UNIVERSITY OF WARSAW

Long-lived particle searches with the ILD experiment

Jan Franciszek Klamka

Doctoral thesis written under the supervision of
prof. dr hab. Aleksander Filip Żarnecki

October 2025

Kochanej babci Jasi

Abstract

This thesis presents the prospects for detecting neutral and charged long-lived particles (LLPs) with the International Large Detector (ILD) assuming its operation at the future e^+e^- International Linear Collider (ILC) with $\sqrt{s}$ = 250 GeV. The main tracker of the ILD is a large time projection chamber (TPC), which allows for almost continuous tracking. The study was performed using full `Geant4` simulation for a set of benchmark scenarios. Multiple background sources were taken into account, both from low-p_T beam-induced processes, as well as Standard Model interactions with high-p_T final states. The analysis incorporated an experiment-orientated approach, in which instead of studying particular points in the parameter space of a specific model, benchmarks were selected based on the experimental signature and kinematic properties of the final state, allowing to test the detector capabilities and reconstruction tools. The study was therefore carried out in a model-independent way, and the results are presented as a function of physical parameters, such as LLP decay lengths or masses.

For neutral LLPs, two classes of benchmarks were selected. The first one involved a little boosted, low-p_T final state, resulting from a small mass difference between a heavy LLP and a dark matter particle created in the LLP decay. The second introduced a very light LLP which, due to its large boost, produced high-p_T, strongly collimated final-state tracks. The signature of a displaced vertex was considered, and the analysis relied on a vertex-finding procedure designed for this work. It was also extended to a study of exotic Higgs decays, where the model-independent approach was tested, and the search was optimised for this specific BSM scenario.

For charged LLPs, the analysis was designed based on the experience of the search for neutral LLPs. The benchmarks involved pair-production of charged LLPs which could decay into a lepton and a dark matter particle, resulting in the signature of a kinked track. Two LLP masses (high and moderate) were considered, with a set of mass differences between LLP and dark matter selected for each of them. The search was performed using a kinked-track finding tool that was optimised for this study.

The results indicate that LLPs can be detected with the ILD for a wide range of lifetimes thanks to the large acceptance of the TPC. The study has shown that for the ILC running at 250 GeV signal production cross sections at the level of 0.1 fb can be probed for most of the scenarios considered, and the Higgs boson branching ratios to LLPs down to 10^{-4} can be reached. This demonstrates the potential of the ILD experiment in detecting the LLPs.

Streszczenie

Niniejsza praca przedstawia perspektywy detekcji neutralnych oraz naładowanych cząstek długożyciowych (ang. *long-lived particles*, LLP) przy użyciu detektora International Large Detector (ILD), zakładając jego działanie przy przyszłym zderzaczu e^+e^- International Linear Collider (ILC) przy energii $\sqrt{s}$ = 250 GeV. Głównym detektorem śladowym ILD jest duża komora projekcji czasowej (ang. *time projection chamber*, TPC) umożliwiająca praktycznie ciągły pomiar toru cząstki. Badania przeprowadzono z wykorzystaniem pełnej symulacji opartej o program `Geant4` dla zestawu scenariuszy referencyjnych. Uwzglednione zostały różne źródła tła, zarówno od oddziaływań wiązki produkujących cząstki o małym p_T, jak i od procesów Modelu Standardowego ze stanami końcowymi o wysokim p_T. W analizie zastosowano podejście eksperymentalne: zamiast badać konkretne punkty w przestrzeni parametrów wybranego modelu teoretycznego, scenariusze dobrano na podstawie sygnatury eksperymentalnej i właściwości kinematycznych stanu końcowego, umożliwiając przetestowanie możliwości detektora oraz narzędzi do rekonstrukcji. Studia przeprowadzono więc w sposób niezależny od modelu, prezentując wyniki w funkcji parametrów fizycznych, takich jak czasy życia cząstek LLP oraz ich masy.

Dla neutralnych LLP wybrano dwie klasy scenariuszy. Pierwsza opisywała słabo pchnięte stany końcowe o niskim p_T, wynikającym z niewielkiej różnicy mas pomiędzy ciężką LLP, a cząstką ciemnej materii powstającą w rozpadzie LLP. Druga obejmowała bardzo lekkie LLP, które ze względu na duże pchnięcie rozpadają się na silnie skolimowane cząstki o dużym p_T. Rozważaną sygnaturą był przesunięty wierzchołek, a analiza opierała się na algorytmie do wyszukiwania wierzchołków, stworzonym na potrzeby tej pracy. Dodatkowo, wyniki rozszerzono o analizę egzotycznych rozpadów bozonu Higgsa, w której przetestowano podejście niezależne od modelu oraz zoptymalizowano procedurę pod kątem konkretnego scenariusza nowej fizyki.

Dla naładowanych LLP analizę zaprojektowano korzystając z doświadczeń zdobytych przy poszukiwaniu cząstek neutralnych. Zakładano produkcję par naładowanych LLP, które mogły rozpadać się na lepton i cząstkę ciemnej materii, co prowadziło do charakterystycznej sygnatury „załamanego" śladu. Rozważano dwie wartości mas LLP (średnią i dużą), dla każdej z nich wybierając kilka wartości różnic mas między LLP a ciemną materią. Poszukiwania prowadzono z użyciem narzędzia do rekonstrukcji załamanych torów, zoptymalizowanego dla tej analizy.

Uzyskane wyniki pokazują, że ILD umożliwia detekcję LLP w szerokim zakresie czasów życia cząstek, głównie dzięki dużej akceptancji TPC. Wykazano, że przy zderzaczu ILC pracującym przy energii 250 GeV możliwe będzie badanie sygnałów o przekroju czynnym na poziomie 0.1 fb dla większości analizowanych scenariuszy, a także osiągnięcie czułości na stosunki rozgałęzień bozonu Higgsa do LLP rzędu 10^{-4}. Wyniki te podkreślają duży potencjał eksperymentu ILD dla poszukiwań cząstek długożyciowych.

Acknowledgements

The following thesis summarises approximately four years of my research at the University of Warsaw on long-lived particle searches within the ILD Concept Group. I have been involved in sensitivity studies for future linear e^+e^- colliders since 2018, when I started my cooperation with prof. Aleksander Filip Żarnecki on the production of dark inert scalars at CLIC for my MSc thesis. I want to stress here that I do believe that a linear Higgs factory is the most exciting, versatile, and promising option for the future of particle physics, and I would love to see it in operation one day.

I would like to thank several people without whom the completion of this thesis would never happen. First of all, I must thank my supervisor, prof. Aleksander Filip Żarnecki, for his constant guidance and help at every stage of my work (even with the smallest and seemingly irrelevant matters), which has continued since my master's studies. Such support is rare, and I am sincerely grateful for it.

I want to thank my wonderful wife Monika for her constant support and patience, which is not an easy task, especially when your fiance is travelling to conferences two weeks before your wedding. I am also grateful to the group of my closest friends and my family, in particular my parents and my brother Filip, as they will always support me no matter what I do or decide.

I would never have ended up here if not for the people I met along the way. Being part of the RL Study Group was a pleasure and an honour, with all the things we have been through ever since the first year of our studies, and all the crazy and inspiring discussions we had. Special thanks also go to Krzysztof for being the best travelling companion I could imagine, as well as to Mateusz and Gosia for being the best officemates I could ask for; you three were often the voice of reason I needed.

I also want to thank all the people in ILD who provided their feedback and comments throughout my work. I am particularly grateful to Daniel Jeans and Jurina Nakajima for their warm welcome at KEK and all the helpful discussions we had. Last but not least, I am very happy to have had the opportunity to be part of a great group of people who contributed to the Early Career Researcher White Paper for the 2026 European Strategy — I am truly proud of our work, whatever the outcome.

The work presented in this thesis was partially supported by the National Science Centre (Poland) under OPUS research project no. 2021/43/B/ST2/01778. This work has also benefited from computing services provided by the ILC Virtual Organization, supported by the national resource providers of the EGI Federation and the Open Science GRID.

Contents

Chapter 1

Introduction

This thesis builds on top of a century of research carried out to understand fundamental constituents of matter in the Universe and interactions between them. Following the discoveries of quantum mechanics and special relativity more than 100 years ago, decades of development of Quantum Field Theory led to the successful description of electroweak interactions, which eventually became the crucial part of the Standard Model (SM) of particle physics. However, this would not be possible without constant interplay and synergy between theoretical and experimental studies. Experiments, initially driving progress in theory by making observations of phenomena requiring explanation, at some point became an indispensable tool to confirm predictions of rapidly evolving models and frameworks.

All these achievements culminated with the 2012 discovery of the Higgs boson in the Large Hadron Collider (LHC) at CERN by the ATLAS and CMS experiments [1, 2]. The Higgs was the last piece of the Standard Model remaining to be confirmed; however, this theory is far from being the final description of the Universe, as it is not able to explain several observations in cosmology and neutrino physics. With the lack of any new measurements in particle physics that would explicitly violate the SM, all eyes are now directed at the Higgs boson. It is the most mysterious and least studied particle of all, that is also linked to parts of the SM where most of the open questions arise.

The LHC was built to achieve as high energies as possible with the aim to directly observe not only the Higgs, but also new, very massive particles that were expected at the TeV scale. This approach was adapted in the past to search for and measure heavy states predicted by the SM, as most of the masses in this theory are free parameters. However, now the situation is slightly different – there is no single model so compelling that it could be used to guide the experiment in a direction where new physics should be looked for and what new particles should we expect at certain scales. Therefore, the paradigm is slowly shifting – instead of chasing the energy frontier, a new convincing path to pinpoint physics Beyond the Standard Model (BSM) now seems to be the pursuit for an unprecedented precision and treatment of the SM as an effective theory.

A natural consequence of these conclusions is the need for a machine that would allow for very precise measurements of the Higgs boson properties – the so-called Higgs factory – and the most promising option is an electron-positron (e^+e^-) collider with the $\mathcal{O}(250\,\mathrm{GeV})$ centre-of-mass energy. Because, as already mentioned, there is no clear indication as to where we should search for new physics, the proposed accelerator concepts involve the possibilities for more than just the Higgs measurements. This includes the top quark studies and electroweak precision measurements – also because they are needed to fully complement and interpret the Higgs properties. Hence, such machines are also referred to as Higgs/electroweak/top factories, although in this thesis they will be simply called the "Higgs factories" for short.

However, these machines will not only allow for precision studies – Higgs factories are characterised by features different from a hadron collider, such as a very clean experimental environment, much lower backgrounds, and access to full information about an event kinematics. Therefore, these accelerators can also be used to complement direct LHC searches for new particles, although in many cases offering a lower mass reach.

Several concepts of future e^+e^- colliders exist, and some of them compete for the same site. The most mature projects are the International Linear Collider (ILC) [3] in Japan, Compact Linear Collider (CLIC) [4, 5] at CERN, Future Circular Collider (FCC-ee) [6] at CERN and Circular Electron-Positron Collider (CEPC) [7] in China. Recently, also a proposal for the Linear Collider Facility (LCF) [8] at CERN has emerged, strongly building on the experiences of the ILC. Since the scientific world cannot afford all of these projects at once, complex decision-making process is needed prior to the realisation of any of these proposals. In particular, every six to seven years, CERN initiates the Update of the European Strategy for Particle Physics (ESPPU) to determine the future of the field in Europe. It should be noted that the previous ESPPU 2020 confirmed that "an e^+e^- Higgs factory is the highest-priority next collider" [9] following the planned upgrade of the LHC to high luminosity (HL-LHC), and the next update, scheduled for early 2026, is expected to point to the preferred option for such a machine. The Physics Briefing Book, aiming to summarise the discussions around the proposed ideas and projects, has just been published at the time of writing these words [10].

During the ESPPU process, different aspects of future projects are scrutinised, including technical feasibility, cost, community involvement, environmental impact, etc. However, one of the most important factors is the physics potential of the proposed machines, which needs to be assessed thorough sensitivity studies. This must be accompanied by the R&D and evaluation of the potential of the detector concepts proposed for operation at future colliders.

In 2021, based on the outcome of 2020 ESPPU, the European Committee for Future Accelerators (ECFA) initiated a series of workshops on physics studies, experiment design, and detector technologies towards a future Higgs factory "to bring together the efforts of various e^+e^- projects". As explained in more detail later, the work presented in this thesis fits perfectly into this task and was presented at each of the ECFA workshops, as well as included in the final report summarising all the research conducted within the ECFA study [11].

The physics potential studies also include the sensitivity in direct new physics searches, as it is an important part of the programme for the future Higgs factory. Of the many models proposed as extensions of the SM to solve its problems, a particular category of new particles received a lot of attention in recent years. In the beginning of the LHC running, most of the signatures considered involved the assumption that the heavy hypothetical states decay promptly, and one can only measure their decay products, which should originate from the region of the primary interaction vertex. However, many models allow and predict states that are characterised by much higher, macroscopic decay lengths, comparable with the size of detectors (of the order of millimetres or more). In light of the absence of BSM signals at the LHC, such scenarios of long-lived particle (LLP) production and the associated unconventional signatures started to gain more interest, as this could explain the lack of observation of any new phenomena.

Following significant efforts to improve experimental methods in reconstruction and data analysis, it was possible to achieve impressive sensitivity to LLPs at the LHC despite many challenges, such as high backgrounds and the need for operation with triggers. However, due to the properties of hadron colliders many limitations persist, whereas an e^+e^- machine could complement the LHC reach thanks to the clean experimental environment, which seems perfectly suited for such exotic searches.

In particular, the International Large Detector (ILD) [12], a detector concept for a future Higgs factory equipped with a time projection chamber (TPC) as its main tracking device, offers great prospects for LLP studies. The TPC allows for almost continuous tracking, which provides the possibility to directly observe in-flight decays within the detector volume. Its other feature is an option to detect new heavy states by precisely measuring the amount of ionisation caused by a passing charged particle. It should also be noted that LLPs provide a very good case study for future experiments by challenging their design and software infrastructure with non-standard signatures, identifying the areas for potential improvements.

This thesis describes the study of the potential of the ILD experiment to detect neutral and charged LLPs. Because the baseline design of the ILD is optimised for operation at the ILC, this accelerator is considered a "reference collider" in this work. ILC is also the most mature Higgs factory concept, with a wide availability of software tools and event samples. Since there is no clear evidence suggesting that any of the proposed new physics models predicting the existence of LLPs is realised by Nature, a model-independent, experiment-focused approach was taken. Instead of analysing particular BSM models, benchmarks that provide the most interesting and kinematically challenging scenarios were selected for the study. Consequently, the results are also presented in the space of physical quantities (not the model parameters), such as cross sections, decay lengths, or masses. In this way, the capabilities of the ILD can be investigated with as few assumptions as possible. The analysis was performed for neutral and charged LLPs using the signatures of displaced vertex and kinked track, respectively. For the neutral LLP production, the analysis was also extended to a particular case of the Higgs boson decaying to dark long-lived scalars, a scenario often considered in the literature.

The thesis is organised as follows. Chapter 2 elaborates on the need for BSM physics and provides a brief introduction to the LLP phenomenology, typical LLP signatures, and the current status of LLP searches. Chapter 3 describes the most advanced proposals of future e^+e^- Higgs factories, and Chapter 4 introduces the concept of the ILD experiment, its design, and the software ecosystem used by the ILD Concept Group and in this work. Chapter 5 provides details on the selection of benchmark scenarios, the generation of event samples, and the methods (and their performance) used in this study for reconstruction of displaced vertices and kinked tracks; it also provides some discussion about the impact of detector design on the neutral LLP reconstruction. Chapter 6 is the main part of the dissertation, where the event selection is described, and the expected exclusion limits on the signal production cross section are presented, for all scenarios considered. For reference in the main text, Appendix A introduces the basic concepts of the SM while Appendix B presents example event displays to help grasp intuition about signal signatures and related challenges, as well as about some of the background sources.

The content of Chapters 5 and 6, as well as the classification of the neutral LLP signatures presented in Section 2.3.1, correspond to the original work carried out by the author. Section 6.2 describes the model-independent study of the neutral LLP production and is almost entirely based on Ref. [13]. This analysis, as well as its extension to exotic Higgs decays discussed in Section 6.3, have been carried out as a part of the ECFA study on Higgs/electroweak/top factories, and were included in its final report [11]. The results for an all-silicon ILD design presented in Section 5.4.1 were previously included in Ref. [14, 15, 16]. The results presented in this thesis have been shown at each of the ECFA workshops, as well as at various other international conferences and workshops, including the International Conference on High Energy Physics (ICHEP), or multiple editions of the European Physical Society Conference on High Energy Physics (EPS-HEP) and the International Workshop on Future Linear Colliders (LCWS).

Chapter 2

Physics Beyond the Standard Model

2.1 The need for Beyond the Standard Model physics

Although the SM is a very comprehensive, profound and well tested theory (for a short introduction of the model and the notation used here, see Appendix A), it is not only unable to explain several phenomena observed in experiments, but also contains some inner problems. The main issue within the theory itself is the lack of mechanism that could generate masses of neutrinos, since in Equation (A.17) the Higgs doublet after the spontanous symmetry breaking takes only the lower component out of $\overline{L}_j$. On the other hand, ϕ_c cannot be used in the same way as for upper quarks, because it would require using a right-handed neutrino field, and those have never been observed. Meanwhile, the existence of neutrino oscillations has been confirmed by many experiments [17, 18, 19, 20], and this clearly indicates that neutrinos are massive particles.

Some phenomena that are not explained by the SM arise from cosmological observations. The first problem is the so-called baryon asymmetry, which is a significant imbalance between the observed amount of matter and antimatter in the Universe. If all charges are conserved, and equal amounts of matter and antimatter have been produced in the Big Bang, some mechanism is needed to explain why the baryonic matter did not completely annihilate into photons and dominates the observed Universe. According to Andrei Sakharov [21], three conditions are necessary for this effect to occur:

1. baryon number must be violated,
2. charge (C) and charge-parity (CP) symmetries must be violated,
3. the Universe needs to depart from thermal equilibrium during its evolution.

In the SM, only a difference between the baryon and lepton numbers ($B - L$) is conserved. The baryon number symmetry is violated in the SM in the non-perturbative regime by the so-called sphaleron, a static saddle point solution to the electroweak potential [22, 23]. In addition, the amount of CP violation, although the asymmetry was discovered in 1964 [24], is insufficient in the SM to explain the observed extent of baryon asymmetry. The last condition means that there must have been a period in the early Universe expansion when its contents was out of thermal equilibrium. Otherwise, if the net baryon number was zero just after the Big Bang, it would never change in the stationary state. The exact realisation of the departure from equilibrium depends on the model, however, it is not possible within the SM. For example, to generate baryon asymmetry through electroweak processes, a *very strong* first order phase transition would be needed, which is the reason why baryogenesis cannot be generated in the SM itself [25, 26].

A very important shortcoming of the SM is the lack of an explanation for the existence of dark matter (DM). In 1933 Fritz Zwicky observed that the dispersion of velocities in the Coma nebulae cluster does not agree with expectations based on visible (luminous) matter distribution in the cluster [27]. He suggested that there might be a significant amount of some invisible form of matter, which could explain the motion of nebulae. Further evidence came from observations of galaxies rotation curves in the 1970s, for which the rotation velocity dependence on the distance from the galaxy centre was different than expected from the distribution of luminous matter [28]. Again, adding a contribution from the DM allows one to describe the observations.

Currently, there are many more indications for the existence of DM. The first example is the distribution of anisotropies in the Cosmic Microwave Background (CMB) spectrum, which strongly depends on evolution of the early Universe. The CMB fluctuations can be perfectly described in the framework of the so-called Lambda Cold Dark Matter (ΛCDM) model [29] that explains the Universe expansion based on Einstein's general relativity theory. ΛCDM contains six free parameters and relies on two components: the cosmological constant, Λ, and the cold DM, where the term "cold" indicates small velocity of particles compared to the speed of light. Moreover, DM models are very successful in explaining formation of large structures in the Universe, such as galaxies or galaxy clusters, although more work is needed in order to solve some smaller scale problems [30, 31]. Another, more direct evidence for DM, is a mass density measurement in a merging galaxy cluster 1E0657-56, the famous Bullet Cluster [32, 33]. Analysis of gravitational lensing suggests a different centre of mass location than would be implied by the visible matter distribution, strongly indicating the presence of a significant invisible dark component [34].

One should note here that there is also some effort to explain astrophysical anomalies with modified theories of gravity (such as modified Newtonian dynamics, MOND) [35, 36]. Although they allow to describe some phenomena, so far the DM models seemed to be more successful and convincing to the community, since MOND fails to explain several important observations and is incompatible with the current understanding of cosmology [37, 38].

There is also a set of issues with the SM for which new physics is not indispensable, but explanations without the BSM physics, for reasons such as the need for strong fine-tuning, seem unnatural. One of the most important such issues is the so-called strong CP problem. In the QCD sector of the SM it is possible to add another term to Equation (A.7) without changing the equations of motion:

$$\mathcal{L}_\theta = \theta_{\text{QCD}} \frac{\alpha_s^2}{32\pi^2} G^a_{\mu\nu} \tilde{G}^{a\mu\nu}, \tag{2.1}$$

where $\tilde{G}^{a\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma} G^a_{\rho\sigma}$ is a dual field strength tensor, $\epsilon^{\mu\nu\rho\sigma}$ is fully anti-symmetric Levi-Civita tensor, and θ_{QCD} is a free parameter. This term would be a source of CP-violation in the theory of strong interactions [39], however, measurements of the electric dipole moment of neutron [40, 41] set an extremely small upper limit on the θ_{QCD} parameter [42]. Therefore, either there is some BSM mechanism responsible for suppression of θ_{QCD} (or cancellation of this term), or its value is very close to zero by pure accident, what seems highly unlikely.

Another fine-tuning issue, known as the hierarchy problem, is connected to the Higgs boson. Assuming that the SM is a part of some greater theory [43], which is valid up to a very high energy scale $\Lambda \sim 10^{15}$-10^{19} GeV, the Higgs field must be participating in such a "theory of everything". Considering quantum corrections to the Higgs propagator from other fields in the SM, which are quadratic in Λ, at very high mass scales the Higgs bare mass parameter would have to be enormous compared to the observed mass of 125 GeV [44, 45]. Without new physics, obtaining the observed Higgs boson mass would require fine tuning of tens of orders of magnitude.

An open question which also lacks an answer (and some could argue that it does not require one) is the apparent asymmetry between fermions and bosons. In principle, there is no fundamental reason for the matter to be made of fermions and the interactions to be mediated by bosons. It is therefore natural to seek an explanation other than "it is what it is". An interesting solution is proposed by Supersymmetric (SUSY) theories that for each SM particle introduce a partner which has a spin number different by a half of an integer. Interestingly, SUSY not only solves the fermion-boson asymmetry but also the hierarchy problem, and provides a natural candidate for a DM particle.

2.2 Searches for BSM physics

Since we know the SM does not explain some of the observed phenomena, and therefore some BSM physics must exist, efforts are ongoing both on theoretical and experimental frontiers to discover new physics. The LHC has been constructed with massive hopes for discovering "new physics", particularly SUSY particles, but also models with extra dimensions, heavy Majorana neutrinos, extensions of the scalar sector, or Technicolour [46, 47]. It is currently leading the efforts in BSM searches with its multi-purpose experiments, ATLAS and CMS (and, to some extent, also LHCb), performing analyses both targeting specific models and using partly model-independent methods. However, thus far no BSM phenomena have been observed, neither at the LHC nor beyond. This section gives a brief overview of methods available for new physics searches in particle physics experiments.

2.2.1 Direct searches

The most natural method for new physics searches is trying to directly observe, measure the new particles. It is the primary method used at hadron colliders, and hence they are often called "discovery machines". A good example is the discovery of the Higgs boson in 2012 through measurements of its decays to photons and Z bosons. It was based on very clean signals, easily distinguishable from the background, with high transverse momentum (p_T) final states, the case for which the LHC is the best suited.

Hadron colliders, despite having potential to reach very high collision energies, suffer from a huge background resulting from high cross sections of QCD processes. An additional source of background are the so-called pile-up interactions. In order to increase luminosity, the number of protons per bunch and the bunch collision rates are maximised, resulting in an overwhelming number of soft interactions between the protons, which produce hundreds of tracks and additional calorimeter deposits. This background creates a very challenging experimental environment, requiring strong reliance on triggers and clear signatures, and making it difficult to observe some processes.

For example, Higgs decays to bottom quarks [48, 49], or to tau leptons [50, 51], were much more challenging to detect, even though the branching ratios for such decays are significantly larger than for the discovery channels. Similarly, Higgs decays to muon pairs have been detected with 3σ significance [52] before the corresponding decays to charm quarks, despite the branching ratio for the latter channel being two orders of magnitude higher. That is because two oppositely-charged, high-p_T muons provide much cleaner signature than jets, even assuming an excellent flavour tagging.

Important features of direct searches are the expected experimental signature and the level of model dependence. The assumed signature can be either more general or very specific — e.g. just two tracks in the final state, or two muons associated with a Z boson production. A direct consequence of this choice is the extent to which a search is model-dependent; a general signature allows to cover a wider range of models since, in principle, different models (or realisations of a model) can give rise to the same signature. On the other hand, narrowing down the assumptions about the final state (such as its isolation, boost, production angle, or flavour of expected particles) can enhance the sensitivity to a particular scenario.

Searches for new physics can also be conducted at e^+e^- colliders. In fact, many current exclusion limits still come from the Large Electron-Positron collider at CERN (LEP), the predecessor of the LHC [53, 54]. An example of a currently running e^+e^- collider, suitable for direct BSM searches, is the B-meson factory, SuperKEKB, and the Belle II experiment operating at this machine [55, 56]. The main advantages of an e^+e^- collider compared to a hadron-hadron machine are much cleaner experimental environment (without the background from QCD processes) and the well-defined initial state, since the interacting beam particles are fundamental. An obvious drawback of such a machine is typically much lower centre-of-mass energy than available in hadron collisions. This will be described in more detail in Chapter 3.

One should note that direct searches are also possible beyond colliders. In particular, beam-dump and fixed-target experiments should be mentioned here, as their high intensity provides good reach in the low-mass regime, probing even very small couplings. Facilities operating at the CERN Super Proton Synchrotron (SPS) — NA62, NA64, or future SHiP — can set competitive constraints on models with right-handed neutrinos or extended gauge sectors [57, 58, 59].

Notable examples are also DM direct detection experiments, such as XENON, PandaX and Lux-Zeplin [60, 61, 62] using xenon, or DEAP and DarkSide [63, 64, 65] using argon. Their strategy is to set up a large tank filled with liquid gas for exposure as an active target and search for an elastic recoil of a DM particle against the gas nucleus. Other examples are experiments searching directly for axions — particles proposed to explain the strong CP problem — such as CAST or ALPS [66, 67] which rely on a coupling with electromagnetic field predicted for axions.

Direct searches[1], mostly for DM, are attempted also in astrophysical observations, in particular in measurements of cosmic rays — for example, in Fermi-LAT, AMS, HESS [68, 69, 70], and many others.

[1] In astrophysics, these are usually referred to as *indirect* searches. However, from the point of view presented in this thesis, they are typical direct searches, since the measured particles should originate directly from the DM annihilation.

2.2.2 Indirect searches

It is also possible to search for new physics indirectly. At particle colliders, this is allowed mostly by properties of QFT and its perturbative nature. For example, one can see from Equation (A.19) that the differential cross section depends on the matrix element, which takes contributions both from processes at the tree level and from higher-order Feynman diagrams. New physics can alter the predictions for differential cross sections by adding new diagrams to the matrix element, and one could see a deviation from the SM predictions by performing very precise measurements.

Similarly, renormalisation schemes provide frameworks for determination of relations between parameters and processes in a theory. In this way, based on precise measurements of some parameters, the values of others can be estimated if they contribute to corrections to the measured quantities. Primary examples for this are *i*) prediction of the top quark mass window before its discovery, which was based on precision electroweak measurements [71], and then analogously, *ii*) estimation of the Higgs mass window thanks to inclusion of top quark mass measurements in the analysis [72].

A modern framework for such indirect searches is given by the Effective Field Theories (EFTs) which allow to interpret the precision measurements in a global way using operators of dimension higher than four. In this way, one can reach and probe energies much higher than those directly accessible at particle colliders. This is because the influence of high-energy phenomena should be visible as deviations from the theory in the observables at low energies, as long as one has sufficient precision. However, EFTs not only allow to collectively include effects from various observables, but also to connect a potential discrepancy between the experiment and the theory with a specific Wilson coefficient. These coefficients correspond to particular higher-dimensional operators in the Lagrangian, so this allows to identify where the deviation might originate from. Furthermore, this can be achieved in an almost model-independent way. Details about EFTs and the corresponding formalism can be found in Appendix A.3.

The EFT frameworks, in particular the Standard Model Effective Field Theory (SMEFT) [73, 74], have already been successfully used to interpret the existing data from the LHC, LEP, SLC and Tevatron, resulting in constraints on Wilson coefficients and the scale of new physics under different assumptions [75, 76, 77, 78]. These results will be significantly improved by future accelerators, especially e^+e^- colliders. The expected precision on selected effective couplings (described in Ref. [79]), projected from fit to SMEFT for different future collider options and HL-LHC, is shown in Figure 2.1. It can be seen that in nearly all cases e^+e^- colliders (CEPC, FCC-ee, ILC, CLIC) can improve the sensitivity of HL-LHC. This precision corresponds to the new-physics energy scales of 10-100 TeV, depending on the operator. Future e^+e^- colliders will be introduced and described in more detail in Chapter 3. Figure 2.1 presents the results compiled for the ECFA Higgs/Top/Electroweak study report [11] and is an update of previous work [79, 80].

2.3 Long-lived particles

Decay widths of the heaviest particles in the SM — top quark, Z, W, and Higgs bosons — span from the order of MeV (Higgs) to several GeV. As a result, they have such small decay lengths that, compared to the typical size of detectors at colliders, their decays are prompt. In most

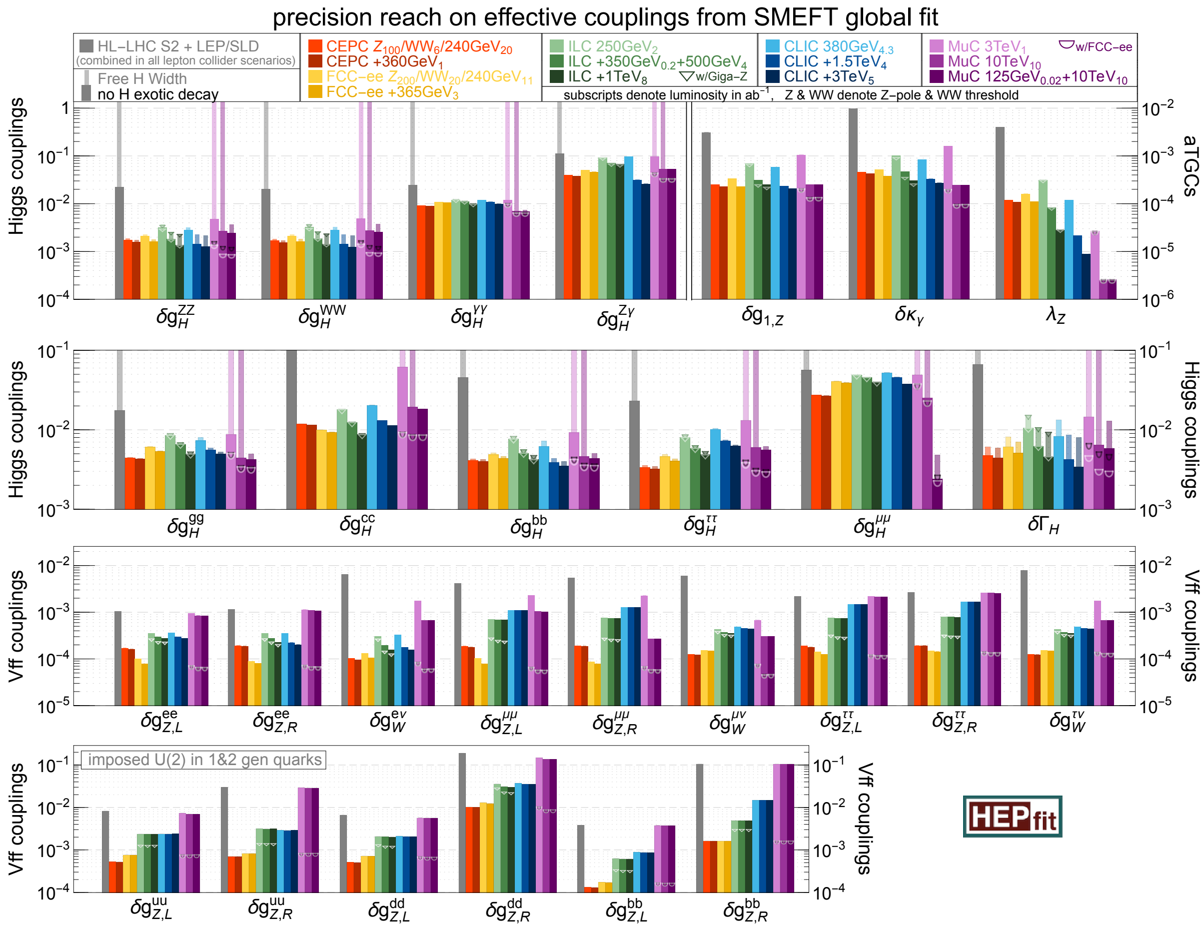


Figure 2.1: Comparison of the 1σ (68% probability) sensitivity to new physics modifications in the interactions of the Higgs and electroweak vector bosons, obtained projecting the SMEFT results into the *effective couplings* of these particles to the other SM fields. Figure taken from Ref. [11]. See text for details.

