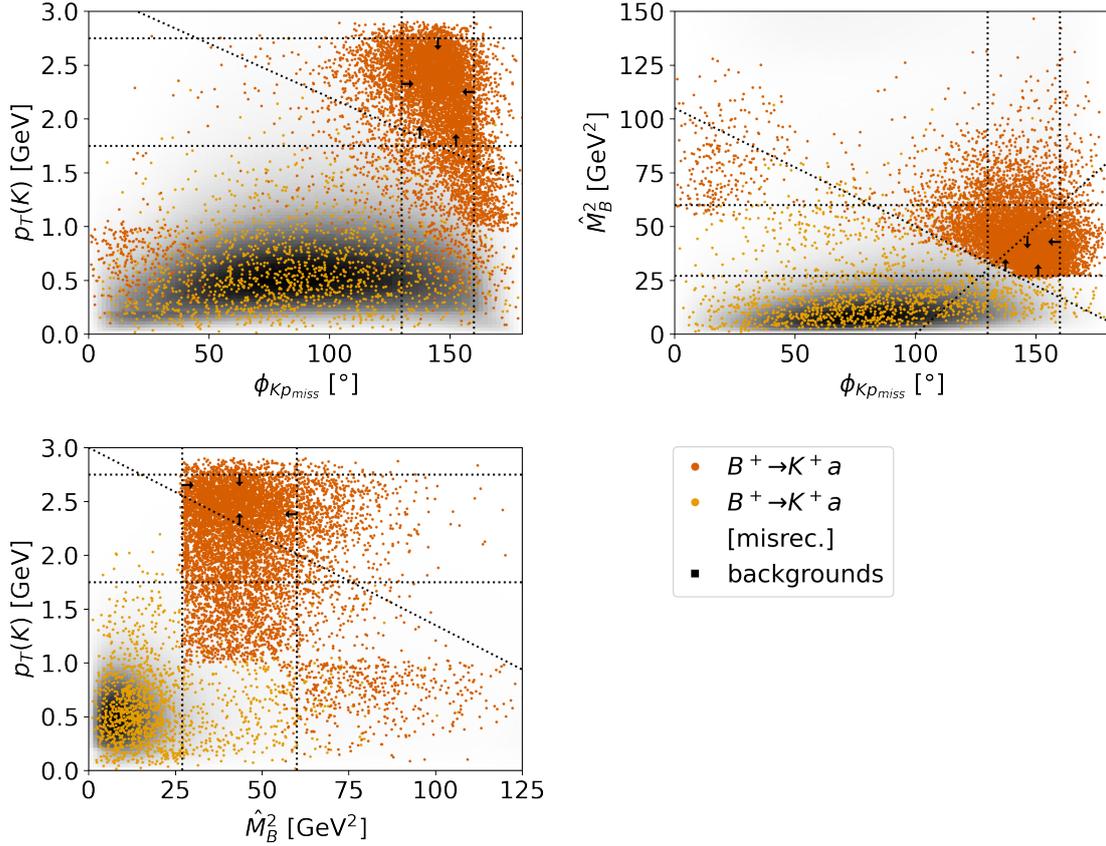


FIGURE 38: The two-dimensional distributions of properly reconstructed (orange-red dots) and misreconstructed (yellow dots) signal events displayed as points for every of the 10,000 generated events, and the distribution of the full background as a two-dimensional histogram (black area) in the three kinematic variables chosen for cutting for a representative selection of benchmark masses: the reconstructed B meson mass \hat{M}_B^2 in the top left, the opening angle between the leading kaon's momentum and the missing momentum, $\phi_{Kp_{miss}}$ in the top right, and the leading kaon's transverse momentum $p_T(K)$ on the lower left. The thin black dotted lines and arrows show the kinematic cuts to be made. Taken from [5].

have already mentioned, the kaon's transverse momentum peaks at high values, and so does the opening angle between the kaon and missing momentum. Both peaks broaden for high ALP masses, where the kinematic endpoint for the kaon momentum is lower and both particles point in the direction of the B meson more to compensate its boost. For ALPs with low masses, $m_a \lesssim 0.5$ GeV, the kinematic distributions strongly resemble that of the (near-)massless ALP, $m_a = 5$ MeV.

The reconstructed B meson mass distribution starts at its true value, $\hat{M}_B^2 = m_B^2 \approx 27.9$ GeV², peaks at slightly higher \hat{M}_B^2 and has a long tail. When all particles are counted as visible and perfectly reconstructed, the distribution is a sharp peak at the B meson mass. The introduction of invisible and missed particles leads to the shift of part of the distribution to higher masses, while the momentum and energy smearing smooths

the distribution and introduces a negligible number of events with a reconstructed B meson mass that is smaller than expected.

In general, the signal distribution looks well-separable from both the background from misreconstructed signal events and the background from generic non-signal events. This can be seen especially well in the two-dimensional plots. It should here be noted that while it looks like there is no background at all in the regions where the signal events are concentrated, this is at least partly due to the depiction of the background. Due to the sheer magnitude of background events, we depict the background as a two-dimensional histogram with linear scaling of the opacity, while the signal (and misreconstructed signal) events are shown with a point for every event in the generated set of samples. The visibility of a low number of events is thus more pronounced for signal events than for background events.

To decide on the cuts shown in the preceding figures, we use the Punzi figure of merit [81] for a 5σ discovery introduced in section II C and optimise the cuts for it. For this, we choose cuts that seem to divide the signal and backgrounds well from each other as seen in the plots in the figures 36, 37 and 38. We express these as functions in one or two of the kinematic parameters, with variable cut-parameters, which we vary to find the optimal Punzi value. Using this procedure, we settle on the selection cuts

$$\begin{aligned}
1.75 \leq p_T(K) \leq 2.75 \text{ GeV}, & \tag{152} \\
p_T(K) \geq 3.2 \text{ GeV} - 0.01 \text{ GeV}/^\circ \cdot \phi_{Kp_{miss}}, & \\
p_T(K) \geq 3.0 \text{ GeV} - 0.0165 \frac{1}{\text{GeV}} \cdot \hat{M}_B^2, & \\
130^\circ \leq \phi_{Kp_{miss}} \leq 160^\circ, & \\
27 \leq \hat{M}_B^2 \leq 60 \text{ GeV}^2, & \\
\hat{M}_B^2 \geq 105 \text{ GeV}^2 - 0.55 \text{ GeV}^2/^\circ \cdot \phi_{Kp_{miss}}, & \\
\hat{M}_B^2 \leq -100 \text{ GeV}^2 + 0.992 \text{ GeV}^2/^\circ \cdot \phi_{Kp_{miss}}. &
\end{aligned}$$

The one- and two-dimensional distributions of the signal and background distributions after applying the cuts are shown in the figures 39 and 40. There, we apply all cuts that are not in the variable(s) on the plot axes, while leaving the shown directions uncut.

The only background events left after the cuts are 39 from the charged B meson and 28 from the neutral B meson backgrounds, none of the continuum backgrounds nor the misreconstructed signal events produce any events that survive the cuts. Of the signal events, we get an efficiency of about 10 %, that is, ca. 10 % of the generated signal events are recognisable as signal events and are within the signal region defined by the cuts.

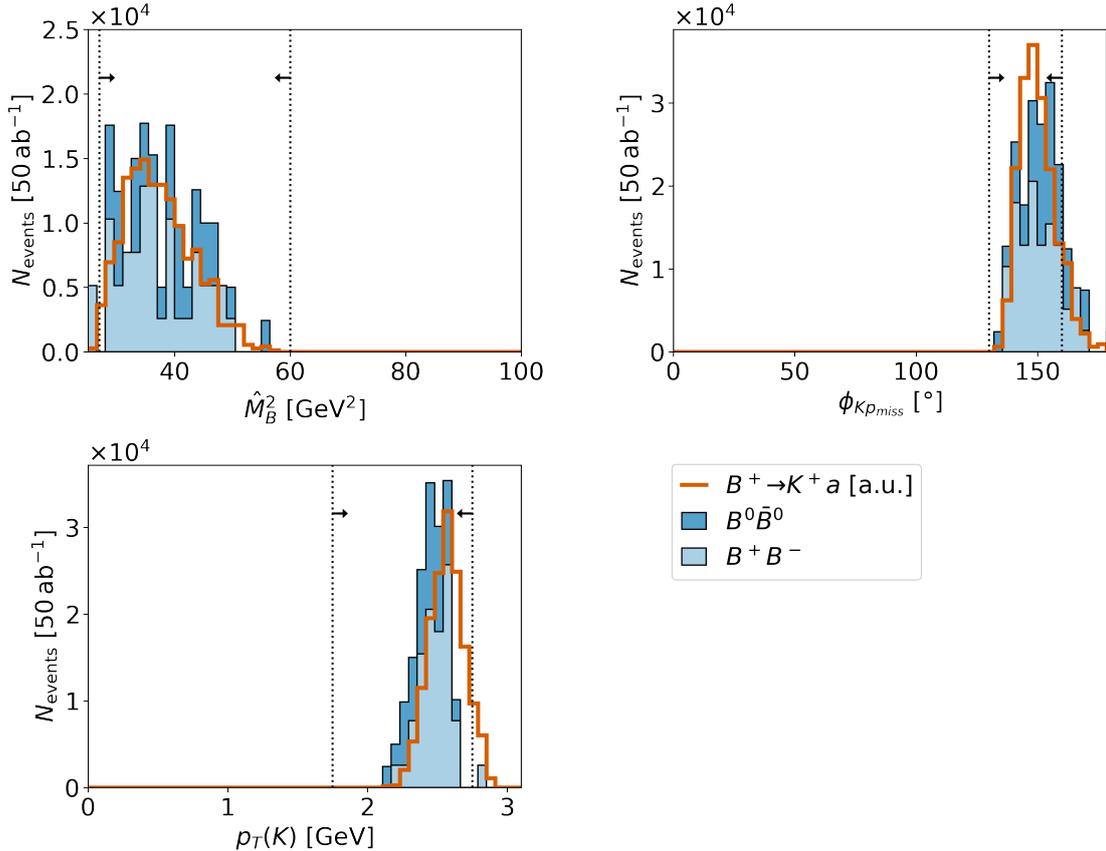


FIGURE 39: The distributions of properly reconstructed (orange-red line) and misreconstructed (yellow line) signal events for a mass of $m_a = 0.3$ GeV and for the different background channels (stacked, in solid colours) with all cuts applied that are not shown in the three kinematic variables chosen for cutting: the reconstructed B meson mass \hat{M}_B^2 in the top left, the opening angle between the leading kaon's momentum and the missing momentum, $\phi_{Kp_{miss}}$ in the top right, and the leading kaon's transverse momentum $p_T(K)$ on the lower left. The background event numbers are normalised to the expected number for a luminosity of 50 ab^{-1} , while the signal events are normalised to be a tenth of the shown background events and the misreconstructed events scale with the signal events. The light dotted lines show the kinematic cuts to be made. Taken from [5].

As we see in table VI, the signal efficiency is smaller for higher masses, falling off to 0 for $m_a > 3$ GeV. A different set of cuts optimised specifically for the higher masses could lead to a better signal efficiency for them. Overall, this is a very strong background reduction that, as we will soon see, leads to strong bounds.

Using the leftover background events and the signal efficiency after cuts, we can now define the lower bound to which we expect this search could probe the ALP parameter space. For this, we normalise the remaining background events with their respective cross sections, see table V. We then determine the minimum number of signal events N_{sig} necessary to reject the background-only hypothesis at 90% CL based on the number of

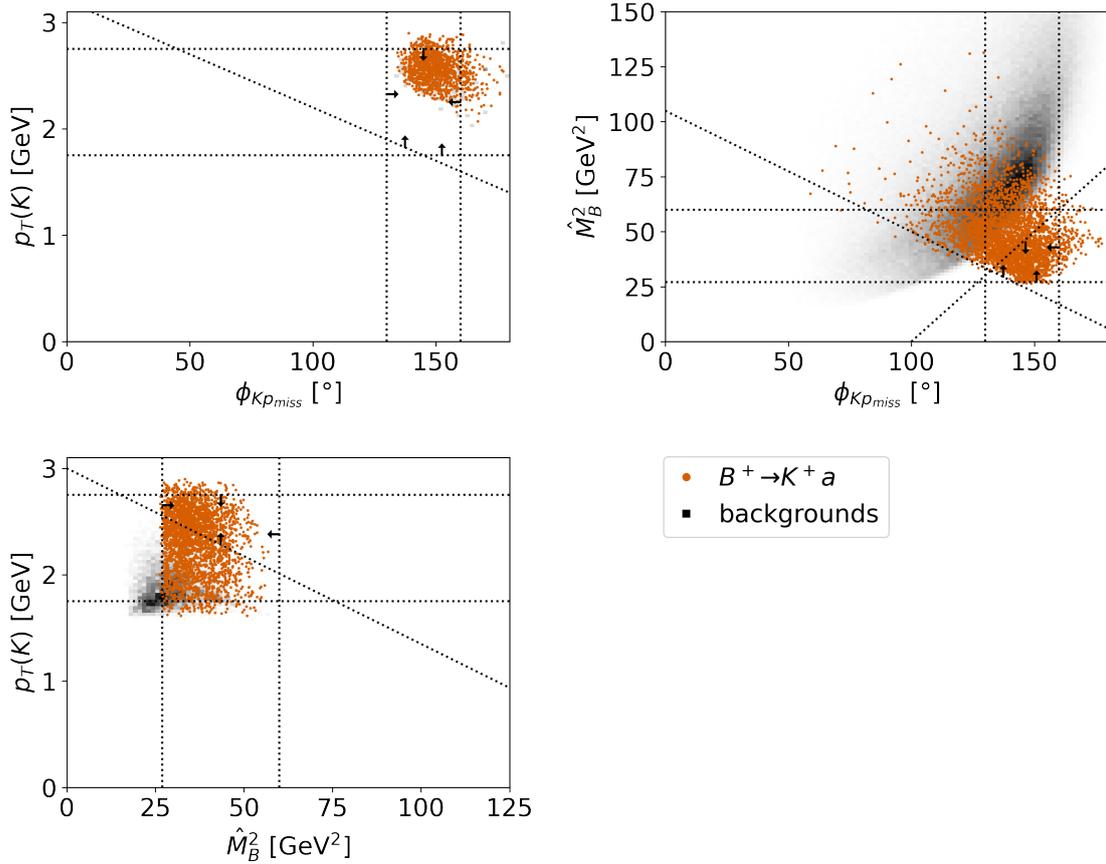


FIGURE 40: The two-dimensional distributions of properly reconstructed (orange-red dots) and misreconstructed (yellow dots) signal events displayed as points for every of the 10,000 generated events, and the distribution of the full background as a two-dimensional histogram (black area) with all cuts applied that are not shown in the three kinematic variables chosen for cutting for a representative selection of benchmark masses: the reconstructed B meson mass \hat{M}_B^2 in the top left, the opening angle between the leading kaon's momentum and the missing momentum, $\phi_{Kp_{miss}}$ in the top right, and the leading kaon's transverse momentum $p_T(K)$ on the lower left. The light dotted lines show the kinematic cuts to be made. Taken from [5].

background events as described in section II C. From this, we can derive an upper bound on the branching ratio and lifetime as

$$N_{sig} \geq N_{B\bar{B}} Br(B^\pm \rightarrow K^\pm a) \langle \mathbb{P}(c\tau) \rangle. \quad (153)$$

In figure 41, we show the corresponding bound in the parameter space of the ALP production branching ratio $Br(B^\pm \rightarrow K^\pm a)$ and its lifetime $c\tau$. We then, in figure 42, interpret the bounds in terms of the couplings c_{WW} and c_{ff} and the ALP mass m_a , using the root finding algorithm `scipy.optimize.fsolve` [141]. For both projections, we use both the full 50 ab^{-1} of Belle II's luminosity, and additionally a smaller value of 0.5 ab^{-1} which roughly corresponds to the luminosity of BaBar and is expected to be

m_a [GeV]	before selection		after selection	
	N_{signal}	$N_{\text{misrec.}}$	N_{signal}	$N_{\text{misrec.}}$
0.005	7802	1442	1091	0
0.3	7823	1453	1022	0
1	7737	1460	770	0
3	7568	1649	0	0

TABLE VI: The number of properly reconstructed and misreconstructed signal events for a representative sample of ALP masses, before and after cuts. Taken from [5].