BSM scenarios, this is also the case for new heavy resonances. Therefore, corresponding searches rely on measuring the decay products of hypothetical BSM states, with the assumption that tracks of daughter particles originate from a region very close to the interaction point (IP) of the beam particles. This is particularly important at hadron colliders, where pile-up interactions can produce numerous tracks with large impact parameters, while cuts on those variables also allow one to reject tracks originating from cosmic rays. Due to theoretical motivations and these additional factors, searches at colliders, in particular at the LHC, are dominated by prompt signatures, and reconstruction algorithms or triggers have been primarily designed with them in mind.

Meanwhile, particles with much higher, macroscopic decay lengths, comparable to the size of a detector (so the order of millimetres or longer), are forbidden neither in the SM, nor in many BSM models, and frequently appear in theoretical scenarios. In fact, the SM contains multiple states with masses below the weak scale whose lifetimes are high enough for them to travel macroscopic distances: muon, pion, kaon, or neutron, just to name a few. Masses and lifetimes of the SM particles are presented in Figure 2.2. Based on Equation (A.20), one can see

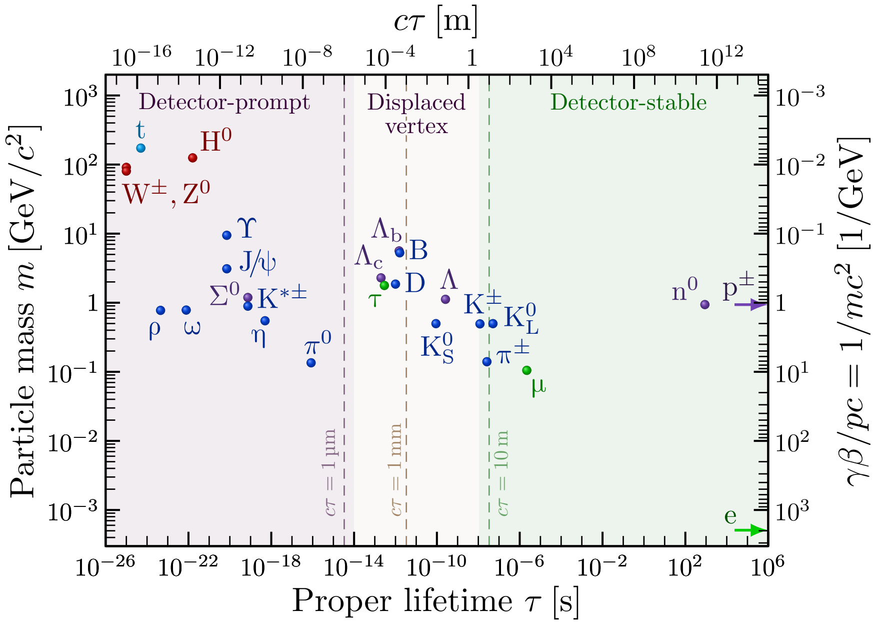


Figure 2.2: Masses of the SM particles shown as a function of their proper lifetime. Figure taken from Ref. [81].

that a small decay width, and therefore a long lifetime, can result from a small matrix element or a small phase space. There are three possible mechanisms for particles to acquire such lifetimes.

1. Small coupling

   A very small coupling between a mediator and its decay products enters the matrix element and therefore decreases the resulting width. Before weak gauge bosons were proposed, weak decays have been described using the Fermi constant. In this description, its small value has been a source of long lifetimes of muons, pions, or kaons.

2. Scale suppression

   A decay can be mediated at a much higher scale than masses of the decaying particle and its products. The large mass of a mediator appears in the denominator in the calculation of matrix element, resulting in its smaller value and decreasing the width. Examples are muon and pion decays in the SM mediated by the W boson, a particle almost three orders of magnitude heavier.

3. Limited phase space

   Kinematic constraints can limit the allowed phase space of the final state, reducing the width and increasing the lifetime. The most common reason is a small mass difference between a particle and (one of) its decay products. In the SM, a good example is the neutron decay, which is nearly mass degenerate with the proton. However, in this case, the scale suppression also plays a role, since the decay is mediated by the much heavier W boson, further increasing the neutron lifetime.

Similar mechanisms can also naturally occur in BSM scenarios. The case of scale suppression can give rise to long lifetimes of the so-called axion-like particles (ALPs). Axions are pseudoscalar Nambu-Goldstone bosons, first proposed to explain the strong CP problem [82, 83], while more recently generic pseudoscalar ALPs also started to receive attention. They are typically introduced by an effective Lagrangian with a cut-off energy scale, up to which the effective theory is valid. Because ALPs can be very light (with masses of the order of few MeV or GeV) their decay modes highly depend on the mass, while partial decay widths depend on ratios of respective couplings to the cut-off scale, having impact on the ALP lifetime [84].

A BSM model that realises the second scenario can be SUSY with small mass difference between Higgsino and Gaugino — the lightest (LSP) and the next-to-lightest (NLSP) supersymmetric particles, respectively [85]. Because the mass splitting between Gaugino and Higgsino is loop-induced, it remains below 355 MeV even for Gauginos much heavier than the Z boson. The resulting decay length of the NLSP in the laboratory frame can be of the order of centimetres in e^+e^- collisions at centre-of-mass energies $\mathcal{O}(100\,\text{GeV})$.

For the third mechanism, where the lifetime is acquired via a small coupling, a natural example is the frozen-in DM [86]. In such models, the DM interacts so weakly that it has never been in thermal equilibrium with the rest of particles in the early Universe. It is produced in some decay or scattering process of particles that are in the equilibrium, until the Universe cools down enough for the production to cease. The freeze-in is contrary to the famous freeze-out mechanism, assumed in most DM models with weakly-interacting massive particles (WIMPs), where DM is in a thermal bath and its abundance decreases until a certain point. Because of their tiny couplings, particles in freeze-in models are referred to as feebly-interacting massive particles (FIMPs). The size of this coupling is often responsible for macroscopic lifetimes of particles in such models.

2.3.1 Collider signatures

Due to their exotic nature, LLPs can lead to a wide variety of possible signatures at colliders [87]. The types of signatures depend on the characteristics of a particle — lifetime, charge, mass, or decay mode — and in some cases also on its production channel. However, it should be noted that the nomenclature or classification of signatures is subject to convention and can serve for the purpose of a particular experiment or analysis, as the exact same production and decay of an LLP can be addressed differently, depending on specific detector features or dominant backgrounds. In this section, signatures most commonly discussed in literature are described, and a general categorisation of possible signatures for neutral LLPs is proposed.

Charged LLPs can lead to signatures of kinked, displaced, disappearing, or strongly ionising tracks.

- If a charged LLP undergoes a one-prong decay, with the other decay product(s) invisible, the resulting signal in the detector will be a track which is bent, or *kinked* (KT) at some point. In the SM, this is the case, e.g., for charged kaon decays, but it is a prediction of some BSM models as well.
- If the lifetime of an LLP is not long enough to produce its own track, the same type of decay can lead to a *displaced track* signature. This is a track which starts near the region

of IP, but has unusually large impact parameter.

- On the other hand, a one-prong decay of charged LLPs can be a source of the *stub* or *disappearing track* signature. If the mass difference between LLP and its invisible decay product is very small, the daughter track can have a very small transverse momentum. In this case, it can curl in the magnetic field so much that it is not possible to reconstruct the associated track. The resulting signature is a track that unexpectedly ends in the middle of a tracking detector.

- The lifetime of LLPs can also be in principle very long, allowing them to travel inside the detector, and making them indistinguishable from long-lived SM particles, such as muons, kaons, or protons. However, if the mass of such an LLP differs significantly from the SM states — in particular, it could be very heavy — one can identify it based on its ionisation of the detector material or time of flight (ToF). These particles are usually referred to as *heavy stable charged particles* (HSCP), and the signature of a strongly ionising track inherited this name.

- One can consider decays other than to one charged and invisible particles. In such a case, the signature can be similar to a displaced vertex (even for two-prong decays, if we assume some BSM models could allow charge violation). More details about displaced vertices are given below for neutral LLPs.

Neutral LLPs, on the other hand, can produce signatures of displaced vertices, displaced jet(s), or displaced photons.

- If a neutral LLP decays in the middle of the tracking detector to charged particles producing displaced tracks originating from a single point, it corresponds to a signature of *displaced vertex* (DV). In the SM, neutral long-lived hadron decays or photon conversions can lead to a DV signature.

- A special case of a DV often highlighted is the *displaced jet* (DJ), when multiple tracks from an LLP decay form a cone-like structure and can be clustered. This signature is particularly useful in LHC searches, as it provides a clean and distinctive signal. Interestingly, some searches even consider more exotic displaced jets in the muon spectrometer, beyond the tracking detectors. However, in this work, all such final states will be collectively referred to as displaced vertices.

- An interesting signature of a neutral LLP decay is a *displaced photon* (DP, also known as a *non-pointing photon*). It could serve as an identifier for such an LLP, particularly if the BSM model does not predict its decays into charged particles, but allows in-flight decays into photons, or a photon and one or more invisible particles. However, detecting this signature is highly challenging, as it requires precise reconstruction of electromagnetic calorimeter clusters to accurately determine original direction of the photon. A displaced photon conversion can help in addressing this challenge and is sometimes referred to as a separate signature.

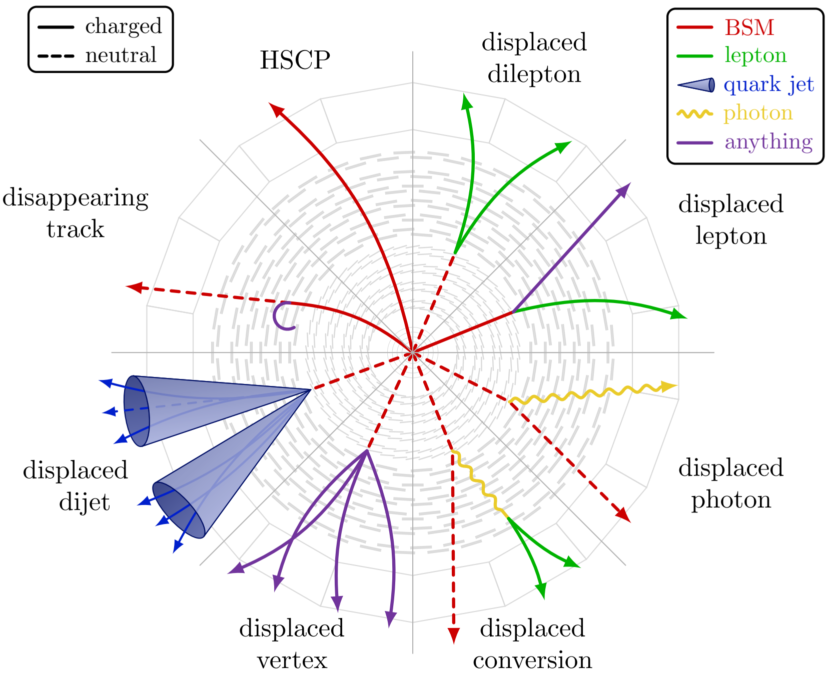


Figure 2.3: Examples of some of possible LLP signatures considered in the CMS experiment. In the context of this thesis, displaced dilepton signature is a special case of a displaced vertex. Similarly, displaced lepton is a special case of kinked or displaced track, if other LLP decay products are invisible. Figure taken from Ref. [88].

- In the most pessimistic scenario, a neutral LLP may also have no visible decay modes. In such a case, possibilities for its reconstruction are highly limited, if not impossible. However, methods commonly used in searches for prompt decays can be adapted, for instance missing energy ($\not{E}$) or mono-X signatures. The mono-X channel assumes fully invisible final state, that can be identified by an initial-state radiation (ISR) of a single particle — for example, a photon at e^+e^- colliders (mono-photon) or a gluon at hadron colliders (mono-jet).

Due to challenges posed by LLPs, even more exotic signatures can be considered. One such example is a *delayed decay*, a scenario in which a particle is produced with an extremely small boost. In this case, an LLP may move at a very low velocity, potentially even remain at rest with respect to the laboratory frame, or get stuck within the detector material. Its delayed decay would then produce an out-of-time signal in subdetectors, occurring between bunch crossings, or even between experimental runs. A schematic drawing of some of the possible LLP signatures in a collider detector is shown in Figure 2.3. The sketch was produced in February 2023 for the purposes of CMS experiment at the LHC and presents signatures considered by the CMS Collaboration at that time.

Some past efforts have aimed to classify LLP signatures at the LHC in a model-independent way, or to map them into generic FIMP scenarios [87, 89]. Since neutral LLPs will be the main focus of this work, a detailed classification of their signatures is further proposed for a production in e^+e^- collisions. It is presented in Tables 2.1 and 2.2 for the single and pair production of LLPs, respectively.

The signatures are categorised according to combinations of possible production and decay channels. Diagrams shown as column and row labels are used only for visualisation – all contributing Faynmann diagrams should be taken into account. Since many BSM models involving LLPs also introduce DM candidates, associated production of an LLP and DM was considered. However, from the experimental point of view, the associated DM production would be invisible in the same way as a neutrino production. Therefore, these two scenarios are treated collectively, and the corresponding branches are marked as "invisible".

The stability of DM in BSM models is typically achieved by imposing some additional symmetry — for example, R-parity in SUSY, but the most common is the Z_2 symmetry. For this reason, the behaviour of an LLP field under this symmetry is additionally marked when the DM has to be odd under Z_2, which is the case when the total number of DM particles in the final state is even. Signatures without any DM particles are also marked, since in principle they can arise in models where DM is also present. For example, the signature of the $e^+e^- \to X \to 2\,SM$ process may be considered if X decays both to SM and DM are allowed.

2.3.2 Searches at the LHC and beyond

LLPs have been extensively searched for at the LHC, but no evidence of their production has been found so far. The ATLAS and CMS experiments have explored vast regions of parameter space, placing strong constraints on very massive neutral LLPs across a wide range of lifetimes. Searches for decays in the muon systems probe decay lengths beyond 500 m for LLPs with masses around 40 GeV [90, 91], while constraints from delayed decays extend proper decay lengths to the order of 10^{12} m [92].

Extremely heavy neutral LLPs have also been investigated: searches for displaced jets have probed masses up to 2.5 TeV [93], while displaced vertex searches with missing momentum signatures have reached up to 1.8 TeV [94, 95]. Recently, lighter LLPs have gained increasing attention. In particular, models with LLP masses as low as 400 MeV have been constrained in Higgs boson decays at the LHC [90, 96]. Additionally, a recent search for displaced muon pairs has probed heavy squark decays into long-lived neutralinos, with a small mass splitting of just 25 GeV [97].

However, in this mass range, LHC experiments face significant background challenges, which can impact searches for both light and heavy LLPs — especially in scenarios with compressed spectra, where the mass differences between particles in the decay chain are small. To improve sensitivity, light LLPs must either originate from the decays of much heavier particles or be analysed under specific model-dependent assumptions.

Significant improvements in this regard have recently been achieved in the case of charged LLPs. Although the kinked track signature does not seem to find much use at the LHC, many searches for charged particles use signatures of disappearing tracks and $\mathrm{d}E/\mathrm{d}x$ from slow heavy states. A combination of these two signatures allowed CMS to set strong limits on the production of charginos (NLSP) decaying to neutralinos (LSP) where the mass difference between the two is only 0.4 GeV or less [98]. A very similar result was also obtained by ATLAS with only slightly lower mass reach [99]. Another frequently used signature is the displaced track, which in the same BSM scenario allowed to constrain the chargino-neutralino mass difference down to the

Table 2.1: Single neutral LLP (X) production signatures. Scenarios with DM odd under Z_2 symmetry: X can be odd (◇) or even (◆) under Z_2. See text for more details.

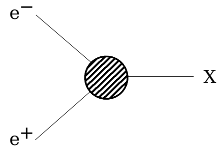
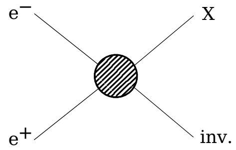
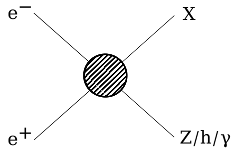
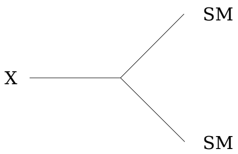
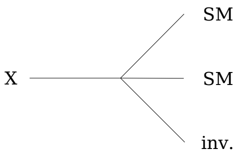
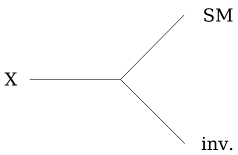
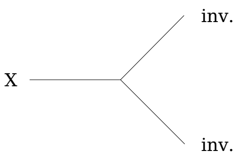
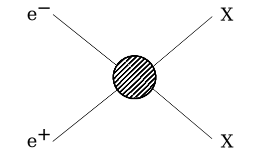
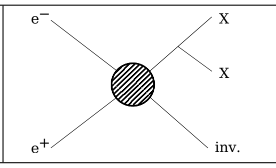
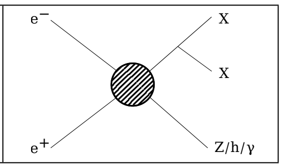
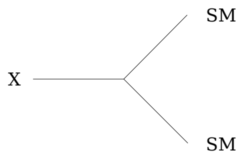
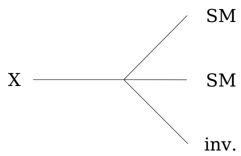
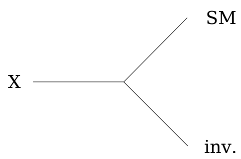
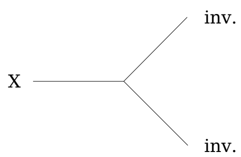
<table>
<tr><th>production / decay</th><th></th><th></th><th></th></tr>
<tr><td></td><td>• 2 tracks / jets / γ from IP or DV
◆</td><td>• displaced vertex
• $\not{E}$</td><td>• displaced vertex
• Z / h / γ
◆</td></tr>
<tr><td></td><td>• 2 tracks / jets / γ from IP or DV
• $\not{E}$</td><td>• displaced vertex
• $\not{E}$
◇</td><td>• displaced vertex
• Z / h / γ
• $\not{E}$</td></tr>
<tr><td></td><td>• 1 track / jet / γ from IP or DV
• $\not{E}$</td><td>• displaced track / jet / γ
• $\not{E}$
◇</td><td>• displaced track / jet / γ
• Z / h / γ
• $\not{E}$</td></tr>
<tr><td></td><td>• Mono-photon
◆</td><td>• Mono-photon</td><td>• Z / h / γ
• $\not{E}$
◆</td></tr>
</table>

Table 2.2: Neutral LLP (X) pair production signatures. Scenarios with DM odd under Z_2 symmetry: X can be odd (◇) or even (◆) under Z_2. See text for more details.

decay \ production			
	• 2 displaced vertices ◆	• 2 displaced vertices • $\not{E}$	• 2 displaced vertices • Z / h / γ ◆
	• 2 displaced vertices • $\not{E}$ ◇	• 2 displaced vertices • $\not{E}$	• 2 displaced vertices • Z / h / γ • $\not{E}$ ◇
	• 2 displaced tracks / jets / γ • $\not{E}$ ◇	• 2 displaced tracks / jets / γ • $\not{E}$	• 2 displaced tracks / jets / γ • Z / h / γ • $\not{E}$ ◇
	• Mono-photon ◆	• Mono-photon	• Z / h / γ • $\not{E}$ ◆

level of 10 GeV [100] and also the very challenging range of 0.3-1.2 GeV for chargino masses up to 180-190 GeV [101, 102].

Using dE/dx and ToF is a very efficient way to search for heavy LLPs with very long lifetimes because their velocity should be significantly smaller than the speed of light. For example, this signature recently allowed ATLAS to exclude staus below 400 GeV and charginos below 0.8-1.3 TeV for any lifetime above 3 ns [103]. At CMS, heavy charged LLPs have been ruled out up to 0.7 TeV for stau hypothesis, or even 2.1 TeV when interpreted as gluino R-hadrons [104].

Another search strategy explored in the last years is to place a detector away from the beam collision point. Initially, this idea originated from the fact that the cross section for inelastic proton-proton scattering is strongly peaked in the forward direction. Therefore, there is a possibility that very weakly interacting particles are produced in sufficient numbers in pp collisions at the LHC, but they escape undetected in the very forward region. This led to a proposal of the FASER experiment [105, 106], that was primarily designed to search for light dark photons — hypothetical vector bosons, Z', charged under the U(1) group. FASER has been constructed in an old service tunnel, approximately 500 m from the ATLAS detector. Despite the lack of any new physics observation, it was successful in observing neutrinos produced in pp collisions at ATLAS [107] and setting competitive limits on the dark photon production [108].

Other proposed "far detector" experiments have emerged in the following. MATHUSLA (MAssive Timing Hodoscope for Ultra-Stable neutraL pArticles) is a proposed tracking detector with an air-filled decay volume placed on the surface near the CMS experiment [109]. It aims to reconstruct DVs from particles with lifetimes too long to observe them in a detector operating at the IP, considering much higher backgrounds. ANUBIS, for which the prototype has already been constructed, will be another tracking device, installed in the ceiling and service shafts above the ATLAS experiment to search for LLPs with decay lengths $\mathcal{O}(10\,\mathrm{m})$ [110, 111]. The LHCb experiment will also host a dedicated LLP far detector — the Compact DEtector for EXotics at LHCb (CODEX-b) — located in its cavern, a demonstrator of which should start collecting data in 2025 [112, 113]. Finally, interesting prospects are given by the Forward Physics Facility planned for the HL-LHC phase. With four complementary detectors located around 600 m from the ATLAS IP, a broad physics programme will be carried out, not only focussing on direct LLP searches, but also involving neutrino, QCD, and astrophysics studies [114].

Light LLPs can also be searched for beyond experiments at (or near) the LHC. Belle II is particularly well-suited to search for LLPs in B decays, and such studies have already been performed on a relatively small dataset for dark scalar production in b → s transitions [115]. Another analysis has been carried out for long-lived heavy neutral leptons (HNLs) [116] and a very good sensitivity to dark photons is also expected [117] in this experiment. The possibility of a far detector, similar to MATHUSLA, operating near Belle II has been considered as well [118, 119]. Great prospects for detecting long-lived FIMPs are also expected with SHiP at the CERN SPS [59]. It will be sensitive to particles predicted by various BSM models in the 0.5-5 GeV mass range, such as HNLs, dark photons, dark scalars, ALPs, DM, or even SUSY particles in some scenarios [120]. The mass range is relatively narrow, but it allows to complement the current and future experiments with unprecedented precision.

The potential for direct LLP searches at future high-energy e^+e^- colliders is the main topic of this thesis and will be covered in more detail in the following chapters. However, it is worth mentioning here the prospects offered by beam-dump experiments at linear e^+e^- machines. In this case, the full intensity of the beam can be used since no particles are being wasted, which allows to obtain sensitivity in the low-mass regime complementary to direct collider searches [121]. Sensitivity competitive to other beyond-collider experiments (such as SHiP or FASER) is expected for ALPs below masses of 10 GeV and for leptophilic scalars in the sub-GeV range [122], as well as for HNLs below 10 GeV [123] or dark photons below 5 GeV [124].

Chapter 3

The e^+e^- Higgs factories

3.1 Motivation for the next collider

As described in Section 2.1, there are strong indications that the SM is not the final "theory of everything" and some BSM physics must exist. However, despite high expectations and the huge success of discovering the Higgs boson, neither the LHC nor any other particle physics experiment has observed any new physics phenomena by the time of writing this thesis. So far, all anomalies could have been explained with well-known SM sources, eventually turned out as statistical fluctuations and disappeared after collecting more data, or are still not fully confirmed and remain under thorough investigation.

For this reason, the paradigm of direct new physics searches that currently prevails in particle physics, in particular with the LHC leading the research frontier, is facing increasing pressure. The lack of conclusive evidence for any of the well-motivated BSM models, or even their definitive exclusion [125], has led the particle physics community to explore alternative approaches. Given the ability of the theory community to continually develop new BSM scenarios that evade existing constraints set by the LHC, the pursuit of direct searches is beginning to feel like an endless quest.

A natural alternative, described in Section 2.2.2 and appendix A.3, is given by precision measurements of the SM, which allow for model-independent indirect BSM searches using interpretations within the EFT framework. Especially the discovery of the Higgs boson opens new possibilities and directions for exploration, as the Higgs remains one of the least studied components of the SM. Due to its unique properties and the central role in the SM, it is expected to participate in any interactions at higher energy scales. Meanwhile, its absolute couplings have not yet been measured, the current precision on the width is only at the level of 50% (assuming equal on- and off-shell effective couplings) and the current constraints still leave much room for unknown decay modes or CP violation. Its potential composite structure or the existence of other fundamental scalars have also not been ruled out [126, 127]. A measurement of the Higgs pair-production would provide the only direct test of the Higgs self-coupling and the shape of the scalar potential in the SM. If new physics phenomena are coupled to the Higgs or occur in its sector, they could be identified via precision measurements of this particle. Nevertheless, one should note that to some extent this is also the case for other parts of the SM, even though the Higgs particle is the most unique.

Although it is possible to study many processes involving the Higgs boson at the LHC, precision of the measurements is limited. The total width and production cross section of the Higgs cannot be measured directly, as the measured quantities are typically cross sections time branching ratios. Therefore, instead of absolute coupling values, one has to consider their SM coupling modifiers, which is a model-dependent and potentially misleading approach that relies on many assumptions [128]. Within this so-called κ framework, coupling modifiers are currently measured at the level of 10% at best and are found to be consistent with the SM, while the self-coupling remains beyond the reach [129, 130]. A significant improvement is expected after the upgrade to the high-luminosity stage of the LHC and enormous effort is being put into increasing the precision. However, it should be noted that the LHC is primarily a discovery machine and was never expected to be optimal for electroweak precision measurements in the first place (see Figure 2.1). The projections for HL-LHC from current LHC results suggest a possibility of a measurement with $4\,\sigma$ significance ($6\,\sigma$ in a perfect scenario where systematics is neglected) and an error on the self-coupling modifier around 50% [131, 132].

Electron-positron colliders offer much better possibilities when it comes to precision, in particular in the case of the Higgs boson studies. Due to their ability to abundantly produce Higgs particles and measure them with high efficiency, e^+e^- colliders at centre-of-mass energies in range 240-380 GeV are often referred to as the *Higgs factories*. The precision of these machines is primarily ensured by two main factors: fully known initial state and a clean experimental environment. In hadron-hadron collisions where the interactions occur in fact between quarks or gluons carrying only fraction of the hadron's momentum (partons), the exact value of longitudinal component of the initial-state momentum is unknown, and one can only rely on the so-called parton distribution functions for the statistical description of its distribution. Because electrons are fundamental, point-like particles, their energy and momentum at the IP are well defined and known, which can be used as an additional constraint when reconstructing the final state. Because at tree level electrons are sensitive only to electromagnetic and weak interactions, the cross sections for processes with hadronic final states are at the level similar to (other) electroweak processes. Therefore, e^+e^- colliders do not suffer from the QCD background and pile-up interactions like pp machines, where hadrons are produced with orders of magnitude higher cross sections. However, e^+e^- machines are obviously not completely background-free: the electric charge of electrons is a source of additional $\gamma\gamma$ interactions (a background described in more detail in Section 3.4).

If the centre-of-mass energy is high enough, e^+e^- collisions open access to the so-called Higgsstrahlung process, $e^+e^- \to Zh$, the "golden channel" for the Higgs boson production. In modern collider proposals, its maximum rate is expected around 240-250 GeV due to the interplay between cross section and luminosity. The main advantage of this process is that it allows reconstructing the Higgs independently of its decay mode. Knowing fully the initial state, one can reconstruct the Higgs mass just by measuring the four-momentum of the Z boson, calculating the so-called recoil mass

$$M^2_{\text{recoil}} = (p_{cms} - p_Z)^2 = s - 2\sqrt{s}E_Z + M_Z^2. \tag{3.1}$$

p_Z, M_Z and E_Z are, respectively, the reconstructed four-momentum, mass and energy of the Z boson, and $p_{cms} = (\sqrt{s}, 0, 0, 0)$ is the four-momentum of the initial state with the centre-of-mass energy $\sqrt{s}$.[1] For the Higgsstrahlung process, the recoil mass corresponds to the mass of the produced Higgs boson. By reconstructing the Z boson using its decays to electrons and muons, one can precisely measure the Higgs mass and the total Higgsstrahlung cross section [133]. The Zh process will also allow for a very precise and relatively model-independent measurement of total Higgs width and determination of absolute couplings to fermions and gauge bosons [134].

Increasing the e^+e^- collision energy to 500-550 GeV opens a direct access to the di-Higgs production and therefore to the Higgs self-coupling determination on the tree level. The measurement is possible in the $e^+e^- \to Zhh$ (via Higgsstrahlung) and $e^+e^- \to hh\nu\nu$ (through vector-boson fusion, VBF) channels, where for the latter the cross section increases with energy. Hence, e^+e^- collisions at even higher energies could further improve self-coupling determination precision. A 550 GeV collider with an integrated luminosity of 4.4 ab^{-1} could measure the trilinear Higgs coupling with a precision of 15%, and operation at a higher repetition rate, providing the luminosity of 8 ab^{-1}, can reduce it to the level of 11%. It should be noted that these estimates correspond to the assumed SM value of self-coupling, $\kappa_\lambda = 1$; if it deviates from this prediction and, e.g., $\kappa_\lambda = 2$, the accuracy will improve to 5%, while at HL-LHC it would stay at 25% [135].

Increasing the energy even more may provide further improvements, as in old studies for CLIC at 1.5 TeV and 3 TeV an accuracy of [-8%,11%] was obtained [136]. Moreover, energies above 1 TeV allow one to constrain not only the trilinear, but also the quartic coupling via triple Higgs boson production [137]. All estimates mentioned here may be updated in the future for both e^+e^- colliders and HL-LHC following the advances in detector technologies and reconstruction techniques.

Very interesting prospects given by the Higgs factories are also the potential top quark measurements. In the LHC, the mass of the top is most precisely measured on the basis of its reconstructed decay products — jets or leptons. The kinematics of these objects cannot be directly obtained from perturbative QCD calculations, and therefore the measurements are interpreted in the context of predictions of Monte Carlo (MC) event generators. However, the top mass is not a physical quantity and depends on a selected renormalisation scheme.

Future e^+e^- colliders will allow the interpretation of measurements within a well-defined scheme [134]. The highest precision of the top quark mass determination is expected in the so-called top threshold scan, where the top-quark pair production cross section is measured at several points around the energy threshold for this process. Although the Yukawa coupling of the top quark can also be constrained in the threshold scan, a direct access to this coupling is available in the $e^+e^- \to t\bar{t}h$ process at centre-of-mass energies above approximately 600 GeV.

Although e^+e^- colliders will focus primarily on precision measurements, they also offer strong potential for direct searches, providing complementarity to the LHC reach. Thanks to their much cleaner experimental environment mentioned above, trigger-less operation can be considered for these colliders. The well-defined initial state is an additional factor improving the overall ability to efficiently identify and reject background processes based on full information about event kinematics. As a result, the Higgs factories are capable of fully excluding BSM models almost up to the kinematic limit. They are particularly well-suited for searches involving light BSM particles or final states with low momenta, which are often challenging at the LHC due to overwhelming

[1] We neglect here the possible transverse component of the initial four-momentum resulting from the finite beam crossing angle.

Table 3.1: Comparison of centre-of-mass energies and integrated luminosities in the FCC-ee [6] and CEPC [140] collider programmes across different running stages.

Running Stage	**FCC-ee** (4 IPs) $\sqrt{s}$ [GeV]	$\int \mathcal{L}\,\mathrm{d}t$ [ab^{-1}]	**CEPC** (2 IPs) $\sqrt{s}$ [GeV]	$\int \mathcal{L}\,\mathrm{d}t$ [ab^{-1}]
Z pole	88, 91, 94	205	91.2	100
WW threshold	157, 163	19.2	160	6
Higgs factory	240	10.8	240	20
Top threshold scan	340–350	0.42	—	—
$\mathrm{t\bar{t}}$ production	365	2.7	360	1

QCD backgrounds and the need for very clean signals to isolate such events [138, 139].

3.2 Circular colliders

Two types of electron-positron colliders are currently being considered in the context of future Higgs factories — circular and linear. The main advantage of circular machines is a very high luminosity they can achieve at low energies, since circulating particle bunches collide multiple times. However, because of synchrotron radiation, this comes with a price of power consumption, which increases fast with energy. As a result, when assuming constant power, the luminosity of circular $\mathrm{e^+e^-}$ colliders decreases very quickly with the centre-of-mass energy, so the energy reach is limited.

Two major projects of such machines have been proposed: Future Circular Collider (FCC-ee) [6] at CERN and Circular Electron-Positron Collider (CEPC) [7] in China. Both proposals offer an implementation in an approximately 100 km circumference underground tunnel to host an $\mathrm{e^+e^-}$ Higgs factory with centre-of-mass energy $\sqrt{s}$ = 240 GeV, potentially followed by a pp high-energy machine running at $\sqrt{s} \simeq$ 75-85 TeV (with an ambition to eventually go beyond that) in the same tunnel. Both would also run in alternative energy stages, at the Z pole with $\sqrt{s}$ = 88-94 GeV, at $\mathrm{W^+W^-}$ threshold with $\sqrt{s}$ = 157-163 GeV and above $\mathrm{t\bar{t}}$ threshold with $\sqrt{s}$ = 360-365 GeV, while only the FCC-ee programme assumes performing a top threshold scan at $\sqrt{s} \simeq$ 350 GeV (although this should also be feasible at CEPC). Currently, four IPs are considered for the FCC-ee and only two for CEPC, however, a possible extension to four IPs at CEPC is under study [140]. A summary of the assumed running scenarios of FCC-ee and CEPC, including the assumed total integrated luminosities, is presented in Table 3.1.