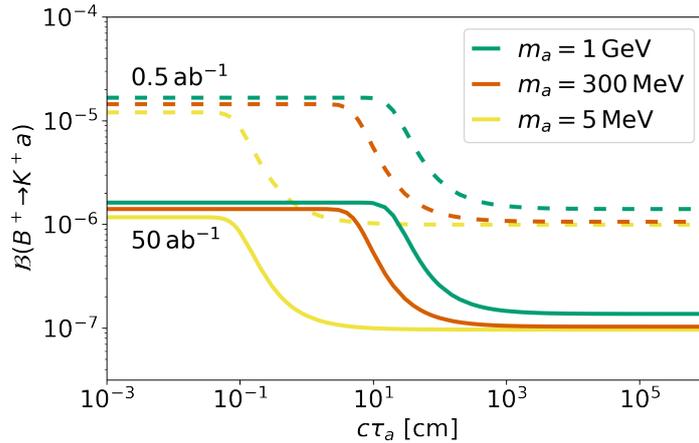


FIGURE 41: The projected 90% CL upper bounds on the production branching ratio $Br(B^\pm \rightarrow K^\pm a)$ of the invisibly decaying ALP $a \rightarrow \text{inv.}$ from missing energy searches at Belle II for luminosities of 0.5 ab^{-1} (shown as dashed lines) and 50 ab^{-1} (shown as solid lines). Taken from [5].

collected soon. In addition to these two lines, we also add a line at the full luminosity with zero background, since we are optimistic that a full experimental analysis would be able to improve upon our cuts somewhat.

All three plots show a transition from a stronger to a weaker bound. This transition corresponds to the coupling or lifetime for which the decay length of the ALP becomes short-lived on the scale of Belle II, meaning that most of the ALPs decay within the detector and thus visibly. Through the angular acceptance of the detector that is large but not fully complete and the fallibility of the detector that misses some tracks, we still get ALP events that appear as invisible decays. Their existence can also still be an excess over the number of such invisible decays we expect from generic events at Belle II, especially because the signal region, as defined by the cuts, is inhabited by so few generic events. Thus, even for short-lived ALPs this analysis can still set a bound on their coupling or production cross section, if not as strong of one as for the long-lived ALPs. In the next section, we will compare these results with a displaced search at Belle II and several other prompt, displaced, and invisible search results.

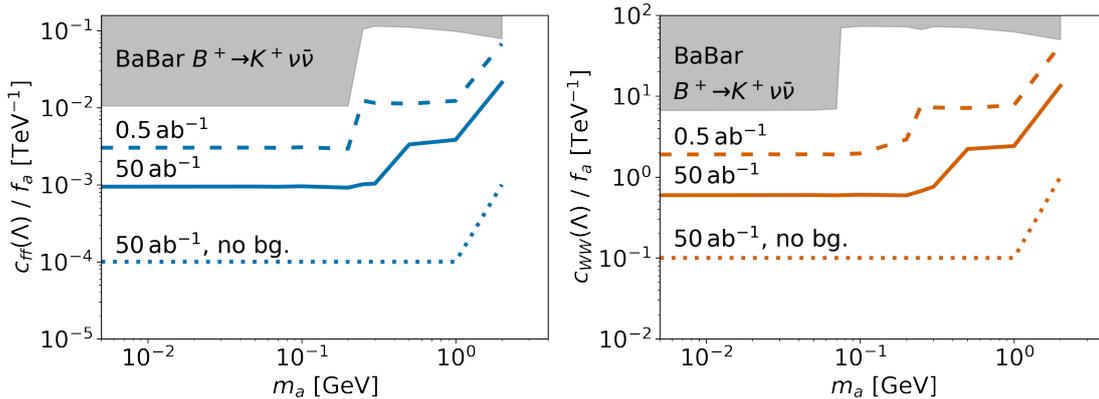


FIGURE 42: The projected 90% CL upper bounds on the ALP UV coupling $c_{ff}(\Lambda)$ (on the left) or $c_{WW}(\Lambda)$ from missing energy searches at Belle II for luminosities of 0.5 ab^{-1} (shown as dashed lines) and 50 ab^{-1} (shown as solid lines), both with background as predicted from our analysis, and 50 ab^{-1} with zero background events (shown as dotted lines). The shaded grey region shows the exclusion from the BaBar bound on $B^\pm \rightarrow K^\pm \nu \bar{\nu}$ [125]. Taken from [5].

3. Comparing LLP Signatures

To compare the results in figure 42, we reinterpreted many results from different signatures and experiments for this ALP model. As the main focus of our study was the missing energy search, I will go through the different bounds quickly, referring to [5] for details. Figure 43 shows the parameter space for the ALP with the limit from the missing energy search described in the last section and other bounds as described in the following.

We show three different types of bounds. The first are bounds from lepton searches, probing the $a \rightarrow \ell^+ \ell^-$ decay. These are shown in figure 43 in green. Here, we have considered bounds from LHCb [88, 89, 147] and show the strongest bound from displaced $B \rightarrow KX$, $X \rightarrow \mu^+ \mu^-$ decays [89], with the larger couplings excluded by searches of prompt and less displaced muons in [88]. We also take into account reinterpreted [146] sterile neutrino bounds from CHARM [148], also on the process $B \rightarrow KX$, $X \rightarrow \mu^+ \mu^-$. We have checked the bounds of additional experiments: BaBar’s search of displaced muon and pion pairs [95], CHARM’s search for displaced muon pairs in kaon decays [85], and NA48/2’s search for $B \rightarrow Ka$, $a \rightarrow \ell^+ \ell^-$ [149, 150]. The regions of parameter space they exclude are already excluded by other experiments. For clarity and readability, we only include the leading bounds in figure 43, so that they are not shown there.

The second type of bounds come from photon searches, which are shown in blue. We use the BaBar analysis of ALPs in $B^\pm \rightarrow K^\pm \gamma \gamma$ decays [144] which is expectedly a much stronger constraint in the c_{WW} scenario, but is still impressively constraining even in

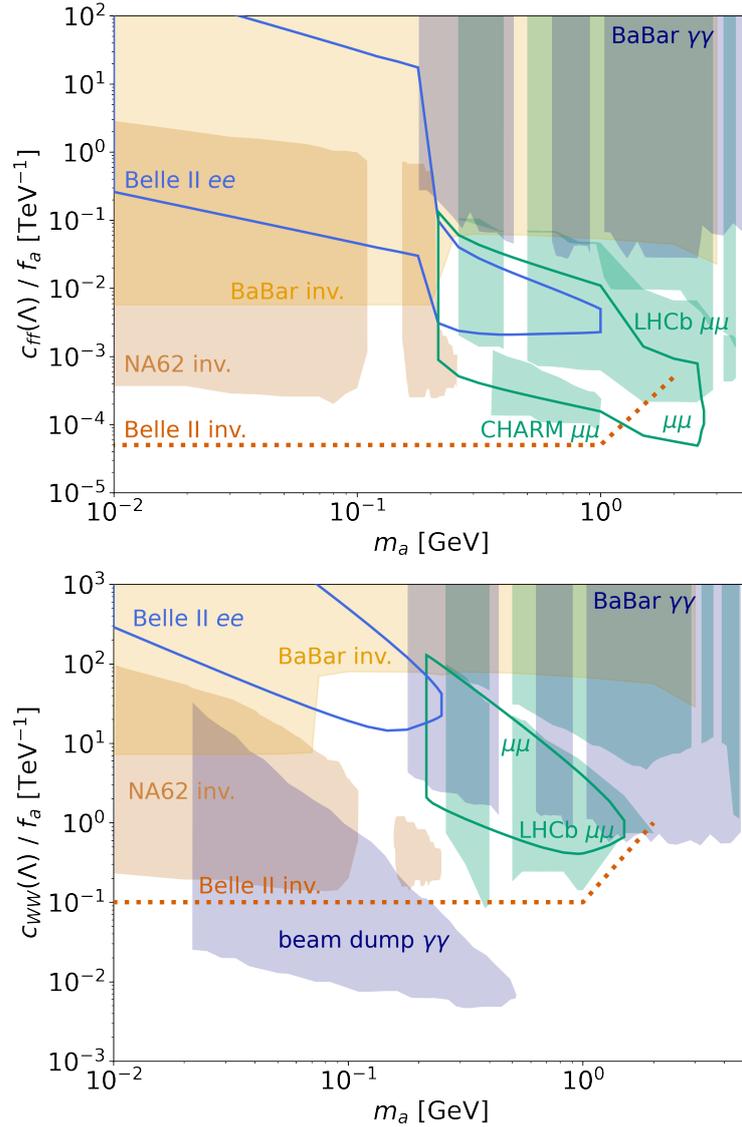


FIGURE 43: Collected bounds on the ALP parameter space of the c_{ff} scenario (top) and the c_{WW} scenario (bottom). The dotted orange line shows the bound for the missing energy search described in section III C 2 for a luminosity of 50 ab^{-1} and zero background events. The blue and green outlined regions are reinterpretations of [1] of displaced decays of the ALP into electron (blue) or muon (green) pairs. The upper tan region is a reinterpretation of the missing energy search for $B^\pm \rightarrow K^\pm \nu \bar{\nu}$ at BaBar [125]. The lower tan region reinterprets the search $K^\pm \rightarrow \pi^\pm \nu \bar{\nu}$ at NA62 [142, 143]. The upper blue region is a reinterpretation of the search for $B^\pm \rightarrow K^\pm \gamma \gamma$ at BaBar [144], while the lower blue region in the c_{WW} scenario plot is a reinterpretation of a beamdump search for diphotons at NuCal, CHARM and E137 [145]. The green regions are reinterpretations of the prompt and displaced searches for $K \rightarrow K X$, $X \rightarrow \mu^+ \mu^-$ at LHCb [88, 89] and CHARM [146]. Taken from [5].

the c_{ff} scenario, given that it operates purely at loop level there. From the beamdump experiments NuCal, CHARM, and E137, we reinterpret the bounds from [145] for the c_{WW} scenario. Unfortunately, they cannot be directly translated into the c_{ff} scenario. In addition, we have also analysed the bounds from E949's $K \rightarrow \pi\gamma\gamma$ search [151] as well as the direct search for $e^+e^- \rightarrow a\gamma$, $a \rightarrow \gamma\gamma$ at Belle II [152], which are both subleading. The Belle II search is expected to become much stronger in the future as more data is taken at Belle II.

The third type of bound we consider are missing energy searches, $a \rightarrow \text{inv.}$, which are shown in yellow/tan. The BaBar search for $B \rightarrow K\nu\bar{\nu}$ [125] allows us to reconstruct limits that are not only shown in figure 43, but also, as its most direct corresponding previous measurement, in the invisible bounds predicted by our analysis in figure 42. A similar Belle study [153] also exists, but their modelling is optimised for three-body decays without resonances, such that we cannot reinterpret it for a decay with an intermediate ALP. This is a problem that also hindered our reinterpretations of the subleading bounds from photon searches at E949 [151] and lepton searches at NA48/2 [149, 150]. The second invisible bound shown comes from the search $K \rightarrow \pi\nu\bar{\nu}$ at NA62 and the collaboration's own interpretation of it as a bound on long-lived particles X produced as $K \rightarrow \pi X$ [142, 143].

The last regions, then, that we have not yet discussed, are the unfilled regions labelled Belle II ee and $\mu\mu$ in blue and green. These correspond to the limits from long-lived ALPs decaying at Belle II into displaced electron or muon pairs. For these, we simply redo the analysis of section III A with the ALP production and decay rates used instead of the dark scalar's.

In conclusion, we find that searches for long-lived particles at Belle II are a good way to probe parameter space that is not yet excluded by other experiments. We find, that a search for $B \rightarrow Ka$, $a \rightarrow \text{inv.}$ has a strong reach for a large mass range, with additional improvements expected from optimising the search for light and heavy ALPs separately. We could also expect the bound to become even stronger when additional kaon resonances and neutral $B \rightarrow K$ decays are included, as we saw in section III A. While the displaced search shows a smaller overall sensitivity to new parameter space than the invisible search, it still probes it and is especially strong in the high mass region where the missing energy bound falters. Due to the loop-induced coupling to fermions, the displaced decays are more suppressed in the c_{WW} scenario. Again, further parameter space is expected to be probeable with the addition of heavier final state particles, like $\tau^+\tau^-$ pairs or D mesons, as we saw in section III A, and also through a dedicated missing energy background analysis for heavier ALPs.

IV. SMEFT

Now, instead of exploring a specific model in a specific experiment, we want to instead use data from several experiments to explore the parameter space of the SMEFT. A recent paper [154] made a fit of SMEFT Wilson coefficients to a collection of top quark observables measured at the LHC. To alleviate some of its blind directions, we will add flavour observables to the fit. We will also analyse how much the flavour structure of the SMEFT, and thus of new physics in the UV, can be probed by a combined fit of top and flavour observables. Due to their high precision, we will use B meson decays. This section is based on my paper [2].

We will first define the operators of the SMEFT, in section IV A, and the WET, in section IV B. Then, we will formalise our notion of the flavour structure and introduce Minimal Flavour Violation (MFV) in section IV C. With this, we will then express the relevant SMEFT operators in this new formalism, in section IV D, and use SMEFT-to-WET matching to do the same with the WET operators, in section IV E. Then, we will quickly introduce our fitting setup and the used observables, in section IV F, before showing and analysing the results in section IV G.

A. SMEFT operators

We assume in this analysis that the new physics couples to the quark sector and that SMEFT operators of dimension six are sufficient to capture it. Additionally, we assume that the Lagrangian of SMEFT conserves CP, which implies that all our Wilson coefficients are real-valued. All SMEFT operators in this section (and this thesis) are defined as in the Warsaw basis [64].

We then focus our analysis on the following 23 effective quark interaction operators, of which 11 contain two quark fields,

$$\begin{aligned}
O_{\phi q}^{(1)pr} &= \left(\phi^\dagger \overleftrightarrow{iD}_\mu \phi \right) (\bar{q}_p \gamma^\mu q_r) & \ddagger O_{uB}^{pr} &= (\bar{q}_p \sigma^{\mu\nu} u_r) \tilde{\phi} B_{\mu\nu} & (154) \\
O_{\phi q}^{(3)pr} &= \left(\phi^\dagger \overleftrightarrow{iD}_\mu^i \phi \right) (\bar{q}_p \gamma^\mu \tau^i q_r) & \ddagger O_{uW}^{pr} &= (\bar{q}_p \sigma^{\mu\nu} u_r) \tau^i \tilde{\phi} W_{\mu\nu}^i \\
O_{\phi u}^{pr} &= \left(\phi^\dagger \overleftrightarrow{iD}_\mu \phi \right) (\bar{u}_p \gamma^\mu u_r) & \ddagger O_{uG}^{pr} &= (\bar{q}_p \sigma^{\mu\nu} T^a u_r) \tilde{\phi} G_{\mu\nu}^a \\
O_{\phi d}^{pr} &= \left(\phi^\dagger \overleftrightarrow{iD}_\mu \phi \right) (\bar{d}_p \gamma^\mu d_r) & \ddagger O_{dB}^{pr} &= (\bar{q}_p \sigma^{\mu\nu} d_r) \phi B_{\mu\nu} \\
\ddagger O_{\phi ud}^{pr} &= \left(\tilde{\phi}^\dagger \overleftrightarrow{iD}_\mu \phi \right) (\bar{u}_p \gamma^\mu d_r) & \ddagger O_{dW}^{pr} &= (\bar{q}_p \sigma^{\mu\nu} d_r) \tau^i \phi W_{\mu\nu}^i \\
& & \ddagger O_{dG}^{pr} &= (\bar{q}_p \sigma^{\mu\nu} T^a d_r) \phi G_{\mu\nu}^a
\end{aligned}$$

the two-quark operators, and 12 contain four quark fields,

$$\begin{aligned}
O_{qq}^{(1)} &= (\bar{q}_p \gamma^\mu q_r) (\bar{q}_s \gamma_\mu q_t) & O_{uu}^{prst} &= (\bar{u}_p \gamma^\mu u_r) (\bar{u}_s \gamma_\mu u_t) \\
O_{qq}^{(3)} &= (\bar{q}_p \gamma^\mu \tau^i q_r) (\bar{q}_s \gamma_\mu \tau^i q_t) & O_{dd}^{prst} &= (\bar{d}_p \gamma^\mu d_r) (\bar{d}_s \gamma_\mu d_t) \\
O_{qu}^{(1)} &= (\bar{q}_p \gamma^\mu q_r) (\bar{u}_s \gamma_\mu u_t) & O_{ud}^{(1)} &= (\bar{u}_p \gamma^\mu u_r) (\bar{d}_s \gamma_\mu d_t) \\
O_{qu}^{(8)} &= (\bar{q}_p \gamma^\mu T^a q_r) (\bar{u}_s \gamma_\mu T^a u_t) & O_{ud}^{(8)} &= (\bar{u}_p \gamma^\mu T^a u_r) (\bar{d}_s \gamma_\mu T^a d_t) \\
O_{qd}^{(1)} &= (\bar{q}_p \gamma^\mu q_r) (\bar{d}_s \gamma_\mu d_t) & \dagger O_{quqd}^{(1)} &= (\bar{q}_p^i u_r) \varepsilon_{ij} (\bar{q}_s^j d_t) \\
O_{qd}^{(8)} &= (\bar{q}_p \gamma^\mu T^a q_r) (\bar{d}_s \gamma_\mu T^a d_t) & \dagger O_{quqd}^{(8)} &= (\bar{q}_p^i T^a u_r) \varepsilon_{ij} (\bar{q}_s^j T^a d_t) .
\end{aligned} \tag{155}$$

the four-quark operators. The \dagger that precedes some of these operators denotes them as non-hermitian. The SMEFT Lagrangian of mass-dimension six operators is then

$$\mathcal{L}_{\text{SMEFT}} = \sum_a \frac{C_a O_a}{\Lambda^2} + \sum_b \left(\frac{C_b \dagger O_b}{\Lambda^2} + \text{h. c.} \right) \tag{156}$$

with the hermitian operators in the first sum and the non-hermitian operators in the second sum. We will not mark the non-hermitian operators with $\dagger O_b$ from this point forward, simply denoting them as O_b .

The Wilson coefficients C_{pr} , C_{prst} of the operators O_{pr} , O_{prst} have generally fully independent elements for different flavour indices $pr(st)$. The symmetric four-quark operators $O_{qq}^{(1)}$, $O_{qq}^{(3)}$, O_{uu} , and O_{dd} , however, fulfil the additional relation

$$C_{prst}^{xx} = C_{stpr}^{xx} \tag{157}$$

due to their identical gauge and Lorentz structures in the two quark bilinears. In the SMEFT Lagrangian, we sum not over all flavour indices, but instead over the independent degrees of freedom of each Wilson coefficient. For these four operators, then, we include only one of the combinations $prst$ and $stpr$.

B. WET operators

For B meson decay observables, specifically in the decays $b \rightarrow s\gamma$, $b \rightarrow sg$, and $b \rightarrow s\ell^+\ell^-$, the WET operators directly involved are

$$\begin{aligned}
\mathcal{O}_7 &= \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \frac{e}{16\pi^2} m_b (\bar{s} \sigma^{\mu\nu} P_R b) F_{\mu\nu} \\
\mathcal{O}_{7'} &= \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \frac{e}{16\pi^2} m_b (\bar{s} \sigma^{\mu\nu} P_L b) F_{\mu\nu}
\end{aligned} \tag{158}$$

$$\begin{aligned}
\mathcal{O}_8 &= \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \frac{g_s}{16\pi^2} m_b (\bar{s} \sigma^{\mu\nu} T^A P_R b) G_{\mu\nu}^A \\
\mathcal{O}_{8'} &= \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \frac{g_s}{16\pi^2} m_b (\bar{s} \sigma^{\mu\nu} T^A P_L b) G_{\mu\nu}^A \\
\mathcal{O}_9 &= \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \frac{e^2}{16\pi^2} (\bar{s} \gamma_\mu P_L b) (\bar{\mu} \gamma^\mu \mu) \\
\mathcal{O}_{9'} &= \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \frac{e^2}{16\pi^2} (\bar{s} \gamma_\mu P_R b) (\bar{\mu} \gamma^\mu \mu) \\
\mathcal{O}_{10} &= \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \frac{e^2}{16\pi^2} (\bar{s} \gamma_\mu P_L b) (\bar{\mu} \gamma^\mu \gamma_5 \mu) \\
\mathcal{O}_{10'} &= \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \frac{e^2}{16\pi^2} (\bar{s} \gamma_\mu P_R b) (\bar{\mu} \gamma^\mu \gamma_5 \mu)
\end{aligned}$$

as defined in EOS basis [59, 60].

In the Standard Model, the processes $b \rightarrow s\gamma$ and $b \rightarrow sg$ are dominated by the operators \mathcal{O}_7 and \mathcal{O}_8 , respectively, while the operators \mathcal{O}_9 and \mathcal{O}_{10} dominate $b \rightarrow sl^+\ell^-$ transitions. The primed operators have opposite chiralities from the unprimed ones. In the Standard model, the unprimed operators have the values [60]

$$\mathcal{C}_{\text{SM}}^7 \sim -0.3, \quad \mathcal{C}_{\text{SM}}^8 \sim -0.2, \quad \mathcal{C}_{\text{SM}}^9 \sim +4.2, \quad \mathcal{C}_{10}^{\text{SM}} \sim -4.3, \quad (159)$$

the operators $\mathcal{O}_{7'}$ and $\mathcal{O}_{8'}$ are suppressed in comparison to \mathcal{O}_7 and \mathcal{O}_8 by a factor of m_s/m_b each, and the operators \mathcal{O}_9 and \mathcal{O}_{10} are absent due to the weak interaction only coupling to left-chirality fermions.

In addition to \mathcal{O}_9 and \mathcal{O}_{10} , there are six more semi-leptonic operators of scalar, pseudoscalar and tensor structures in the WET. Since we use the WET as a low-energy theory of the SMEFT and only use operators that are produced from its matching, these operators are not considered. For the scalar and pseudoscalar operators \mathcal{O}_S , $\mathcal{O}_{S'}$, \mathcal{O}_P , and $\mathcal{O}_{P'}$, this is due to the fact that they match only to semi-leptonic SMEFT operators, which we neglect in this analysis. For the tensor operators \mathcal{O}_T and \mathcal{O}_{T5} , it is because they are produced only from SMEFT operators of higher mass dimensions than six [155, 156], leaving us with no contribution from the ones we consider.

Beyond these operators, four more WET operators can contribute to the above processes, these being the four-quark operators

$$\begin{aligned}
\mathcal{O}_1 &= \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* (\bar{c} \gamma^\mu P_L T^A b) (\bar{s} \gamma_\mu P_L T^A c) \\
\mathcal{O}_{1'} &= \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* (\bar{c} \gamma^\mu P_R T^A b) (\bar{s} \gamma_\mu P_L T^A c) \\
\mathcal{O}_2 &= \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* (\bar{c} \gamma^\mu P_L b) (\bar{s} \gamma_\mu P_L c)
\end{aligned} \quad (160)$$

$$\mathcal{O}_{2'} = \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* (\bar{c} \gamma^\mu P_R b) (\bar{s} \gamma_\mu P_L c).$$

In the Standard Model, \mathcal{O}_1 and \mathcal{O}_2 are produced via a W boson exchange and are quite sizeable, [60]

$$\mathcal{C}_{\text{SM}}^1 \sim -0.3, \quad \mathcal{C}_{\text{SM}}^2 \sim 1 \quad (161)$$

and contribute to the observables through operator mixing and non-local contributions.

Due to the connection between the top and B observables in this analysis, our WET Wilson coefficients are run down from the weak scale to the scale of B decays at $\mu_B \sim m_B$. Thus, we need to take into account the operator mixing from the running. Under the RG evolution [55, 157], the four-quark operators \mathcal{O}_1 and \mathcal{O}_2 mix into the dipole operators \mathcal{O}_7 and \mathcal{O}_8 and the semi-leptonic operator \mathcal{O}_9 , while the four-quark operators $\mathcal{O}_{1'}$ and $\mathcal{O}_{2'}$ similarly mix into $\mathcal{O}_{7'}$, $\mathcal{O}_{8'}$, and $\mathcal{O}_{9'}$. Due to their lepton current's axial-vector nature, $\mathcal{O}_{10'}$ and \mathcal{O}_{10} do not receive any mixing contributions. In the Standard model, these are significant contributions, making up approximately 50% of $\mathcal{C}_9(m_B)$, 10% of \mathcal{C}_7 , and 3% of \mathcal{C}_8 .

The non-local contributions are given by the production of intermediate resonant mesons in decays of the shape $B \rightarrow K \ell^+ \ell^-$, for example $B \rightarrow K J/\psi \rightarrow K \ell^+ \ell^-$. The best way to treat these contributions to minimise systematic uncertainties is an area of active research [158–165], and as such, we avoid observables that have a large non-local contribution in this analysis, using only the processes $B \rightarrow X_s \gamma$ and $B_s \rightarrow \mu^+ \mu^-$, but not $B \rightarrow K \mu^+ \mu^-$.

C. Flavour structure and Minimal Flavour Violation

Before we get to matching the two EFTs to each other, let us first specify what we mean by flavour structures and define Minimal Flavour Violation (MFV) [166, 167], the specific flavour structure we will focus on in this analysis. Then, in the next section, IV D, we can first explore its impact on the considered SMEFT operators before matching them to the WET and seeing the impact on the WET operators in section IV E.

A flavour structure is the pattern of the Wilson coefficient matrices in flavour space. For example, a flavour-diagonal flavour structure allows only Wilson coefficients with quark bilinears coupling quarks within the same generation, while a flavour-universal structure is given by a flavour-diagonal structure in which all flavours couple with the same value.

The Standard Model is invariant under the flavour transformations

$$U(3)_Q \times U(3)_U \times U(3)_D \quad (162)$$

in its gauge interactions, but its flavour symmetry is broken by the Yukawa couplings Y_U and Y_D in the quark's mass terms and Higgs couplings. We can construct the flavour structure that we call Minimal Flavour Violation [166, 167] by requiring that any flavour symmetry breaking beyond the Standard Model is also propagated by the same Yukawa matrices. We can treat the Yukawa matrices as fictitious fields, so-called spurions, that transform under flavour symmetry as

$$Y_U : (3, \bar{3}, 1) \quad Y_D : (3, 1, \bar{3}) \quad (163)$$

and construct operators with them that then do not violate flavour symmetry explicitly. For example, the current $\bar{q}_L u_R$ is made up of the left-handed quark doublet that is charged under $U(3)_Q$ and the right-handed up quark that is charged under $U(3)_U$, such that

$$q_L : (3, 1, 1) \quad u_R : (1, 3, 1) \quad \bar{q}_L u_R : (\bar{3}, 3, 1). \quad (164)$$

The current thus breaks flavour symmetry. With the up-type Yukawa spurion Y_U , we can restore the flavour symmetry, as

$$\bar{q}_L Y_U u_R : (\bar{3} \times 3, 3 \times \bar{3}, 1) \ni (1, 1, 1) \quad (165)$$

that is, the combination contains a flavour-symmetric term as $3 \times \bar{3} = \bar{3} \times 3 = \mathbf{1} + \mathbf{8}$.