The main highlight of each $\mathrm{e^+e^-}$ circular collider programme that particularly stands out compared to other proposals is the so-called Tera-Z stage. A high-luminosity circular machine running at the Z pole, $\sqrt{s} \simeq$ 91 GeV can deliver more than 10^{12} Z bosons. Besides testing very high-energy scales of new physics (although limited to scenarios accessible via measurements of Z couplings to fermions), this provides unprecedented statistics also for flavour physics (b/c quark and τ lepton studies), almost two orders of magnitude higher than current B factories, such as Belle II or BaBar [141, 142].

Several dedicated detectors have been proposed for the FCC-ee as four interaction regions are planned at this collider: CLD, IDEA, and the most recent, ALLEGRO. The final choice of technologies for each of these options is still under consideration; however, CLD (CLIC-like detector) [143] seems to be the most advanced concept, because it was adapted to FCC-ee conditions from the mature detector design for CLIC (introduced in the next section). For CEPC, two generic options for a baseline detector have been considered and recently also a new detector concept has been presented [140]. It should be noted that the ILD experiment (a detector concept that provides the main framework for this thesis, also introduced in the following section and described in Chapter 4) published a design optimised for operation at a circular machine [144] which is considered for both FCC-ee and CEPC.

3.3 Linear colliders

An alternative approach is to build a collider in a linear tunnel, to avoid the synchrotron radiation losses that limit the energy at circular machines. This allows one to reach higher centre-of-mass energies while retaining the clean experimental environment of an e^+e^- collider. Such a design not only makes precision Higgs and electroweak measurements possible (including channels not available at low energies) but also opens direct access to processes involving new heavy BSM states. A further key advantage is flexibility: the tunnel can be extended to reach higher energies in later stages, and the planned energy stages can be adjusted if early discoveries indicate that a different collision energy would be more optimal than the baseline. Such adaptability, achieved without major changes to the accelerator technology (though such changes are not excluded, as discussed later), is a unique feature of linear colliders.

In addition, a crucial feature of linear e^+e^- machines is the possibility of colliding polarised beams. This allows one to reduce systematic uncertainties and partially compensate for lower luminosity. A further advantage of polarisation is the ability to better suppress and control some backgrounds, which can help both in precision measurements and direct BSM searches. Finally, in case of a discovery, polarisation enables identification of the properties of a BSM model, such as the coupling structure.

Several proposals of linear machines exist, differing not only by the assumed location, but also by the accelerating technology. The two most advanced projects are the International Linear Collider (ILC) [3] originally proposed in Japan, and the Compact Linear Collider (CLIC) [4, 5] envisioned at CERN. Recently, also a proposal for a Linear Collider Faciliy (LCF) [8] at CERN was presented, which takes advantage from the mature design of the ILC but also takes into account recent progress in technology. The linear colliders are described in more detail here, with particular focus placed on the ILC concept, since it is the most mature collider proposal, and for this reason it was used as the reference collider for studies presented in this thesis. The results are also applicable to a large extent to the LCF due to almost identical design with the ILC, however, this concept was proposed when the work described here was already in its final stages.

3.3.1 International Linear Collider

The ILC, with its Technical Design Report (TDR) published already in 2013 [145], is based on a well-proven technology of superconducting radio-frequency (RF) cavities, already industrialised

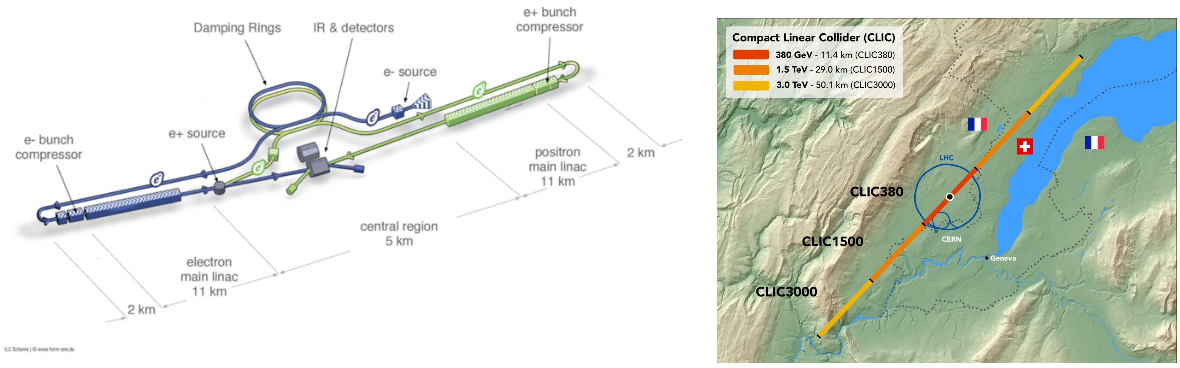


Figure 3.1: Left: layout of the ILC concept as presented in 2013 TDR [145]. Right: Schematic view of CLIC accelerator site at CERN in its original baseline staging scenario [148].

in projects such as the European X-Ray Free-Electron Laser (XFEL) Facility [146, 3]. The schematic layout of the ILC accelerator is presented in Figure 3.1 (left). The latest baseline implementation plan for the ILC (H-20) [147] assumes staged operation at three centre-of-mass energies: 250 GeV (for measurements of the Higgsstrahlung process), 350 GeV (dedicated to the $t\bar{t}$ threshold scan) and 500 GeV (allowing access to the $t\bar{t}$ continuum, $t\bar{t}h$ and Zhh processes). The 250 GeV Higgs factory stage requires a 20.5 km long underground tunnel, while its further extension to 31 km allows for an operation at 500 GeV.

The H-20 scenario assumes collecting 2 ab^{-1}, 200 fb^{-1} and 4 ab^{-1} at the three consecutive stages, respectively, shared between different beam polarisation setups as shown in Table 3.2, since ±80% electron and ±30% positron beam polarisation is assumed at the ILC. The evolution of luminosity in the full 22-year programme of the ILC, divided into the corresponding energy stages in the H-20 scenario, is presented in Figure 3.2. Some of the ILC beam parameters, relevant for analysis described later in this work, are shown in Table 3.3.

The H-20 scenario could be complemented by two additional running stages. An upgrade to the energy of 1 TeV, that was considered as a potential running stage for the ILC already since its TDR [145], would require a 40 km tunnel and should allow to collect 8 ab^{-1} in 10 years. Another running option considered is the "Giga-Z" run, which assumes ILC operation at the Z pole, producing $5 \cdot 10^9$ Z bosons in 1.5 years for precise measurements of the Z couplings to fermions [149].

Two detector concepts have been proposed for the ILC: International Large Detector (ILD) and Silicon Detector (SiD) [12, 150]. They both rely on highly granular calorimeters, required for the Particle Flow reconstruction [151], but have very different tracking systems and overall size. The main tracker of the ILD is a large gaseous time projection chamber (TPC), whereas the SiD is equipped with all-silicon detectors. SiD is also a much smaller experiment (3 m less in diameter), and it requires a stronger magnetic field to compensate for its size. The work presented in this thesis was carried out within the framework of the ILD concept and this experiment is described in more detail in Chapter 4.

Table 3.2: Integrated luminosities per beam helicity configuration resulting at the ILC in scenario H20. Adapted from Ref. [3].

$\text{sgn}(P(e^-), P(e^+)) =$	$(-,+)$	$(+,-)$	$(-,-)$	$(+,+)$
$\sqrt{s}$	$[\text{fb}^{-1}]$	$[\text{fb}^{-1}]$	$[\text{fb}^{-1}]$	$[\text{fb}^{-1}]$
250 GeV	900	900	100	100
350 GeV	135	45	10	10
500 GeV	1600	1600	400	400

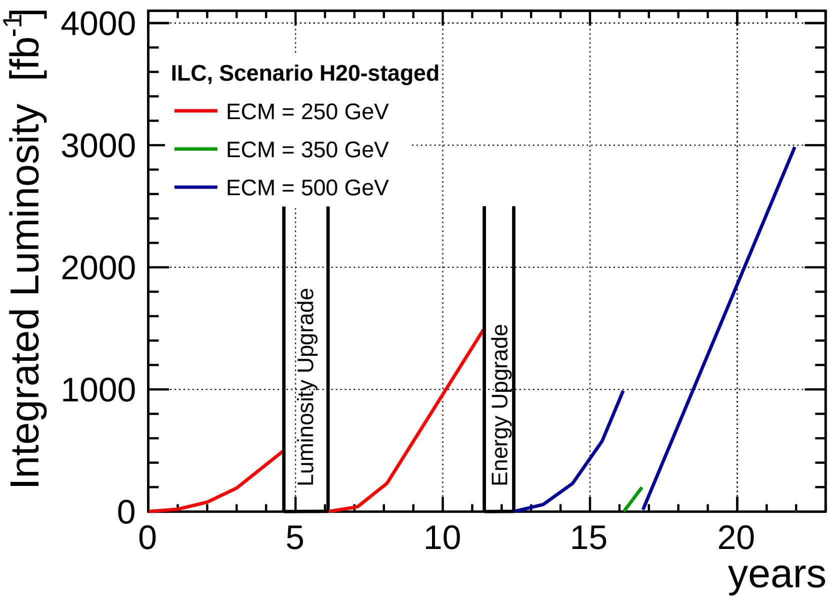


Figure 3.2: The nominal 22-year running program for the staged ILC, starting operation at 250 GeV with the current baseline beam parameters (H-20 scenario) for the 250 GeV runs. Taken from Ref. [152].

3.3.2 Compact Linear Collider

In contrast to the ILC, CLIC relies on the technology of normal-conducting ("warm") acceleration structures, where the power is transferred to the main linac from low-energy, high-current drive beams. This allows for higher gradients and therefore CLIC could reach much higher centre-of-mass energies with smaller or similar tunnel footprint than ILC. This technology and the target acceleration gradients have been demonstrated in a test facility at CERN [148]. However, it has not been tested on the larger scales yet and lacks proper industrialisation the superconducting RF cavities have undergone.

CLIC implementation plan assumes staged operation, with the first stage involving $t\bar{t}$ threshold scan around $\sqrt{s} \simeq 350$ GeV and a Higgs factory operation (for Higgstrahlung measurements) at 380 GeV in an 11 km long tunnel. In this way, CLIC would also allow for $t\bar{t}$ production studies already at its first stage. The next stage is planned at 1.5 TeV for top Yukawa ($t\bar{t}h$) and Higgs self-coupling (hZZ and hh$\nu\nu$) measurements, which would require 29 km tunnel. Also an upgrade to the energy of 3 TeV with a site length of 50 km has been considered, allowing for more precise

Table 3.3: Summary table of some of the ILC accelerator parameters in the initial 250 GeV staged configuration and possible upgrades. A higher luminosity at 250 and 500 GeV can be obtained by doubling the number of bunches per pulse. A 500 GeV machine could also be operated at 250 GeV with 10 Hz repetition rate, bringing the maximum luminosity to $5.4 \cdot 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. Adapted from Ref. [3].

Quantity	Symbol	Unit	Initial	$\mathcal{L}$ Upgrade	$\sqrt{s}$ Upgrades	
Centre of mass energy	$\sqrt{s}$	GeV	250	250	500	1000
Luminosity	$\mathcal{L}$	$10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$	1.35	2.7	1.8/3.6	4.9
e^+ (e^-) polarisation	$P_-(P_+)$		80 %(30 %)	80 %(30 %)	80 %(30 %)	80 %(20 %)
Repetition frequency	f_{rep}	Hz	5	5	5	4
Bunches per pulse	n_{bunch}	1	1312	2625	1312/2625	2450
Site length	L_{site}	km	20.5	20.5	31	40

Higgs self-coupling determination as well as direct exploration of BSM physics at higher energies. The schematic footprint of CLIC near CERN and Geneva lake with its three implementation stages is shown in Figure 3.1 (right). One should note that since 2025 the 3 TeV scenario is no longer in the baseline implementation plan [5]. The possibility of a run at the Z pole, or at other centre-of-mass energies (250 and 550 GeV), has also been explored at CLIC [153, 5].

CLIC detector (CLICdet) concept [154] has been inspired by the ILC detectors, however, it was optimised to the operation parameters and experimental requirements of the CLIC accelerator at $\sqrt{s}$ =3 TeV. In particular, it was decided the TPC has to be replaced with an all-silicon tracker, mainly due to beam conditions at the highest energy stage. The current 2025 CLIC baseline design [5] assumes implementation with two IPs and two detectors that could operate in parallel (sharing the luminosity), or one at a time.

3.3.3 Linear Collider Facility

The most recent proposal for a linear Higgs factory is the Linear Collider Facility (LCF) [8] at CERN. It was first proposed in 2024 and presented in March 2025 as the input to 2025-26 Update of the European Strategy for Particle Physics (EPPSU), jointly with the Linear Collider Vision (LCV) report [135]. Linear Collider Vision is an initiative that builds on combined efforts of the whole linear collider community, in particular ILC and CLIC. The LCF was proposed as both a more affordable and flexible option for CERN than the FCC-ee. Its design is strongly inspired by the ILC and relies on the same technology, although several upgrades to the ILC baseline are assumed, resulting in increased energy and luminosity. The improvements would allow to collect $3\,\mathrm{ab}^{-1}$ and $8\,\mathrm{ab}^{-1}$ at 250 GeV and 550 GeV energy running, respectively; a Giga-Z stage at the Z pole and a dedicated $\mathrm{t\bar{t}}$ threshold scan are also foreseen.

The LCV initiative document [135] also describes a set of potential far-future upgrades for the LCF or another linear collider option at CERN. They consider the use of highly innovative technologies that are currently under development and would allow a possible increase in energy even to 10 TeV, or in luminosity to more than $10^{36}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ (about two orders of magnitude above the baseline ILC luminosity). These options are briefly introduced in the following section.

For completeness, the centre-of-mass energies and total integrated luminosities assumed for

Table 3.4: Comparison of centre-of-mass energies, and integrated luminosities of the ILC [3], CLIC [5], and LCF [8, 135] at different running stages. High-energy runs are considered possible extensions for each collider, but they are not in the baseline. Z pole run is included in the baseline only at the LCF.

	ILC		CLIC		LCF	
Run Stage	$\sqrt{s}$ [GeV]	$\int \mathcal{L}\,dt$ [ab^{-1}]	$\sqrt{s}$ [GeV]	$\int \mathcal{L}\,dt$ [ab^{-1}]	$\sqrt{s}$ [GeV]	$\int \mathcal{L}\,dt$ [ab^{-1}]
Z pole	91	0.1	91	0.008-0.14/yr.	91	0.1
Higgs factory	250	2	380	4.3	250	3
Top threshold scan	350	0.2	350	0.1	350	0.2
Beyond $t\bar{t}$ thr.	500	4	1500	4	550	8
High energy run	1000	8	3000	5	depends on technology	

the ILC, CLIC and LCF proposals at different running scenarios are presented in Table 3.4. Instantaneous luminosity as a function of centre-of-mass energy for different e^+e^- collider options (both linear and circular) is shown in Figure 3.3. For details, see the captions of Table 3.4 and Figure 3.3.

3.3.4 Novel accelerator technologies

More recently, other proposals for linear e^+e^- colliders have also emerged: the Cool Copper Collider (C^3) [155] and the Hybrid Asymmetric Linear Higgs Factory (HALHF) [156]. They incorporate technologies not as mature as the projects discussed so far, but present interesting options for the future. C^3 is the successor to the CLIC technology, featuring even higher achievable accelerating gradients on a more compact footprint, reaching 550 GeV with a site length of only 8 km. It is based on normal-conducting RF structures cooled cryogenically, which results in lower power consumption compared to CLIC. HALHF, on the other hand, is a novel concept that combines conventional RF cavity acceleration with plasma wakefield acceleration (PWFA). Although PWFA technology is still very far from industrial maturity in general, one of the main concerns in its context is the challenge of positron acceleration. HALHF addresses this issue by proposing an asymmetric Higgs factory, where electrons are accelerated to very high energies using PWFA and positrons to lower energies using superconducting cavities. However, if PWFA becomes available also for positrons, it could drive the c.m.s. energy of a linear machine even to 10 TeV [157].

An alternative direction to explore would be to mitigate the biggest disadvantage of a linear collider — beam loss after single collision resulting in the relatively low luminosity. The possibility of its increase by almost two orders of magnitude is offered by the concept of beam and energy recovery. Two technologies for a possible upgrade have been considered in the LCV, ReLiC [158] and ERLC [159], which would deliver the luminosity of 135-140 $\cdot 10^{34}$ cm^{-2}s^{-1} at a 250 GeV collider.

These future upgrade projects not only explore interesting and highly innovative technologies, but also demonstrate how linear colliders can evolve through advanced accelerator concepts rather than rely solely on larger infrastructure. For more details on the status of current R&D,

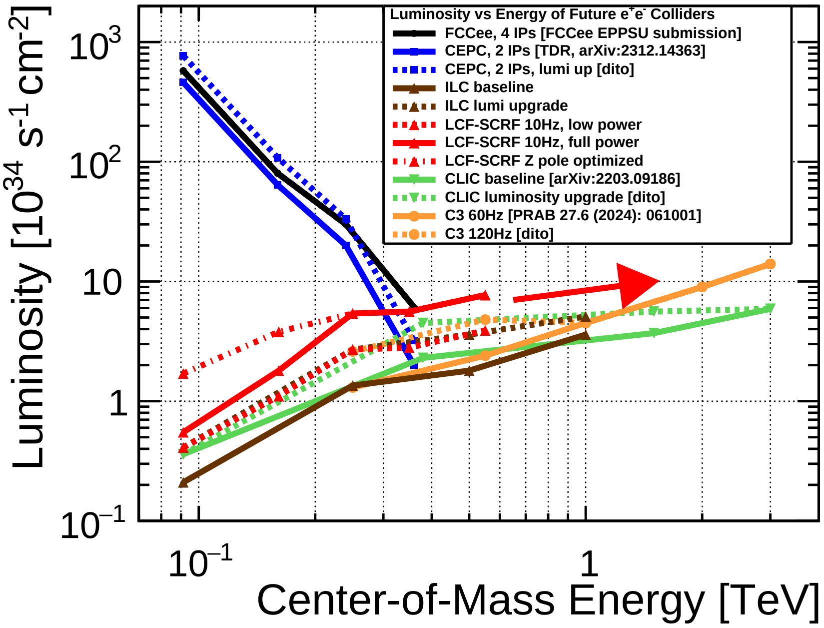


Figure 3.3: Instantaneous luminosity as a function of the centre-of-mass energy for various e^+e^- colliders. The LCF is drawn only up to 550 GeV, since the luminosity at higher energies will depend on the yet-to-be-chosen technology for these energies. However, they can be expected to be similar to the high-energy stages of CLIC and C^3, as indicated by the red arrow. Adapted from Ref. [135].

see Ref. [135] and references therein.

3.4 Beam-induced backgrounds at linear colliders

At linear e^+e^- colliders the beams need to be very strongly focused at the IP, to reach the required luminosity. Consequently, the charge density at the IP is very high. Therefore, besides quasi-real photons radiated by individual electrons (as described by Weizsäcker-Williams approximation), bunches are accompanied by synchrotron radiation caused by an interaction with the opposite beam, the so-called beamstrahlung. Both quasi-real (γ^W) and beamstrahlung (γ^B) photons emitted by one of the bunches can interact with those from the opposite bunch. Hence, each bunch crossing (BX) at a linear collider is a source of low-p_T hadron photoproduction ($\gamma\gamma \rightarrow$ hadrons) and incoherent e^+e^- pairs due to beam-induced photon interactions.

At the 250 GeV (500 GeV) ILC, around 1.55 (1.05) $\gamma\gamma \rightarrow$ hadrons events are expected on average per BX. In addition, $\mathcal{O}(10^5)$ incoherent pairs are produced per BX. However, they are predominantly emitted in the forward direction and only a small fraction of them enters the detector and can be measured. Both of these processes occur in the detector simultaneously with physical hard events (and hence they are typically called *overlay* events). However, with an order of 10^{11} BXs expected at the ILC per year, these events are also a potential source of a standalone background if BSM signals with soft final states are considered.

It should be underlined that the amount of overlay background strongly depends on parameters of the accelerator. For example, at CLIC, where the bunch repetition rate is much higher, these events have a huge impact on physics, and dedicated mitigation techniques must be implemented.

3.5 Potential of future Higgs factories for LLP searches

The experimental features of future Higgs factories are perfectly suited to LLP searches. As discussed in Section 2.3, the common factors leading the acquisition of macroscopic lifetimes in BSM models are small couplings, reduced phase space, or scale suppression. e^+e^- colliders have already demonstrated a strong sensitivity to such scenarios in many studies involving prompt searches [11]. Their clear experimental environment, triggerless operation[2] and well-defined initial state at Higgs factories could help overcome most common challenges associated with LLP detection and complement or improve the LHC reach.

The potential sensitivity of future Higgs factories to LLP detection has been extensively explored in the past, and this section aims to briefly describe some selected results. However, only a limited number of studies have been performed using full detector simulation which is essential to properly assess the experimental challenges posed by exotic signatures of LLPs. The majority of existing works are phenomenological analyses based on theory estimates. Occasionally, they also rely on fast simulation techniques, such as the `Delphes` package [160], that allow to take into account reconstruction efficiencies in a parametric way. However, the input efficiencies assumed for a detector are taken from full simulation studies, which typically only involve prompt decays and may not fully capture the detector-level effects or unexpected behaviour of reconstruction algorithms when facing the exotic displaced signatures. Therefore, in such analyses, the backgrounds may be underestimated — particularly because sometimes they are not taken into account at all, as the authors assume background-free regime — and the results should be interpreted with caution.

Phenomenology studies

Most studies are based on phenomenological calculations, using only Monte Carlo generators and making assumptions on detector size and dimensions (rarely also on reconstruction efficiencies). Such an analysis of sensitivity to sterile neutrinos has been performed for FCC-ee, CEPC and ILC [161]. All of these colliders have the best sensitivity to masses between 20 and 80 GeV, and the best limits can be obtained with the Z pole run at the FCC-ee.

In another study, the sensitivity to ALP decays to leptons and photons at the FCC-ee and CLIC has been explored [162]. Associated production with h, Z and γ has been tested, finding that both colliders present similar prospects for ALP masses in 10^{-3}-10^{2} GeV range. The FCC-ee performs better in the $e^+e^- \to a\gamma$ production channel, while the high-energy CLIC can extend the reach up to masses of the order of 1 TeV. An interesting analysis has also been performed for ALPs to compare sensitivities of the ILD at the ILC, Belle II and dedicated far detectors at an

[2]This is a definitive assumption only for linear machines; circular colliders, due to much higher collision rate, may require some form of a trigger system.

e^+e^- collider [119]. The study finds that the best results can be obtained with a regular search at the ILC performed using the detector located at the IP.

A scenario commonly explored in the context of future Higgs factories is the Higgs decays to LLPs, which is motivated by neutral-naturalness [163] and Hidden Valley [164] models. A phenomenological study has been conducted for the FCC-ee and CEPC considering these two models, using sub-GeV and $\mathcal{O}(10\,\text{GeV})$ scalar LLP decays to muons and jets [165]. The analysis finds almost identical sensitivity for both colliders, improving the bounds from the LHC.

Fast simulation studies

Fast simulation techniques can improve the reliability of phenomenology estimates to some extent. Studies of exotic Higgs decays to LLPs, similar to those described in the previous paragraph, have been carried out in a mass range 2.5–50 GeV and with the partial help of fast simulation techniques [166]. The results indicate that using hadronic decays of the LLPs in the range of decays lengths 10^{-3}–1 m, the Higgs branching ratios can be constrained down to the level of 10^{-4}, and in some scenarios even down to 10^{-5}.

Dedicated analyses have been also conducted to evaluate the FCC-ee sensitivity to long-lived ALPs [167], heavy neutral leptons (HNLs) [168] and Higgs decays to LLPs [169]. The ALP search relies on associated production with photons, with the ALPs decaying either invisibly or into photon pairs, showing that the FCC-ee could improve the current limits by two to three orders of magnitude in the mass range of 0.1–100 GeV. The HNL study explores the Z pole run and the semi-leptonic final state, suggesting that in this way the FCC-ee could probe HNL masses up to around 65 GeV, improving limits from prompt searches by two orders of magnitude. Finally, the analysis concerning exotic Higgs decays studies long-lived dark scalar decays to b jets, with the LLP masses between 20 and 60 GeV. The study suggests the FCC-ee has the required sensitivity in this range, allowing to probe Higgs branching ratios to LLPs as low as 10^{-4} for decay lengths around 1 m.

Another `Delphes` analysis studied ALP production in $\gamma\gamma$ interactions at different centre-of-mass energies at FCC-ee and ILC [170]. The FCC-ee can achieve an excellent sensitivity in the Z pole run for ALP masses below 90 GeV. On the other hand, ILC is superior for higher masses, almost up to the kinematic limit — so even 1 TeV if one takes into account the ILC running at this energy.

Full simulation studies

Only a few existing studies on the potential for LLP detection at future colliders are based on full simulation of the detector response. These analyses have been performed for CLIC, CEPC, and the SiD detector for the ILC, and all of them concern exotic Higgs decays to LLPs. The CLIC analysis studies the Hidden Valley model at 350 GeV and 3 TeV, where the Higgs decays to dark pions that hadronise leading to a displaced jet signature [171]. The resulting expected limits show that CLIC could exclude such decays down to branching ratios below 10^{-4} for dark pion masses of 25-50 GeV and their lifetimes 1-300 ps ($c\tau \approx$ 0.3-90 mm). The CEPC study incorporates artificial neural networks to approach the Higgs decays to LLPs and their decays to jets and neutrinos [172]. They find sensitivity at the level of $\mathrm{BR}(h \to \chi\chi) \simeq 10^{-6}$ in the range of χ lifetimes 1-10^3 ps ($c\tau \approx$ 0.3-300 mm) for LLP masses: 1, 10 and 50 GeV. Finally, the SiD analysis concerns Higgs decays to dark photons in the mass range of 0.1-10 GeV [173]. With the assumption that the selection is tight enough to operate in the background-free regime, and using dark photon decays both to jets and leptons, the authors find that the SiD could be sensitive to branching ratios down to the level of 10^{-5} in the range of decay lengths between 1 mm and 1 m.

Chapter 4

The ILD experiment

The work presented in this thesis has been carried out in the framework of the International Large Detector (ILD) concept and this chapter describes its baseline design and software ecosystem, including the techniques used for simulation and reconstruction. The ILD was first proposed [174] as one of the experiments for the ILC and a design suitable for operation at a circular machine has been presented recently as well [144], although its potential operation in the Z pole run is still under study [175]. The ILD is a general multipurpose detector, which relies on almost continuous tracking and highly granular calorimeters, and is optimised for the concept of particle-flow reconstruction [151]. The detector has a barrel shape and a layered structure, typical of collider detectors (such as ATLAS or CMS), with vertex and tracking detectors closest to the IP, surrounded by calorimeters, and followed by muon spectrometers as the outermost part. Visualisation of the concept is shown in Figure 4.1.

4.1 Detector design

4.1.1 Tracking detectors

In its baseline design, the tracking system of the ILD is composed of four subdetectors: vertex detector (VTX), silicon internal tracker (SIT), time projection chamber (TPC) and silicon external tracker (SET). The dimensions and main parameters of each of the subdetectors, as well as technologies used for the respective systems, are presented in Table 4.1. The layout of the inner tracking detectors is shown in Figure 4.2 (left). Quadrant of the full ILD cross section, including the calorimeters, magnetic coil, and muon spectrometer, is presented in Figure 4.2 (right).

Silicon detectors

The VTX is a very precise silicon pixel detector consisting of three double layers, for which the point resolution $\sigma_{r\phi} = \sigma_z = 3.0\,\mu$m (where $r\phi$ indicates the transverse plane and z the longitudinal direction) is the key target parameter for flavour tagging and track separation. The VTX thickness must also be very low to minimise multiple scattering in the material that could lead to track misreconstruction. The first layer is placed as close to the IP as possible, with a radius r_{in} = 16 mm, to improve the efficiency of flavour tagging. The considered technologies for

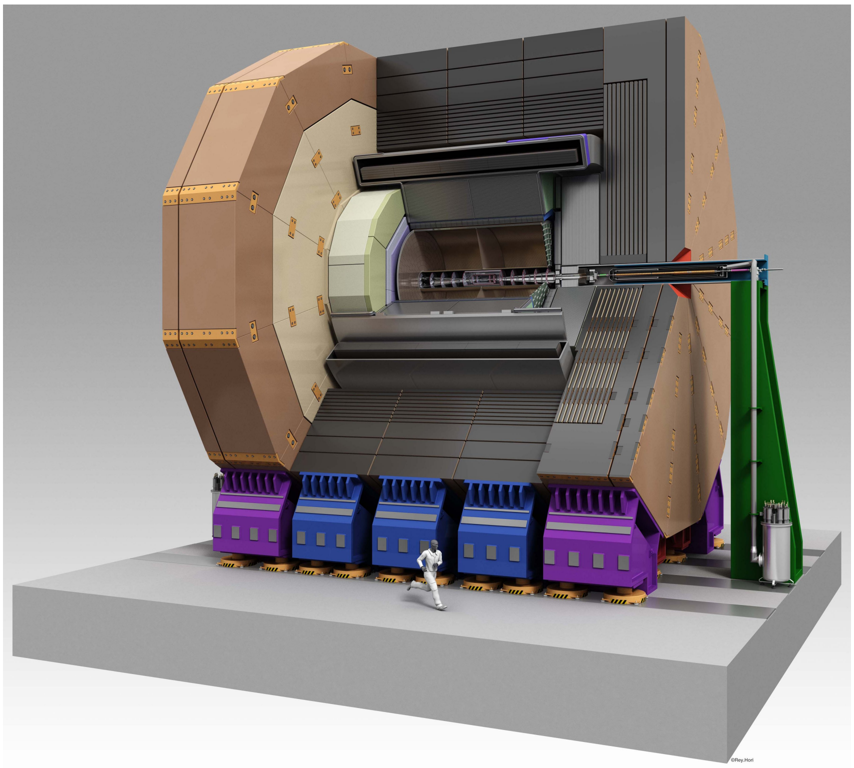


Figure 4.1: Visualisation of the ILD concept as presented in 2013 ILC TDR [145].

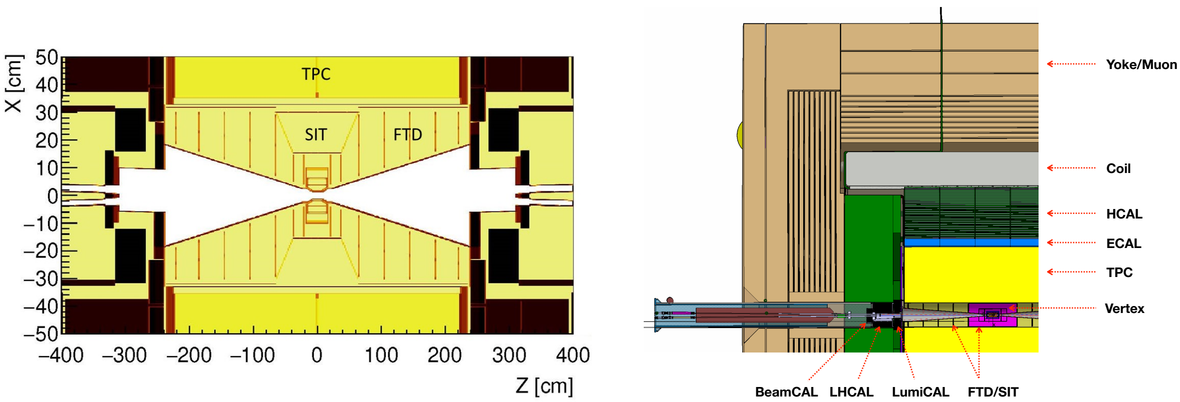


Figure 4.2: Left: layout of the ILD inner tracking systems in the x–z coordinates, with the SIT and FTD surrounding VTX, and encapsulated by the TPC. Right: r–z view of an ILD quadrant. The IP is located in the bottom-right region of the figure. Adapted from Ref. [12].

Table 4.1: Main parameters of the ILD tracking detectors for the barrel and endcap parts. Numbers correspond to the sensitive parts of the subdetectors. The inner and outer radii, r_{in} and r_{out}, are calculated as the orthogonal distance from the IP to the subdetector plane. z_{min} and z_{max} for the endcap refer to minimum and maximum distances of subdetector planes from the IP in the z coordinate; for the barrel z_{max} refers to the system's length in the z coordinate calculated from the IP (in one direction only). Resolutions are given for the transverse plane ($r\phi$), direction along the beam (z) and time (t), and ϕ_{stereo} is the relative tilt stereo angle between the FTD strips. Table adapted from Ref. [12].