We construct three spurions that are singlets in all but one flavour direction

$$\mathcal{A}_Q : (3 \times \bar{3}, 1, 1) \quad \mathcal{A}_U : (1, 3 \times \bar{3}, 1) \quad \mathcal{A}_D : (1, 1, 3 \times \bar{3}). \quad (166)$$

in terms of Yukawa matrices and to first order as

$$\mathcal{A}_Q = a \mathbf{1} + b Y_U Y_U^\dagger + c Y_D Y_D^\dagger + \dots \quad (167)$$

and based on it

$$\begin{aligned} \mathcal{A}_U &= a \mathbf{1} + Y_U^\dagger \mathcal{A}_Q Y_U = a \mathbf{1} + b Y_U^\dagger Y_U + c Y_U^\dagger Y_D Y_D^\dagger Y_U + \dots \\ \mathcal{A}_D &= a \mathbf{1} + Y_D^\dagger \mathcal{A}_Q Y_D = a \mathbf{1} + b Y_D^\dagger Y_U Y_U^\dagger Y_D + c Y_D^\dagger Y_D + \dots \end{aligned} \quad (168)$$

where \mathcal{A}_Q governs currents between left-handed quarks, \mathcal{A}_U those between right-handed

up-type quarks, and \mathcal{A}_D those between right-handed down-type quarks. Here, we only take into account the leading terms in Y_U and Y_D , and label the coefficient of the universal diagonal term a , the coefficient of the leading Y_U contribution as b and the coefficient of the leading Y_D contribution as c .

With these definitions, we can now describe the flavour structures of general quark currents as

$$\begin{aligned}
(\mathcal{A}_Q)_{pr} (\bar{q}_p \gamma_\mu q_r) & & (\mathcal{A}_U)_{pr} (\bar{u}_p \gamma_\mu u_r) & & (169) \\
(\mathcal{A}_Q Y_U)_{pr} (\bar{q}_p \sigma_{\mu\nu} u_r) & & (\mathcal{A}_D)_{pr} (\bar{d}_p \gamma_\mu d_r) & & \\
(\mathcal{A}_Q Y_D)_{pr} (\bar{q}_p \sigma_{\mu\nu} d_r) & & (Y_U^\dagger \mathcal{A}_Q Y_D)_{pr} (\bar{u}_p \gamma_\mu d_r) & & .
\end{aligned}$$

The Yukawa matrices Y_U and Y_D are diagonalisable to the Yukawa matrices $Y_u = \text{diag}(y_u, y_c, y_t)$, $Y_d = \text{diag}(y_d, y_s, y_b)$ with the help of unitary matrices $\mathcal{U}_{R,L}$ and $\mathcal{D}_{R,L}$,

$$Y_U = \mathcal{U}_L Y_u \mathcal{U}_R^\dagger \quad Y_D = \mathcal{D}_L Y_d \mathcal{D}_R^\dagger. \quad (170)$$

With these definitions, the CKM matrix is given by

$$V = \mathcal{U}_L^\dagger \mathcal{D}_L. \quad (171)$$

Without loss of generality, we will work in the up-quark mass basis, where

$$q_p = \begin{pmatrix} u_{Lp} \\ V_{pr} d_{Lr} \end{pmatrix} \quad u_p = u_{Rp} \quad d_p = d_{Rp} \quad (172)$$

and the Yukawa matrices are

$$Y_U = Y_u \quad Y_D = V Y_d. \quad (173)$$

We will call this setup as up-alignment. As long as there are no additional sources of flavour breaking, the choice of basis has no impact on observables and is not accessible by experiment. This is the case here, as we assume the flavour breaking Yukawa couplings of the Standard Model as our only sources of flavour breaking. It should be noted that in the following we will label specific elements of the CKM matrix by the generation number of the participating quarks rather than the quark's abbreviation, that is, V_{33} instead of V_{tb} , to avoid confusion with the variable flavour indices p, r, s, t .

D. MFV in SMEFT

Now that we have introduced MFV, let us see what impact it has on the Wilson coefficients of the SMEFT operators introduced in section IV A. For every operator with a quark current as shown in equation (169), we express the Wilson coefficient by the corresponding combination of spurions. For example, the two-quark operator $O_{\phi u}$ is defined as

$$O_{\phi u} = \left(\phi^\dagger \overleftrightarrow{D}_\mu \phi \right) (\bar{u}_p \gamma^\mu u_r) \quad (174)$$

with a right-handed up-type current, such that following equation (169) its Wilson coefficient is

$$\begin{aligned} C_{\phi u} &= (\mathcal{A}_U)_{pr} \quad (175) \\ &= a_{\phi u} \delta_{pr} + b_{\phi u} \left(Y_U^\dagger Y_U \right)_{pr} + c_{\phi u} \left(Y_U^\dagger Y_D Y_D^\dagger Y_D \right)_{pr} + \dots \\ &= a_{\phi u} \delta_{pr} + b_{\phi u} \left(Y_u^\dagger Y_u \right)_{pr} + c_{\phi u} \left(Y_u^\dagger V Y_d (V Y_d)^\dagger Y_u \right)_{pr} + \dots \end{aligned}$$

in up-alignment MFV. Now, we simplify this by neglecting the Yukawa couplings of light quarks (y_c or lighter) and keeping only the leading contribution in the Yukawa couplings,

$$\begin{aligned} C_{\phi u} &= a_{\phi u} \delta_{pr} + b_{\phi u} y_t^2 \delta_{p3} \delta_{r3} + c_{\phi u} y_t^2 y_b^2 V_{33} V_{33}^* \delta_{p3} \delta_{r3} + \mathcal{O}(y_c^2) \quad (176) \\ &= a_{\phi u} \delta_{pr} + b_{\phi u} y_t^2 \delta_{p3} \delta_{r3} + \mathcal{O}(y_b^2) . \end{aligned}$$

In the last step, we also neglect the y_b^2 term as it only adds a sub-leading contribution to the $p = 3, r = 3$ component of the Wilson coefficient and does not add a new flavour structure. These are approximations that we will be using for the other operators, too. The full set of two-quark operators with their flavour structure degrees of freedom under up-alignment MFV is shown in table VII. Due to its common occurrence, we will label the combination $a + by_t^2$ of MFV coefficients as A in the following sections.

Since the top-quark Yukawa coupling is not a small parameter, neglecting higher powers $\mathcal{O}(y_t^n)$ of it is not as well-justified as for the other quarks. We only include them in cases where they add additional flavour structures, which is only the case for the four-quark operators we will look at next. For a full study of $\mathcal{O}(y_t^n)$ terms, see [168, 169].

Four-quark operators are made up of two quark bilinears. In addition to using one flavour structure as seen in equation (169) for each of the bilinears and multiplying them, we can also cross the flavour structure between the bilinears, giving us a structure $(pt)(sr)$ as well as $(pr)(st)$. We will label this second structure $\tilde{\mathcal{A}}$ to distinguish it from \mathcal{A} . The

	$C_{\phi q}^{(1)}_{pr}$	$C_{\phi q}^{(3)}_{pr}$	$C_{\phi u}_{pr}$	$C_{\phi d}_{pr}$	$C_{\phi ud}_{pr}$	C_{uX}_{pr}	C_{dX}_{pr}
$p = r < 3$	a	a	a	a	0	0	0
$p = r = 3$	$a + by_t^2$	$a + by_t^2$	$a + by_t^2$	a	$(a + by_t^2) y_b y_t V_{33}$	$(a + by_t^2) y_t$	$(a + by_t^2) y_b V_{33}$
$p > r$	$cy_b^2 V_{p3} V_{r3}^*$	$cy_b^2 V_{p3} V_{r3}^*$	0	0	0	0	0
$p < r = 3$	$cy_b^2 V_{p3} V_{33}^*$	$cy_b^2 V_{p3} V_{33}^*$	0	0	0	$cy_b^2 y_t V_{p3} V_{33}^*$	$ay_b V_{p3}$
#	3	3	2	1	1	2	2

TABLE VII: Shown are the components of the Wilson coefficients of two-quark SMEFT operators under up-alignment MFV, as explained in the text, and the total degrees of freedom #, that is, number of independent parameters in the operator's flavour structure. Taken from [2]

bilinears are then defined by the flavour structures

$$\begin{aligned}
& \left((\mathcal{A}_Q)_{kl} (\mathcal{A}_Q)_{mn} + (\tilde{\mathcal{A}}_Q)_{kn} (\tilde{\mathcal{A}}_Q)_{ml} \right) \left(\bar{q}^k \gamma^\mu q^l \right) \left(\bar{q}^m \gamma_\mu q^n \right) \quad (177) \\
& \left((\mathcal{A}_U)_{kl} (\mathcal{A}_U)_{mn} + (\tilde{\mathcal{A}}_U)_{kn} (\tilde{\mathcal{A}}_U)_{ml} \right) \left(\bar{u}^k \gamma^\mu u^l \right) \left(\bar{u}^m \gamma_\mu u^n \right) \\
& \left((\mathcal{A}_D)_{kl} (\mathcal{A}_D)_{mn} + (\tilde{\mathcal{A}}_D)_{kn} (\tilde{\mathcal{A}}_D)_{ml} \right) \left(\bar{d}^k \gamma^\mu d^l \right) \left(\bar{d}^m \gamma_\mu d^n \right) \\
& \left((\mathcal{A}_Q)_{kl} (\mathcal{A}_U)_{mn} + (\tilde{\mathcal{A}}_Q Y_U)_{kn} (Y_U^\dagger \tilde{\mathcal{A}}_Q^\dagger)_{ml} \right) \left(\bar{q}^k \gamma^\mu q^l \right) \left(\bar{u}^m \gamma_\mu u^n \right) \\
& \left((\mathcal{A}_Q)_{kl} (\mathcal{A}_D)_{mn} + (\tilde{\mathcal{A}}_Q Y_D)_{kn} (Y_D^\dagger \tilde{\mathcal{A}}_Q^\dagger)_{ml} \right) \left(\bar{q}^k \gamma^\mu q^l \right) \left(\bar{d}^m \gamma_\mu d^n \right) \\
& \left((\mathcal{A}_U)_{kl} (\mathcal{A}_D)_{mn} + (Y_U^\dagger \tilde{\mathcal{A}}_Q Y_D)_{kn} (Y_D^\dagger \tilde{\mathcal{A}}_Q^\dagger Y_U)_{ml} \right) \left(\bar{u}^k \gamma^\mu u^l \right) \left(\bar{d}^m \gamma_\mu d^n \right) \\
& \left((\mathcal{A}_Q Y_U)_{kl} (\mathcal{A}_Q Y_D)_{mn} + (\tilde{\mathcal{A}}_Q Y_D)_{kn} (\tilde{\mathcal{A}}_Q Y_U)_{ml} \right) \left(\bar{q}^k u^l \right) \left(\bar{q}^m d^n \right),
\end{aligned}$$

When expanding the products of spurions, we label the flavour parameters as (aa) , $(\tilde{a}\tilde{a})$, (ab) , etc., such that

$$\begin{aligned}
(\mathcal{A}_Q)_{kl} (\mathcal{A}_Q)_{mn} &= (aa) \delta_{kl} \delta_{mn} + (ba) y_t^2 \delta_{k3} \delta_{l3} \delta_{mn} + (ab) y_t^2 \delta_{kl} \delta_{m3} \delta_{n3} + \dots \quad (178) \\
(\tilde{\mathcal{A}}_Q)_{kn} (\tilde{\mathcal{A}}_Q)_{ml} &= (\tilde{a}\tilde{a}) \delta_{kn} \delta_{ml} + (\tilde{b}\tilde{a}) y_t^2 \delta_{k3} \delta_{n3} \delta_{ml} + (\tilde{a}\tilde{b}) y_t^2 \delta_{kn} \delta_{m3} \delta_{l3} + \dots
\end{aligned}$$

and similarly for the other combinations. The full set of four-quark operators with their flavour structure degrees of freedom under up-alignment MFV is shown in table VIII.

We find that in MFV, two-quark operators are described by at most 3 independent degrees of freedom, and often less, while four-quark operators have a richer flavour structure with up to 9 degrees of freedom. We note that overall, operators with left-handed quarks have the highest number of degrees of freedom while operators with, and even more so with exclusively, down quarks have the least degrees of freedom. This is due to the suppression through Y_u and Y_d in both right-handed quark fields, and due to the additional suppression of V_{pr} in the down-quark case.

	$C_{prst}^{(1)}, C_{prst}^{(3)}$	C_{prst}^{uu}	C_{prst}^{dd}
$p = r = s = t < 3$	$(aa) + (\bar{a}\bar{a})$	$(aa) + (\bar{a}\bar{a})$	$(aa) + (\bar{a}\bar{a})$
$3 > p = r \neq s = t < 3$	(aa)	(aa)	(aa)
$3 > p = t \neq r = s < 3$	$(\bar{a}\bar{a})$	$(\bar{a}\bar{a})$	$(\bar{a}\bar{a})$
$3 = p = r > s = t$	$(aa) + (ba)y_t^2$	$(aa) + (ba)y_t^2$	(aa)
$3 = p = t > r = s$	$(\bar{a}\bar{a}) + (\bar{b}\bar{a})y_t^2$	$(\bar{a}\bar{a}) + (\bar{b}\bar{a})y_t^2$	$(\bar{a}\bar{a})$
$p = r = s = t = 3$	$(aa) + (\bar{a}\bar{a}) + 2((ba) + (\bar{b}\bar{a}))y_t^2 + \mathcal{O}(y_t^4)$	$(aa) + (\bar{a}\bar{a}) + 2((ba) + (\bar{b}\bar{a}))y_t^2 + \mathcal{O}(y_t^4)$	$(aa) + (\bar{a}\bar{a})$
$3 > p = r \neq s \neq t$	$(ac)y_b^2 V_{s3} V_{t3}^*$	0	0
$3 > p = t \neq r \neq s$	$(\bar{a}\bar{c})y_b^2 V_{s3} V_{r3}^*$	0	0
$3 = p = r \neq s \neq t$	$((ac) + (bc)y_t^2)y_b^2 V_{s3} V_{t3}^*$	0	0
$3 = p = t \neq r \neq s$	$((\bar{a}\bar{c}) + (\bar{b}\bar{c})y_t^2)y_b^2 V_{s3} V_{r3}^*$	0	0
#	9	5	2

	$C_{prst}^{(1)}, C_{prst}^{(8)}$	$C_{prst}^{(1)}, C_{prst}^{(8)}$	$C_{prst}^{(1)}, C_{prst}^{(8)}$	$C_{prst}^{(1)}, C_{prst}^{(8)}$
$p = r = s = t < 3$	(aa)	(aa)	(aa)	0
$3 > p = r \neq s = t < 3$	(aa)	(aa)	(aa)	0
$p = r < s = t = 3$	$(aa) + (ab)y_t^2$	(aa)	(aa)	0
$3 = p = r > s = t$	$(aa) + (ba)y_t^2$	$(aa) + (ba)y_t^2$	$(aa) + (ba)y_t^2$	0
$p = r = s = t = 3$	$(aa) + ((ab) + (ba) + (\bar{a}\bar{a}))y_t^2 + \mathcal{O}(y_t^4)$	$(aa) + (ba)y_t^2$	$(aa) + (ba)y_t^2$	$((aa) + (\bar{a}\bar{a}) + (ab)y_t^2 + ((ba) + (\bar{a}\bar{b}) + (\bar{b}\bar{a}))y_t^2) \times y_b y_t V_{33} + \mathcal{O}(y_t^4)$
$p \neq r \neq s = t < 3$	$(ca)y_b^2 V_{p3} V_{r3}^*$	$(ca)y_b^2 V_{p3} V_{r3}^*$	0	0
$p \neq r \neq s = t = 3$	$((ca) + (cb)y_t^2)y_b^2 V_{p3} V_{r3}^*$	$((ca) + (\bar{a}\bar{a}))y_b^2 V_{p3} V_{r3}^*$	0	0
$p < r = s = t = 3$	$((ca) + ((cb) + (\bar{c}\bar{a}))y_t^2) \times y_b^2 V_{p3} V_{33}^* + \mathcal{O}(y_t^4)$	$((ca) + (\bar{a}\bar{a}) + (\bar{a}\bar{b})y_t^2) \times y_b^2 V_{p3} V_{33}^*$	0	$((\bar{a}\bar{a}) + (\bar{a}\bar{b})y_t^2)y_b y_t V_{p3}$
$r < p = s = t = 3$	$((ca) + ((cb) + (\bar{a}\bar{c}))y_t^2) \times y_b^2 V_{33} V_{r3}^* + \mathcal{O}(y_t^4)$	$((ca) + (\bar{a}\bar{a}) + (\bar{b}\bar{a})y_t^2) \times y_b^2 V_{33} V_{r3}^*$	0	0
$s < p = r = t = 3$	0	0	0	$((aa) + (ba)y_t^2)y_b y_t V_{s3}$
#	8	6	2	3

TABLE VIII: Shown are the components of the Wilson coefficients of four-quark SMEFT operators under up-alignment MFV, as explained in the text, and the total degrees of freedom #, that is, number of independent parameters in the operator's flavour structure. Taken from [2]

E. Matching SMEFT to WET

Now that we have defined the structure of our SMEFT Wilson coefficients in up-alignment MFV, we want to know how the SMEFT (and UV) flavour structure of MFV imprints itself onto the Wilson coefficients of the WET operators. For this, we use SMEFT-to-WET matching [56, 170–172] following [171] to first match the WET operators to their SMEFT counterparts and then translate their Wilson coefficients into the MFV parameters accordingly. While the CKM matrix also gains contributions from SMEFT operators and UV new physics, we neglect this here, referring instead to [173].

From the two-quark SMEFT operators there are both tree-level and loop-level contributions to the WET operators considered, while the four-quark operators only contribute at loop level. In table IX, we show the flavour coefficients with which the SMEFT operators contribute to the WET operators at tree level. These are only operators with unsuppressed $b \rightarrow s$ currents.

	\mathcal{C}_7	\mathcal{C}_8	\mathcal{C}_9	\mathcal{C}_{10}
$C_{\phi q}^{(1)}$	—	—	$b_{\phi q}^{(1)} y_t^2$	$b_{\phi q}^{(1)} y_t^2$
$C_{\phi q}^{(3)}$	—	—	$b_{\phi q}^{(3)} y_t^2$	$b_{\phi q}^{(3)} y_t^2$
C_{dB}	$b_{dB} y_t^2$	—	—	—
C_{dW}	$b_{dW} y_t^2$	—	—	—
C_{dG}	—	$b_{dG} y_t^2$	—	—

TABLE IX: Tree-level contributing two-quark SMEFT operators and the WET operators they contribute to as well as the MFV flavour coefficients they contribute with. Taken from [2].

	$t\bar{t}$	single top	tW	tZ	$t\bar{t}Z$	$t\bar{t}W$	#
$C_{\phi q}^{(1)}$	—	—	—	$a_{\phi q}^{(1)}, A_{\phi q}^{(1)}$	$a_{\phi q}^{(1)}, A_{\phi q}^{(1)}$	—	2
$C_{\phi q}^{(3)}$	—	$a_{\phi q}^{(3)}, A_{\phi q}^{(3)}$	$A_{\phi q}^{(3)}$	$a_{\phi q}^{(3)}, A_{\phi q}^{(3)}$	$a_{\phi q}^{(3)}, A_{\phi q}^{(3)}$	$a_{\phi q}^{(3)}$	2
$C_{\phi u}$	—	—	—	$A_{\phi u}$	$a_{\phi u}, A_{\phi u}$	—	2
$C_{\phi d}$	—	—	—	—	$a_{\phi d}$	—	1
C_{uB}	—	—	—	$A_{uB} y_t$	$A_{uB} y_t$	—	1
C_{uW}	—	$A_{uW} y_t$	$A_{uW} y_t$	$A_{uW} y_t$	$A_{uW} y_t$	—	1
C_{uG}	$A_{uG} y_t$	—	$A_{uG} y_t$	—	$A_{uG} y_t$	$A_{uG} y_t$	1

TABLE X: Loop-level contributing two-quark SMEFT operators and the WET operators they contribute to as well as the MFV flavour coefficients they contribute with. Taken from [2].

The loop-induced decays in two-quark operators come from either loops with a W boson and a t quark or loops with a Z boson and a b or s quark. In the tW loop, the flavour change is mediated by the Standard Model flavour changing W boson, while in the b/sZ loop the flavour change comes from a new physics effect leading to a flavour changing neutral current SMEFT operator. Both diagrams are shown, with the vertices where the SMEFT operators can be inserted marked with unfilled circles, in figure 44a and 44b. In table X, we show the contributions from the two-quark operators via these two different loops to the WET operators separately, each with the MFV coefficients that the matching brings with it. We detail the specific mixing relations more closely in [2].

For four-quark operators, there are no tree-level contributions as none of the WET operators we take into account are four-quark operators. Instead, we deal with two different types of loop-induced contributions, as shown in the figures 44c and 44d. We differentiate here between operators where the b and s come from the same bilinear and from different ones. The first case corresponds to two quark currents with one going through the loop and the other being the b - s transition. In the other case, the currents flow from the b quark into the loop and out of the loop into the s quark.

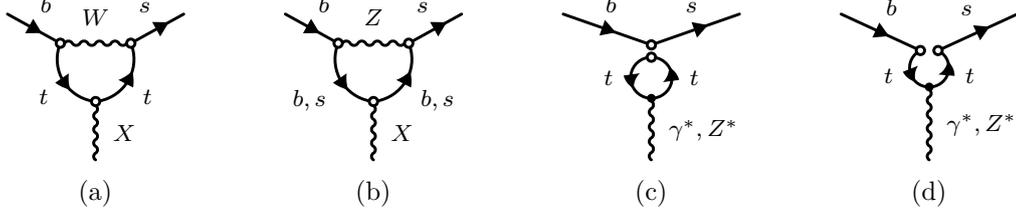


FIGURE 44: Feynman diagrams showing the loops through which the SMEFT operators match onto the WET. In (a), we have a tW loop, in (b) a b/sZ loop, and in (c) and (d) four-quark loops in which flavour either changes through the loop or separately. X is either an on-shell photon or gluon, when matching onto the WET operators \mathcal{O}_7 and \mathcal{O}_8 , or virtual photons that couple to a lepton pair for the purposes of matching onto \mathcal{O}_9 and \mathcal{O}_{10} . Taken from [2].

The operators $O_{qq}^{(1)}$ and $O_{qq}^{(3)}$ give three different quark current combinations,

$$\begin{aligned}
 (\bar{u}_L u_L)(\bar{u}_L u_L), (\bar{d}_L d_L)(\bar{d}_L d_L) &: C_{qq}^{(1)} + C_{qq}^{(3)} \\
 (\bar{u}_L u_L)(\bar{d}_L d_L) &: C_{qq}^{(1)} - C_{qq}^{(3)} \\
 (\bar{u}_L d_L)(\bar{d}_L u_L) &: C_{qq}^{(3)}.
 \end{aligned} \tag{179}$$

Only the second and third of these combinations are inserted at one-loop in the matching of SMEFT operators to the $b \rightarrow s$ WET operators, such that it is useful to redefine

$$\begin{aligned}
 A_{qq}^{(1/3)} &= (aa)_{qq}^{(1/3)} + (ba)_{qq}^{(1/3)} y_t^2, & A_{qq}^{(-)} &= A_{qq}^{(1)} - A_{qq}^{(3)}, \\
 \tilde{A}_{qq}^{(1/3)} &= (\tilde{a}\tilde{a})_{qq}^{(1/3)} + (\tilde{b}\tilde{a})_{qq}^{(1/3)} y_t^2, & \tilde{A}_{qq}^{(-)} &= \tilde{A}_{qq}^{(1)} - \tilde{A}_{qq}^{(3)}, \\
 B_{qq}^{(1/3)} &= (ba)_{qq}^{(1/3)} + (\tilde{b}\tilde{a})_{qq}^{(1/3)}, & B_{qq}^{(-)} &= B_{qq}^{(1)} - B_{qq}^{(3)}.
 \end{aligned} \tag{180}$$

In the WET operators \mathcal{O}_9 and \mathcal{O}_{10} we also notice contributions of the form

$$\begin{aligned}
 F_{qq}^{(-)} &\equiv \tilde{A}_{qq}^{(-)} + B_{qq}^{(-)} y_t^2 &= (V_{33} V_{32}^*)^{-1} \sum_p \left(C_{qq}^{(1),33pp} - C_{qq}^{(3),33pp} \right) V_{p2}^* V_{p3} \\
 F_{qq}^{(3)} &\equiv A_{qq}^{(3)} + B_{qq}^{(3)} y_t^2 &= (V_{33} V_{32}^*)^{-1} \sum_p C_{qq}^{(3),3pp3} V_{p2}^* V_{p3} \\
 F_{qu}^{(1)} &\equiv ((\tilde{a}\tilde{a})_{qu}^{(1)} + (ba)_{qu}^{(1)}) y_t^2 &= (V_{33} V_{32}^*)^{-1} \sum_p C_{qu}^{(1),pp33} V_{p2}^* V_{p3},
 \end{aligned} \tag{181}$$

from the SMEFT operators $O_{qq}^{(1)}$, $O_{qq}^{(3)}$, and $O_{qu}^{(1)}$, while the WET operators \mathcal{O}_7 and \mathcal{O}_8 get contributions

$$\begin{aligned}
 \mathcal{C}_7 &: \left(F_{quqd}^{(1)} + \frac{4}{3} F_{quqd}^{(8)} \right) y_b y_t = (V_{33} V_{32}^*)^{-1} \sum_p \left(C_{quqd}^{(1),p333} + \frac{4}{3} C_{quqd}^{(8),p333} \right) V_{p2}^* V_{p3} \\
 \mathcal{C}_8 &: \left(F_{quqd}^{(1)} - \frac{1}{6} F_{quqd}^{(8)} \right) y_b y_t = (V_{33} V_{32}^*)^{-1} \sum_p \left(C_{quqd}^{(1),p333} - \frac{1}{6} C_{quqd}^{(8),p333} \right) V_{p2}^* V_{p3},
 \end{aligned} \tag{182}$$

	\mathcal{C}_7	\mathcal{C}_8	\mathcal{C}_9	\mathcal{C}_{10}
SM	-0.337	-0.183	4.27	-4.17
$(a_{\phi q}^{(-)}, b_{\phi q}^{(-)})$	(0, -0.008)	(0, 0.025)	(-0.01, -2.07)	(0.1, 24.73)
$(a_{\phi q}^{(3)}, b_{\phi q}^{(3)})$	(-0.034, 0.061)	(-0.017, 0.091)	(0.25, -4.18)	(-0.82, 48.67)
$(a_{\phi u}, b_{\phi u})$	(0, 0)	(0, 0)	(0.01, 0.01)	(-0.1, -0.1)
$A_{\phi ud}$	-0.033	-0.015	0	0
A_{uB}	-0.188	0	0.148	0
A_{uW}	0	0.024	0.115	-0.440
A_{uG}	0	-0.055	0	0
(a_{dB}, b_{dB})	(-0.056, 19.814)	(0, -0.005)	(0, 0)	(0, 0)
(a_{dW}, b_{dW})	(0.059, -10.796)	(0.118, 0.064)	(0, 0)	(0, 0)
(a_{dG}, b_{dG})	(0, 0)	(-0.016, 5.816)	(0, 0)	(0, 0)
$F_{qq}^{(-)}$	0	0	-0.1	0.59
$F_{qq}^{(3)}$	0	0	-0.12	0.7
$F_{qu}^{(1)}$	0	0	-0.01	-0.59
$F_{quqd}^{(1)}$	-0.019	-0.028	0	0
$F_{quqd}^{(8)}$	-0.025	0.005	0	0

TABLE XI: The numerical one-loop matching relations between SMEFT and WET in up-alignment MFV for a new physics scale of $\Lambda = 1$ TeV. Taken from [2].

with

$$F_{quqd}^{(1/8)} \equiv (aa)_{quqd}^{(1/8)} + ((ab)_{quqd}^{(1/8)} + (ba)_{quqd}^{(1/8)} + (\widetilde{b\bar{a}})_{quqd}^{(1/8)})y_t^2 \quad (183)$$

from the SMEFT operators $O_{quqd}^{(1)}$ and $O_{quqd}^{(8)}$. For more details, see [2].

In table XI, we show the numerical results for the SMEFT-to-WET matching relations. It has been calculated with `wilson` [67]. In addition to the matching itself, we also take the SMEFT running from the top scale of our high-energy observables to the Z scale where we match into account, as well as the running of the WET operators [55, 157] between the matching scale and the B meson scale of the flavour observables. This is not shown here, but can be found in addition to more details on the matching and running in [2].

F. Observables

With all this in place, we can now have a detailed look at the observables we work with. For the top sector, we refer mostly to [154], but we will have a closer look at the flavour sector here.

The two flavour observables we take into account for our fit are the branching ratios of

the processes $B_s \rightarrow \mu^+ \mu^-$ and $B \rightarrow X_s \gamma$. In the WET, they are given by [60]

$$\begin{aligned} Br(B_s \rightarrow \mu^+ \mu^-) &= (3.57 - 1.71\mathcal{C}_{10} + 0.21\mathcal{C}_{10}^2) \times \left(1 \pm 1.2\% \Big|_{f_{B_s}} \pm 1.5\% \Big|_{\text{CKM}}\right) \times 10^{-9} \\ Br(B \rightarrow X_s \gamma) &= (3.26 - 15.17\mathcal{C}_7 - 0.77\mathcal{C}_8 + 1.66\mathcal{C}_7\mathcal{C}_8 + 1.36\mathcal{C}_7\mathcal{C}_{8'} \\ &\quad + 18.03(\mathcal{C}_7^2 + \mathcal{C}_{7'}^2) + 0.20\mathcal{C}_8^2 + 0.09\mathcal{C}_{8'}^2) \times (1 \pm 5\%) \times 10^{-4} \end{aligned} \quad (184)$$

where the constant factor contains the Standard Model contributions $\mathcal{C}_a^{\text{SM}}$ of each of the operators, so that the shown part \mathcal{C}_a is only the contribution coming through matching from the SMEFT, and thus from new physics. As we can see, both processes probe different operators, which also match to SMEFT differently, such that we can expect them to probe different directions in flavour space and to break blind directions in the SMEFT.