Barrel system						
System	r_{in}	r_{out} [mm]	z_{max}	technology	comments	
VTX	16	60	125	silicon pixel sensors	3 double layers at r_0 $\sigma_{r\phi,z}$ = 3.0 μm σ_t = 2–4 μs	= 16 mm, 37 mm, 58 mm (layers 1–6)
SIT	153	303	644	silicon pixel sensors	2 double layers at $\sigma_{r\phi,z}$ = 5.0 μm σ_t = 0.5–1 μs	r = 155 mm, 301 mm (layers 1–4)
TPC	329	1770	2350	MPGD readout	220 layers	$\sigma_{r\phi} \approx$ 60–100 μm
SET	1773	1776	2300	silicon strip sensors	1 double layer at	r = 1774 mm
End cap system						
System	z_{min}	z_{max} [mm]	r_{out}	technology	comments	
FTD	220 645	371 2212	153 300	silicon pixel sensors silicon strip sensors	2 discs 5 double discs	$\sigma_{r\phi,z}$ = 3.0 μm $\sigma_{r\phi}$ = 7.0 μm ϕ_{stereo} = 7°

VTX are MAPS (Monolithic Active Pixel Sensor)rapidly developed in recent years, DEPleted Field Effect Transistor (DEPFET) pixels, and Fine Pixel CCD (FPCCD).

The VTX is surrounded by additional layers of silicon detectors. The recent progress in development of CMOS sensors will allow the usage of this technology also in two double layers of the SIT which surrounds the VTX and bridges the gap to TPC. The SET is one double layer of silicon strip detector surrounding the TPC, however, its exact design is still being discussed. For example, its use as the first layer of ECAL or implementation using technology with high timing resolution, σ_t, of the order of 10 ps are considered.

The VTX and SIT are extended in the forward direction by the FTD silicon discs perpendicular to the beam axis. The angular coverage of the FTD starts already at 4.8 degrees with respect to the beam. The first two discs are very close to VTX (see Figure 4.2, left) and could be implemented based on similar pixel sensor technology (MAPS, DEPFET, or FPCCD) providing the space point resolution of 3–5 μm. Additional five micro-strip sensors extend the FTD in the forward direction, occupying radii up to r_{out} = 300 mm near the inner wall of the TPC. The total of seven FTD layers span from z_{min} = 220 mm to z_{max} = 2212 mm along the beam axis, with respect to the IP, in each direction.

Time projection chamber

The large TPC as the main tracker is a very unique feature of the ILD. It not only allows to detect more than 200 points on a charged particle's trajectory providing almost continuous tracking, but also enables particle identification with $\mathrm{d}E/\mathrm{d}x$ measurements. In the baseline design, the readout structure consists of 220 layers of *pad rows* (or pad rings) with individual pads measuring 6×1 mm. The hit position in the TPC in the r–ϕ coordinates is determined based on the centroid of signal in a given pad row. Several candidate technologies based on Micro-Pattern Gaseous Detectors (MPGDs) are being considered for ionisation signal readout and amplification: Gas Electron Multiplier (GEM) foil, Micro Mesh Gaseous Structure (Micromegas).

An alternative readout option, not included in the baseline design, are the GridPix chips. This technology is based on Timepix3 pixelated readout ASIC (application-specific integrated circuit) and is currently under study. It still looks like a very challenging choice, as it assumes collecting the signal on a highly granular array of silicon pixels of 50 μm size, with amplification implemented as for the baseline pad-row readout with Micromegas [176]. However, it would also allow for further improvement of track reconstruction precision and $\mathrm{d}E/\mathrm{d}x$ [1].

4.1.2 Calorimeters

The electromagnetic calorimeter (ECAL) is compact and highly-segmented in order to be well-suited for the particle-flow reconstruction. The absorption material is tungsten and the sensitive layers are silicon sensors with pads of about $5 \times 5\,\mathrm{mm}^2$ size, assuming 30 layers with the option to decrease it to 26 or 22 if cost reduction is needed. Further savings and reductions in the number of readout channels can be made by replacing some of the layers with scintillator strips, and silicon photomultipliers (SiPMs) for readout, which results in slightly lower granularity.

The hadronic calorimeter (HCAL) surrounding the ECAL is planned with steel absorber and based on similar principles of high granularity and optimisation for particle flow. Two technologies are under consideration: the SiPM-on-tile and resistive plate chambers (RPCs). The former, with scintillator tiles of $3 \times 3\,\mathrm{cm}^2$ size, provides slightly lower granularity, but offers readout with an analogue signal (hence the acronym, AHCAL). The latter option gives high granularity of $1 \times 1\,\mathrm{cm}^2$, however, also less precise energy measurement with its semi-digital readout (SDHCAL).

The very forward region of the detector is equipped with additional specialised calorimeters that improve hermeticity of the detector and allow to measure integrated luminosity. High density silicon-tungsten LumiCAL provides the possibility to measure the electrons from Bhabha scattering, while the Luminosity Hadronic Calorimeter (LHCAL) closes the gaps in the end-cap hadronic calorimeter system and extends its reach to lower angles. Finally, the BeamCAL

[1] In highly-granular detectors, another promising method for particle identification is cluster counting ($\mathrm{d}N/\mathrm{d}x$) [177].

allows for fast beam position monitoring and angular coverage at even lower angles, down to 6 milliradians.

4.2 Software ecosystem

4.2.1 Event generation

The main generator used in the linear e^+e^- collider community, and hence also by members of the ILD group, to generate Monte Carlo samples of particle collision events is `Whizard` [178, 179]. It provides the calculation of the tree-level matrix elements and integration over the kinematic phase space, enabling the generation of fully simulated events with realistic observable distributions for the partonic final-state considered. One can also internally call PYTHIA 6 [180] or PYTHIA 8 [181] (linked as an external library) to simulate parton showering and hadronisation. Processes can be simulated in a simple way by defining the initial and final states, with the option to set the produced particles as unstable and specify their decays. The beam polarisation can be taken into account, as well as the proper beam spectra of different collider concepts, as parametrised with CIRCE [182].

For the BSM processes, `Whizard` provides a variety of implemented models. However, one can also link `Whizard` with models defined through the SARAH package [183], or take advantage of the UFO interface [184] (e.g., for models developed with FeynRules [185]), which allows the direct use of models in the UFO format. Using a UFO model is the most convenient option; it is defined as a set of Python scripts that contain information about all particles, the types of interactions between them, and all parameters. By loading the model in a `Whizard` script, the user gets access to all of the model's contents and can define processes with its particles and refer to its parameters.

Unfortunately, `Whizard` has no implemented mechanism for generating displaced decays — in its output every final-state particle originates from the $(0, 0, 0)$ coordinate. This issue can be partially handled using the `Pythia` interface in `Whizard` by setting the LLP as a non-resonant particle and fixing its width to a certain value. However, this solution works properly only for scalar particles whose PDG code is already implemented within `Pythia` 6.

4.2.2 Detector simulation

As in the majority of current high-energy physics experiments, simulation of the particle interactions with the detector material and of the response of its active components is performed with the `Geant4` package [186]. In this framework, the trajectory of a particle is calculated by iteratively solving certain differential equations and propagating particle in steps. All effects such as ionisation, multiple scattering, or secondary particle production through interactions with the material budget are taken into account. Detector geometry description is handled with the `DD4hep` toolkit [187]. It is based on `Geant4` libraries, while allowing for much simpler use, and provides a single framework to manage geometry, materials, visualization, readout, alignment, calibration, reconstruction, etc. The linear collider-specific extensions of `DD4hep` have been implemented in the `lcgeo` package, which included compact detector descriptions (with all detector parameters, stored in the XML format) or scripts to run the simulation. Currently, `lcgeo` is used

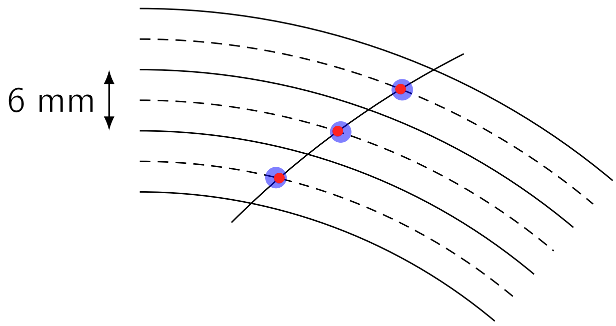


Figure 4.3: Schematic representation of simulated hits in the ILD TPC. Black solid lines are boundaries between pad rows and dashed lines show boundaries between the virtual volumes. Red dots represent the simulated hits placed in the middle of the pads, while blue dots are the smeared hits after taking into account detector resolution and digitisation. Figure taken from Ref. [176].

beyond the linear collider community, for example for the FCC-ee studies, and is now known as `k4geo`, being a part of the `Key4hep` software stack [188].

In the simulation process, `Geant4` calculates the exact trajectory of a particle and outputs the detector *hits* — coordinates of the point at which the interaction with the detector material occurred, together with the deposited energy, particle type, its momentum at this point, etc. The detector in `Geant4` is defined as a set of virtual sensitive volumes, that particles can interact with. After a particle enters the virtual volume, a detector hit can be placed in several ways: for example, at some point within a volume or exactly at the boundary between the volumes, depending on how the user wants to simulate the response of a specific detector type. These *simulated hits* created by `Geant4` exist only at the level of the MC simulation and do not correspond exactly to the hits in a real experiment, as detector resolution is not yet taken into account at this stage.

An interesting and important aspect for the work presented in this thesis is the case of TPC, where in reality the entire chamber filled with gas is a sensitive material. However, its simulation is performed in a more complex manner. The TPC in the ILD simulation is divided into 440 cylindrical sensitive volumes, coaxial with the chamber, each corresponding to a half of a readout pad row. In this way, if a particle crosses the boundary between the volumes, the hit is recorded at the point of crossing with this boundary, exactly in the middle of the row. This idea is presented schematically in Figure 4.3. However, one should note a clear issue related to this procedure — if a particle with low p_T, travelling almost tangentially to a virtual volume, enters it but does not cross the boundary in the middle of the pad, it will not create a hit, whereas in a real TPC it would be a source of signal in this pad row. These are the so-called *missing hits*. On the other hand, it is unclear to what extent one could use the signal in multiple pads along the row to reconstruct a track. An example event display of the simulated hits in the ILD, illustrating the missing hits problem, is presented in Figure 4.4. More examples of event displays can be found in Appendix B — in some of them, the missing hits problem is visible as well.

`Geant4` and `DD4hep` provide an alternative method to simulate displaced decays compared to the one described in Section 4.2.1. Although a predefined list of particles is implemented within

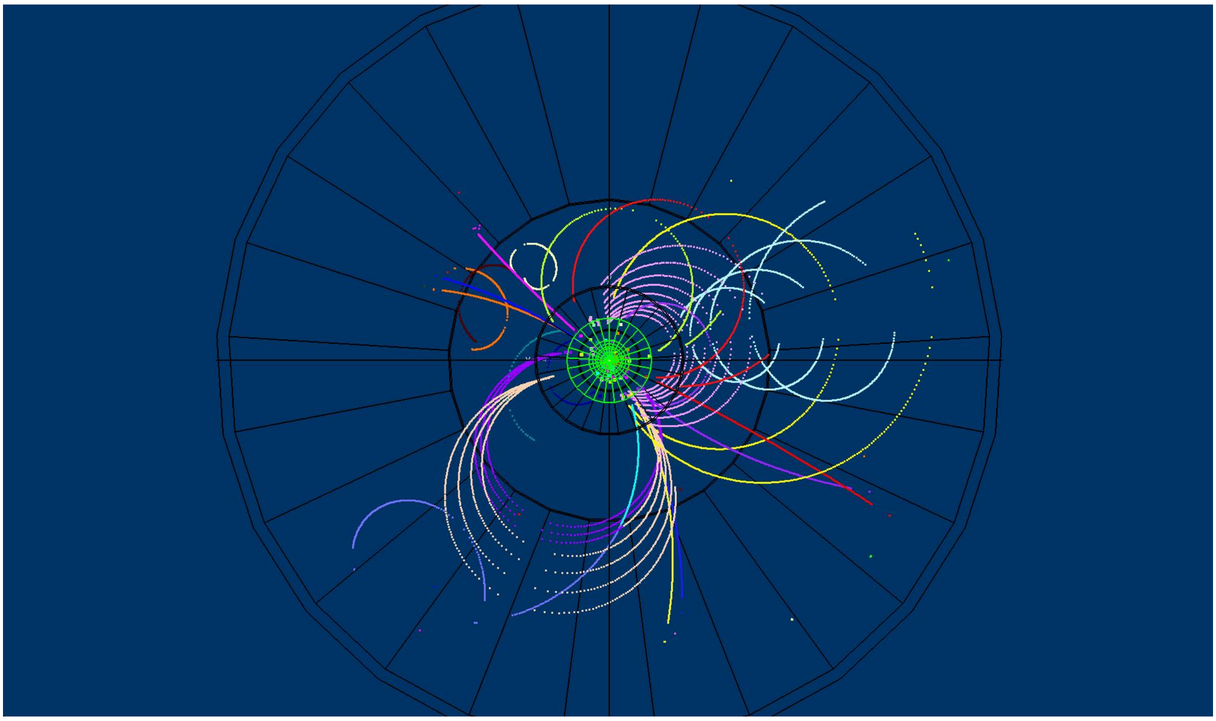


Figure 4.4: Simulated hits in an example event display of the $q\overline{q}$ production process in the ILD. The event presents the missing hit problem: when a particle travels parallel to a pad row, it does not produce any hits (see Section 4.2.2 and Figure 4.3). This is visible for most of the curling tracks. Only TPC cage is shown in the display for better visibility.

`Geant4`, `DD4hep` (and, in particular, `lcgeo` or `k4geo`) allows a text file containing a custom particle list to be passed as an input to the simulation. This list, in addition to PDG codes or masses of particles, must contain their lifetimes that can be arbitrarily specified by the user. If the PDG code and the mass of a particle in the list match the corresponding values of some MC particles at the generator level in the input event sample, `Geant4` will propagate these particles in simulation and decay them according to the exponential law with the mean lifetime provided in the particle list.

4.2.3 Reconstruction chain

Currently, most of the future collider experiments are migrating with their software ecosystems into `Key4hep` — and this includes not only experiments for future Higgs factories, but also Electron-Ion Collider or Muon Collider communities [188]. However, for more than two decades, the standard tools for linear e^+e^- colliders (since there were no other proposals for most of that time) have been developed within the `iLCSoft` software stack [189], originally built and maintained by and for the ILC community, later also joined by the CLICdp Collaboration. `iLCSoft` is now integrated into the `Key4hep` stack and is its important part. However, the results presented in this thesis have been performed only within the `iLCSoft` framework, since `Key4hep` was still in the early stages of development when these studies began.

The idea behind `iLCSoft` was to keep it as universal and modular as possible. As a result, most of the tools are written using the `MARLIN` framework (Modular Analysis and Reconstruction for the Linear collider) [190]. It allows for a dispersed reconstruction by subsequent operation of the so-called *processors* written in C++. The processors are arranged in a chain and each of them performs a single task, for example digitisation of hits in one of the subdetectors, or more complex, such as reconstruction of particle-flow objects (PFOs) using PandoraPFA libraries [151]. Although not exactly relevant in the context of this thesis, the particle-flow approach to calorimetry is key to meeting jet energy resolution goals for the ILD and ILC. It assumes that four-vectors of all visible particles in an event are reconstructed and the energy of a jet is calculated as the sum of energies of its individual particles. This requires highly granular calorimeters and a very efficient matching of tracks with deposits in the calorimeters, which allows the resolution of contributions from charged and neutral particles and their full reconstruction (so the formation of PFOs).

All data is transferred sequentially from one processor to the next, allowing processors further down the chain to access the output data produced earlier (more details on the data format is given in Section 4.2.5). Based on this approach, the standard reconstruction chain for the ILD has been organised [191]. First, the raw hits simulated by `Geant4` and `DD4hep` are taken as input, and their digitisation is performed, which is the simulation of the expected response of the detector, the readout electronics and the data acquisition chain (DAQ). The reconstructed hits are then used for tracking and other track-related operations. Subsequently, the digitisation of hits in the ECAL, HCAL, forward calorimeters, and muon detectors is performed, followed by running of particle flow algorithms that use the data from previous steps. BeamCal reconstruction is optionally performed next, and the chain ends with the reconstruction of high-level objects. In this section, some details are given on the digitisation, tracking, and the high-level reconstruction, since they are most significant in the context of the work presented in this thesis. Track reconstruction is the crucial aspect and is therefore described separately in a dedicated Section 4.2.4.

If the respective flag in the reconstruction chain is set, the beam-induced backgrounds can be overlaid on the input samples by a dedicated processor before the first step of reconstruction (digitisation) begins. For each type of $\gamma\gamma \rightarrow$ hadrons background ($\gamma^{B/W}\gamma^{B/W}$), the events are added from the corresponding overlay samples according to the Poisson distribution with the mean value specified for each background type. A similar procedure is used for the incoherent e^+e^- pairs, but their distribution is already taken into account at the stage of sample generation, so in this case exactly one background event is overlaid per input sample event. Thanks to this approach, the overlay does not need to be simulated each time together with a given physics sample; instead, it is simulated separately and taken into account during the reconstruction stage.

Digitisation

The simulated digitisation, by relying on the specific detector characteristics and its expected resolution, aims to produce *reconstructed* tracker hits similar to those expected in a real experiment for reconstruction of tracks or $\mathrm{d}E/\mathrm{d}x$ and ToF measurements. The pairs of hits measured in the silicon strips are combined into three-dimensional space points. In order to take into account the detector resolution, displacement (offset) of individual simulated hits is applied by a random number drawn from the Gaussian distribution with the standard deviation of the distribution equal to the respective resolution. In the TPC, smearing in both longitudinal, z, and tangential, r–ϕ, directions is applied depending on the hit position, which corresponds to resolutions related to electron drift and readout granularity.

Tracking

Track reconstruction for the ILD is divided into several steps with different processors performing pattern recognition and fitting. Track segments are reconstructed separately in different subdetector parts and merged at the final stage. The `SiliconTracking` processor performs pattern recognition in all inner silicon detectors (VTX, SIT and FTD), whereas the `ForwardTracking` processor searches for tracks only in the FTD. To improve the performance in the forward direction, both of these processors use the FTD hits, and the best subset of track segments identified by the algorithms is then selected for further use. The `Clupatra` processor is used to find tracks in the TPC. In the last step, the `FullLDCTracking` processor makes an attempt to match and combine the tracks produced by all processors, and refit them globally [192].

At this point in the ILD reconstruction chain, other processors that operate on tracks are also run, and the most important in the context of this thesis are `V0Finder` and `KinkFinder`. `V0Finder` attempts to identify and reconstruct the so-called V^0 particles, which is a collective term for long-lived neutral hadrons and converting photons; the algorithm tests the mass hypotheses of different particle species (γ, K_S and Λ) for every pair of tracks that originate sufficiently close to each other and pass certain kinematical criteria. On the other hand, `KinkFinder` is an analogous tool, but aimed at searching for kinked tracks. Its strategy is similar — it attempts to find the kink and reconstruct the vertex, and identifies the particle by testing mass hypotheses for $\pi^{\pm}$, $K^{\pm}$, $\Sigma^{\pm}$ and Ξ^{-}. However, its advantage is an option to categorise and output separately kinked track candidates that do not match any of the SM particles, which makes it useful also in the case of BSM searches.

High-level reconstruction

The last step in the reconstruction chain contains high-level processors that cannot be used earlier, mainly because they require PFOs as input. This includes finding resonances with the $\gamma\gamma$ final state, additional particle identification based on likelihood profiles, vertex finding using the `LCFIPlus` flavour tagging processor [193], or TOF estimation.

The studies presented in this thesis partly take advantage of the `RecoMCTruthLinker` processor which allows to link the reconstructed objects, such as tracks and PFOs, to the generator-level information. PFOs are already related to clusters and the simulated hits to MC particles, therefore it is sufficient to link tracks and clusters to simulated hits. The relation is defined by a weight that measures the strength of the link between two objects. For tracks, the weight is the ratio between the number of hits used for track fitting and the number of all hits created by an MC particle. For example, if 70% of simulated hits created by a particle (after digitisation) was used to fit a track, the weight of link between the two will be equal to 0.7.

4.2.4 Track reconstruction at the ILD

The general procedure for track reconstruction at the ILD and its position in the reconstruction chain were briefly described in the previous section. In this section, more details are given on track parametrisation, tracking algorithms and pattern recognition used at the ILD, and the resulting tracking performance.

Track parametrisation and the Kalman filter

Charged particles in the magnetic field move along a helical trajectory, which can be fully described by a set of parameters. The parametrisation used throughout this thesis follows the definitions in Ref. [194], which is the standard for the software used in the ILD. It features five track parameters: $d_0, z_0, \phi_0, \Omega, \tan\lambda$, defined in two planes: xy and sz. x, y, z are regular coordinates in the Cartesian system with the origin in the IP and z parallel to the beam axis, whereas s is the arc length in the xy plane between two arbitrary points. The track is therefore represented by a straight line in the sz plane. The parameters are defined at a point $\mathbf{P^0}$ (that can be any point on the helix), with respect to the *reference point*, $\mathbf{P^r}$ (or its z component, P_z^r), which can also be completely arbitrary. In practice, the reference point is usually set either to $\mathbf{P^r} = (\mathbf{0}, \mathbf{0}, \mathbf{0})$, or to the position of the first or the last hit of the track. $\mathbf{P^0}$ is then set to the point of closest approach (p. c. a.) from the reference point to the helix. In this way, the choice of a specific reference point corresponds to particular values of parameters defined at $\mathbf{P^0}$, which can be used, e.g., to calculate the track momentum at $\mathbf{P^0}$. In the following, p.c.a. will be used as $\mathbf{P^0}$.

The parameters, presented schematically in Figure 4.5, are defined below.

- ϕ_0 is the azimuthal angle of the momentum of the particle at $\mathbf{P^0}$
- d_0 is the signed impact parameter in the xy plane. Its sign is defined with respect to the reference point: in the projection such as in Figure 4.5, while looking from $\mathbf{P^r}$ to $\mathbf{P^0}$, d_0 is positive (negative) if the particle travels from left to right (right to left).
- Ω corresponds to the curvature of the helix, with the relation

  $$R = \frac{1}{|\Omega|},$$

  where R is the radius of the helix projection to the xy plane. The sign of Ω is related to the direction of a particle. If the magnetic field is exactly aligned with the z axis, with its longitudinal component $B_z > 0$, $\Omega > 0$ for a clockwise motion of the particle.
- z_0 is the track position in z at p. c. a. with respect to P_z^r, and therefore it is the signed impact parameter in the sz plane.
- $\tan\lambda$ is the slope of the straight line representing the track in the sz plane ($\mathrm{d}z/\mathrm{d}s$), which makes it constant along the track.

These parameters can be directly used to calculate the p_x, p_y, p_z components of the track momentum and charge q, knowing the magnetic field strength B_z and assuming it is parallel to the

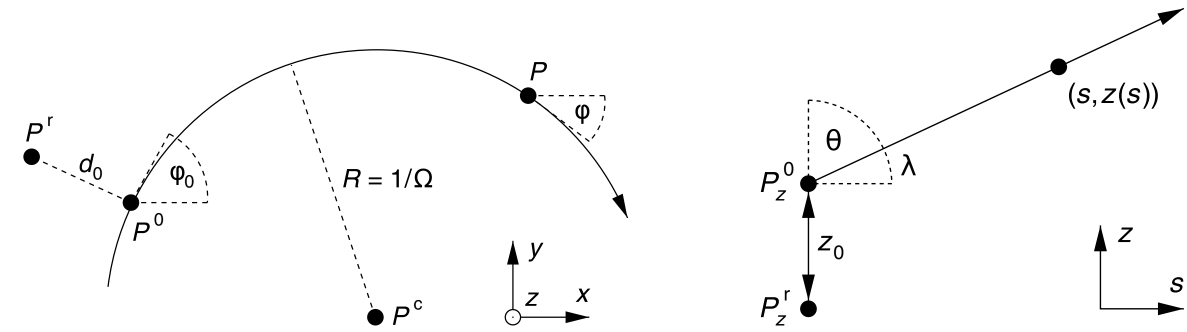


Figure 4.5: The projection of a helix section onto the xy (left) and the sz (right) planes. See text for details. Taken from Ref. [194].

z axis:

$$p_T = \alpha \left| \frac{B_z}{\Omega} \right|,$$

where $\alpha = c \cdot 10^{-15}$ is for the unit conversion and c is the speed of light. Now

$$\begin{aligned} p_x &= p_T \cos \phi_0, \\ p_y &= p_T \sin \phi_0, \\ p_z &= p_T \tan \lambda, \end{aligned}$$

and q corresponds simply to $\operatorname{sgn} \Omega$. Therefore, charge determination is only possible for a well-defined direction of motion of a particle. If the direction is not clear, two equivalent sets of parameters can be used to describe the helix.

While pattern recognition allows to identify and assign the hits to particular tracks, in order to obtain the set of parameters describing the helix, one needs to fit them to the set of measured hits. However, the energy loss of the particle in the detector material must be taken into account, which makes the simple least-square method highly ineffective. Therefore, the standard fitting procedure in particle physics experiments, including the ILD, is based on the Kalman filter [195]. It is an iterative method designed for systems with a small disturbance, originally invented to track ballistic missiles. The vector of helix parameters defines the *track state* at a measurement point (hit) and is propagated to subsequent sites by a nonlinear function, taking into account small noise. Therefore, the parameters can differ at each measurement point, which is the main advantage compared to the standard χ^2 fit. The Kalman filter for the ILD was implemented in the `KalTest` package [196, 197], which is now used beyond the ILC community, e.g., for FCC-ee (CLD) or CEPC studies. In the `LCIO` data format, which is the standard in the community (see Section 4.2.5), four track states are stored for each track: two corresponding to first and last hits in the fit, and two corresponding to helix extrapolations to the IP region and to the calorimeter.

Tracking processors at ILD

The `SiliconTracking` processor relies on considering triplets of hits and testing whether they can fit a helix. The search for triplets is performed in bins of polar and azimuthal angles; once the triplet has been identified, an attempt is made to extend it by adding more hits that pass the

χ^2 criterion and hence fit the helix. The resulting tracks are sorted according to their total χ^2 and those with higher values, that share the hits with tracks with lower χ^2, are discarded. This procedure is first performed in the VTX and SIT (effectively considered as one subdetector) and subsequently in all FTD layers. In the last step, the track segments found in these two cases are merged based on their angular separation and on the result of refitting with hits from both tracks taken into account.

The `ForwardTracking` algorithm is subsequently used to search for tracks just in the FTD but in a more sophisticated manner [198]. It is based on the cellular automaton (CA) model in which hits in two or more adjacent layers are connected into segments (or *cells*). First, all possible segments passing a set of criteria are created, and they are assigned an initial state value zero. Then, an inward search for *neighbours* of segments is performed; in order to be neighbours, the segments must share at least one hit and the state value, as well as pass some geometric criteria typical of tracks. The search is continued iteratively and in each iteration the states of segments for which a neighbour was found are incremented by one; the algorithm stops when no further changes are possible. Now, the state of each segment should be equal to the number of the corresponding detector layer (meaning that it is fully connected to segments on layers inward from it); otherwise, it is removed. This procedure is performed for both 2-hit and 3-hit segments, and the segment chains remaining after the removal are the candidate tracks to be fitted and accepted (or rejected) based on their compatibility and χ^2.

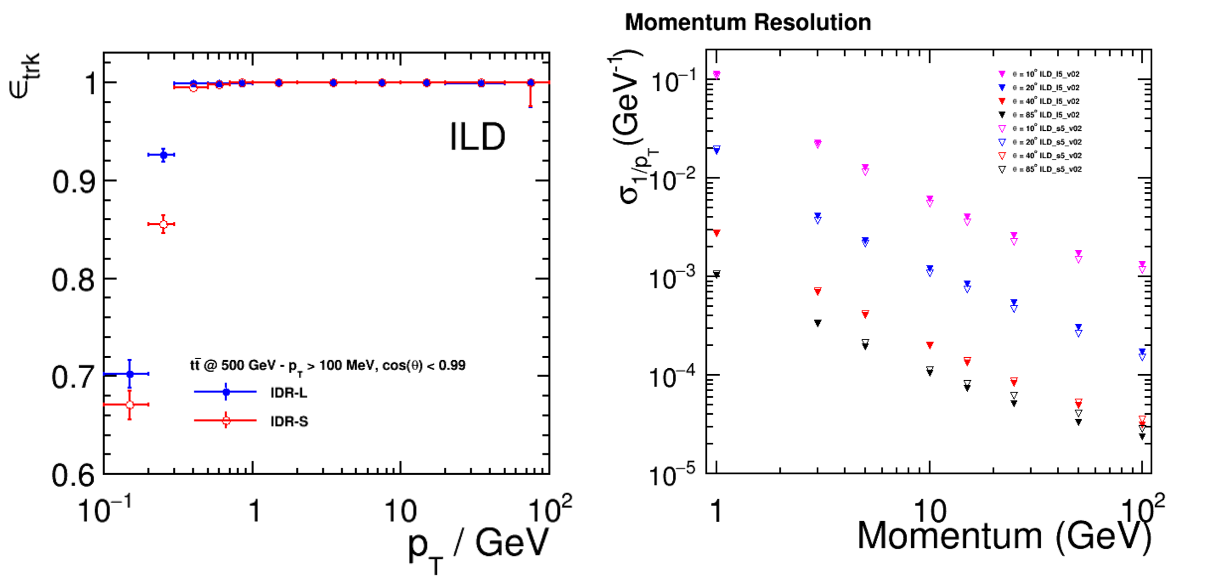
Since in this way two different processors can output the forward tracks, to select those of the highest quality the `TrackSubsetProcessor` is used next. This is done with the Hopfield Neural Network (HNN) [199], which is a recurrent network where all neurons are connected to each other. Each neuron corresponds to a track and has a state between 0 and 1 — the higher the state value, the bigger the acceptance probability. The state of a neuron is increased by the quality of the corresponding track (χ^2 of the fit) and decreased by incompatibility with other tracks (shared hits, proximity), so a strong link between two neurons reduces their states. HNN is trying to minimise the energy function, which is a sum of the qualities and incompatibilities across all neurons, by iteratively updating the states of neurons until they converge or the iteration limit is reached. Tracks with state values above 0.5 are accepted and saved.

The `Clupatra` algorithm was designed as a standalone pattern recognition tool for tracks in the TPC. This task should be rather simple, since the tracks are often visible just by looking at the raw TPC hits with bare eyes. However, due to the large number of hits, the standard methods described above are not optimal for this problem. Instead, `Clupatra` uses the nearest-neighbour (NN) clustering algorithm, where hits are merged into a cluster if they are sufficiently close to each other. Because the track density can be high close to the inner TPC wall, the hits are clustered in small ranges of pad rows, moving inwards from outer rows. Such clean *seed clusters* are further extended using the Kalman filter (hits are added if they pass the χ^2 criterion) both inwards and outwards. To account for highly collimated tracks (e.g., from γ conversion or τ decay), NN-clustering is rerun on leftover hits and the clusters are split based on hit multiplicities in particular pad rows and refitted. Finally, the track segments are merged based on the helices radii, the distance between their centres, and $\tan\lambda$. This is particularly important for very low-p_T tracks (curlers) since they often get split — mainly because of the missing hit problem described in Section 4.2.2.

The last step in tracking is to perform global merging, refitting, and the final selection of tracks with the `FullLDCTracking` processor. It first makes two attempts (with tight and loose criteria on track parameters) to merge segments from the TPC and silicon trackers. Tracks formed from exactly two segments are accepted. The algorithm also handles all the remaining tracks that have not been merged and attempts to combine them, including an option for forced merging of TPC and silicon segments. In the last step, the algorithm tries to assign leftover hits to the reconstructed tracks based on their distance from the helix extrapolation.

Tracking performance

Example results on tracking performance are presented in Figure 4.6. Shown are the efficiency of track reconstruction in $t\bar{t}$ events including the beam-induced backgrounds (left) and the inverse transverse momentum resolution for single muons (right). The efficiency is calculated for prompt tracks that originate less than 10 cm from the beam axis, have $p_T > 100$ MeV and are within the TPC acceptance (polar angle $\cos\theta < 0.99$); the result is nearly 100% efficiency above 300 MeV. The asymptotic resolution of $2 \cdot 10^{-5}$ GeV^{-1} is reached at momenta around 100 GeV, which meets the goal set by the Higgs mass measurement in the Higgsstrahlung process. Thanks to the small material budget of the TPC, good resolution at low momenta can be achieved [12].



(a) (b)

Figure 4.6: Left: Track finding efficiency for prompt (vertex radii less than 10 cm) tracks in $t\bar{t}$ events at 500 GeV as a function of as a function of transverse momentum p_T. Right: Tracking resolutions for single muons for the large and small ILD detector models. Inverse transverse momentum resolution σ_{1/p_T} as a function of momentum. The results are given both for the large ILD model (IDR-L, `ILD_l5_v02`), which is the baseline design, and for its alternative, smaller version (IDR-S, `ILD_s5_v02`). Figures taken from Ref. [12].

4.2.5 Data format, processing and production

The event data format used both at ILD and widely in the linear collider community is the `LCIO` [200, 201]. It aims to introduce a lightweight common way to store and exchange data in particle physics, providing classes for MC truth, simulation, reconstruction, and high-level information (such as relations between data elements), relevant for collider experiments. The `LCIO` implementations in C++, Java, and Fortran exist, while its Python bindings (`pyLCIO`) and `ROOT` [202] dictionaries are also available. The work presented in this thesis is fully based on the `LCIO` format and the analysis has been carried out with the help of `pyLCIO` and `ROOT` packages.