In addition to their potential to break blind directions in the top fit [154], we choose these two observables because they are precisely measured,

$$\begin{aligned} Br(B_s \rightarrow \mu^+ \mu^-)_{\text{exp}} &= (2.69 \pm 0.37) \times 10^{-9} \\ Br(B \rightarrow X_s \gamma)_{\text{exp}} &= (3.49 \pm 0.19) \times 10^{-4}, \end{aligned} \quad (185)$$

by several experiments, and also because they introduce few additional nuisance parameters to the fit. We use for $B_s \rightarrow \mu^+ \mu^-$ the combined value [174] of the measurements from ATLAS [175], CMS [176], and LHCb [177]. For $B \rightarrow X_s \gamma$, we use the world average [38, 178] built on the experimental measurements by BaBar [179–181], Belle [182, 183], and CLEO [184]. In addition to the experimental uncertainties above, we also consider theory uncertainties. For $B_s \rightarrow \mu^+ \mu^-$, we add 1.5% for the uncertainty of the theory elements and 1.2% for the uncertainty of the B_s decay constant. For $B \rightarrow X_s \gamma$, the theory uncertainty is taken from [185] as 5%. Additional details can be found in [2] and particularly the appendices A and B therein.

In the top sector, we take into account the same observables as the analysis [154]. These are:

- the production cross section of $t\bar{t}$ in 87 measurements,
- the charge asymmetry of $t\bar{t}$ in five measurements,
- the production cross sections of $t\bar{t}Z$ and $t\bar{t}W$ in four measurements,
- the production cross sections of a single t in the s - or t -channel in 13 measurements,
- the production cross sections of tZ and tW in seven measurements
- the helicity fractions of the W in t decay in eight measurements.

	$t\bar{t}$	single top	tW	tZ	$t\bar{t}Z$	$t\bar{t}W$	#
$C_{\phi q}^{(1)}$	—	—	—	$a_{\phi q}^{(1)}, A_{\phi q}^{(1)}$	$a_{\phi q}^{(1)}, A_{\phi q}^{(1)}$	—	2
$C_{\phi q}^{(3)}$	—	$a_{\phi q}^{(3)}, A_{\phi q}^{(3)}$	$A_{\phi q}^{(3)}$	$a_{\phi q}^{(3)}, A_{\phi q}^{(3)}$	$a_{\phi q}^{(3)}, A_{\phi q}^{(3)}$	$a_{\phi q}^{(3)}$	2
$C_{\phi u}$	—	—	—	$A_{\phi u}$	$a_{\phi u}, A_{\phi u}$	—	2
$C_{\phi d}$	—	—	—	—	$a_{\phi d}$	—	1
C_{uB}	—	—	—	$A_{uB} y_t$	$A_{uB} y_t$	—	1
C_{uW}	—	$A_{uW} y_t$	$A_{uW} y_t$	$A_{uW} y_t$	$A_{uW} y_t$	—	1
C_{uG}	$A_{uG} y_t$	—	$A_{uG} y_t$	—	$A_{uG} y_t$	$A_{uG} y_t$	1

TABLE XII: Contributions of the two-quark SMEFT operators to the top observables in terms of the MFV coefficients, as well as the total number of probeable degrees of freedom for each Wilson coefficient. Independently probed coefficients are separated by commas. Taken from [2].

As we saw in the last section for the matching of the SMEFT onto the WET coefficients, for calculating the cross sections of the top observables we need to take into account both single and double operator insertions, as well as taking into account interference terms between the Standard Model and SMEFT amplitudes.

Using the dependence of the observables on the considered SMEFT parameters, we can determine their dependence on the MFV parameters. In table XII, we show all two-quark operators and the combinations of MFV parameters they contribute to the different top observables. For the more involved contributions of the MFV parameters of the SMEFT four-quark operators, we refer to [2].

G. Fit and fit results

For the numerical analysis, we work with the statistical framework `sfitter` [186–188], just like the previous paper establishing the top quark dataset we use [154]. We follow the Rfit scheme [189] for our error handling, which treats theoretical and non-data-driven systematic uncertainties as a flat central core region and experimental statistical and data-driven systematic uncertainties as half-Gaussian tails attaching at the edges of this core region. When showing contours, we choose the contours that, in a completely Gaussian case, correspond to 68% (solid) and 95% (dashed) CL regions. We show blue contours for the fit result with only top observables, green contours for the case of top observables and the process $B_s \rightarrow \mu^+ \mu^-$, and a solid orange region (of two opacities for the two intervals) for the fit with all observables, that is, also including $B \rightarrow X_s \gamma$.

The first result we find is that, as expected, the top and bottom observables probe different spaces of parameters,

$$\text{bottom: } O_{\phi d}, O_{\phi ud}, O_{dB}, O_{dW}, O_{dG} \quad (186)$$

$$\begin{aligned}
\text{top: } & O_{\phi u}, O_{uB}, O_{uW}, O_{uG}, O_{uu}, O_{qu}^{(8)}, O_{qd}^{(1)}, O_{qd}^{(8)}, O_{ud}^{(1)}, O_{ud}^{(8)} \\
\text{top \& bottom: } & O_{\phi q}^{(1)}, O_{\phi q}^{(3)}, O_{qq}^{(1)}, O_{qq}^{(3)}, O_{qu}^{(1)}
\end{aligned}$$

though the operators $O_{quqd}^{(1)}$, $O_{quqd}^{(8)}$ and O_{dd} still remain unprobed through all of our observables. The operators in the first row are probed only in bottom observables and contain right-handed down-quark currents that induce tree-level effects in the bottom observables. The operators in the second row are probed dominantly by top quarks, with the operators with right-handed top-quarks also contributing to the bottom observables at loop-level. The third row contains those operators which are probed similarly well by both sets of observables. These operators all have a left-handed quark current, thus correlating their top and bottom quark couplings through these operators to each other. We will focus on these operators that are best constrained in this section.

In particular we test flavour universality in the sectors of left-handed quark current two-quark operators, $O_{\phi q}^{(1)}$ and $O_{\phi q}^{(3)}$, and four-quark operators, $O_{qq}^{(1)}$ and $O_{qq}^{(3)}$.

For the two-quark operators, the relevant degrees of freedom are

$$a_{\phi q}^{(-)}, b_{\phi q}^{(-)}, a_{\phi q}^{(3)}, b_{\phi q}^{(3)}. \quad (187)$$

The observables, expressed in these parameters, come out to be

$$\begin{aligned}
\sigma_t \text{ [pb]} &= 126 + 15.1 [a_{\phi q}^{(3)} + A_{\phi q}^{(3)}] + 0.5 [a_{\phi q}^{(3)} + A_{\phi q}^{(3)}]^2 \\
\sigma_{tW} \text{ [pb]} &= 75.3 + 9.1 A_{\phi q}^{(3)} + 0.27 (A_{\phi q}^{(3)})^2 \\
\sigma_{tZ} \text{ [pb]} &= 0.78 + 0.17 [a_{\phi q}^{(3)} + A_{\phi q}^{(3)}] + 0.01 a_{\phi q}^{(-)} + 0.10 (a_{\phi q}^{(3)})^2 + 0.02 (A_{\phi q}^{(3)})^2 \\
\sigma_{t\bar{t}Z} \text{ [pb]} &= 0.679 + 0.023 a_{\phi q}^{(3)} - 0.070 A_{\phi q}^{(-)} + 0.008 (a_{\phi q}^{(-)})^2 + 0.004 [2a_{\phi q}^{(3)} + a_{\phi q}^{(-)}]^2 \\
\sigma_{t\bar{t}W} \text{ [pb]} &= 0.446 + (0.054 + 0.008) a_{\phi q}^{(3)} + 0.062 (a_{\phi q}^{(3)})^2,
\end{aligned} \quad (188)$$

for the top observables and

$$\begin{aligned}
Br(B \rightarrow X_s \gamma) &= \left(3.26 + 0.36 a_{\phi q}^{(3)} - 0.76 b_{\phi q}^{(3)} \right) \times 10^{-4} \\
Br(B_s \rightarrow \mu^+ \mu^-) &= \left(3.57 - 41.0 \left(2b_{\phi q}^{(3)} + b_{\phi q}^{(-)} \right) + 117.8 \left(2b_{\phi q}^{(3)} + b_{\phi q}^{(-)} \right)^2 \right) \times 10^{-9}
\end{aligned} \quad (189)$$

for the bottom observables. We can see that while in the top observables, the flavour-violating MFV coefficients b only appear in the combination $A = a + by_t^2$, the bottom observables probe it directly, too. In figure 45, we show three two-dimensional representation of the four-dimensional fitting space.

In all three plots we can clearly see that the addition of the flavour observables has

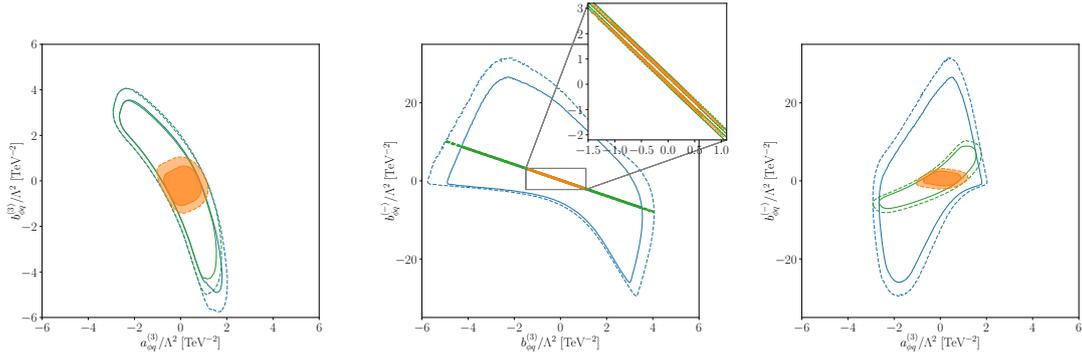


FIGURE 45: Two-dimensional projections of the four-dimensional likelihood distribution we gain from fitting the considered observables to the left-handed quark-current two-quark Wilson coefficients $C_{\phi q}^{(-)}(m_t) = C_{\phi q}^{(1)} - C_{\phi q}^{(3)}$ and $C_{\phi q}^{(3)}(m_t)$ in up-alignment MFV. On the left, we show the test of flavour universality in the weak triplet interaction, $a_{\phi q}^{(3)}$, $b_{\phi q}^{(3)}$. In the middle, we show the comparison between the flavour-breaking coefficient in the charged current $b_{\phi q}^{(3)}$ and the corresponding neutral current coefficient $b_{\phi q}^{(-)}$. On the right, we show the interplay of the flavour-diagonal charged current $a_{\phi q}^{(3)}$ and the neutral flavour-breaking up-quark current $b_{\phi q}^{(-)}$. The blue contour corresponds to the fit with only the top data taken into account, the green contours take into account the $B_s \rightarrow \mu^+ \mu^-$ as well, and the orange region shows the fit region with all observables included in the fit. Taken from [2].

helped resolve blind directions of the fit with only the top observables. In particular, we can see that the addition of flavour observables strongly confines our fit results in areas almost perpendicular to the least constrained top fit regions. For a more detailed discussion of the results and their implications, see [2]. Therein, we also take a closer look at fits in the four-quark operator sector.

Overall, we find that the addition of flavour observables to the top fit strongly constrains the remaining parameter region in fits of two Wilson coefficients. The fit results also imply that there are flavour breaking effects as defined in MFV. In the future, an inclusion of further flavour observables like neutral B meson mixing or semi-leptonic B decays would be very interesting to include, though the handling of their uncertainties and non-local effects is non-trivial.

V. CONCLUSIONS

In the course of this thesis, we have explored two approaches of searching for new physics, through long-lived particle searches at colliders and through EFT fits.

We have seen that searches for displaced decays can illuminate a larger amount of parameter space of many models than only searching for prompt decays. Specifically, we have made these analyses for long-lived axion-like particles and dark scalars. We have seen this to be the case particularly for e^+e^- colliders in the form of our analyses based on Belle II and the ILC. Our analysis [1] has caught the interest of researchers of the Belle II collaboration, who are now establishing a working group on long-lived particle searches. Hopefully in the near future we will see an increasing number of experimental analyses for long-lived particle searches, particularly at B factories and other e^+e^- colliders which we have found to have a great potential in exploring long-lived particle signatures.

We have also seen that far detectors, while clearly very promising for long-lived particle searches at the LHC, do not improve the potential for long-lived particle searches at e^+e^- colliders significantly. We have come to this conclusion specifically for the two detectors Belle II and ILC, which both also have very large solid angles which has a large impact on this result. Our study of several realistic, and one unrealistically large, far detector designs have shown improvements of only up to order one, unlike those expected at far detectors proposed for the LHC. This is due to the cleaner background of e^+e^- colliders in comparison to hadron colliders. While far detectors could still be an interesting addition to these detectors, for example through connecting them to the main detector's trigger and allowing for the capture of more elusive events that might show up as missing energy in the main detector but shed light on the missing particle in the far detector, this result has confirmed our earlier finding that Belle II is a detector with a great potential for detecting long-lived particles signatures.

Our third finding came from the comparison of invisible and displaced decays, that is, missing energy searches and searches for displaced vertices in charged particle tracks. We compared, in the same model, the search strategies for these two searches, including a dedicated background analysis for the missing energy search, and found that the reach of missing energy searches is very strong over a larger range of long-lived particle masses, but that displaced decay searches can probe the coupling down to similar or even lower values for high masses, and give us additional information about the long-lived particle. This analysis happened for the case of axion-like particles with specific UV couplings and at the Belle II detector. It would be very interesting to see further studies of the complementarity of displaced and missing energy searches in other models and at other

detectors.

In the EFT approach, we found that including flavour observables in a fit of top quark observables on the SMEFT parameter space allowed us to resolve blind directions that cannot easily be probed in the high-energy observables. We used this better resolution to explore the flavour structure of SMEFT operators that have strong contribution to both classes of observables. To do this, we imposed Minimal Flavour Violation on the operators and compared the matching to its different flavour structure coefficients. We found that the flavour-breaking coefficients of left-handed quarks are strongly constrained in fits of a pair of Wilson coefficients. This opens up many areas of future research, through the addition of more flavour observables and the exploration of bigger groups of Wilson coefficients, as well as through imposing and comparing different types of flavour structures. We conclude that the flavour structure of UV physics really can be explored through SMEFT fits of a wide range of observables, particularly including flavour observables.

In general, we have found these to be two interesting and fruitful approaches to explore the possibilities of new physics beyond the Standard Model of particle physics, and I am eager to see the future work that is done using both approaches.