The ILD concept group performs the mass production of all relevant SM and some BSM processes using the computational resources of the European Grid Infrastructure and Open Science Grid. The SM samples at 250 GeV ILC (considering different polarisations) include two-, three-. four-, five-, and six-fermion production, as well as inclusive Higgs production in the e^+e^- collisions. In addition, two- and four-fermion production in $\gamma\gamma$ scattering has also been processed. Two-fermion and di-Higgs samples are available for the energy of 500 GeV. The overlay samples have been generated either with `Pythia` and a dedicated "Barklow generator" through the `Whizard` interface ($\gamma\gamma \to$ low-p_T hadrons) [203], or with `GuineaPig` (e^+e^- pairs) [204]. However, only detector response simulation has been performed for these events by the ILD group since they are typically reconstructed together with other physics events as the accompanying background.

The approach to sample generation in the ILD assumes the production of as many events as expected in the real experiment, and in some cases the produced event multiplicities even exceed those values. For the generated parton-level samples, the detector response simulation is further performed, and the events are later passed through the standard reconstruction chain of the ILD group (for more details see Sections 4.2.2 and 4.2.3). However, the full information about the events (all detector hits and calorimeter deposits) is stored only for a small fraction of the produced samples — these are the so-called REC samples. The majority of events is saved in the so-called DST files, where only the high-level information (such as tracks, reconstructed PFOs, or vertices) is kept. For example, one $q\bar{q}$ event in a DST file has an average size of around 0.12 MB, compared to 3.24 MB for a REC file. Considering that a total of 639.9 million $q\bar{q}$ events have been generated by the ILD (this corresponds to the luminosity of 5 ab^{-1}), only a small fraction of $q\bar{q}$ events could be stored in REC format due to storage limitations.

The grid resources are available to all members of the ILD concept group, and hence also to the author of this thesis. They can be interfaced with the `ILCDIRAC` framework [205], an extension to the `DIRAC` tool created for the Linear Collider community, allowing to manage distributed computing and job scheduling with simple Python scripts. All SM samples used for the work presented in this thesis have been produced by the ILD members, except for additional processing performed by the author when directly specified (for example, the reconstruction for standalone overlay samples). The most computationally expensive processing of the signal and background samples performed by the author has been done on the grid with the help of `ILCDIRAC`.

Chapter 5

Benchmark scenarios and reconstruction of displaced signatures

5.1 Benchmark scenarios and the sample production

From many possible signatures discussed in Section 2.3.1 two signatures of long-lived particles have been selected for detailed study in this work: displaced vertices and kinked tracks. As explained in more detail in the next chapter, the analysis described in this thesis was conducted from the experimental perspective and was meant to be model-independent to as large extent as possible. However, in order to test the experimental sensitivity, one has to generate the samples and for that some specific models are needed, e.g. in the UFO format (see Section 4.2.1). Therefore, for each of these signatures, a set of benchmarks have been selected and corresponding models identified allowing to produce samples with the respective topologies.

For each of the benchmark scenario considered, samples of 10,000 events were generated with `Whizard` [178] for the ILC running at $\sqrt{s} = 250\,\mathrm{GeV}$. The full simulation of the detector response was based on the ILD-L (`ILD_l5_o1_v2`) detector model [12] implemented in the `DD4hep` [187] framework and interfaced with `Geant4` [186]. Event samples were overlaid with beam-induced backgrounds and processed on the grid with the `ILCDIRAC` interface [205].

Reconstruction was performed using iLCSoft v02-02-03 [189] and based on the Marlin framework [190]. The standard ILD reconstruction chain [191] was used, however, selection cuts based on impact parameters d_0, z_0 (in the XY plane and Z coordinate, respectively [194]) of the reconstructed tracks have not been applied at the final step of track reconstruction, in the `FullLDCTracking` processor [192]. For this reason, the SM background samples had to be re-processed, which resulted in a reduced number of analysed events, compared to most of the ILD studies (see more in Section 6.2.2).

Three factors that can limit sensitivity to LLP production have been identified: soft (but very numerous) beam-induced background, hard[1] SM background, and detector or reconstruction limitations. Soft signals can be mimicked by the first of these backgrounds and the boosted topologies by the second one, while both scenarios are challenging for detectors and reconstruction techniques. Therefore, with the choice of benchmarks described below, all three potential

[1] Throughout the thesis, *soft* refers to low-p_T, and *hard* to high-p_T.

limitations are being tested.

5.1.1 Neutral LLPs

The main focus in this thesis is given to the neutral LLP production. The first class of analyzed scenarios involves low-p_T, little boosted tracks in the final state, whose combined momentum do not extrapolate close to the interaction point (IP). The model used to generate events with this signature was the Inert Doublet Model (IDM) [206], which introduces four additional scalars: $H^{\pm}$, A, and H, which could be both lighter and heavier than the SM-like 125 GeV scalar h. The lighter neutral scalar H is stable and is a DM candidate. The neutral scalars can be pair-produced in e^+e^- collisions, with a predominant further decay of $A \rightarrow Z^{(*)}H$; this process is shown in Figure 5.1 (left). If the mass splitting, $\Delta m_{AH} = m_A - m_H$, between A and H is small, the Z boson is highly virtual and A can be long-lived due to the small decay phase space available. Because H is invisible and escapes undetected, only the $Z^{(*)}$ decay products are expected to be measured inside the detector. The small mass splitting and the sizable mass m_A of the A boson implies that the final-state tracks are soft and are not pointing to the IP. We consider four benchmark scenarios with fixed LLP mass $m_A = 75$ GeV and generated decay length $c\tau = 1$ m, and different $\Delta m_{AH} = 1, 2, 3, 5$ GeV. For this model, the implementation in SARAH was used, interfaced with `Whizard 3.0.1`.

The second type of scenario under consideration was a production of a very light LLP, decaying into strongly boosted and almost colinear tracks. It was modelled using the photon-associated axion-like particle (ALP) production process, as shown in Figure 5.1 (right). In many scenarios, ALPs having macroscopic lifetimes for $\mathcal{O}$(GeV) mass scales could be copiously produced in e^+e^- collisions [119]. For this signature, we select four benchmarks as well, with ALP masses $m_a = 0.3, 1, 3, 10$ GeV and decay lengths $c\tau = 10 \cdot m_a$ mm/GeV to maintain large number of decays inside the detector volume.[2] The samples were generated with `Whizard 3.1.2` with the implementation of the model in the UFO format.

In both classes of scenarios, only decays to muons (of ALPs and $Z^{(*)}$) in the final state have been generated in order to speed up the simulation process, as muons do not produce showers in calorimeters in contrast to electrons or pions. As discussed in more detail in Section 6.1, this choice does not limit the generality of the analysis and its results. Taking advantage from the fact that in both cases the LLPs can be treated as scalars, the vertex displacement was produced at the generator level with `Pythia`, using the method described in Section 4.2.1.

In order to avoid confusion that these particular models are being studied — while they have only been used to produce the samples — throughout the thesis the scenario with low-p_T and little boosted tracks is referred to as *heavy scalars* scenario, and the strongly boosted and collimated track scenarios as *light pseudoscalars* scenario.

[2] Final results of this study will be presented for a wide range of LLP decay lengths. However, by generating signal samples with decay lengths optimal from the detector acceptance point of view we are able to preserve high statistical precision of the result also after sample reweighting to different LLP lifetimes; see Sec. 6.2.4 for details.

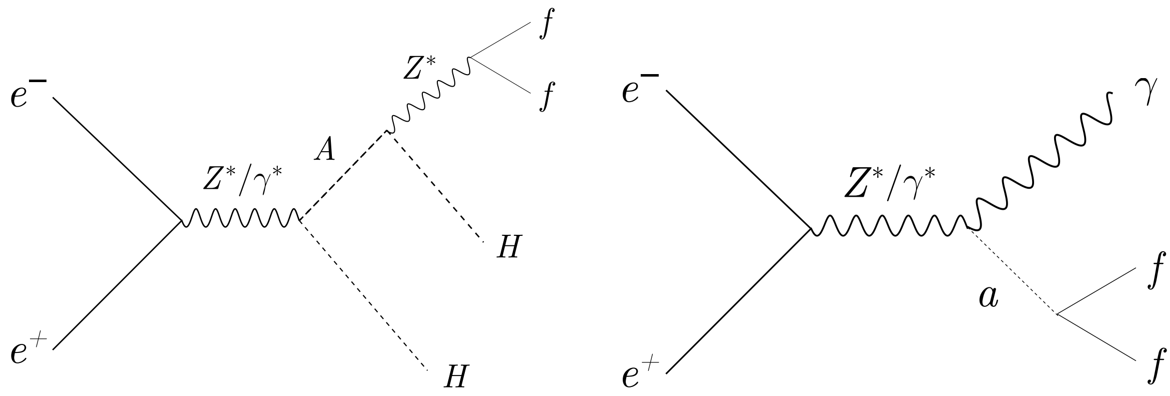


Figure 5.1: Leading Feynman diagrams corresponding to the processes used to simulate the considered benchmark scenarios to study neutral LLPs. Left: IDM neutral scalar pair-production with A as the LLP; H is invisible and Z boson is virtual for small Δm_{AH}. Right: Associated production of an ALP, a, with a photon where ALP can be long-lived.

5.1.2 Charged LLPs

The analysis methods developed for neutral LLPs analysis can also, to a certain degree, be extended to other signatures. Taking advantage of this observation, charged long-lived state production has also been considered in the study. For this purpose, a class of scenarios was selected to analyse both high-mass LLPs with small- and large-mass splittings, as well as lighter particles. To generate events, a model with an additional long-lived, vector-like fermion F and a scalar DM candidate s was used. It has been proposed as a simple "LHC-friendly" minimal freeze-in extension of the SM, where the particle s is a FIMP, and in one of the model instances[3] it couples to F via SM leptons [207]. This *leptonic* variant was used to generate the desired signatures — the pair-production mechanism of F and s in e^+e^- collisions is presented in Figure 5.2. This time, a total of six benchmark scenarios have been considered for the tests; two values of the F mass, $m_F = 60$ GeV, 110 GeV, with three values of DM masses: $m_s = m_F - 1$ GeV, $m_s = m_F/2$, and $m_s = 12$ keV, for each value of m_F. The selected benchmarks involve both slow and boosted LLPs and cover a range of mass differences. The $m_s = 12$ keV value was inspired by the model used to generate the samples, as this is the smallest DM mass allowed by the cosmological bounds [207].

In the case of charged LLPs, decays of F to electrons and muons were generated using `Whizard 3.0.1` with the implementation of the model in the UFO format. Since the LLP here is not scalar, `Pythia` could not have been used and the decays were prompt at the generator level. Instead, the displaced decays were simulated by `Geant4` using `DD4hep` (for details, see Section 4.2.2).

[3]The F transformation properties are different for quarks and leptons and in the original paper they were considered as separate models. Although such considerations are beyond the scope of this study, it is interesting to note that a significant fraction of the parameter space in the leptonic model is still not excluded by the LHC (and will not be by the HL-LHC). In particular, the allowed parameter space corresponds to the masses of F higher than 250 GeV and decay lengths $c\tau \sim 0.06$-300 m, which makes this model well suited for searches in high-energy e^+e^- colliders.

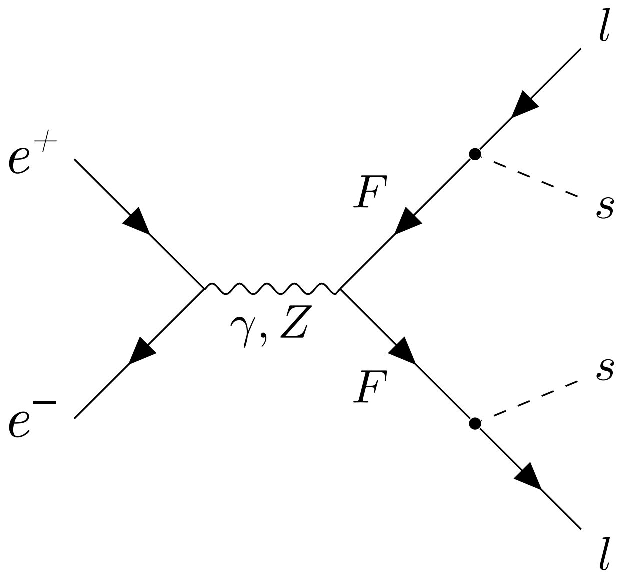


Figure 5.2: Leading Feynman diagram corresponding to the process used to simulate the considered benchmark scenarios to study charged LLP production. Long-lived dark fermion F decays to a lepton and a DM candidate.

5.2 Displaced vertex signature

5.2.1 Track reconstruction

As already mentioned above, the default cuts on track impact parameters, d_0 and z_0, used in the standard reconstruction chain in the `FullLDCTracking` processor had to be removed, as they have been strongly limiting the acceptance for high track displacements. The resulting efficiency of track matching to the MC truth is shown in Figure 5.3 for scenarios with the heavy scalars and the light pseudoscalars.

The tracks are considered matched if the angle (in three dimensions) between the reconstructed and MC momenta is smaller than 0.1, and if the respective charges are the same. The TPC is not sensitive to the track direction and for one helix two hypotheses are possible, corresponding to two oppositely charged particles with opposite momenta. For tracks with a large impact parameter, originating from highly displaced vertices and going inward, it is likely that they will be reconstructed in the wrong direction. This is taken into account in the efficiency calculation by checking the distance from the first and last hits to the MC vertex, and using the momentum corresponding to the track state for the hit closer to the vertex. The MC information is used here only to show pattern recognition and track fitting performance for displaced tracks; the track direction misreconstruction issue is properly taken into account at the vertex-finding stage. For more details on helix parametrisation and track states refer to Section 4.2.4.

In the case of heavy scalars the efficiency is above 80% almost throughout the whole TPC volume even in the most challenging $\Delta m_{AH} = 1\,\text{GeV}$ scenario. However, it is visibly lower for the LLP decays before the inner TPC wall. This is because it is more challenging to reconstruct the momentum of a curling low-p_T particle using a few hits in silicon layers, than with hundreds densely packed hits in the TPC. It shows that the TPC greatly enhances the acceptance for the soft displaced tracks.

In the second class of scenarios, thanks to the much higher track momenta, the efficiency is reaching nearly 100% in almost the entire detector volume. The reduced performance in some regions (not including the very edges of the detector acceptance) is due to the high collimation of the tracks. Hit point resolution in a TPC is lower than in silicon detectors, and therefore, for very small opening angles, hits from the two tracks cannot be resolved close to the LLP decay vertex. In such a case, the longer the tracks, the higher the chance they eventually become separable at some point, and the created hits could be used to reconstruct individual tracks. For this reason, for $m_a = 0.3\,\mathrm{GeV}$, the efficiency drops to about 50% (one of the tracks can still be reconstructed using the merged hits) already around the radii of 150 cm, and for heavier LLPs it remains high until the end of acceptance. The efficiency in the central part of the detector is slightly reduced because the hit resolution depends on the drift length and is the worst in this part of the TPC. Therefore, the tracks need to be longer for separation to be possible.

5.2.2 Displaced vertex finding

The algorithm for vertex finding, from now on called the `LLPFinder`, has been designed specifically for this analysis and implemented as a `MARLIN` processor. It was partly inspired by the `V0Finder` tool, but meant to be more general. The main assumption was also that it should use as few constraints as possible to be compatible with the model-independent approach, which ruled out, e.g., the use of kinematic fits.

Description of the algorithm

The algorithm tries to find a vertex considering each possible pair of tracks in an event. However, the track direction is not well defined in the TPC, and hence two hypotheses for each track are possible — a negatively charged particle going in some direction or a positively charged particle travelling in the opposite way. This also results from the helix parametrisation described in Section 4.2.4. Therefore, if one considers a pair of tracks, and assumes that they originate from a single vertex and have opposite charges, two hypotheses for each pair are possible. Since the momenta and charges resulting from the track parameters are defined by the *track states* in the reference points at the first and last hits of the tracks, the hypothesis selected is the one for which the distance between the corresponding hits is smaller. This is illustrated in the plot on the left of Figure 5.4.

Next, using the parameters corresponding to the selected hypothesis, the algorithm reconstructs a vertex using the distance d_{pca} between p. c. a.'s of track helices. This is done using `HelixClass::getDistanceToHelix` function from the `HelixClass` available in the `MarlinUtil` package. The function first calculates the distance between the centres of the helix projections in the xy plane and determines if the circles are separate, if the circles are one inside another, or if there are two intersections. In the first two cases, the p. c. a. is found on each circle individually, together with its z component. In the last case, the process is slightly more complicated. The $(x_{\mathrm{sect}}, y_{\mathrm{sect}})$ coordinates of intersection points are found using the parameters of only one of the helices, the one with higher momentum. For each track, two hypotheses (corresponding to each of the intersections) are found for the z_{sect} component based on the charge and the $\tan\lambda$

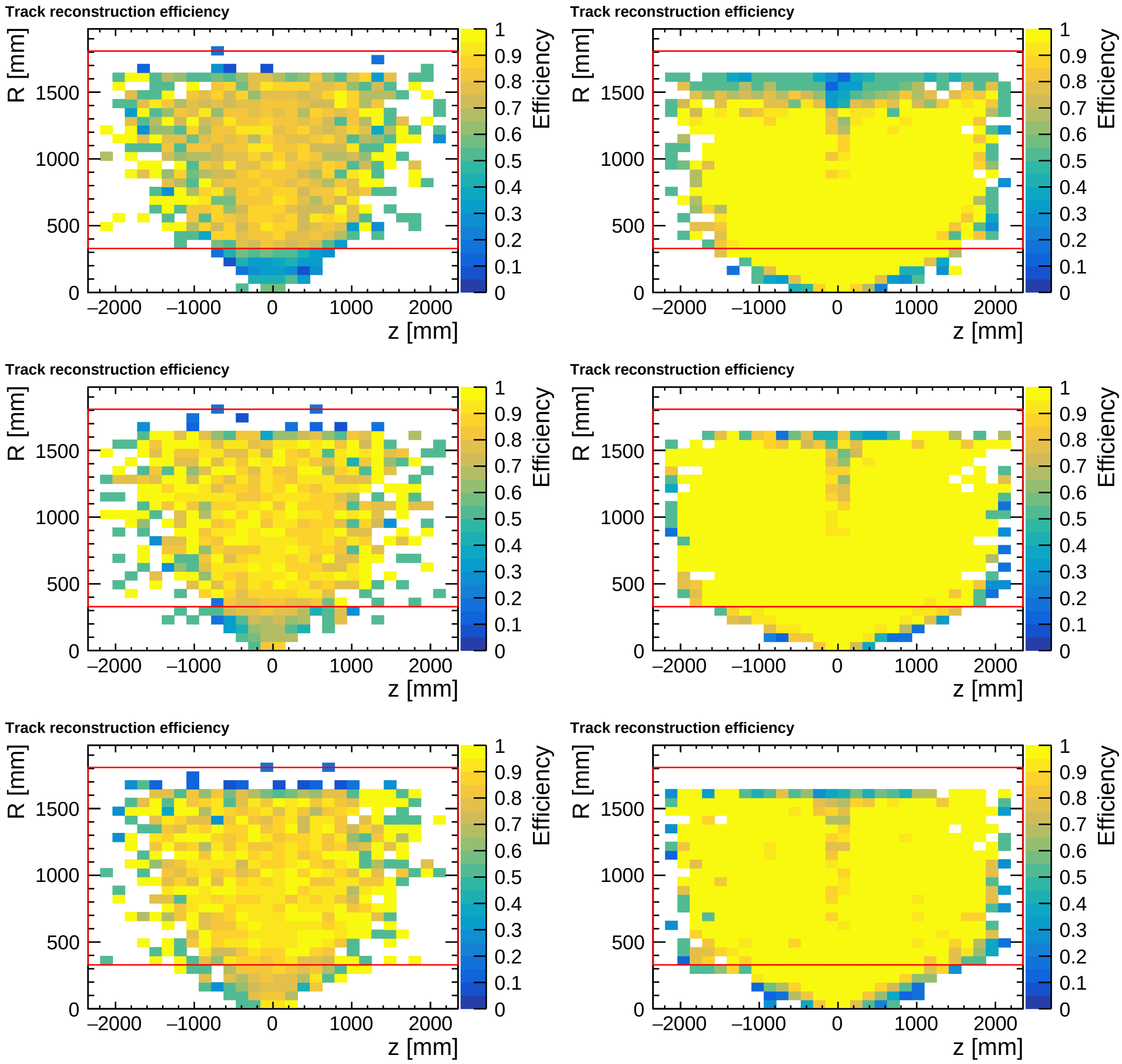


Figure 5.3: Track reconstruction efficiency as a function of the radius and the z coordinate of the LLP decay vertex position. Left: heavy scalar production scenarios with Δm_{AH} = 1, 3, 5 GeV (top to bottom). Right: light pseudoscalar production scenarios with m_a = 0.3, 3, 10 GeV (top to bottom). The red box corresponds to the TPC boundaries.

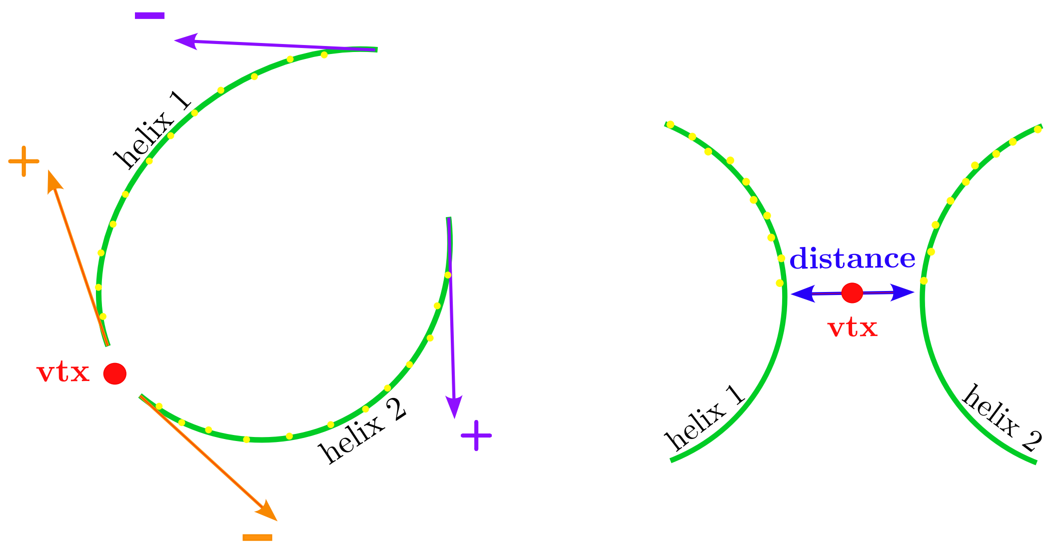


Figure 5.4: Schematic layout of the strategy in the vertex finding algorithm. Parametric helices are represented by the green lines, hits by yellow points, and the vertex candidate by the red dot. Left: two hypotheses for track directions and charges, each shown with different colour (orange and purple). Here, the orange variant would be used, as the corresponding hits are closer. Right: The placement of the vertex in the midpoint between p. c. a.'s of two helices. Note that this is a 3D representation, not the projection on the xy plane.

parameter. The intersection with a smaller distance between the two z_{sect} hypotheses is selected, and the intersection points on each helix are used as the p. c. a.'s. Finally, the vertex position is calculated as the midpoint between the p. c. a.'s, and the total momentum as the sum of the individual track momenta at these points. The placement of the vertex is schematically shown in the plot on the right of Figure 5.4. It should be noted that this approach, widely used in the whole reconstruction chain, may bring limitations in the case of certain topologies (e.g. very forward tracks). However, it provides high vertex-finding efficiencies (as shown later) and was therefore used for the presented study as well.

Nevertheless, for the study presented in this thesis, the class member functions in the `HelixClass` responsible for the initialisation of the helix and the calculation of its properties (`HelixClass::Initialize_Canonical`), as well as the `HelixClass::getDistanceToHelix` mentioned above, had to be modified. They repeatedly assumed that the reference points of the track helices are in the IP, at the (0, 0, 0) coordinate, while the `LLPFinder` in most cases uses reference points at the first and last hits of the track.

The exception to using hits as reference points is vertices with tracks that are not very highly displaced ($d_0 < 100$ mm for both tracks), for which the reference point in the IP is used. In this case, the algorithm also checks if there are no hits on tracks at radii significantly smaller than the reconstructed vertex ($r_{1,2} < 0.9R_{vtx}$, where $r_{1,2}$ are radii of the innermost hits of the two tracks).

For highly collimated track pairs, it can also happen that the closer pair of hits selected out of two possible hypotheses for vertex reconstruction is not the pair closest to the true vertex. This can be the case if the hits close to the production vertex merge, so the starting point of one of the tracks is far from the origin of the other track, while the endpoints of both tracks are still relatively close to each other. For this reason, after the vertex is reconstructed, the algorithm checks if its distance from the pair of hits used for its reconstruction is indeed smaller than the distance from the other pair. If not, the vertex position finding is performed again, using the hypothesis that was initially rejected.

The LLPFinder performance

Following the procedure presented here, additional cuts on the quality of vertices can be applied in the `LLPFinder`. The selection used later in this study is described in detail in Chapter 6. Figure 5.5 presents the vertex finding efficiency as a function of the decay position of the LLP, for heavy scalars and light pseudoscalars respectively, only with the requirement that the distance $d_{pca} < 25\,\mathrm{mm}$. The vertex is considered correctly reconstructed if it is less than 30 mm from the true MC vertex. The total efficiencies obtained for all scenarios, with the same requirements, are shown in Table 5.1 together with purities, calculated as the ratio of the number of correct vertices to all vertices found in the signal sample. The values in brackets were obtained considering only decays inside the TPC volume.

For the heavy-scalar case, it can be observed that the efficiency remains high almost within the whole detector volume. The vertex-finding performance in the silicon tracker is higher than the track reconstruction efficiency in the same region because even poorly reconstructed tracks can be successfully used for vertex-finding. The efficiency also increases with the mass splitting, which means it depends on the final-state boost.

In the case of light LLPs, correct vertex reconstruction can be very challenging due to the merging of hits close to the vertex position. As a result, the reconstructed vertex can be very far from its true position and therefore can be classified as false. This results in the low efficiency and purity for the $m_a = 300\,\mathrm{MeV}$ scenario. The performance is much better in the silicon detectors, thanks to the much higher hit-point resolution. The reconstruction efficiency, also inside the TPC volume, improves significantly with increasing mass of the LLP and decreasing collimation of the final-state tracks.

The efficiency is also presented in Figure 5.6 as a function of the radius of the LLP decay vertex and the opening angle θ_{vtx} between the final-state MC particles in the laboratory frame. For heavy scalars, the performance is good for almost any angle and radius value, except for small radii when θ_{vtx} is large. In the scenario with small mass difference, where the final-state tracks are much softer, this effect becomes even more significant, and the efficiency decreases also for smaller angles and higher radii.

For the light pseudoscalar case, the opening angle between the decay products is strongly constrained by kinematics and can become extremely small. Vertex reconstruction becomes particularly challenging for angles below approximately $\theta_{vtx} \simeq 0.006\,\mathrm{rad} \simeq 0.35^{\circ}$. However, the efficiency is influenced by several other factors, such as the decay position along the z-axis, so there is no sharp threshold below which reconstruction universally fails. For higher LLP masses, above the minimum angle allowed by kinematics, the performance is very good up to the radius of 1.6 m.

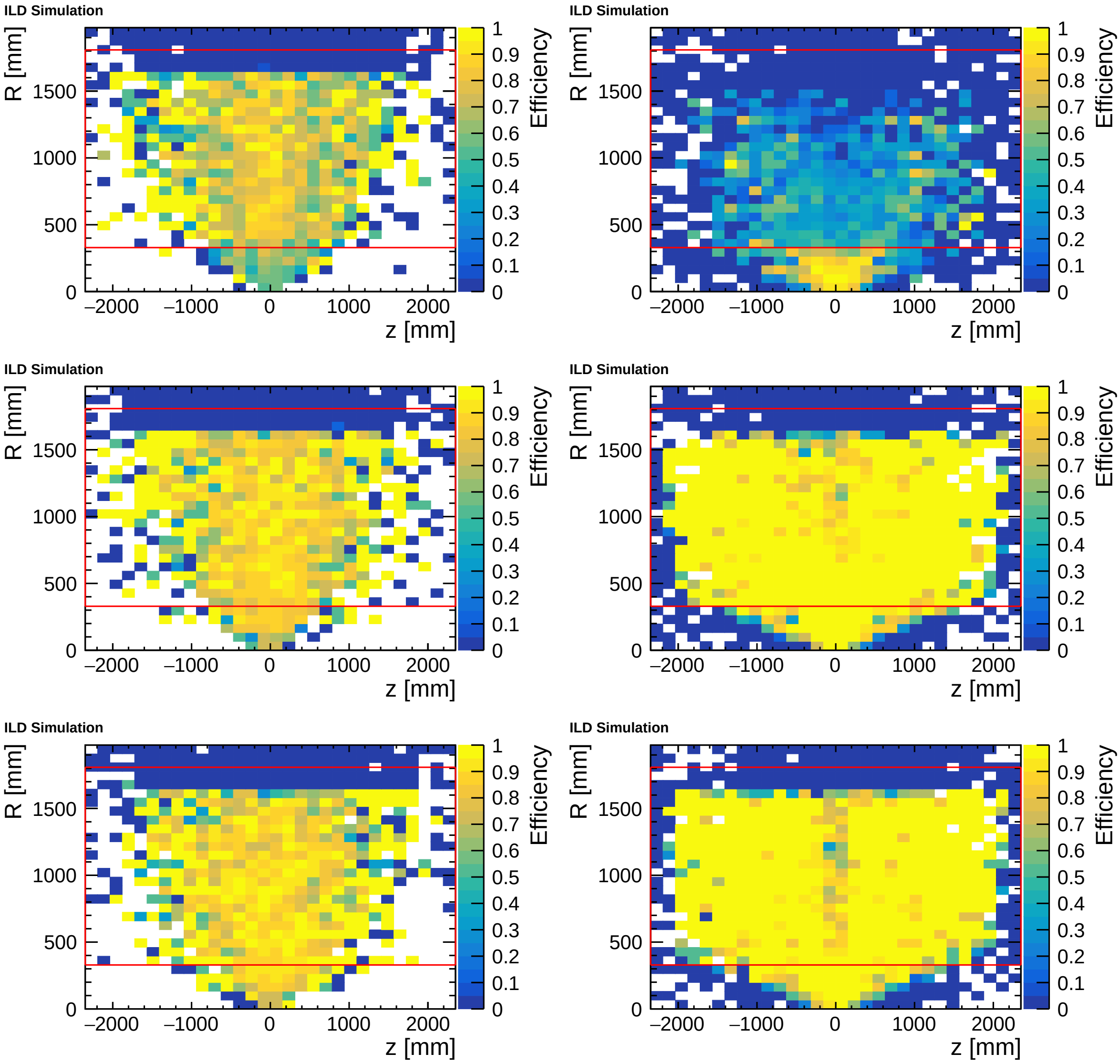


Figure 5.5: Displaced vertex finding efficiency as a function of the radius and the z coordinate of the LLP decay vertex position. Left: heavy scalar production scenarios with $\Delta m_{AH} = 1, 3, 5\,\mathrm{GeV}$ (top to bottom). Right: light pseudoscalar production scenarios with $m_a = 0.3, 3, 10\,\mathrm{GeV}$ (top to bottom). Only $d_{pca} < 25\,\mathrm{mm}$ requirement was applied. The red box corresponds to the TPC boundaries.

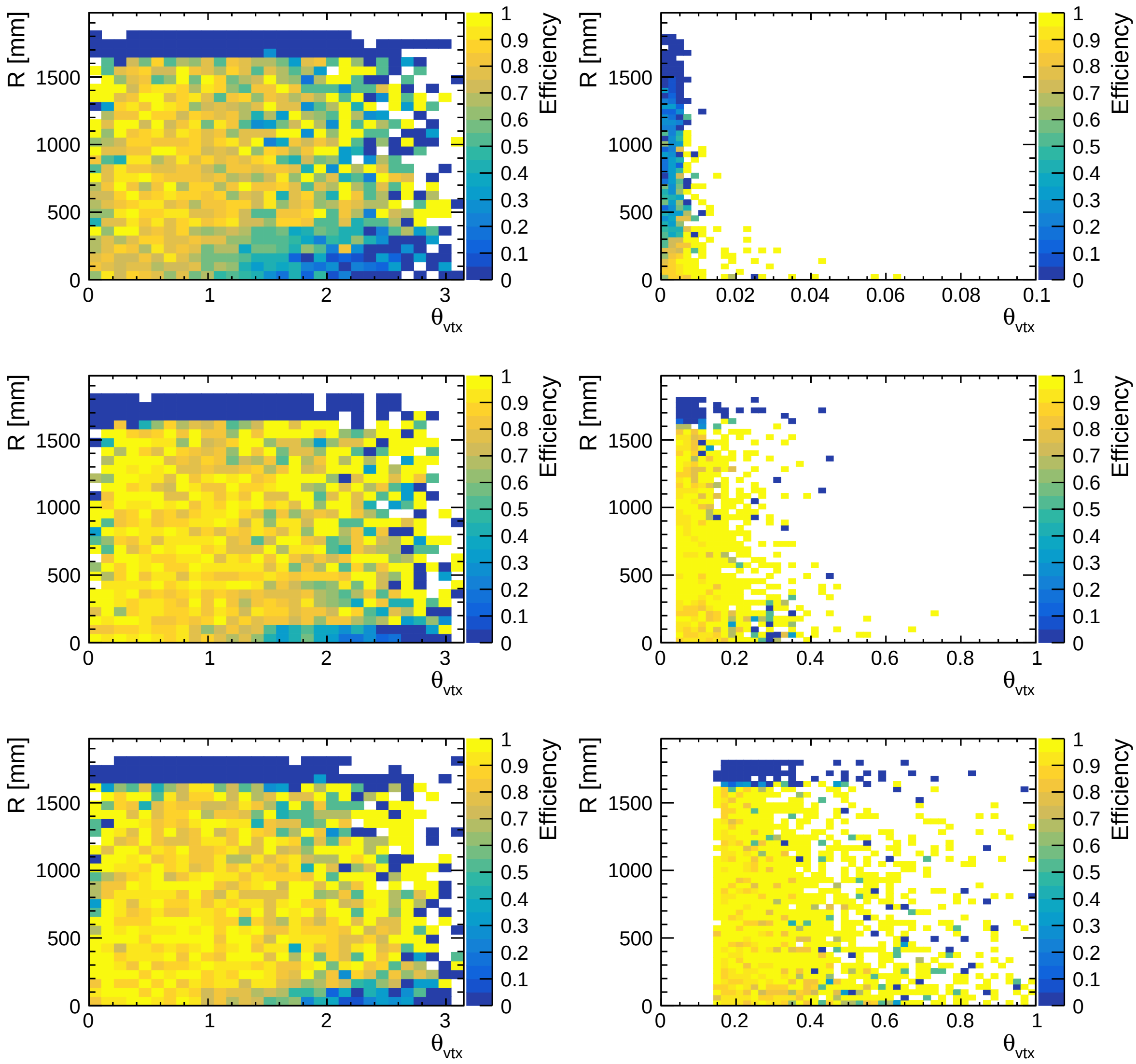


Figure 5.6: Displaced vertex finding efficiency as a function of the radius of the true LLP decay vertex position and the opening angle between final-state muons. Left: heavy scalar production scenarios with $\Delta m_{AH} = 1, 3, 5$ GeV (top to bottom). Right: light pseudoscalar production scenarios with $m_a = 0.3, 3, 10$ GeV (top to bottom). Note that a range only up to $\theta_{vtx} = 0.1$ and 1 is shown for light LLPs to indicate the region of small angles. Only $d_{pca} < 25$ mm requirement was applied.

Table 5.1: The efficiencies and purities of the displaced vertex finding procedure, both for heavy scalars and light pseudoscalars for all considered scenarios, after requiring only $d_{pca} < 25$ mm. The values in brackets represent those corresponding only to the TPC volume. See text for details.

Δm_{AH} [GeV]	1	2	3	5
Efficiency [%]	68.2 (72.6)	76.2 (77.5)	79.3 (80.2)	81.1 (81.3)
Purity [%]	79.8 (76.9)	87.7 (85.2)	92.7 (91.3)	95.0 (94.4)
m_a [GeV]	0.3	1	3	10
Efficiency [%]	50.3 (28.3)	81.3 (74.8)	90.7 (90.5)	89.4 (88.9)
Purity [%]	61.4 (42.3)	83.5 (81.7)	92.7 (96.6)	92.5 (94.8)

5.3 Kinked track signature

5.3.1 Track reconstruction

Track reconstruction for kinked track signature has been performed in the same way as for displaced vertices. The resulting efficiency of the displaced (daughter) track matching to the MC truth is shown in Figure 5.7, both for the scenarios with low- and high-mass dark fermion pair production.

As for the displaced vertices, the tracks are considered matched if the 3D angle between the reconstructed and MC momenta is smaller than 0.1 and the charge is the same. However, in this case, the weight in the relation between the true and reconstructed track was also required to be at least 0.8 (see the high-level reconstruction in Section 4.2.3 for details) and the distance between the first hit and the true vertex had to be smaller than 200 mm. This is to avoid mismatch between parent and daughter for small kink angles — since their charges are the same, and the directions very similar, it could happen that one is matched to another or vice versa.

In all scenarios the efficiency is high in nearly the entire detector volume, but visibly lower in the case of heavier LLPs. This is because the daughter track reconstruction efficiency decreases with the kink angle (cf. kink vertex reconstruction for different angles in Figure 5.9, Section 5.3.2), and for higher masses this angle tends to be larger. A slight drop in efficiency can be noted in the middle of silicon tracker — this corresponds to decays, where the daughter particles produced very few hits in the silicon detectors and these hits were not merged with the segment in the TPC. Hence, these tracks failed to pass the requirements for matching with MC.

One can note that the overall tracking efficiency for displaced tracks here is slightly lower than in the case of displaced vertices, even for scenarios with lower LLP masses and large mass differences. One of the reasons is that in this case, for small kink angles, the two tracks can overlap and be misreconstructed as a single one, which is very unlikely for two oppositely-charged tracks originating from a displaced vertex (at least in the analysed scenarios). However, the decreasing tracking efficiency for higher kink angles that was mentioned above is another factor contributing to this effect.

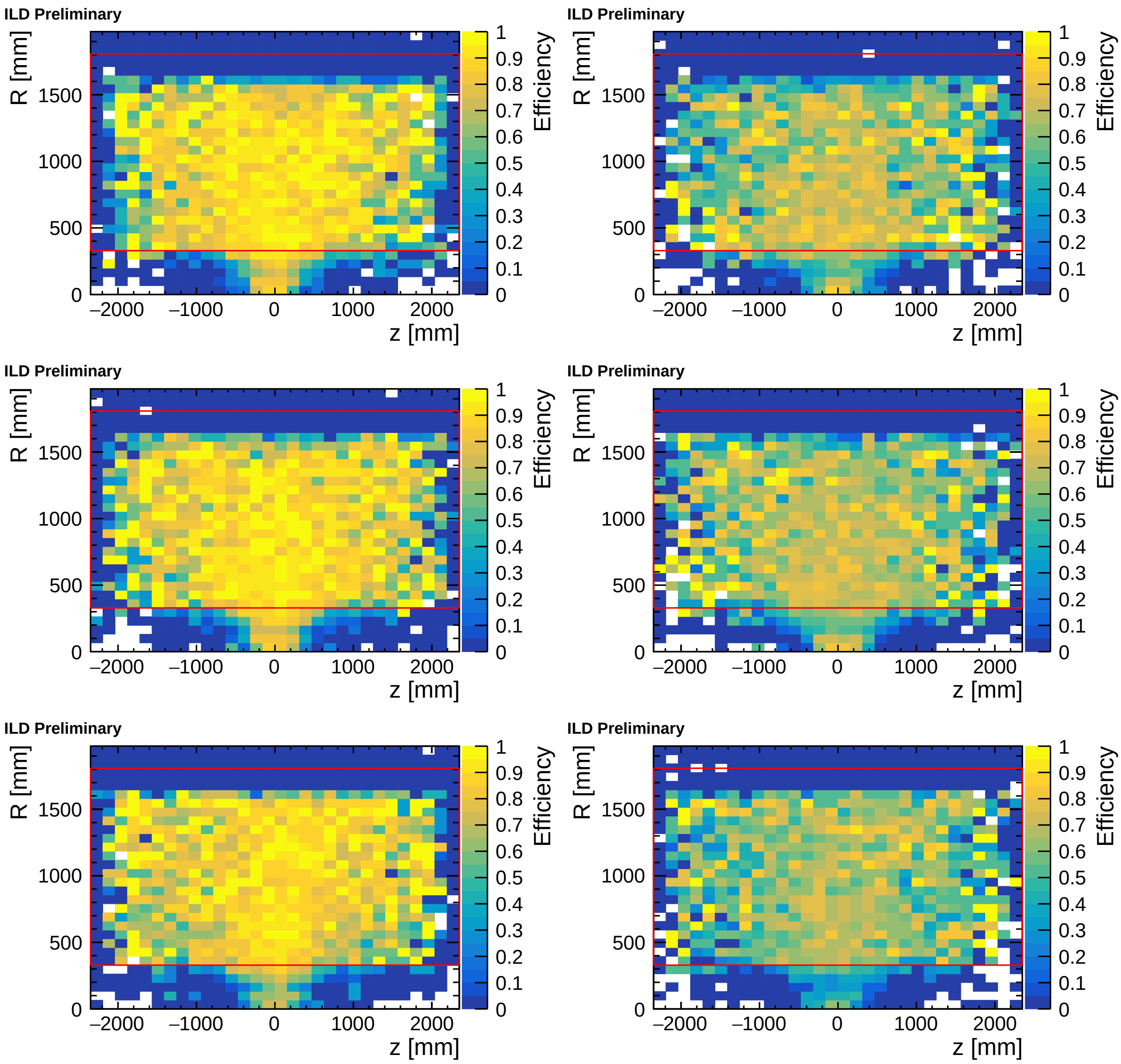


Figure 5.7: Daughter track reconstruction efficiency as a function of the radius and the z coordinate of the LLP decay vertex position for the dark fermion production. Left: $m_F = 60\,\text{GeV}$ with $m_s = 12\,\text{keV}$, $30\,\text{GeV}$, $59\,\text{GeV}$ (top to bottom). Right: $m_F = 110\,\text{GeV}$ with $m_s = 12\,\text{keV}$, $55\,\text{GeV}$, $109\,\text{GeV}$ (top to bottom). The red box corresponds to the TPC boundaries.

5.3.2 Kink vertex finding

The algorithm allowing to perform identification and reconstruction of kink vertices, called the `KinkFinder`, is already available in the `iLCSoft` software stack. Its most important feature in the context of presented work is the ability to classify the vertices into three categories: `KinkVertices`, `SplitVertices`, and `ProngVertices`, where each of them is a separate output collection in the `LCIO` file. The first category corresponds to the vertices identified as originating from SM decays of long-lived hadrons: $K^{\pm}, \pi^{\pm}, \Sigma^{+}, \Xi^{-}$. Kinks created when a track was split in the reconstruction into two segments, which can happen, e.g., due to scattering in the detector material, are stored in the `SplitVertices` collection. Finally, vertices that do not fit into any of these categories are assigned to the `ProngVertices` collection, which allows to identify preliminary candidates for BSM signals.

Description of the algorithm

In `KinkFinder`, each pair of tracks is considered iteratively, similar to the approach in `LLPFinder`. For each pair, a set of quality cuts is applied on the difference in track momenta, minimum number of hits, and on the radii of innermost hits. Based on the distances in z-coordinate from the end point of the inner track to both ends of the outer track, the kink is classified as either "normal" or "flipped" — the latter indicating that the daughter particle propagates inward.

The distance Δr_k between the two helices is initially calculated using the position of the outermost hit on the inner track and the corresponding point on the outer helix at the same z-coordinate. This distance is then corrected by searching along the z axis to find the point where Δr_k is minimised. The kink vertex is placed at the z position where this minimum occurs, between the corresponding points on the two helices. Additional cuts are applied to Δr_k, as well as to the distance Δr_{xy} in the xy plane between the last hit of the parent and the first hit of the daughter track, depending on the location of the vertex (VTX, SIT or TPC). All cuts are described in detail in Section 6.4.1.

In the next step the algorithm calculates invariant masses for possible SM decay hypotheses assuming different decay product masses, and for each particle type assigns a measure of likelihood that the corresponding hypothesis is correct. This measure is based on the difference between the reconstructed and expected particle mass, as well as on the decay time normalised to the proper lifetime of a given particle. Further tests of momentum, charge, and hit pattern consistency are used to classify the vertices as originating from split segments.

The KinkFinder optimisation

Several minor modifications were made to the `KinkFinder` code to enhance the algorithm performance. The most significant improvement was the proper handling of the flipped track case. As mentioned earlier, `KinkFinder` accounted for the possibility that the daughter track propagates inward (i.e., opposite to the parent's direction). However, it previously ignored such kink candidates at a later stage by not saving them in the output collection. This functionality was now implemented.

Additionally, one of the early quality cuts, applied to the ratio of the track momenta, was based on the assumption that the daughter track could have at most a slightly higher momentum than the parent. However, in the case of LLP decays occurring deep within the detector volume, the momentum resolution for both parent and daughter tracks is significantly worse than for prompt decays. Moreover, both this early cut and the later classification of kink candidates as split tracks were controlled by the same `KinkFinder` input parameter, making it difficult to tune the two criteria independently. In addition, the rejection of kinks with high daughter-track momentum was found to severely limit sensitivity to certain signal scenarios and was therefore removed. Instead, the original cut was replaced by a dedicated requirement on the momentum ratio, specifically designed to reject split tracks, which are treated as background in this analysis.

The search for the minimum Δr_k by scanning over z axis was also optimised. Previously, the number of steps in the scanning procedure was fixed to 10 steps, so the change involved parametrisation with an additional input parameter to the algorithm. The exception for the flipped track case was also added here.

Some modifications have also been made to the part of the code where the cuts are defined based on the vertex position in the detector (VTX, SIT, or TPC). In the case of SIT, an additional buffer was added to the cut on Δr_{xy} if the kink is exactly between the last layer of SIT and TPC. Originally, the cut was set to the difference between the radii of the last SIT layer and the first TPC pad-row. However, sometimes parent tracks can produce a few hits in the TPC, which do not get assigned to any of the tracks, resulting in a slight misreconstruction of the kink position — therefore, a small margin had to be added to the cut. Another change involved decays between the two double SIT layers. It can happen that the SIT hits from the second double-layer are lost in the reconstruction of the daughter track (and then it contains only TPC hits). The modification introduced an increase of the cut on Δr_{xy} to the difference between the radii of the first SIT layer and the first TPC pad-row.

Finally, the algorithm previously identified the positions of the first and last hits on a track by iterating over the hit list directly, rather than using the reference points provided by the `TrackState` class. This has been updated to improve the performance.

The KinkFinder performance

The kink vertex finding efficiency is presented in Figure 5.8, both for the low- and high-mass scenarios, as a function of the true dark fermion decay position in the detector. The results were obtained following a set of cuts applied on the reconstruction level. A vertex is considered correctly reconstructed if it is less than 30 mm from the true MC vertex.

Performance depends only weakly on the decay position of the LLP in the detector. For low-mass scenarios, smaller decay angles result in the kink vertex reconstruction being slightly more challenging in the central part of the TPC. For higher LLP masses, the efficiencies are visibly lower, in particular for the benchmark with small mass difference.

Figure 5.9 shows the reconstruction efficiency also as a function of the LLP decay radius, R_{kink}, and the kink opening angle, θ_{kink}. In general, the algorithm performs best at low to moderate opening angles, with efficiency slightly decreasing with growing angles, although the sensitivity remains robust even in that region.

For light LLPs in scenarios with a very large mass difference, the efficiency is mildly reduced at the smallest opening angles compared to other cases. This is due to both parent and daughter tracks having high and nearly equal momenta, which makes it more difficult to resolve the two

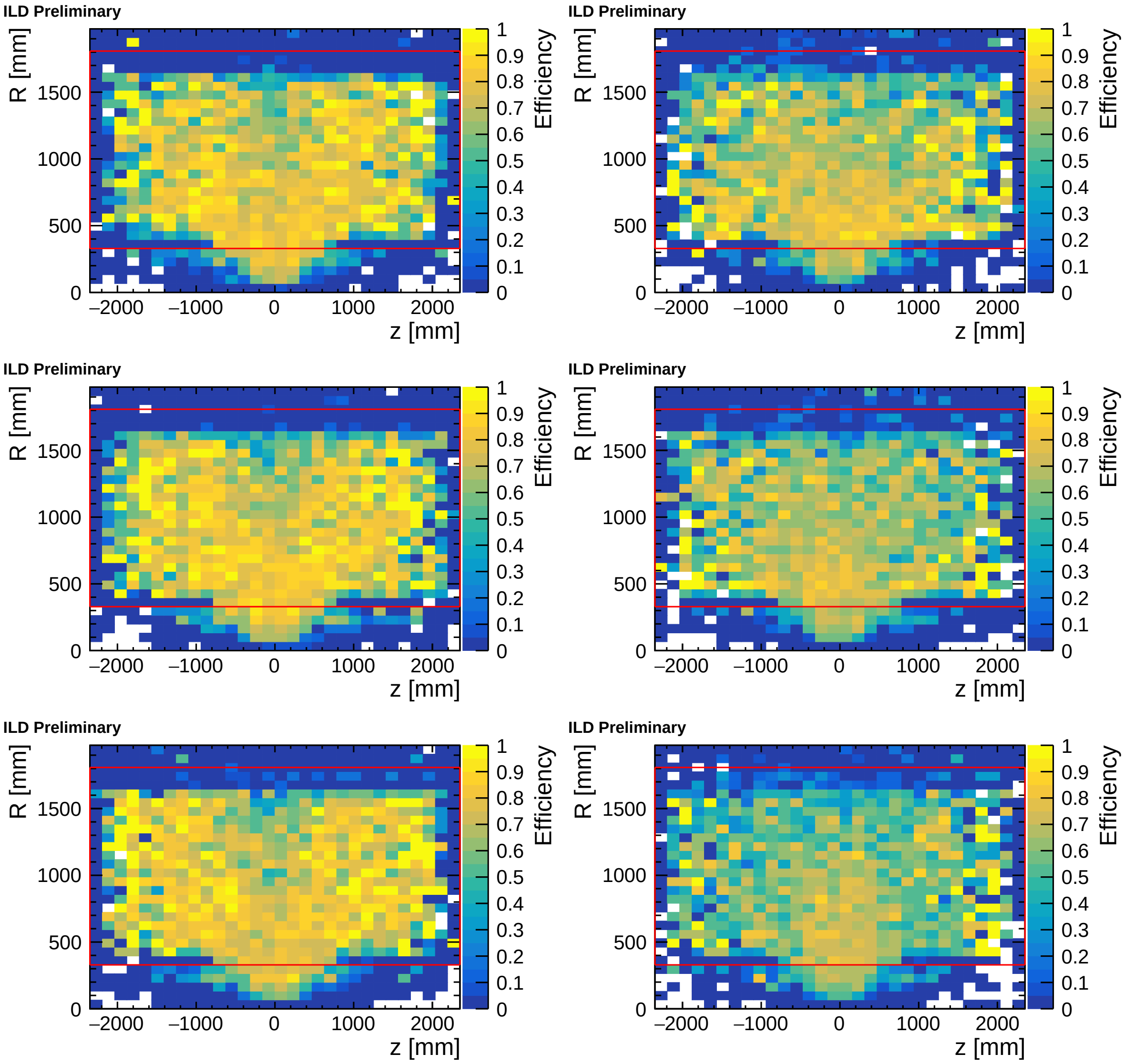


Figure 5.8: Kink vertex finding efficiency as a function of the radius and the z coordinate of the LLP decay vertex position for the dark fermion production. Left: m_F = 60 GeV with m_s = 12 keV, 30 GeV, 59 GeV (top to bottom). Right: m_F = 110 GeV with m_s = 12 keV, 55 GeV, 109 GeV (top to bottom). The red box corresponds to the TPC boundaries.

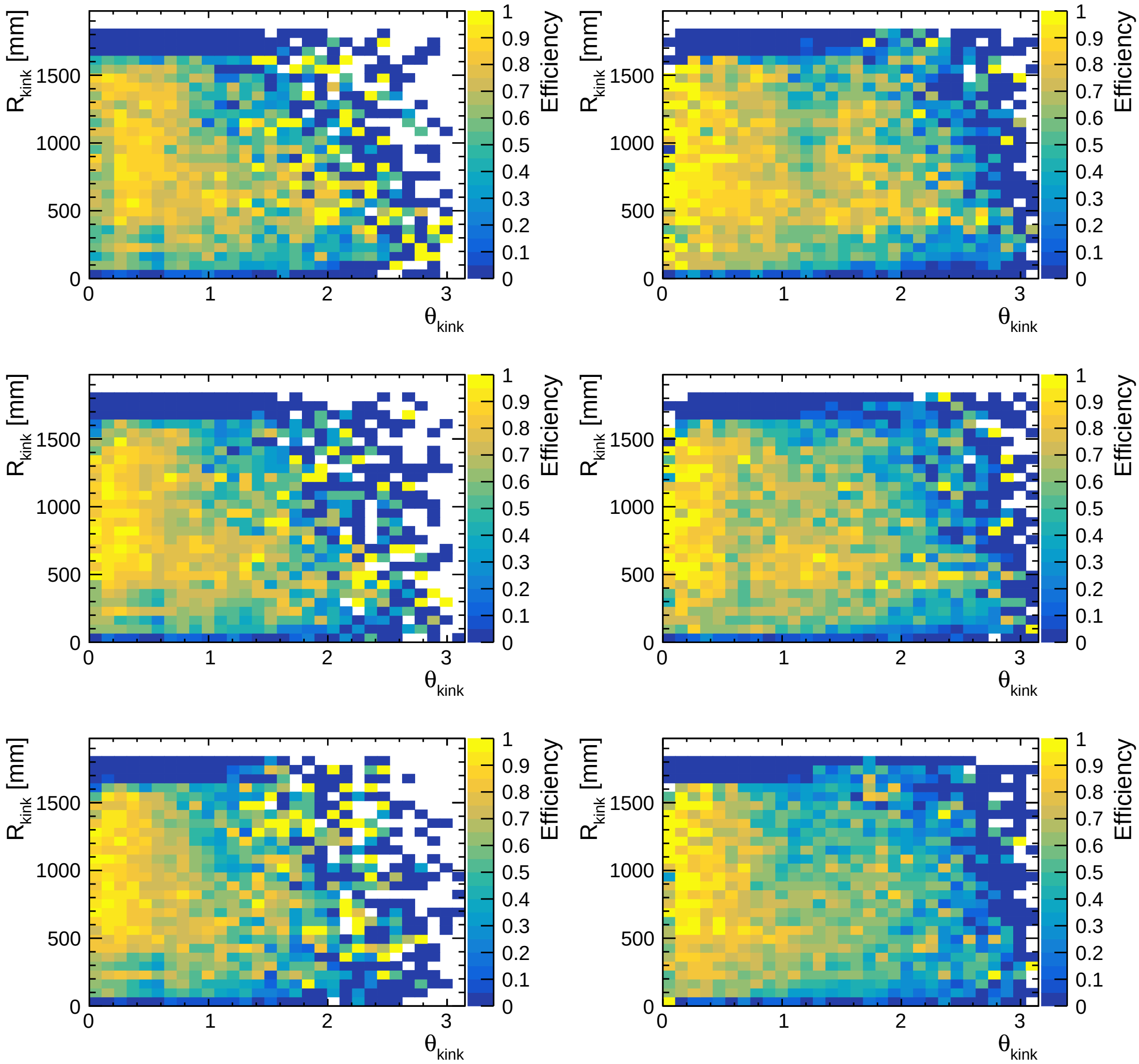


Figure 5.9: Kink vertex finding efficiency as a function of the radius of the true LLP decay vertex position and the kink angle for the dark fermion production. Left: m_F = 60 GeV with m_s = 12 keV, 30 GeV, 59 GeV (top to bottom). Right: m_F = 110 GeV with m_s = 12 keV, 55 GeV, 109 GeV (top to bottom).

Table 5.2: The efficiencies and purities of the kink vertex finding procedure for all considered scenarios of dark fermion pair-production, after preliminary cuts applied in the `KinkFinder` processor. The values in brackets represent those corresponding only to the TPC volume. See text for details.

m_s	12 keV	$m_F/2$	$m_F - 1$ GeV
$m_F = 60$ GeV			
Efficiency [%]	65.2 (69.1)	67.1 (70.9)	64.6 (69.7)
Purity [%]	96.8 (96.5)	96.5 (96.1)	89.8 (89.2)
$m_F = 110$ GeV			
Efficiency [%]	61.1 (65.9)	58.6 (62.4)	55.4 (58.9)
Purity [%]	94.5 (93.6)	95.0 (94.1)	82.5 (81.2)

tracks and to distinguish them from a split track. A similar drop in efficiency is observed for heavy LLPs with small mass differences, but this is connected to limited statistics in that region, resulting in visible fluctuations.

It is important to note that in all scenarios a large fraction of kink vertices with large opening angles can still be successfully reconstructed. The sensitivity to such decays can be a source of difficulties and a limiting factor in kinked-track searches at detectors relying only on silicon trackers or in the busy environment of the LHC.

The corresponding total efficiencies for all scenarios considered are shown in Table 5.2, together with purities, analogously to Table 5.1. Values in brackets were calculated taking into account only decays inside the TPC. The performance is good for all analysed benchmarks, with the purity slightly worse for scenarios with small mass differences.

5.4 Impact of the tracker design

5.4.1 All-silicon tracker

In order to directly quantify the impact of the continuous TPC tracking on the detector acceptance in detecting displaced signals, additional tests have been performed, using scenarios with displaced vertices and soft tracks in the final state. For this purpose, an alternative ILD detector model (`ILD_l5_v09`) was used in which the TPC and SET were replaced by the all-silicon Outer Tracker from the CLIC detector model (CLICdet) [154]. To make the CLICdet tracking system compatible with the ILD geometry, an additional barrel layer was added and the spacing between endcap layers was increased. Track reconstruction in this modified model was performed using the Conformal Tracking (CT) algorithm [208], a pattern recognition tool originally developed for CLICdet. The parameters of CT have been optimised by the author for the all-silicon ILD design using single muon, $t\bar{t}$, and displaced heavy-scalar production events.

Since the performance of the displaced track reconstruction is the main driver of the LLP sensitivity, it was used to compare the two detector designs. Figure 5.10 presents the track reconstruction efficiency as a function of the true LLP decay vertex distance from the beam axis, for various heavy scalar production scenarios. The efficiency is compared between the

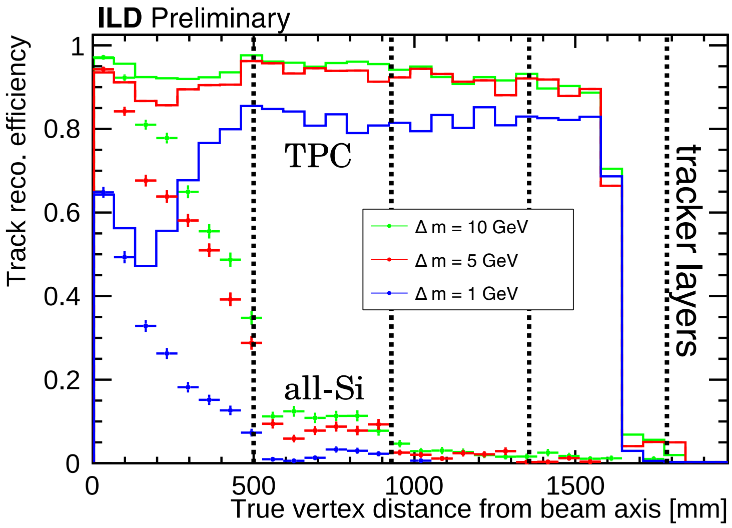


Figure 5.10: Track reconstruction efficiency as a function of distance from the beam axis for the heavy scalar samples in the all-silicon ILD design (points), compared to efficiency in the baseline TPC-equipped ILD (solid lines). Different colours correspond to particular scalar mass-splitting scenarios in the test samples. Vertical dashed lines show the position of barrel layers of the outer silicon tracker in all-silicon model.

standard ILD design and the alternative version with an all-silicon tracker. For decays occurring close to the interaction point, the performance is nearly the same. However, with increasing displacement, the efficiency in the all-silicon tracker drops sharply (falling nearly to zero around a radius of 1 m) while the baseline ILD design maintains high efficiency throughout most of the tracking volume.

The reason is primarily due to the limited number of layers in the silicon tracker. Since a minimum of four hits are required to reconstruct a track, the efficiency begins to decline significantly at around 500 mm, where only four barrel layers remain available. Moreover, even when four hits are present, they do not always allow to produce tracks of sufficient quality.

It is worth noting that these tests have been performed using the heavy scalar production, and other, more boosted final states could alter the results slightly. However, even in the case of almost collimated tracks for the light LLP production, high efficiency is obtained in nearly the whole detector volume (see Figure 5.3). This cannot be achieved in the all-silicon detector even considering its excellent hit point resolution.

5.4.2 TPC with pixel readout

Considering the factors limiting the reconstruction efficiency for displaced vertices in the classes of scenarios studied here, one can specify at least two main issues:

- reduced reconstruction performance of the low-p_T tracks,
- hit merging close to the decay vertex for highly collimated tracks.

In the first case, the problem could be partially solved by improving the pattern recognition. However, the issue of poor hit resolution for particles travelling along the pad rows (manifesting itself in the simulation as the missing hits problem; for more details, see Section 4.2.2) is purely related to the TPC readout design. The reduced hit resolution (or missing hits) can lead to splitting of the track segments or hinder the track momentum estimation near the decay vertex if the hits are missing close to it. For highly collinear tracks, the challenges originate entirely from the lower detector resolution — see, e.g., Figure 5.5 (top right).

In both cases, the sensitivity would benefit from higher detector granularity. This could be achieved using the concept of GridPix, a TPC read-out system in which the signal is collected on an array of 50 μm size silicon pixels [176]. Such an upgrade would increase the number of hits per track by more than two orders of magnitude. The prototype for GridPix has already been constructed and studies on this concept are ongoing.

Within the work carried out so far on the GridPix, `DD4hep` and the relevant software packages in `iLCSoft` have been modified in order to implement the simulation for this pixel readout and to adapt pattern recognition accordingly, although for the latter the performance as good as in the baseline ILD has not been achieved, yet. One of the conclusions of the corresponding study was that the reason is solely due to pattern recognition shortcomings, not detector design [176].

Several technical challenges arise in the context of the simulation of a TPC with pixel readout. The most significant one is the computation time needed to process the events, in particular those with busy environments such as in the $t\bar{t}$ production, which is typically used to evaluate the tracking performance. An alternative, faster simulation mode is available, using 990 μm thick virtual volumes and parabolic interpolation to generate the remaining hits. However, this approach is not well-suited for studies of the low-momentum tracks, as it only enhances the limitations of simulation using the cylindrical virtual volumes.

Additional improvements to the `Clupatra` algorithm for pixel TPC have been implemented by the author of this thesis, and preliminary tests have been performed for single muons originating from the IP. However, the computational time needed to process the samples has impeded further studies. Optimisation of the simulation performance goes beyond the scope of this work and is left for future studies. Nevertheless, the potential of the GridPix readout to bring further improvements in the ILD sensitivity to LLP detection should not be overlooked and is hence mentioned here.

Chapter 6

ILD sensitivity to the long-lived particle detection

This chapter contains the main results of the work presented in this thesis. Its core is Section 6.2, which summarises the search for displaced vertices and is largely based on the Ref. [13]. Section 6.3 presents an extension of this study, the results of which were included in the Ref. [11], and Section 6.4 describes a separate analysis for kinked tracks which, however, to a large extent takes advantage of the experiences from the displaced vertex search.

6.1 Analysis strategy

The prospects for LLP observation at future e^+e^- colliders are studied here from an experimental point of view. Two scenarios of neutral LLP production are discussed in this thesis — light LLP, leading to boosted, high-p_T SM final state, and (relatively) heavy LLP, which decays to a heavy DM particle and to SM particles. For charged LLPs, also two types of benchmarks with kinked track signature were considered: moderate- and high-mass. A set of mass differences between the LLP and a DM candidate was considered to produce both soft and boosted final states, similarly to the neutral LLP case. For both of the analysed signatures, benchmark scenarios (discussed in detail in Section 5.1) are not selected based on any particular theoretical model and existing experimental constraints in its parameter space, but rather as those providing signatures interesting from the experimental perspective — potentially challenging or not accessible at the LHC.

In the search for neutral LLPs, following model-independent approach nothing is assumed about the final state except for the presence of at least one displaced vertex inside the TPC volume, as the most generic case. Hence, other activity in the detector is allowed, but it is ignored in order to remain fully model-agnostic. Only two-prong vertices were considered, as it can be assumed that such a topology will be the most experimentally challenging one. Although multi-prong topology may bring its own analysis challenges, finding a displaced vertex in the TPC with multiple tracks in the final state is not expected to be less efficient than in case of having just two tracks. The background levels, those coming from random track intersections in particular, should also be much smaller in the case of multi-prong final states. The vertex-finding procedure was already described in Section 5.2.2, and the selection of vertex candidates will be covered in Sections 6.2.1 and 6.2.2 of this chapter.

For charged LLPs, the analysis was inspired by the search for displaced vertices and partially meant as its extension. Therefore, the same strategy was adopted, and the considered signature was at least one kinked track in the SIT or TPC — which, in terms of strategy, is the main difference from the displaced vertex search where only TPC is considered. On the other hand, the selection in the charged LLP search had to be slightly less inclusive, as for the soft final states, without a dedicated approach, background from low-p_T events would be even more overwhelming and challenging than for displaced vertices (see more in Section 6.4.1). In principle, for charged LLPs one could also consider the $\mathrm{d}E/\mathrm{d}x$ signature. However, this strategy in a 250 GeV e^+e^- collider would only be efficient in the case of very heavy particles (with a mass above 100 GeV), as they are sufficiently slow to produce higher ionisation, allowing for their distinction from charged long-lived SM hadrons.

Particular models were used to generate samples for the study; however, we do not make any model-specific assumptions in the analyses to make the results relevant for any model providing the considered or similar signatures. The selection (described in detail in Sections 6.2.1, 6.2.2 and 6.4.1) is therefore *not* optimized for a particular model or scenario and uses only variables related to displaced vertices and kinked tracks.

As mentioned in Section 5.1.1, for the neutral LLP study only decays to muon pairs were generated for ALPs and $Z^{(*)}$. This allowed for faster detector response simulation, since muons do not produce showers inside calorimeters as opposed to electrons or pions. However, it should be noted that this choice does not make our search less general. The analysis relies exclusively on tracks reconstructed from tracker hits, without considering calorimeter or muon spectrometer deposits. Furthermore, displaced multi-track final states are also included in the analysis (vertices formed by more than two tracks are not rejected), despite the fact that we do not include them explicitly in the signal samples. As a result, we find that both two- and multi-track displaced vertices contribute significantly to the SM background (see Sections 6.2.1 and 6.2.2 for details). Since they involve not only muons, but also electrons and pions, it confirms that using muon decays for signal samples does not produce any significant bias.