ACKNOWLEDGEMENTS

I want to thank Susanne Westhoff for taking me on as first a Master and then a Ph.D. student and for her support throughout this time, not only in our projects and research, but also for her understanding, patience, and help throughout all of the changes that the pandemic has brought.

I want to thank my collaborators, Anastasiia Filimonova, Sebastian Bruggisser, Danny van Dyk, Torben Ferber, Finn Tillinger, and all the GAZELLE folks, for working with me on interesting projects and for finding a way to fit my changing timezones into our meeting schedules.

I want to acknowledge support from the DFG graduate school *Particle physics beyond the Standard Model* (GRK1940) for giving me the opportunity to complete this Ph.D. and become a part of the wider particle physics community through the conferences and workshops they allowed me to attend.

My thanks also go to Tilman Plehn, Stephanie Hansmann-Menzemer, and Werner Aeschbach for taking the time to be on my examination committee.

I want to thank Karla Tame Narvaez, for being a wonderful friend. Without you, my years in Heidelberg wouldn't have been the same, and even if we're going to different cities now, know that you, Ulises, and Luk will always be welcome wherever I am.

I want to thank my family, Mama and Judith. Thank you for always being there for me, even when we're across the world from each other, and for your warmth in welcoming me home every time. I'm very lucky to have such a wonderful mother and sister as you.

I also want give thanks to the rest of my family, to Victoria and Gustavo, Diana, Sebastian and Bongo, to Christoph, to my aunts, uncles and cousins, and to my grandparents, Elke and Heinz. Thank you for being there for me.

Lastly, I want to deeply thank my husband Carlos, for his incredible support of me through this thesis and through the years since I have met him, for the joy he brings me daily, and for the way that he lights up my life. I love you with all my heart. I look forward to whatever our next adventure may be and where it takes us.

-
- [1] A. Filimonova, R. Schäfer, and S. Westhoff, *Phys. Rev. D* **101**, 095006 (2020), arXiv:1911.03490 [hep-ph].
- [2] S. Bruggisser, R. Schäfer, D. van Dyk, and S. Westhoff, *JHEP* **05**, 257 (2021), arXiv:2101.07273 [hep-ph].
- [3] M. Borsato *et al.*, *Rept. Prog. Phys.* **85**, 024201 (2022), arXiv:2105.12668 [hep-ph].
- [4] S. Dreyer *et al.*, (2021), arXiv:2105.12962 [hep-ph].
- [5] T. Ferber, A. Filimonova, R. Schäfer, and S. Westhoff, (2022), arXiv:2201.06580 [hep-ph].
- [6] R. Schäfer, F. Tillinger, and S. Westhoff, (2022), arXiv:2202.11714 [hep-ph].
- [7] A. Aryshev *et al.* (ILC International Development Team), (2022), arXiv:2203.07622 [physics.acc-ph].
- [8] R. Schafer, “Searching for light dark sectors in rare meson decays,” Master Thesis, 2019.
- [9] S. Chatrchyan *et al.* (CMS), *Phys. Lett. B* **716**, 30 (2012), arXiv:1207.7235 [hep-ex].
- [10] G. Aad *et al.* (ATLAS), *Phys. Lett. B* **716**, 1 (2012), arXiv:1207.7214 [hep-ex].
- [11] S. P. Martin, *Adv. Ser. Direct. High Energy Phys.* **18**, 1 (1998), arXiv:hep-ph/9709356.
- [12] H. E. Haber and G. L. Kane, *Phys. Rept.* **117**, 75 (1985).
- [13] X. Tata, *Eur. Phys. J. ST* **229**, 3061 (2020), arXiv:2002.04429 [hep-ph].
- [14] B. Mukhopadhyaya, *Eur. Phys. J. ST* **229**, 3143 (2020).
- [15] D. Guadagnoli, *Symmetry* **13**, 1999 (2021).
- [16] J. Virto, *PoS BEAUTY2020*, 007 (2021), arXiv:2103.01106 [hep-ph].
- [17] A. Keshavarzi, K. S. Khaw, and T. Yoshioka, *Nucl. Phys. B* **975**, 115675 (2022), arXiv:2106.06723 [hep-ex].
- [18] A. Tewsley-Booth (Muon g-2), *Rev. Mex. Fis. Suppl.* **3**, 020717 (2022), arXiv:2202.11172 [hep-ex].
- [19] A. Strumia and F. Vissani, (2006), arXiv:hep-ph/0606054.
- [20] G. Bertone, D. Hooper, and J. Silk, *Phys. Rept.* **405**, 279 (2005), arXiv:hep-ph/0404175.
- [21] J. Alimena *et al.*, *J. Phys. G* **47**, 090501 (2020), arXiv:1903.04497 [hep-ex].
- [22] B. Shuve, “Theory Overview of Long-Lived Particles at the LHC,” (2017), https://indico.cern.ch/event/607314/contributions/2542308/attachments/1447888/2231430/LHC-LLP_Shuve.pdf.
- [23] H. Russell, <https://hrussell.web.cern.ch/hrussell/graphics.html>.
- [24] H. Russell, “An experimental introduction to long-lived particle searches at the LHC,” (2017), https://indico.cern.ch/event/607314/contributions/2542309/attachments/1447873/2231444/20170424_LLPs.pdf.
- [25] T. Cohen, *PoS TASI2018*, 011 (2019), arXiv:1903.03622 [hep-ph].
- [26] E. Fermi, *Ric. Sci.* **4**, 491 (1933).
- [27] E. Fermi, *Zeitschrift für Physik* **88**, 161 (1934).
- [28] L. Nanni, *Adv. Stud. Theor. Phys.* **13**, 281 (2019), arXiv:1803.07147 [physics.hist-ph].
- [29] W. Pauli, *Phys. Today* **31N9**, 27 (1978).
- [30] L. Hadjiivanov, (2018), arXiv:1812.11629 [physics.hist-ph].
- [31] M. E. Peskin, in *2016 European School of High-Energy Physics* (2017) pp. 1–70, arXiv:1708.09043 [hep-ph].

- [32] C.-N. Yang and J. Tiomno, Phys. Rev. **79**, 495 (1950).
- [33] E. R. Caianiello, Nuovo Cim. **8**, 535 (1951).
- [34] J. Tiomno and J. A. Wheeler, Rev. Mod. Phys. **21**, 144 (1949).
- [35] E. C. G. Sudarshan and R. e. Marshak, Phys. Rev. **109**, 1860 (1958).
- [36] R. P. Feynman and M. Gell-Mann, Phys. Rev. **109**, 193 (1958).
- [37] A. Lesov, *The Weak Force: From Fermi to Feynman*, Other thesis (2009), arXiv:0911.0058 [physics.hist-ph].
- [38] R. Workman *et al.* (Particle Data Group), To be published (2022).
- [39] T. D. Lee, M. Rosenbluth, and C.-N. Yang, Phys. Rev. **75**, 905 (1949).
- [40] S. L. Glashow, Nucl. Phys. **22**, 579 (1961).
- [41] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).
- [42] A. Salam, Conf. Proc. C **680519**, 367 (1968).
- [43] G. Arnison *et al.* (UA1), Phys. Lett. B **122**, 103 (1983).
- [44] G. Arnison *et al.* (UA1), Phys. Lett. B **134**, 469 (1984).
- [45] G. Arnison *et al.* (UA1), Phys. Lett. B **129**, 273 (1983).
- [46] M. Banner *et al.* (UA2), Phys. Lett. B **122**, 476 (1983).
- [47] A. G. Clark (UA2), in *Asia Pacific Physics Conference* (1983).
- [48] G. Sauvage and J. Schacher (UA2), in *International Europhysics Conference on High-Energy Physics* (1983) pp. 472–505.
- [49] J. L. Rosner, in *Cronin Fest: Celebration of James Cronin’s 75th Birthday* (2006) arXiv:hep-ph/0610100.
- [50] G. Rajasekaran, Resonance J. Sci. Educ. **19**, 18 (2014), arXiv:1403.3309 [physics.hist-ph].
- [51] M. D. Schwartz, *Quantum Field Theory and the Standard Model* (Cambridge University Press, 2014).
- [52] A. V. Manohar, (2018), 10.1093/oso/9780198855743.003.0002, arXiv:1804.05863 [hep-ph].
- [53] Y. Aoki *et al.*, (2021), arXiv:2111.09849 [hep-lat].
- [54] “Flavour Lattice Averaging Group,” <http://flag.unibe.ch/>.
- [55] E. E. Jenkins, A. V. Manohar, and P. Stoffer, JHEP **01**, 084 (2018), arXiv:1711.05270 [hep-ph].
- [56] E. E. Jenkins, A. V. Manohar, and P. Stoffer, JHEP **03**, 016 (2018), arXiv:1709.04486 [hep-ph].
- [57] “WCxf basis for the JMS basis,” <https://wxcf.github.io/assets/pdf/WET.JMS.pdf>.
- [58] J. Aebischer *et al.*, Comput. Phys. Commun. **232**, 71 (2018), arXiv:1712.05298 [hep-ph].
- [59] “WCxf basis for EOS,” <https://wxcf.github.io/assets/pdf/WET.EOS.pdf>.
- [60] D. van Dyk *et al.*, “EOS — A HEP program for flavour observables,” <https://eos.github.io/>.
- [61] W. Buchmuller and D. Wyler, Nucl. Phys. B **268**, 621 (1986).
- [62] I. Brivio and M. Trott, Phys. Rept. **793**, 1 (2019), arXiv:1706.08945 [hep-ph].
- [63] T. Cohen, N. Craig, X. Lu, and D. Sutherland, JHEP **03**, 237 (2021), arXiv:2008.08597 [hep-ph].
- [64] B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek, JHEP **10**, 085 (2010), arXiv:1008.4884 [hep-ph].
- [65] M. Neubert, (2019), 10.1093/oso/9780198855743.003.0001, arXiv:1901.06573 [hep-ph].

- [66] R. Alonso, E. E. Jenkins, A. V. Manohar, and M. Trott, *JHEP* **04**, 159 (2014), arXiv:1312.2014 [hep-ph].
- [67] J. Aebischer, J. Kumar, and D. M. Straub, *Eur. Phys. J. C* **78**, 1026 (2018), arXiv:1804.05033 [hep-ph].
- [68] We thank Sebastian Bruggisser and Lara Grabitz for sharing their numerical code for the calculation of the renormalisation group evolution of couplings in the ALP EFT with us.
- [69] D. O’Connell, M. J. Ramsey-Musolf, and M. B. Wise, *Phys. Rev. D* **75**, 037701 (2007), arXiv:hep-ph/0611014.
- [70] M. Bauer, P. Foldenauer, and J. Jaeckel, *JHEP* **07**, 094 (2018), arXiv:1803.05466 [hep-ph].
- [71] C. Blanco, M. Escudero, D. Hooper, and S. J. Witte, *JCAP* **11**, 024 (2019), arXiv:1907.05893 [hep-ph].
- [72] P. Langacker, *Rev. Mod. Phys.* **81**, 1199 (2009), arXiv:0801.1345 [hep-ph].
- [73] B. Holdom, *Phys. Lett. B* **166**, 196 (1986).
- [74] A. M. Abdullahi *et al.*, in *2022 Snowmass Summer Study* (2022) arXiv:2203.08039 [hep-ph].
- [75] P. Agrawal *et al.*, *Eur. Phys. J. C* **81**, 1015 (2021), arXiv:2102.12143 [hep-ph].
- [76] M. Bauer, M. Neubert, S. Renner, M. Schnubel, and A. Thamm, *JHEP* **04**, 063 (2021), arXiv:2012.12272 [hep-ph].
- [77] R. D. Peccei and H. R. Quinn, *Phys. Rev. Lett.* **38**, 1440 (1977).
- [78] S. Weinberg, *Phys. Rev. Lett.* **40**, 223 (1978).
- [79] F. Wilczek, *Phys. Rev. Lett.* **40**, 279 (1978).
- [80] H. Georgi, D. B. Kaplan, and L. Randall, *Phys. Lett. B* **169**, 73 (1986).
- [81] G. Punzi, eConf **C030908**, MODT002 (2003), arXiv:physics/0308063.
- [82] J. Conrad, *Astropart. Phys.* **62**, 165 (2015), arXiv:1407.6617 [astro-ph.CO].
- [83] W. Altmannshofer *et al.* (Belle-II), *PTEP* **2019**, 123C01 (2019), [Erratum: *PTEP* 2020, 029201 (2020)], arXiv:1808.10567 [hep-ex].
- [84] M. Pepe Altarelli and F. Teubert, *Int. J. Mod. Phys. A* **23**, 5117 (2008), arXiv:0802.1901 [hep-ph].
- [85] M. W. Winkler, *Phys. Rev. D* **99**, 015018 (2019), arXiv:1809.01876 [hep-ph].
- [86] N. Gubernari, A. Kokulu, and D. van Dyk, *JHEP* **01**, 150 (2019), arXiv:1811.00983 [hep-ph].
- [87] K. Schmidt-Hoberg, F. Staub, and M. W. Winkler, *Phys. Lett. B* **727**, 506 (2013), arXiv:1310.6752 [hep-ph].
- [88] R. Aaij *et al.* (LHCb), *Phys. Rev. Lett.* **115**, 161802 (2015), arXiv:1508.04094 [hep-ex].
- [89] R. Aaij *et al.* (LHCb), *Phys. Rev. D* **95**, 071101 (2017), arXiv:1612.07818 [hep-ex].
- [90] A. A. Alves, Jr. *et al.* (LHCb), *JINST* **3**, S08005 (2008).
- [91] A. Kachanovich, U. Nierste, and I. Nišandžić, *Eur. Phys. J. C* **80**, 669 (2020), arXiv:2003.01788 [hep-ph].
- [92] J. F. Kamenik and C. Smith, *JHEP* **03**, 090 (2012), arXiv:1111.6402 [hep-ph].
- [93] I. Boiarska, K. Bondarenko, A. Boyarsky, V. Gorkavenko, M. Ovchinnikov, and A. Sokolenko, *JHEP* **11**, 162 (2019), arXiv:1904.10447 [hep-ph].
- [94] J. A. Bailey *et al.*, *Phys. Rev. D* **93**, 025026 (2016), arXiv:1509.06235 [hep-lat].
- [95] J. P. Lees *et al.* (BaBar), *Phys. Rev. Lett.* **114**, 171801 (2015), arXiv:1502.02580 [hep-ex].

- [96] K. Bondarenko, A. Boyarsky, T. Bringmann, M. Hufnagel, K. Schmidt-Hoberg, and A. Sokolenko, *JHEP* **03**, 118 (2020), arXiv:1909.08632 [hep-ph].
- [97] G. Lanfranchi (NA62), PoS **EPS-HEP2017**, 301 (2017).
- [98] A. Ariga *et al.* (FASER), *Phys. Rev. D* **99**, 095011 (2019), arXiv:1811.12522 [hep-ph].
- [99] J. L. Feng, I. Galon, F. Kling, and S. Trojanowski, *Phys. Rev. D* **97**, 055034 (2018), arXiv:1710.09387 [hep-ph].
- [100] S. Alekhin *et al.*, *Rept. Prog. Phys.* **79**, 124201 (2016), arXiv:1504.04855 [hep-ph].
- [101] V. V. Gligorov, S. Knapen, M. Papucci, and D. J. Robinson, *Phys. Rev. D* **97**, 015023 (2018), arXiv:1708.09395 [hep-ph].
- [102] J. A. Evans, *Phys. Rev. D* **97**, 055046 (2018), arXiv:1708.08503 [hep-ph].
- [103] F. Abudinén *et al.* (Belle-II), (2022), arXiv:2206.15227 [hep-ex].
- [104] F. Abudinén *et al.* (Belle-II), (2022), arXiv:2206.07453 [hep-ex].
- [105] F. Abudinén *et al.* (Belle-II), (2022), arXiv:2207.00509 [hep-ex].
- [106] G. Aielli *et al.*, *Eur. Phys. J. C* **80**, 1177 (2020), arXiv:1911.00481 [hep-ex].
- [107] M. Drewes, *Int. J. Mod. Phys. E* **22**, 1330019 (2013), arXiv:1303.6912 [hep-ph].
- [108] K. Bondarenko, A. Boyarsky, D. Gorbunov, and O. Ruchayskiy, *JHEP* **11**, 032 (2018), arXiv:1805.08567 [hep-ph].
- [109] A. Ringwald, in *49th Rencontres de Moriond on Electroweak Interactions and Unified Theories* (2014) pp. 223–230, arXiv:1407.0546 [hep-ph].
- [110] D. J. E. Marsh, in *13th Patras Workshop on Axions, WIMPs and WISPs* (2018) pp. 59–74, arXiv:1712.03018 [hep-ph].
- [111] M. Duerr, T. Ferber, C. Garcia-Cely, C. Hearty, and K. Schmidt-Hoberg, *JHEP* **04**, 146 (2021), arXiv:2012.08595 [hep-ph].
- [112] M. B. Gavela, R. Houtz, P. Quilez, R. Del Rey, and O. Sumensari, *Eur. Phys. J. C* **79**, 369 (2019), arXiv:1901.02031 [hep-ph].
- [113] M. Bauer, M. Neubert, and A. Thamm, *JHEP* **12**, 044 (2017), arXiv:1708.00443 [hep-ph].
- [114] D. J. Lange, *Nucl. Instrum. Meth. A* **462**, 152 (2001).
- [115] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *JHEP* **07**, 079 (2014), arXiv:1405.0301 [hep-ph].
- [116] E. Conte, B. Fuks, and G. Serret, *Comput. Phys. Commun.* **184**, 222 (2013), arXiv:1206.1599 [hep-ph].
- [117] E. Conte, B. Dumont, B. Fuks, and C. Wymant, *Eur. Phys. J. C* **74**, 3103 (2014), arXiv:1405.3982 [hep-ph].
- [118] B. Dumont, B. Fuks, S. Kraml, S. Bein, G. Chalons, E. Conte, S. Kulkarni, D. Sengupta, and C. Wymant, *Eur. Phys. J. C* **75**, 56 (2015), arXiv:1407.3278 [hep-ph].
- [119] E. Conte and B. Fuks, *Int. J. Mod. Phys. A* **33**, 1830027 (2018), arXiv:1808.00480 [hep-ph].
- [120] J. Y. Araz, M. Frank, and B. Fuks, *Eur. Phys. J. C* **80**, 531 (2020), arXiv:1910.11418 [hep-ph].
- [121] J. Y. Araz, B. Fuks, and G. Polykratis, *Eur. Phys. J. C* **81**, 329 (2021), arXiv:2006.09387 [hep-ph].
- [122] J. Y. Araz, B. Fuks, M. D. Goodsell, and M. Utsch, *Eur. Phys. J. C* **82**, 597 (2022), arXiv:2112.05163 [hep-ph].
- [123] G. Alguero, J. Y. Araz, B. Fuks, and S. Kraml, (2022), arXiv:2206.14870 [hep-ph].

- [124] W. R. Inc., “Mathematica, Version 13.1,” Champaign, IL, 2022.
- [125] J. P. Lees *et al.* (BaBar), Phys. Rev. D **87**, 112005 (2013), arXiv:1303.7465 [hep-ex].
- [126] K. Fujii *et al.*, (2017), arXiv:1710.07621 [hep-ex].
- [127] H. An, B. Echenard, M. Pospelov, and Y. Zhang, Phys. Rev. Lett. **116**, 151801 (2016), arXiv:1510.05020 [hep-ph].
- [128] G. Van Rossum and F. L. Drake, *Python 3 Reference Manual* (CreateSpace, Scotts Valley, CA, 2009).
- [129] T. Kluyver, B. Ragan-Kelley, F. Pérez, B. Granger, M. Bussonnier, J. Frederic, K. Kelley, J. Hamrick, J. Grout, S. Corlay, P. Ivanov, D. Avila, S. Abdalla, C. Willing, and J. development team, in *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, edited by F. Loizides and B. Schmidt (IOS Press, 2016) pp. 87–90.
- [130] R. Frederix, S. Frixione, V. Hirschi, D. Pagani, H. S. Shao, and M. Zaro, JHEP **07**, 185 (2018), [Erratum: JHEP 11, 085 (2021)], arXiv:1804.10017 [hep-ph].
- [131] K. Choi, S. H. Im, C. B. Park, and S. Yun, JHEP **11**, 070 (2017), arXiv:1708.00021 [hep-ph].
- [132] M. Chala, G. Guedes, M. Ramos, and J. Santiago, Eur. Phys. J. C **81**, 181 (2021), arXiv:2012.09017 [hep-ph].
- [133] J. Martin Camalich, M. Pospelov, P. N. H. Vuong, R. Ziegler, and J. Zupan, Phys. Rev. D **102**, 015023 (2020), arXiv:2002.04623 [hep-ph].
- [134] T. Kuhr, C. Pulvermacher, M. Ritter, T. Hauth, and N. Braun (Belle-II Framework Software Group), Comput. Softw. Big Sci. **3**, 1 (2019), arXiv:1809.04299 [physics.comp-ph].
- [135] “Belle II Analysis Software Framework (basf2): release-06-00-00,” (2021).
- [136] S. Jadach, B. F. L. Ward, and Z. Was, Comput. Phys. Commun. **130**, 260 (2000), arXiv:hep-ph/9912214.
- [137] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, Comput. Phys. Commun. **191**, 159 (2015), arXiv:1410.3012 [hep-ph].
- [138] S. Jadach, J. H. Kuhn, and Z. Was, Comput. Phys. Commun. **64**, 275 (1990).
- [139] V. Bertacchi *et al.* (Belle II Tracking Group), Comput. Phys. Commun. **259**, 107610 (2021), arXiv:2003.12466 [physics.ins-det].
- [140] “Measurement of the tracking efficiency and fake rate with $e^+e^- \rightarrow \tau^+\tau^-$ events,” (2020), BELLE2-NOTE-PL-2020-014.
- [141] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Í. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, Nature Methods **17**, 261 (2020).
- [142] E. Cortina Gil *et al.* (NA62), JHEP **03**, 058 (2021), arXiv:2011.11329 [hep-ex].
- [143] E. Cortina Gil *et al.* (NA62), JHEP **06**, 093 (2021), arXiv:2103.15389 [hep-ex].
- [144] J. P. Lees *et al.* (BaBar), Phys. Rev. Lett. **128**, 131802 (2022), arXiv:2111.01800 [hep-ex].
- [145] B. Döbrich, J. Jaeckel, and T. Spadaro, JHEP **05**, 213 (2019), [Erratum: JHEP 10, 046

- (2020)], arXiv:1904.02091 [hep-ph].
- [146] B. Döbrich, F. Ertas, F. Kahlhoefer, and T. Spadaro, Phys. Lett. B **790**, 537 (2019), arXiv:1810.11336 [hep-ph].
- [147] R. Aaij *et al.* (LHCb), JHEP **10**, 156 (2020), arXiv:2007.03923 [hep-ex].
- [148] F. Bergsma *et al.* (CHARM), Phys. Lett. B **157**, 458 (1985).
- [149] J. R. Batley *et al.* (NA48/2), Phys. Lett. B **677**, 246 (2009), arXiv:0903.3130 [hep-ex].
- [150] J. R. Batley *et al.* (NA48/2), Phys. Lett. B **697**, 107 (2011), arXiv:1011.4817 [hep-ex].
- [151] A. V. Artamonov *et al.* (E949), Phys. Lett. B **623**, 192 (2005), arXiv:hep-ex/0505069.
- [152] F. Abudinén *et al.* (Belle-II), Phys. Rev. Lett. **125**, 161806 (2020), arXiv:2007.13071 [hep-ex].
- [153] J. Grygier *et al.* (Belle), Phys. Rev. D **96**, 091101 (2017), [Addendum: Phys.Rev.D 97, 099902 (2018)], arXiv:1702.03224 [hep-ex].
- [154] I. Brivio, S. Bruggisser, F. Maltoni, R. Moutafis, T. Plehn, E. Vryonidou, S. Westhoff, and C. Zhang, JHEP **02**, 131 (2020), arXiv:1910.03606 [hep-ph].
- [155] R. Alonso, B. Grinstein, and J. Martin Camalich, Phys. Rev. Lett. **113**, 241802 (2014), arXiv:1407.7044 [hep-ph].
- [156] O. Catà and M. Jung, Phys. Rev. D **92**, 055018 (2015), arXiv:1505.05804 [hep-ph].
- [157] J. Aebischer, M. Fael, C. Greub, and J. Virto, JHEP **09**, 158 (2017), arXiv:1704.06639 [hep-ph].
- [158] M. Beneke, T. Feldmann, and D. Seidel, Nucl. Phys. B **612**, 25 (2001), arXiv:hep-ph/0106067.
- [159] J. Lyon and R. Zwicky, (2014), arXiv:1406.0566 [hep-ph].
- [160] M. Ciuchini, M. Fedele, E. Franco, S. Mishima, A. Paul, L. Silvestrini, and M. Valli, JHEP **06**, 116 (2016), arXiv:1512.07157 [hep-ph].
- [161] B. Capdevila, S. Descotes-Genon, L. Hofer, and J. Matias, JHEP **04**, 016 (2017), arXiv:1701.08672 [hep-ph].
- [162] S. Jäger, M. Kirk, A. Lenz, and K. Leslie, Phys. Rev. D **97**, 015021 (2018), arXiv:1701.09183 [hep-ph].
- [163] C. Bobeth, M. Chruszcz, D. van Dyk, and J. Virto, Eur. Phys. J. C **78**, 451 (2018), arXiv:1707.07305 [hep-ph].
- [164] A. Arbey, T. Hurth, F. Mahmoudi, and S. Neshatpour, Phys. Rev. D **98**, 095027 (2018), arXiv:1806.02791 [hep-ph].
- [165] N. Gubernari, D. van Dyk, and J. Virto, JHEP **02**, 088 (2021), arXiv:2011.09813 [hep-ph].
- [166] A. J. Buras, P. Gambino, M. Gorbahn, S. Jager, and L. Silvestrini, Phys. Lett. B **500**, 161 (2001), arXiv:hep-ph/0007085.
- [167] G. D'Ambrosio, G. F. Giudice, G. Isidori, and A. Strumia, Nucl. Phys. B **645**, 155 (2002), arXiv:hep-ph/0207036.
- [168] T. Feldmann and T. Mannel, Phys. Rev. Lett. **100**, 171601 (2008), arXiv:0801.1802 [hep-ph].
- [169] A. L. Kagan, G. Perez, T. Volansky, and J. Zupan, Phys. Rev. D **80**, 076002 (2009), arXiv:0903.1794 [hep-ph].
- [170] J. Aebischer, A. Crivellin, M. Fael, and C. Greub, JHEP **05**, 037 (2016), arXiv:1512.02830 [hep-ph].

- [171] W. Dekens and P. Stoffer, *JHEP* **10**, 197 (2019), arXiv:1908.05295 [hep-ph].
- [172] T. Hurth, S. Renner, and W. Shepherd, *JHEP* **06**, 029 (2019), arXiv:1903.00500 [hep-ph].
- [173] S. Descotes-Genon, A. Falkowski, M. Fedele, M. González-Alonso, and J. Virto, *JHEP* **05**, 172 (2019), arXiv:1812.08163 [hep-ph].
- [174] (2020).
- [175] M. Aaboud *et al.* (ATLAS), *JHEP* **04**, 098 (2019), arXiv:1812.03017 [hep-ex].
- [176] A. M. Sirunyan *et al.* (CMS), *JHEP* **04**, 188 (2020), arXiv:1910.12127 [hep-ex].
- [177] R. Aaij *et al.* (LHCb), *Phys. Rev. Lett.* **118**, 191801 (2017), arXiv:1703.05747 [hep-ex].
- [178] Y. S. Amhis *et al.* (HFLAV), *Eur. Phys. J. C* **81**, 226 (2021), arXiv:1909.12524 [hep-ex].
- [179] B. Aubert *et al.* (BaBar), *Phys. Rev. D* **77**, 051103 (2008), arXiv:0711.4889 [hep-ex].
- [180] J. P. Lees *et al.* (BaBar), *Phys. Rev. Lett.* **109**, 191801 (2012), arXiv:1207.2690 [hep-ex].
- [181] J. P. Lees *et al.* (BaBar), *Phys. Rev. D* **86**, 052012 (2012), arXiv:1207.2520 [hep-ex].
- [182] A. Limosani *et al.* (Belle), *Phys. Rev. Lett.* **103**, 241801 (2009), arXiv:0907.1384 [hep-ex].
- [183] T. Saito *et al.* (Belle), *Phys. Rev. D* **91**, 052004 (2015), arXiv:1411.7198 [hep-ex].
- [184] S. Chen *et al.* (CLEO), *Phys. Rev. Lett.* **87**, 251807 (2001), arXiv:hep-ex/0108032.
- [185] M. Misiak, A. Rehman, and M. Steinhauser, *JHEP* **06**, 175 (2020), arXiv:2002.01548 [hep-ph].
- [186] R. Lafaye, T. Plehn, and D. Zerwas, (2004), arXiv:hep-ph/0404282.
- [187] R. Lafaye, T. Plehn, M. Rauch, and D. Zerwas, *Eur. Phys. J. C* **54**, 617 (2008), arXiv:0709.3985 [hep-ph].
- [188] R. Lafaye, T. Plehn, M. Rauch, D. Zerwas, and M. Duhrssen, *JHEP* **08**, 009 (2009), arXiv:0904.3866 [hep-ph].
- [189] A. Hocker, H. Lacker, S. Laplace, and F. Le Diberder, *Eur. Phys. J. C* **21**, 225 (2001), arXiv:hep-ph/0104062.