Similar assumptions were made for the charged LLP analysis, although the F decays to electrons were generated in addition to muons. This was necessary since electrons can naturally be a source of kinked tracks due to bremsstrahlung and hence produce additional vertices in the signal samples. However, the remaining features of the displaced vertex search hold for the kinked track analysis, in particular the lack of bias related to using particular lepton flavours to generate the samples.

In both cases the analysis is performed in an approach similar to an anomaly search, where the signal can take any form of tracks forming displaced vertices or kinked tracks and a discovery would be implied by an excess in the number of reconstructed vertices. As shown in more detail in the following sections of this chapter, the backgrounds hardly exhibit any process-dependent behavior, except for a distinction between low- and high-momenta events, and between processes with different numbers of objects expected in the final state. The vast majority of background sources originate from random coincidences of tracks, single long-lived SM particles, and secondary interactions with the detector material. Therefore, any SM process with a high cross section and a large activity inside the detector is a potential background source.

In many low-background searches for rare processes, particularly with exotic signatures, cosmic rays are often considered as an additional source of background. In light of the lack of proper tools to model their impact, such events were not included in this study. However, one should note it is very unlikely that a cosmic muon will interact with the gas in the TPC, producing a displaced or kink vertex exactly in coincidence with an e^+e^- bunch crossing. In addition, there are efficient ways to veto against such background using timing, ToF, or muon spectrometer deposits, as described, for example, in Ref. [90, 209, 210].

6.2 Search for displaced vertices

6.2.1 Beam-induced background reduction

Beam-induced backgrounds are relevant for signal scenarios which involve very soft final states, such as those considered in this study in the search for displaced vertices. Figure 6.1 (left) presents a distribution of the displaced vertices as a function of the distance R from the beam axis. There is an overwhelming number of vertices found in the overlay samples, most of which are fake (e.g. originate from randomly intersecting tracks). Other result from particle interactions in the detector: the first small peak above 100 mm aligns with the first layer of SIT, while the second one, above 300 mm, with the last SIT layer and the inner TPC wall. Also peaks around the outer TPC wall and the SET are visible, above 1.7 m. Figure 6.1 (right) shows the distribution of the total transverse momentum, p_T^{vtx}, of tracks forming displaced vertex candidate, compared for the overlay events and two of the considered signal scenarios. It is visible that vertices in the overlay sample occupy the same kinematic region. These distributions show that $\gamma\gamma \to$ hadrons and incoherent e^+e^- pairs events cannot be neglected as a background source in this study, and will require a dedicated mitigation strategy. It should be noted that numbers of vertices presented in Figure 6.1 are not normalized and represent the absolute numbers of vertices in the MC samples. For the background, this corresponds to around $3.6 \cdot 10^{-6}$ of the total luminosity. For the signal, we do not refer to cross section predictions of any model, to which the distributions could be normalized. The purpose of these plots is only to illustrate the respective kinematic properties of signal and background events.

Soft hadrons can be produced in photon-photon interactions both from beamstrahlung (γ^B) and from the Weizsäcker-Williams spectrum (γ^W), so four channels are possible. Number of events occurring per single BX for each of them is a random number from the Poisson distribution with a given mean, μ_P, which is used in the standard ILD reconstruction chain when overlaying the background on signal events. Table 6.1 presents these expected values for the 250 GeV ILC, together with the number N_{evt} of events used in this study for each channel. The incoherent e^+e^- pairs are included in the analysis as well. To do so, 99,000 BXs have been generated with GuineaPig [204]. For each of these BXs, only the incoherent pairs that would potentially hit any tracking detector were written out and were fully simulated. From this pool of BXs, one was picked at random for each physics event, and overlayed before reconstruction. This set of BXs was also used as a standalone background, similar to the hadron photoproduction.

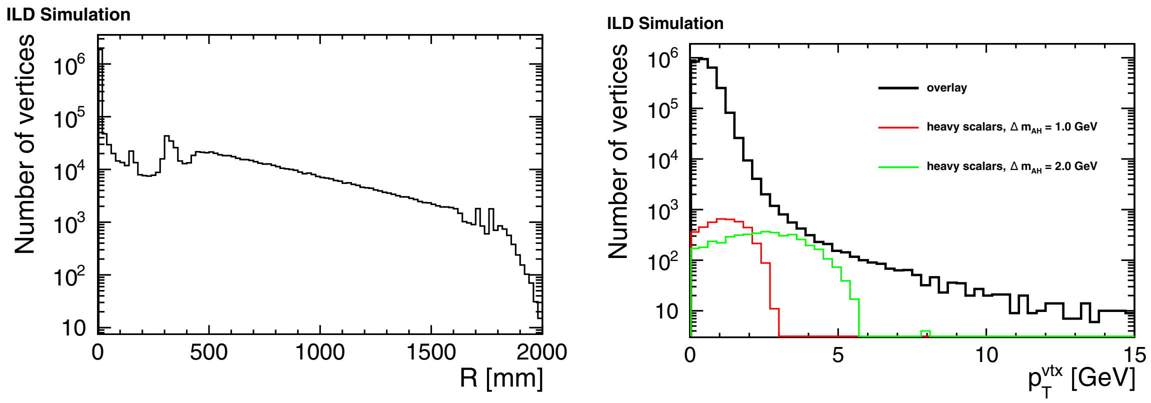


Figure 6.1: Left: Number of displaced vertices found in the overlay sample as a function of distance from the beam axis. Right: Total transverse momentum of tracks coming from a displaced vertex for the overlay (black) and scalar pair-production with Δm_{AH} = 1 GeV (red) and Δm_{AH} = 2 GeV (green). All histograms are normalized to the number of simulated MC events and correspond to no selection applied at all, just except for the required maximum distance between track helices.

Table 6.1: Average number, μ_P, of overlay events expected per single BX for different channels and the number, N_{evt}, of corresponding MC events used in the analysis.

channel	μ_P	N_{evt}
$\gamma^B\gamma^B$	0.8298	1,418,000
$\gamma^B\gamma^W$	0.2972	889,000
$\gamma^W\gamma^B$	0.2975	911,000
$\gamma^W\gamma^W$	0.1257	519,000
tot. $\gamma\gamma \rightarrow$ hadrons	1.55	3,837,000

Preliminary cuts

Highly displaced tracks, not pointing to the IP region, are not considered in the standard event reconstruction at the ILD, and therefore a new, dedicated selection procedure had to be implemented. It was designed to be as general as possible, accounting for the background properties, as described in detail below. Three sets of preliminary cuts are applied to reduce the number of fake vertices of different origin. The first set of "quality cuts" is based on variables describing track pair geometry and is aimed at reduction of nonphysical vertices coming from reconstruction inaccuracies or random coincidences.

- Sometimes a charged particle will scatter at unusually large angle at one point in the detector, and the standard track-reconstruction will split the track in two. Track splitting might also occur during reconstruction, when merging of a curling track segments fails. In this analysis, this would be interpreted as a displaced vertex at the point where the track was split. However, the two track segments will be close to colinear at the split-point, and the curvature of the segments will be close to identical. Therefore, the track pair opening angle θ_{vtx} is restricted to $\cos\theta_{vtx} > -0.6$ and the ratio of the curvature of the tracks to

$|\Omega_1| / |\Omega_2| < 0.94$, assuming $|\Omega_1| < |\Omega_2|$ (following definitions of helix parametrisation described in Section 4.2.4).

- In some cases, reconstruction of one of the segments of a curling track fails completely, which can lead to formation of poor, short tracks, often with unexpectedly large p_T[1]. Such tracks are more likely to pick up a second track in the event to form a vertex. To reject such vertices, it was required that both tracks should have the number of degrees of freedom in the track-fit exceeding 40, a number increased to 70 for the highest p_T track of the vertex, if the p_T was above 1.5 GeV.

- To further suppress vertices coming from accidental intersections of uncorrelated tracks, the vertex should be close to or before the first point measured on the track. To account for tracks with high p_T, almost perpendicular to the beam axis, as well as low-p_T tracks curling many times along the Z axis, two complementary sets of cuts are used, depending on the distance $z_{12} = |z_{\text{first}} - z_{\text{last}}|$ between the first and the last track hit in the Z coordinate. For the tracks with $z_{12} \leq 100\,\text{mm}$, the ratio of azimuthal angular distance ϕ_1 from the vertex to the first hit, relative to the total angle ϕ_{arc} from the first to last hit of the track, ϕ_1 / ϕ_{arc}, should not exceed 0.05 (see Figure 6.2). For the track with $z_{12} > 100\,\text{mm}$, the requirement is used for a difference $z_1 = z_{\text{first}} - z_{\text{vtx}}$ in Z coordinate between the first hit and the vertex, relative to the z_{12}. Value of $\text{sgn}(p_z) \cdot z_1 / z_{12}$ must be higher than -0.01, where p_z is the Z component of the track momentum. Note that z_1 can be both positive and negative depending on the track direction, so with a perfect reconstruction of the vertex position, this expression should always be positive.

An important source of the background are interactions of both charged and neutral particles with the detector material producing (one or more) secondary, displaced tracks emerging from the displaced vertex. As two hypotheses of the direction for each track are considered, a charged particle producing a single secondary track can look like a vertex as well, with one of the tracks going back into the IP. It is worth noting that this is congruent with the signature of *charged* LLPs, the kinked track, for which a dedicated analysis is covered in Section 6.4. Interestingly, it seems that the `LLPFinder` and the vertex-finding procedure designed in this study are sensitive to such signals as well. However, this analysis is focussing on the neutral LLPs and these vertices are regarded solely as a part of the background. To remove vertices from secondary interactions, a second set of requirements is imposed.

- The search region is restricted to the vertex radii in the 0.4-1.5 m range, since secondary interactions occur mainly (but not only) with the inner and outer walls of the TPC.

- Vertices resulting from the charged particles interaction with the detector material are first rejected by requiring no additional tracks that end within 30 mm of the vertex and no tracks with $d_0 < 10\,\text{mm}$ that pass 30 mm from the vertex.

- If at least one of the tracks at a vertex has $d_0 < 50\,\text{mm}$ its first hit (the one closer to the reconstructed displaced vertex) is required to have a smaller radius R_f than its last

[1] In the standard reconstruction, these are rejected by cuts on d_0 and z_0 parameters, which were removed for the purpose of this analysis.

hit R_l. This cut is targeting the "kink-like" signatures: the vertex and secondary track reconstructed as an effect of scattering of charged particles.

The neutral particles producing two or more secondary tracks are considered as mostly irreducible background, partly mitigated by the first cut. That is due to the model-independent approach, in which exclusion of scenarios with LLP decays to more than two charged particles (displaced jet signatures in particular) is not intended even though they are not analysed as such.

Finally, with a third set of preselection cuts, removed are vertices corresponding to decays of V^0 particles (a collective term for long-lived neutral hadrons and converting photons).

- Matching of vertex candidates with the output of a dedicated ILD processor for V^0 identification is performed. The LLP vertex candidate is rejected if there is a vertex reconstructed by the processor closer than 30 mm.

- Vertices with an invariant mass $M_{inv}^{\pi\pi}$ (assuming the tracks are pions) in a window of ±50 MeV around K^0 mass and $M_{inv}^{\pi p/p\pi}$ (assuming one track is a pion and the other is a proton and vice versa) around Λ^0 mass are removed.

- For the assumption that both tracks are electrons, vertices with a mass M_{inv}^{ee} less than 150 MeV are rejected. A wide window aims to account for poorly reconstructed and almost colinear tracks from highly displaced photon conversions.

Cuts on the invariant mass of a pair of tracks for different track mass hypotheses are required for more efficient background suppression because `V0Finder` (the processor for V^0 identification) has finite efficiency and prioritises purity. With the overwhelming number of overlay events, the remaining V^0 particles not identified by the algorithm still constitute a substantial background for rare processes.

The preliminary selection is summarised in Table 6.2, which shows the vertex finding rates in the overlay events, and its efficiencies for selected signal scenarios, after consecutive sets of cuts. For the background, a ratio of number of events with at least one vertex found to the number of all events in the MC sample is shown. For the signal, a ratio of number of correctly identified vertices to all events is presented, where the vertex is considered “correct” if it is reconstructed within 30 mm of the true decay vertex. The first row in Table 6.2 corresponds to the requirement of the maximum 25 mm distance between track helices, imposed at the beginning of the vertex finding procedure (see Section 6.2.3). The following rows correspond to the subsequent sets of cuts described above.

Final selection

The preliminary cuts described above allow for background reduction by a factor of around 500, down to the level of 0.21%. However, due to the huge expected number of approximately 10^{12} BXs in the experiment, a suppression factor at the level of at least 10^{-9} or higher is needed to achieve a good sensitivity to rare BSM signals. Since it is not possible to generate an event sample of a size corresponding to the expected number of BXs in the experiment, a direct determination of the reduction factor is not possible, and it can only be estimated assuming factorisation, as a product of the subsequent selection cut efficiencies. However, this will only provide a reliable

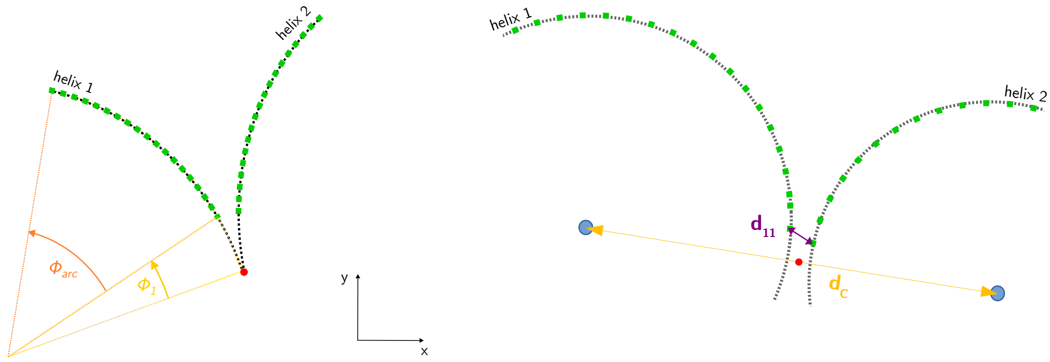


Figure 6.2: Schematic illustration of variables used for the overlay selection. Helix curves are represented by black dotted lines and detector hits by small green rectangles. Left: Projection of helices into XY plane and the angular variables ϕ_1 and ϕ_{arc} used in the preselection. Right: Variables d_C and d_{11} that combined are used in the final selection. Small blue circles represent centres of projections of helices into XY plane. Note that d_{11} is a three-dimensional distance. Red dot corresponds to a vertex candidate.

Table 6.2: Summary of vertex-finding rates for the beam-induced backgrounds and finding efficiencies in selected signal scenarios after subsequent sets of preliminary cuts. For details, please refer to the description of cuts in Section 6.2.1.

	Overlay	Signal efficiency [%]			
Cut set	rate [%]	$\Delta m_{AH} = 1\,\mathrm{GeV}$	$\Delta m_{AH} = 2\,\mathrm{GeV}$	$m_a = 0.3\,\mathrm{GeV}$	$m_a = 1\,\mathrm{GeV}$
Dist. between helices	37.1	72.5	77.5	28.2	74.8
Quality cuts	1.7	58.3	64.6	17.0	56.7
Secondary interact.	0.6	52.6	57.9	14.2	48.4
V^0 particles	0.2	41.4	55.2	7.4	48.4

result if the variables used in the selection are statistically independent. Therefore, uncorrelated variables were selected for the final selection cuts so that the total reduction factor can be obtained by multiplying all cut efficiencies.

Two variables are used. The first is just the total transverse momentum p_T^{vtx} of a pair of tracks, since the background is expected to occupy the region of very low p_T, as shown in the right plot of Figure 6.1. Tracks in the LLP signals should also come out of a single point, so distance d_{11} (in three dimensions) between their first hits in the TPC should be small. On the other hand, helix-circles (projections of track helices onto the XY plane) of these tracks should not overlap, so the distance d_C between their centres should be large. The schematic visualisation of d_{11} and d_C is shown in Figure 6.2 (right). To reduce the number of selection criteria and avoid correlation between these two variables, a new variable $d_h \equiv 2.2 d_{11} - d_C$ was defined, the distribution of which is shown in Figure 6.3 for the signal (left) and background (right) as a function of p_T^{vtx}, after the preliminary selection. The factor 2.2 was chosen as the best for the most optimal signal-background separation, it corresponds to the optimal linear discriminant in the d_{11}-d_C plane. As visible in the distribution on the left, for true vertices the

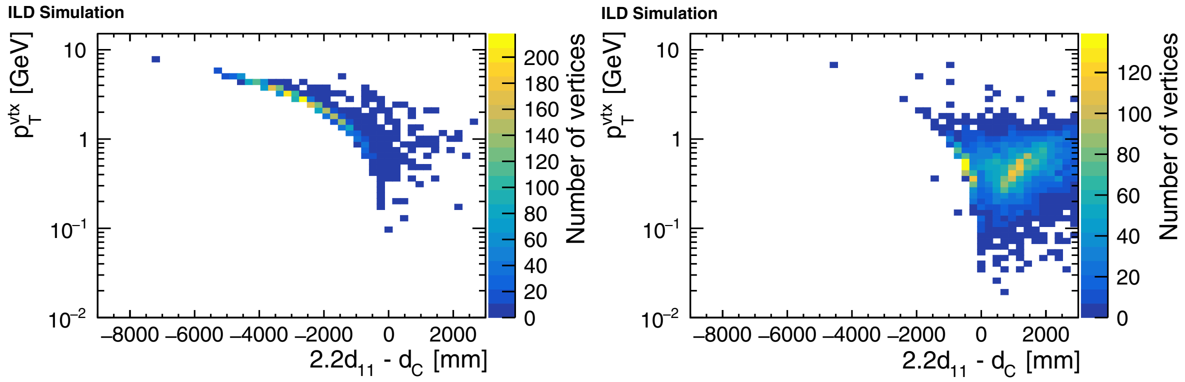


Figure 6.3: Distribution of p_T^{vtx} as a function of d_h variable for a heavy scalars signal scenario with Δm_{AH} = 2 GeV (left) and all unweighted overlay background events that survive the preliminary cuts (right). For the signal (left) these variables show strong correlation, but for the background (right) only very small correlations are visible for a few events at the lowest values of d_h. Both variables are used for the final selection aimed to suppress the background from overlay events; see text for details.

d_h and p_T^{vtx} variables are strongly correlated. However, only little to no correlation should be observed for any fake vertices from random coincidences remaining in the overlay sample, as can be seen in the right plot.

$d_h < -2000$ mm and $p_T^{vtx} > 1.9$ GeV were chosen as the optimal cuts. To mitigate the effect of limited statistics, the reduction factor for each of these variables was estimated from distributions fitted to the respective histograms of p_T^{vtx} and d_h. The functions used were, respectively, log-normal distribution for the p_T and a sum of Crystal Ball functions, $c_1(x) + c_2(-x)$, for the d_h. This was performed for all overlay events that survive the preliminary selection, weighted by the corresponding expected values μ_P shown in Table 6.1. The distributions together with the fitted functions are presented in Figure 6.4 — it is visible that the fits describe the distributions well. Cuts on the p_T^{vtx} and d_h result in selection efficiencies of $1.73 \cdot 10^{-3}$ and $3.48 \cdot 10^{-5}$, respectively. As each of those factors is determined on the sample of events after the preliminary cuts, this can be combined with the preliminary selection efficiency, which gives the total reduction factor of $1.26 \cdot 10^{-10}$.

6.2.2 Backgrounds from physical events

The SM processes from hard e^+e^- scattering (listed in Table 6.3 together with their cross sections at 250 GeV ILC) are also taken into account. To increase the production efficiency, the samples have been generated with ±100% polarisation and later reweighted, according to the assumed beam helicity sharing in the ILC H-20 scenario, with ±80% electron and ±30% positron beam polarization [3].

Preliminary tests have shown that the main sources of displaced vertices in this case are, similar to overlay events, the interactions with the detector material and decays of long-lived hadrons, which happen predominantly inside the jets. Therefore, one could expect that the vertex finding rate for background, defined as a fraction of events with at least one identified vertex to all events, does not depend on a physical process itself and neither on the beam

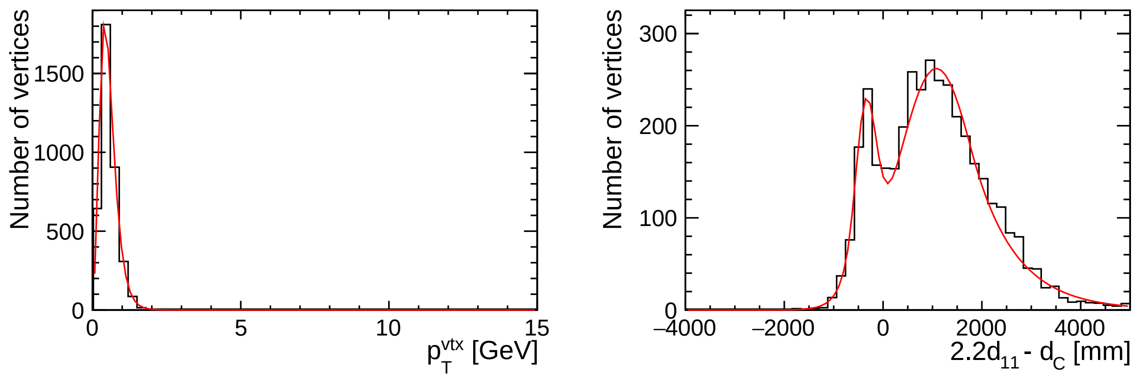


Figure 6.4: Fits to distributions of p_T^{vtx} (left) and d_h (right) used to extract the final selection efficiencies. The vertices are weighted by the corresponding values μ_P shown in Table 6.1.

Table 6.3: Cross sections for the SM processes at $\sqrt{s}$ = 250 GeV considered in this study, for different setups of the beam polarization, assuming ±100% polarization, and different photon sources for γγ processes.

sgn(P(e$^-$), P(e$^+$))	(−, +)	(+, −)	(−, −)	(+, +)
channel	σ [pb]			
qq	127.966	70.417	0	0
qqqq	28.66	0.97	0	0
qq$\ell\nu$	29.043	0.261	0.191	0.191
qq$\ell\ell$, qq$\nu\nu$	2.715	1.817	1.156	1.157
$\ell\ell$	21.214	16.363	0	0
process	BB	BW	WB	WW
$\gamma^{B/W}\gamma^{B/W}$	42.15	90.338	90.12	71.506

polarization, but rather on a number of jets produced.

As mentioned in Section 5.1, some parameters of standard track reconstruction have been modified to increase the efficiency for displaced signals, and therefore SM samples had to be reprocessed. Because rerunning track reconstruction for all SM samples is very time and resource consuming, the selection efficiencies are obtained directly for the most significant processes (those with the highest cross sections), two-fermion ($e^+e^- \rightarrow q\bar{q}$) and four-fermion ($e^+e^- \rightarrow q\bar{q}q\bar{q}$) hadronic channels. In addition, it is also calculated for di-lepton production, since this channel can be a source of jets in hadronic τ lepton decays and contain a significant amount of photon conversions. Then, to estimate the level of background for a complete set of SM processes, the $q\bar{q}$ and $\ell\ell$ vertex finding rates (ϵ_{qq} and $\epsilon_{\ell\ell}$, respectively) were used for other e^+e^- channels. For example, the rate assumed for the $qq\ell\nu$ process was $\epsilon_{qq\ell\nu} = \epsilon_{qq} + \epsilon_{\ell\ell}/2$.

The hard scattering of $\gamma^{B/W}\gamma^{B/W}$ was taken into account separately and simulated as well, since this background might, in principle, contain different sources of (fake) displaced vertices. Six-fermion and Higgs boson production, as well as $e^\pm\gamma$ interactions, were neglected because of much smaller cross sections.

First, $q\overline{q}$ is considered as the dominant background, and the selection established for the reduction of overlay events is applied. To further reduce background from coincidences of random tracks within high p_T hadronic jets, a separate cut on $d_{11} < 50\,\mathrm{mm}$ is applied. This whole set of cuts will be further referred to as the *standard* selection. It yields the vertex finding rate of $(7.99 \pm 0.68) \cdot 10^{-4}$ for the $q\overline{q}$ sample. The rate obtained for the qqqq sample was found to be $(1.486 \pm 0.094) \cdot 10^{-3}$, which is a factor of 1.86 greater, so it seems to confirm the assumption of rate dependence on the number of jets. For hard $\gamma^{B/W}\gamma^{B/W}$ scattering, the rate of $(2.13 \pm 0.28) \cdot 10^{-5}$ was obtained and for the $\ell\ell$ channel $(1.27 \pm 0.16) \cdot 10^{-4}$. As expected, the contribution from the di-lepton production is minor and consists only 2% of the background expected after selection. All uncertainties are statistical, calculated as $\sigma_i = \sqrt{\epsilon_i(1-\epsilon_i)/N_i}$, where ϵ_i is the respective vertex finding rate, and N_i is the total number of events in a sample.

Fully hadronic or leptonic decays of V^0 particles are efficiently identified and rejected by the cuts described so far. However, this is not the case for semileptonic K^0_L decays, which involve neutrinos in the final state and therefore lead to a smeared and shifted invariant mass distribution. A similar effect can also arise from large reconstruction uncertainties, particularly in the case of short or collimated tracks. These effects are visible in the distributions shown in Figure 6.5 for the hypotheses that the tracks are electrons (left) and pions (right), after the standard selection.

To further improve background rejection and final sensitivity to LLPs, a *tight* selection with additional cuts was also defined. In this case, cuts in the electron- and pion-track mass hypotheses are extended to exclude all candidates with $M^{ee}_{inv}, M^{\pi\pi}_{inv} < 700\,\mathrm{MeV}$. The window in the cut on Λ^0 mass is reduced from $\pm 50\,\mathrm{MeV}$ to $\pm 20\,\mathrm{MeV}$, since now it can significantly improve the signal selection efficiency without much increase in the background.

In tight selection, an additional criterion exploits the fact that final-state tracks in signal processes tend to be well isolated from other activity in the event, whereas displaced vertices from background processes are typically reconstructed within jets. The isolation measure is defined as a scalar sum $p_{cone} = \Sigma_i |\vec{p}_i|$ of the track momenta in a cone spanned by an opening angle β of $\cos\beta > 0.98$, around the combined momentum p_{vtx} of the track pair at the vertex. Tracks whose first or last hit is 30 mm within the vertex candidate are not included in the sum to avoid excluding vertices with more than two tracks in the final state. The distribution of p_{cone} as a function of p_{vtx} is presented in Figure 6.6 (left) for the $q\overline{q}$ background and $\Delta m_{AH} = 2, 3\,\mathrm{GeV}$ signal scenarios after the standard selection. To suppress background from displaced vertices reconstructed inside hadronic jets it is required that $p_{cone} < 1\,\mathrm{GeV}$. This is a conservative requirement, since $p_{cone} = 0$ in the vast majority of signal events.

Because of the isolation criterion, the tight selection might be seen as partly model-dependent (in principle, some models could predict an LLP production within or close to hadronic jets). For this reason, it is considered separately from the standard set of cuts. However, it is important to note that the vast majority of models currently considered do not predict an LLP production in association with close-by high-momentum tracks.

Tight cuts on the $M^{ee}_{inv}, M^{\pi\pi}_{inv}$, and on the isolation, make the cut on Λ^0 mass redundant. This is visible in Figure 6.6 (right), where the mass difference distribution after the tight selection, but without the cut on the Λ^0 mass window, is shown for hypotheses that one track is a pion and the other is a proton.

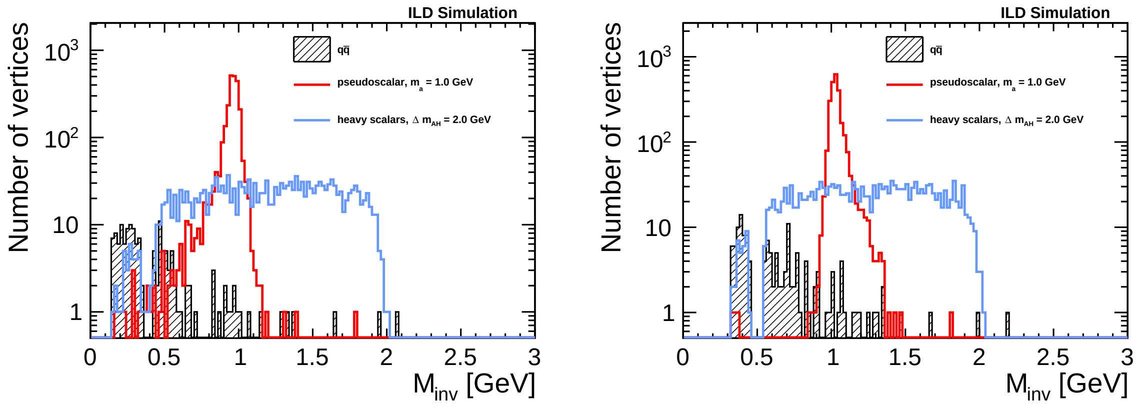


Figure 6.5: Distributions of the invariant mass of tracks coming out of a displaced vertex, assuming tracks are electrons (left) and pions (right). Black histograms corresponds to $q\bar{q}$ sample and different colours to various signal scenarios — pseudoscalar with $m_a = 1\,\text{GeV}$ (red) and scalars with $\Delta m_{AH} = 2\,\text{GeV}$ (azure). All histograms are normalized to the number of simulated MC events.

As already mentioned, the SM background samples had to be reprocessed due to the modified reconstruction criteria. For this reason, $q\bar{q}$ events visible in Figures 6.5 and 6.6 correspond only to the luminosity of $1.75\,\text{fb}^{-1}$ and the histograms have not been rescaled for better readability. However, only the shapes of the distributions should be compared in these plots and not the relative numbers of vertices, since no particular cross sections are assigned to the analysed signal scenarios. The total number of background events expected in the actual experiment is taken into account later.

The tight selection reduces the displaced vertex finding rate by more than an order of magnitude; see Tab. 6.4 for the selection rate comparison. It was also found that tight selection does not lead to any improvement in the removal of overlay events. The price of an improvement in background rejection is a loss of sensitivity to the most challenging signal scenarios, with $\Delta m_{AH} = 1\,\text{GeV}$ and, in particular, $m_a = 300\,\text{MeV}$, which suffer from the cuts on the invariant mass.

6.2.3 Vertex finding results

The resulting vertex finding efficiency after selection described in the previous sections, as a function of the true LLP decay vertex radius and z coordinate, is shown in Figure 6.7 (left) for the $\Delta m_{AH} = 2\,\text{GeV}$ heavy scalar scenario. The vertex is considered "correct" if it is reconstructed within 30 mm of the true vertex. In Figure 6.7 (right) the efficiency is presented as a function of the true decay vertex radius for all the heavy scalar scenarios considered. For the plots presented, cuts on the vertex position were not applied, to indicate differences between the efficiency for LLP decays inside the TPC region and the silicon tracker ($R < 329\,\text{mm}$). It is noticeable that for soft final states the vertex finding efficiency is higher within the TPC volume than in the silicon tracker, which is a direct consequence of a higher track reconstruction efficiency for highly displaced soft tracks (for more details, see Sections 5.2.1 and 5.4.1).

The corresponding vertex finding results for light pseudoscalar production are presented in

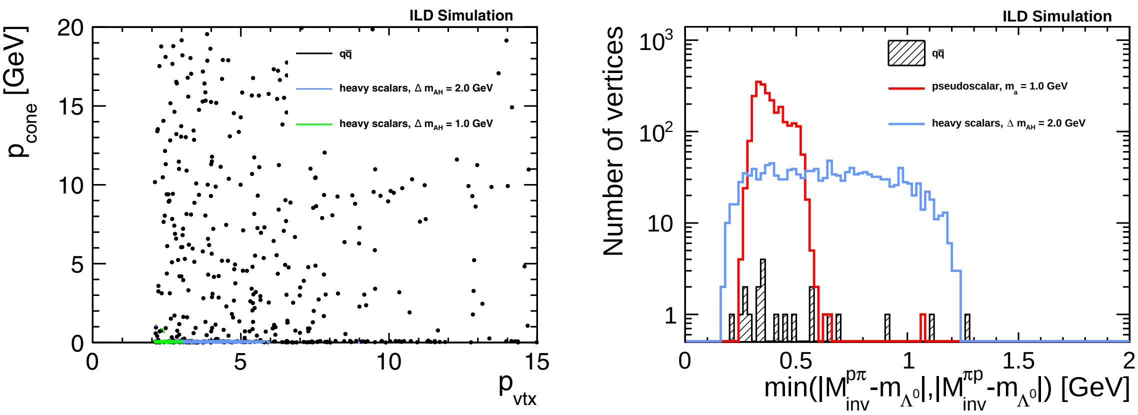


Figure 6.6: Left: momentum p_{vtx} of the track pair system at a displaced vertex shown as a function of the sum of magnitudes of track momenta, p_{cone}, in a cone surrounding the total momentum of track pair coming out of the vertex. Black dots correspond to the $q\bar{q}$ background, green to scalars with Δm_{AH} = 1 GeV and azure to scalars with Δm_{AH} = 2 GeV. Right: distribution of a difference between m_{Λ^0} and invariant mass of a track pair assuming that one of them is a pion and the other is a proton, where the smaller of two possible values is taken. Black histogram correspond to the $q\bar{q}$ background, red to pseudoscalar with m_a = 1 GeV and azure to scalars with Δm_{AH} = 2 GeV. All histograms are normalized to the number of simulated MC events.

Figure 6.8, for the m_a = 1 GeV scenario (left), and compared for all the benchmarks considered (right). Here, in contrary to the heavy scalar case, for the low-mass scenarios the efficiency is higher in the region of silicon detectors (vertex detector, VTX, and silicon internal tracker, SIT), closer to the beam pipe. This is the effect of very high collinearity of tracks from the LLP decay mentioned in Section 6.2.3; in the TPC, hits close to the decay vertex merge, while in a silicon tracker higher resolution allows for a better track separation and more accurate vertex reconstruction. However, it should be noted, that the TPC still greatly enhances the detector acceptance, since in an all-silicon tracker the reconstruction is limited by a number of layers, regardless of the final state kinematics. The further decrease in efficiency visible in the central part of the detector, compared to the results presented in Figure 5.5, is the outcome of cuts on ϕ_1/ϕ_{arc} for z_{12} < 100 mm. If hits close to the vertex merge, one of the tracks has a starting point somewhere along the second track, which looks like those two tracks are intersecting, rather than coming out of a single point.

The final vertex finding efficiencies inside the TPC, after both standard and tight sets of cuts, for the considered scalar and pseudoscalar production scenarios, are presented in Table 6.5. For the scalars, the efficiency strongly depends on their mass splitting (Z^* virtuality), which is responsible for the final state boost. Good performance is achieved for $\Delta m_{AH} \geq 2$ GeV, while for the Δm_{AH} = 1 GeV scenario it is affected mainly by the $p_T^{vtx} > 1.9$ GeV cut used to suppress the overlay background. For the light pseudoscalars the efficiency dependence is opposite, i.e. it decreases with the final state boost. This is due to the high collinearity of tracks coming from decays of very light particles, the reconstruction of which is more efficient in the silicon detectors (VTX and SIT) because of their better two-hit separation. High performance was obtained for

Table 6.4: The vertex finding rates directly obtained for the high-p_T SM backgrounds inside the TPC region after different sets of cuts. Large statistical uncertainties for tight selection rates result from small number of MC events remaining after the selection. The total number, N_{bg}, of background events expected after both selections, taking into account all channels presented in Table 6.3, is also shown in the last row. N_{bg} is calculated assuming the total luminosity of $2\,\text{ab}^{-1}$ (for details, see Sec. 6.2.4).

Background channel	Standard selection	Tight selection
qq	$(7.99 \pm 0.68) \cdot 10^{-4}$	$(2.30 \pm 1.15) \cdot 10^{-5}$
qqqq	$(1.486 \pm 0.094) \cdot 10^{-3}$	$(3.57 \pm 1.46) \cdot 10^{-5}$
$\ell\ell$	$(1.27 \pm 0.16) \cdot 10^{-4}$	$(4.18 \pm 2.95) \cdot 10^{-6}$
$\gamma^{B/W}\gamma^{B/W}$	$(2.13 \pm 0.28) \cdot 10^{-5}$	$(1.06 \pm 0.61) \cdot 10^{-6}$
N_{bg}	128,804	3,763

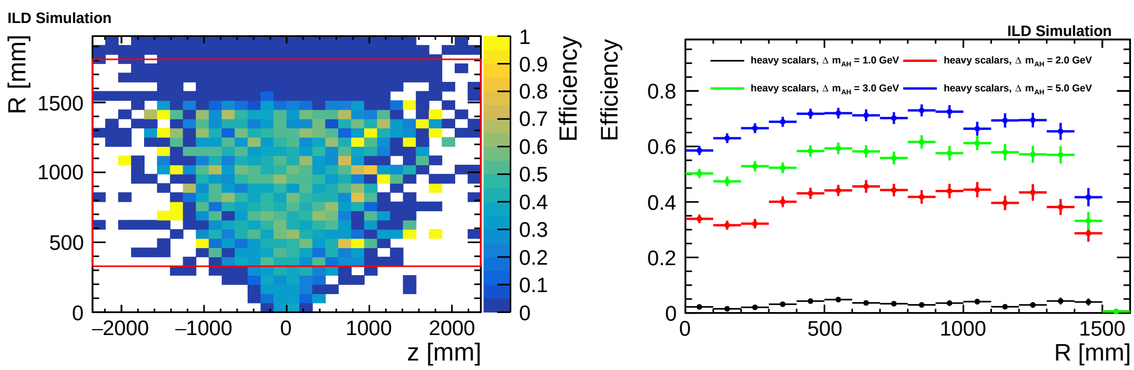


Figure 6.7: Vertex finding efficiency in the scalar pair-production scenario after the standard selection, but without the cut on the vertex radius. Left: the efficiency as a function of the true LLP decay vertex position in the detector for $\Delta m_{AH} = 2\,\text{GeV}$; the TPC volume is shown with the red box. Right: the efficiency as a function of the true LLP decay vertex radius for all scenarios considered in the analysis.

$m_a = 1\,\text{GeV}$ and larger masses; the results for the $m_a = 300\,\text{MeV}$ benchmark are limited by the cuts used to suppress background from semileptonic K^0 decays and misreconstructed photon conversions. It is important to note that the sensitivity in both $\Delta m_{AH} = 1\,\text{GeV}$ and $m_a = 300\,\text{MeV}$ scenarios could be significantly enhanced by a dedicated, model-dependent approach, using additional information about the missing energy or the hard photon produced in association with an ALP.

6.2.4 Model-independent cross section limits

Based on the obtained signal reconstruction efficiencies and expected background levels, the 95% C.L. limits on the signal production cross section are calculated. The total integrated luminosity of $2\,\text{ab}^{-1}$ collected in 10 years was assumed, with a relative share per beam helicity specified in the ILC H-20 scenario [3]. This corresponds to $9.7 \cdot 10^{11}$ BXs collected in the experiment, assuming

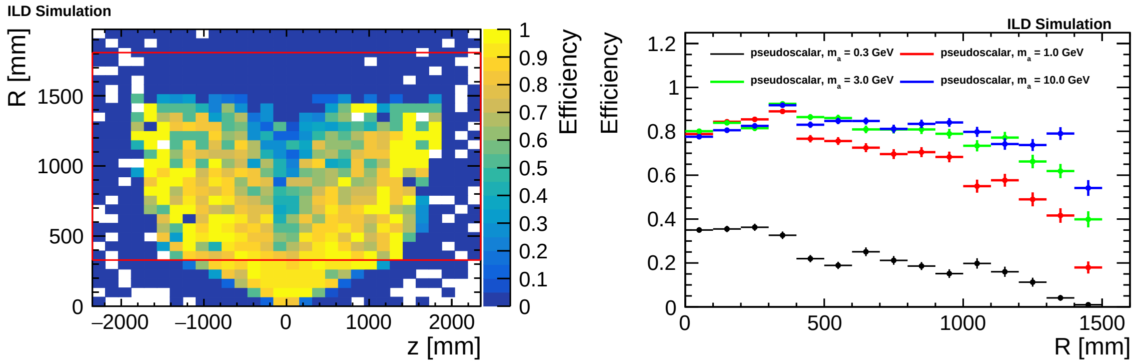


Figure 6.8: Vertex finding efficiency in the light pseudoscalar production scenario after the standard selection, but without the cut on the vertex radius. Left: the efficiency as a function of the true LLP decay vertex position in the detector for $m_a = 1$ GeV; the TPC volume is shown with the red box. Right: the efficiency as a function of the true LLP decay vertex radius for all scenarios considered in the analysis.

Table 6.5: The vertex finding efficiency inside the TPC region obtained in the analysis after different sets of cuts, both for scalars pair-production and light pseudoscalars for all considered scenarios.

Δm_{AH} [GeV]	1	2	3	5
Efficiency (standard) [%]	3	33.2	43.4	51.1
Efficiency (tight) [%]	0.4	28.3	40.7	50.2
m_a [GeV]	0.3	1	3	10
Efficiency (standard) [%]	7.4	48.4	61.7	65.8
Efficiency (tight) [%]	–	47.3	61.7	65.8

running with nominal luminosity. With 1.55 hadronic overlay events per BX on average, and incoherent e^+e^- pairs overlay in each BX (for details, see Section 6.2.1), this gives the total number of approximately $2.47 \cdot 10^{12}$ of the overlay events. The luminosity used for the hard $\gamma^B\gamma^B$ ($\gamma^B\gamma^W, \gamma^W\gamma^B$) interactions was 43% (53%) of the total integrated e^+e^- luminosity [211]. The final number N_{bg} of overlay and hard SM background events that remain after standard selection is 312 and 128,804, respectively, as calculated using the reduction factors obtained in Sections 6.2.1 and 6.2.2.

To estimate limits for different lifetimes without generating new samples and processing them, signal event reweighting is performed in order to estimate the limits for a set of LLP lifetimes for every benchmark scenario. For any given mean lifetime τ', an event with LLP decay after a time t_i is reweighted by a ratio $w_i = P(t_i, \tau')/P(t_i, \tau_0)$ of probabilities P from the exponential distribution, where the mean lifetime τ_0 was used to generate an event sample.

With the number N_p of events containing vertices that match to the MC vertex and pass through the selection, a fraction $\epsilon_{sel} = N_p/N_{\text{total}}$ must be calculated to obtain the limit. However, for a very long τ' considered, $c\tau' \gg 1$ m, the number of decays with a large decay length $ct_i \sim \mathcal{O}(c\tau')$ in a generated sample can be much smaller than the expected number of these events,

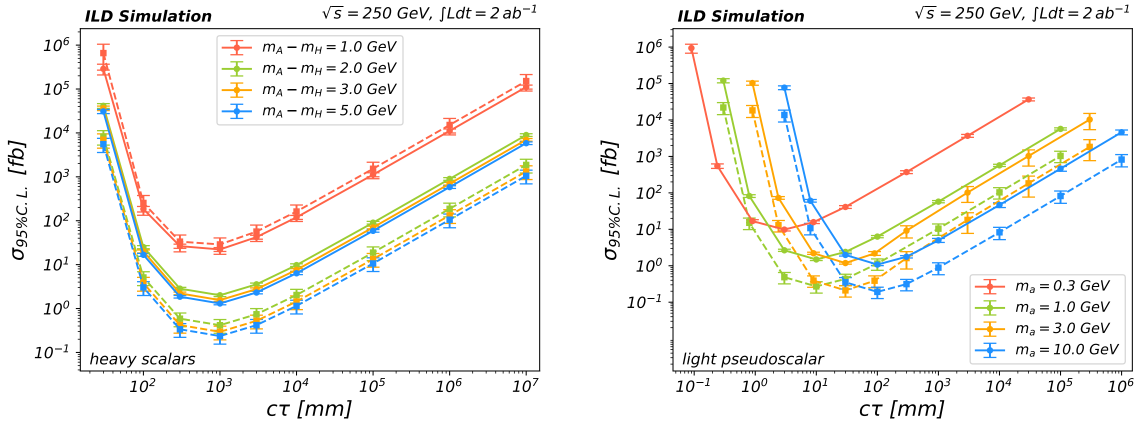


Figure 6.9: Expected 95% C.L. upper limits on the signal production cross section for the considered benchmarks and different LLP mean decay lengths, for the scalar pair-production (left) and the light pseudoscalar production (right) at $\sqrt{s}$ = 250 GeV. Solid lines corresponds to the standard selection and dashed lines to the tight set of cuts. The uncertainties are statistical.

leading to large event weights and statistical uncertainties. Therefore, a cutoff distance L_{max} = 3 m was defined, above which finding a vertex is impossible. Then, with w_{max} being the total probability that an LLP decays before L_{max} (calculated from the exponential distribution for a given τ'), $N_{\text{total}} = \Sigma_i w_i / w_{max}$, where only events with LLP decays before L_{max} are considered, and $N_p = \Sigma_i w_i$, where selection cuts are also applied.

The 95% C.L. limit $\sigma_{95\%\text{C.L.}}$ corresponds to the number $N_{lim} = 1.96\sqrt{N_{bg}}/\epsilon_{sel}$, where the factor 1.96 corresponds to the unified frequentist interval [212] for the measured number of events equal to the nominal expectation (which is equivalent to the CLs method), assuming $\epsilon_{sel} N_{lim} \ll N_{bg}$ and the Gaussian distribution for the background uncertainty. The limits for each of the considered benchmarks in the scalar pair-production (left) and the light pseudoscalar production (right) scenarios are presented in Figure 6.9. A range of mean decay lengths $c\tau$ is considered and two sets of limits are presented — for the standard and the tight selection. Also indicated in the figure are the statistical uncertainties resulting from the MC event selection for the signal and all background sources considered. Due to the relatively small size of the reprocessed MC event samples, the dominant contribution comes from estimates of physical event backgrounds, as described in Section 6.2.2.

In the case of scalar pair-production, good sensitivity can be observed in 0.3-10 m range for $\Delta m_{AH} \geq 2$ GeV. This can be extended by the tight selection, which improves the limits by an order of magnitude. For the $\Delta m_{AH} = 1$ GeV scenario, the tight selection slightly worsens the sensitivity, with the limits reaching the level of tens of femtobarns in the 0.3-3 m range.

For the light pseudoscalars, limits of the order of femtobarns can be obtained with the standard selection for $m_a \geq 1$ GeV in a 3-1000 mm range, depending on a scenario. Tight selection improves the limit by even an order of magnitude, allowing to reach the level of femtobarns (or lower) even up to $c\tau \simeq 10$ m. However, it causes a complete loss of sensitivity to the m_a = 300 MeV benchmark, which can be constrained down to $\mathcal{O}(10\,\text{fb})$ level using the standard selection.

6.3 Higgs decays to long-lived particles

An interesting feature of the model-independent analysis presented in the previous section is that it can be adapted and used as a basis for a search dedicated to a more specific scenario. In such a case, the sensitivity can be further improved by taking into account the properties of not only the displaced vertex signature, but also a full well-defined final state, which allows to optimise the selection procedure. In order to verify this approach, the analysis for neutral LLPs has been extended to a study of a particular BSM scenario, in which the SM Higgs boson decays to long-lived dark scalars. Such scenarios are motivated by neutral-naturalness and Hidden Valley models, and have been extensively studied in the context of future Higgs factories (see Section 3.5). However, the analysis is still performed without a framework of a specific BSM model so that the results are relevant for the whole class of models that predict such decays. Therefore, the limits can still be presented as a function of physical parameters — the cross sections and the branching ratios.

For the purpose of this analysis, a dedicated set of benchmarks with SM-like Higgs boson decays to long-lived scalars has been considered. To generate event samples, the Two-Real-Singlet model [213] was used, with the finite lifetime of one of the additional scalars, S, set at the generator level in the same way as for the heavy scalars and the light pseudoscalar (see Section 5.1.1). The samples were generated using `Whizard 3.1.2` with the implementation of the model in the UFO format. Four benchmarks were selected, two low-mass ($m_S = 0.4$ GeV and 2 GeV, generated with $c\tau = 10$ mm) and two high-mass ($m_S = 50$ GeV and 60 GeV, generated with $c\tau = 1$ m) scenarios. In this way, both regimes of slow and highly boosted LLPs are taken into account here as well.

The production channel considered was the Higgsstrahlung process, $e^+e^- \to hZ$, at 250 GeV ILC. In order to exploit a clean signal of one or two displaced vertices, Z decays to neutrinos were generated. For the dark scalar decays, a similar approach was taken here as in the model-independent study and only decays to muons were generated. In Hidden Valley models, the class of scenarios most commonly considered in searches for exotic Higgs decays, the dominant decay channel of the dark (pseudo)scalars is a bottom-quark pair. However, using muons in simulation allows to greatly speed up the simulation process, while a pair of highly displaced b-jets would be, in principle, even cleaner signature than a pair of tracks, since only decays inside the TPC are considered (as described below). This strategy was explained in more detail in Section 6.1.

6.3.1 Optimised background reduction

Since this analysis is meant as an optimisation of the model-independent search, all of the cuts described in Sections 6.2.1 and 6.2.2 are also used here, including the restriction to decays only within the TPC volume. Because Z decays to neutrinos, and one of the LLPs can escape the acceptance, the signature of at least one displaced vertex is considered. Therefore, in addition to the selection designed for the model-independent study and described so far, two more criteria are applied.

- Unlike scenarios analysed so far, tracks that form displaced vertices in the signal processes are expected to have a relatively high boost. Therefore, a requirement on their combined transverse momentum, $p_T^{vtx} > 10$ GeV, was applied to allow to fully neglect low-p_T beam-induced processes as a standalone background. This cut is also very efficient in rejecting vertices in high-p_T SM processes, since also in this case the sources of displaced vertices (such as V^0 particles or secondary interactions) typically result in a lower combined p_T of the tracks forming the vertex. Hard $\gamma\gamma$ interactions also become negligible after this requirement.

- Events containing additional prompt tracks with transverse momentum greater than 2 GeV are rejected. The threshold in p_T is necessary because the low-p_T $\gamma\gamma \rightarrow$ hadrons and e^+e^- pairs overlayed on the signal processes contain prompt tracks, therefore such events should not be rejected. A track is considered prompt here if one of its impact parameters, d_0 or z_0, is smaller than 10 mm and if its first hit has a radius smaller than 20 mm. The radius requirement is changed to less than 155 mm (first SIT layer) if the position of the first hit along the z axis is greater than 215 mm from the IP (slightly less than the first FTD layer), since the first SIT layer encapsulates the first two layers of the FTD (see the left plot in Figure 4.1 and Table 4.1).

The approach for the background estimation and the samples used are the same as for the model-independent search. Because here the selection is very efficient, and the size of background samples is limited due to the necessity of their reprocessing, for some of the processes considered, the remaining number of events after the cuts is zero. Therefore, the final selection efficiency was estimated in a similar way to the approach used for overlay events, where the vertex finding rates were factorised (see Section 6.2.1). Since the cut on additional prompt tracks concerns tracks and variables completely unrelated to the displaced vertex and the quantities that describe it, this criterium is fully independent of the rest of selection. Therefore, total efficiency is estimated as a product of the finding rate for the standard selection combined with the $p_T^{vtx} > 10$ GeV requirement, with the criterion of no prompt tracks as a separate cut.

The tight selection is not considered in this analysis, because when combined with the $p_T^{vtx} > 10$ GeV cut, it removes all events from the $q\overline{q}$ sample already prior to applying the criterion on additional prompt tracks. It also strongly reduces the sensitivity to the scenario with $m_s =$ 400 MeV.

The efficiencies corresponding to different sets of cuts are presented in Table 6.6. The efficiency for the prompt-track criterion is calculated as a ratio of events with vertices that survive the cut to all events with identified vertices in a sample. The qqqq and hard $\gamma^{B/W}\gamma^{B/W}$ processes become negligible after selection — the former does not survive the requirement on the lack of prompt tracks in an event, and the latter is already rejected by the cut on the transverse momentum of tracks at the vertex. The only remaining background comes from the di-jet and di-lepton processes. What can evade the cut aimed at removing events with prompt tracks, is the case when a hard ISR photon is produced, strongly decreasing the energy of one of the incident electrons. Then it can happen that the transverse momentum of none of the tracks coming from the IP exceeds 2 GeV.

The expected number of background events after the standard selection and the additional cuts is 8.36 ± 4.11, considering all the processes listed in Table 6.3, where the uncertainty is statistical and is combined between different channels. The size of the error results from a small

Table 6.6: The vertex finding rates obtained for different simulated background samples in the analysis. Results are shown separately for the standard selection from the model-independent study extended by a cut on p_T^{vtx} and for the requirement of no prompt tracks in an event. Combined efficiencies are also presented, as well as statistical uncertainties.

Background channel	Standard selection and $p_T^{vtx} > 10\,\text{GeV}$	Prompt tracks rejection	Combined
qq	$(2.82 \pm 0.68) \cdot 10^{-4}$	$(1.34 \pm 0.95) \cdot 10^{-4}$	$(3.78 \pm 2.82) \cdot 10^{-8}$
qqqq	$(4.28 \pm 0.94) \cdot 10^{-4}$	—	—
$\ell\bar{\ell}$	$(2.00 \pm 0.21) \cdot 10^{-4}$	$(1.35 \pm 0.78) \cdot 10^{-3}$	$(3.42 \pm 1.37) \cdot 10^{-7}$
$\gamma^{B/W}\gamma^{B/W}$	—	0.1972 ± 0.0044	—

number of events in the MC samples that survive the selection despite its factorisation.

6.3.2 Vertex finding results

The results of the selection procedure for signal are presented in Figure 6.10 as a function of the radius, R, for all scenarios considered in the analysis of Higgs decays and as a function of the LLP decay position in the R-z plane for selected benchmarks. The performance visible in these plots exhibits a very similar behaviour to the corresponding scenarios in the model-independent analysis. For low-mass scenarios, the relation between efficiency and LLP decay position is almost identical to that for light pseudoscalar production, including the visible resemblance between the $m_S = 400\,\text{MeV}$ and the $m_a = 300\,\text{MeV}$ scenarios. For high masses of the dark scalar, the results correspond to heavy-scalar production scenarios, as expected due to relatively small boost of LLP in these benchmarks. The dependence on the radius is particularly similar to the $\Delta m_{AH} = 5\,\text{GeV}$ scenario (especially for small radii), which is the least extreme case of the considered benchmarks for heavy scalar production. The difference in efficiency between the two high-mass scenarios results from the requirement on the opening angle between two tracks forming a vertex, $\cos\theta_{vtx} > -0.6$, introduced in the preliminary selection for the model-independent analysis (see Section 6.2.1). The lowest kinematically allowed opening angles for $m_s = 60\,\text{GeV}$ are significantly higher than for $m_s = 50\,\text{GeV}$, leading to only a limited part of the phase space passing the cut.

The total efficiencies obtained after the full selection procedure for all considered benchmarks are presented in Table 6.7. They are calculated as in Section 6.2.3, considering the ratios of the number of "correct" vertices both to all decays in the detector acceptance, as well as only to decays inside the TPC volume. Again, there is a strong correspondence visible between these results and those presented in Table 6.5, in particular for low-mass scenarios. However, for high masses of the dark scalar S, the total efficiency is slightly decreased compared to the scenario with $\Delta m_{AH} = 5\,\text{GeV}$ due to the increased threshold for p_T^{vtx}.

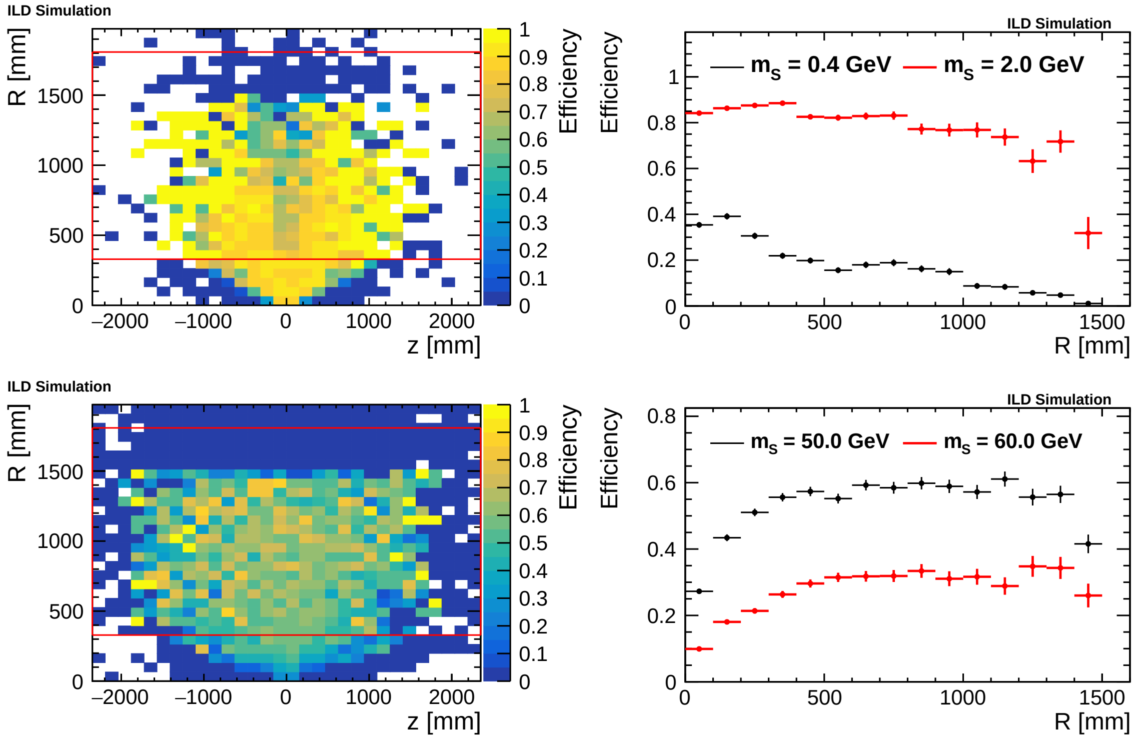


Figure 6.10: Displaced vertex finding efficiency, in different scenarios of exotic Higgs decays, after the full selection but without the cuts on the radius and the distance d_{11} between the first hits on the tracks. Left: the efficiency as a function of the true LLP decay vertex position in the detector for m_S = 2 GeV (top) and 50 GeV (bottom); the TPC volume is shown with the red box. Right: the efficiency as a function of the true LLP decay vertex radius for all benchmakrs considered in the analysis, for low-mass (top) and high-mass (bottom) scenarios.

6.3.3 Cross section and branching ratio limits

The corresponding 95% C.L. limits on the signal production cross section (left) and on the branching ratio BR(h → SS) = $\sigma_{95\%\,\mathrm{C.L.}}/\sigma_{h\nu\nu}$ (right) are presented in Figure 6.11. A procedure analogous to the one used in the model-independent analysis was also applied here to obtain the limits for a range of lifetimes (for details, see Section 6.2.4), this time taking into account that there are two LLPs in the final state and the nominator in the event weight is a product of the corresponding probabilities. w_{max} in this case was defined as a product of total probabilities that each of the LLPs decays before L_{max}, where both LLPs had to decay before L_{max} for the event to be considered in the sum $N_{\mathrm{total}} = \Sigma_i w_i / w_{max}$ (again, to avoid too large event weights and uncertainties).

The limits are obtained by assuming the background-only hypothesis with a mean value N_{bg} equal to 8.36 and by averaging over the expected limits for an ensemble of experiments with different observed numbers of events and the same mean value [212]. The uncertainty on the expected number of background events is taken into account by calculating the additional limits assuming the mean value of $N_{bg} \pm 1\,\sigma$. The limits obtained in this way are presented in Figure 6.11 by a shaded band around the corresponding solid lines. It can be noted that such an approximately 50% variation in the expected number of background events has a relatively

Table 6.7: The displaced vertex finding efficiency for all scenarios of exotic Higgs decays after the full selection, considering all decays in the detector acceptance and only decays inside the TPC.

m_S [GeV]	0.4	2	50	60
Efficiency (total) [%]	5.1	14.5	18.5	7.4
Efficiency (TPC) [%]	7.8	52.4	34.8	18.5

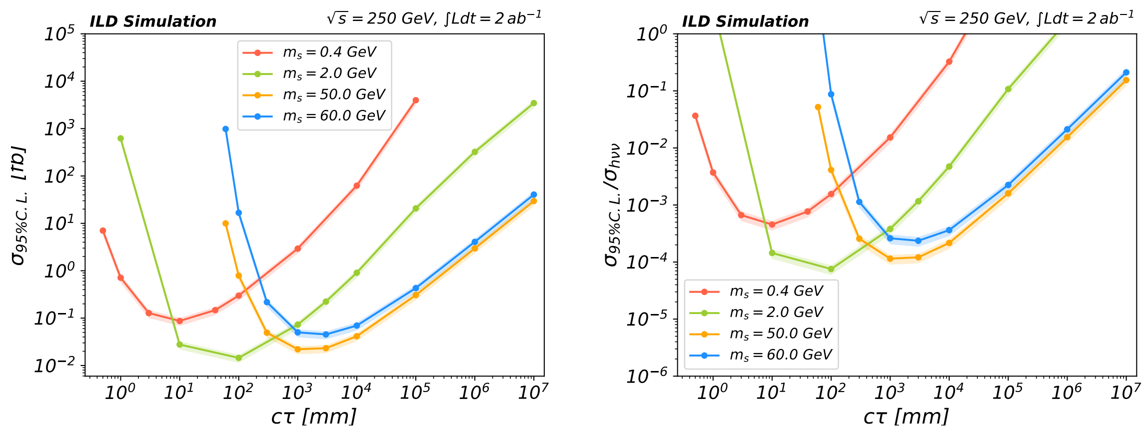


Figure 6.11: Expected 95% C.L. upper limits on the signal production cross-section (left) and the branching ratio (right) for the considered benchmarks and different LLP mean decay lengths, for the Higgs decays to long-lived scalars at $\sqrt{s}$ = 250 GeV. Shaded bands correspond to the limits calculated assuming $\pm 1\,\sigma$ on the expected number of background events (see text for details).

small impact on the resulting 95% C.L. limits.

In terms of cross sections, a clear improvement was achieved with respect to the general model-independent analysis, although one has to note that in this case two LLPs are produced per event. The limits are stronger by 1-2 orders of magnitude than the corresponding ones for light and heavy (pseudo)scalars, depending on the scenario. Particular progress is visible in the light LLP regime, where the results for m_S = 400 MeV are more than two orders of magnitude better than the corresponding ones for a light pseudoscalar production with m_a = 300 MeV.

The limits on branching ratio can provide more insight into the sensitivity to the specific scenario of exotic Higgs decays to LLPs. However, direct comparisons, e.g. with the LHC results, are ambiguous and not easily interpretable because of a different approach typically used in LHC searches. They are not at the same level of generality as the analysis presented in this work, because all the corresponding LHC studies that the author is aware of — if not fully model-dependent — are signature- or final-state driven. They target a particular final-state in the analysis, for example, a pair of muons or a pair of jets (sometimes of a specific type, e.g. b-jets) from the LLP decay. As a result, it is unclear how such limits should be compared to a search presented here, in which any final state was allowed (despite the simulated event samples involved only decays to muons). Nevertheless, the author believes that the results presented here provide valuable information about the sensitivity of the ILD operating at the ILC to exotic Higgs decays. For this reason, some comments on the prospects for the ILD with respect to selected constraints from ongoing experiments are given below, and finally a direct comparison with selected results is also presented.

The ILD could provide an improvement particularly by extending the current limits toward longer lifetimes. LLPs below masses of 2 GeV in Higgs decays are excluded by CMS down to branching ratios of the order of 10^{-4} for decay lengths up to approximately 1 mm [214]. At ATLAS, dark scalars with a mass of 5 GeV were the lightest considered [215, 91]. In this case, the branching ratios of 1% are excluded for $c\tau$ = 0.05-2.9 m and the best limits reach the level of 10^{-3} around $c\tau$ = 0.2-0.3 m. For $c\tau \gtrsim 10$ mm all these constraints for light LLPs could be improved by the ILD.

The strongest existing limits on Higgs decays to LLPs have been obtained by CMS assuming dark photon, Z_D, decays to muons (with a branching ratio of the order of 10%) and the LLP masses of 10-60 GeV [216]. In this scenario, $BR(h \to Z_D Z_D)$ below 10^{-4} can be probed for decay lengths between 0.1 mm and 4 m, almost down to 10^{-5} around 10-100 mm. Such an impressive result has been obtained using the combination of searches in the inner tracker and muon chambers. However, for other signatures or models, slightly worse limits were obtained. For LLPs with the mass of 50 GeV the sensitivity only down to 10^{-4} was reached around $c\tau \sim 10$ mm assuming Higgs decays to dark scalars, S, and $BR(S \to ee) = BR(S \to \mu\mu) = 0.5$ [217]. For the 55 GeV mass, the limits are even weaker, probing Higgs branching ratios to dark scalars down to the level of 10^{-3} for $c\tau \sim 10$ mm using displaced jets [93] and $c\tau \sim 4$ m in searches for clusters in the muon spectrometer [90]. The corresponding results from ATLAS are less competitive, testing branching ratios slightly below 10^{-3} in the least challenging scenarios [91].

As mentioned above, despite the significant differences between the approaches in this work and in the LHC searches, a comparison of the limits on the branching ratios obtained in this analysis to the selected CMS results (the most competitive ones) is presented in Figure 6.12, for low-mass (left) and high-mass (right) scenarios. The LHC searches for LLP decays to muons were selected, with the exception of the LLP mass of 400 MeV, where such an analysis was not available, and the e^+e^- final-state was chosen instead as the most similar signature for this mass.

In the low-mass regime, the limits obtained in this thesis were compared to the search for Higgs decays to dark scalars [90] and dark photons [218] with masses of 400 MeV and 2 GeV, respectively, corresponding to masses considered in this work. In the search for dark scalars, decays inside the muon spectrometer were targeted, allowing to obtain the limits for relatively long lifetimes, which could be improved by the ILD in a similar range. It should be noted that while the results for electrons are shown here, many different final-states were considered in the original analysis, although for all of them the reach down to the branching ratios of approximately 10^{-3} was achieved. In the case of dark photons, in which the di-muon final state was used, the LHC reach could be complemented by the ILD in the region of higher decay lengths thanks to the vast TPC acceptance. A recent study (mentioned above) [214] slightly improves this result from CMS — however, only for smaller masses and shorter lifetimes.

For high LLP masses, the comparison is shown only for a mass of 60 GeV, with two different interpretations, to highlight the importance of model assumptions in LHC searches. The presented results correspond to Higgs boson decays to dark photons in the Hidden Abelian Higgs Model (HAHM), with subsequent decays to muons, for which the best existing limits were achieved [216], stronger than those obtained in this thesis. However, when this search is rein-

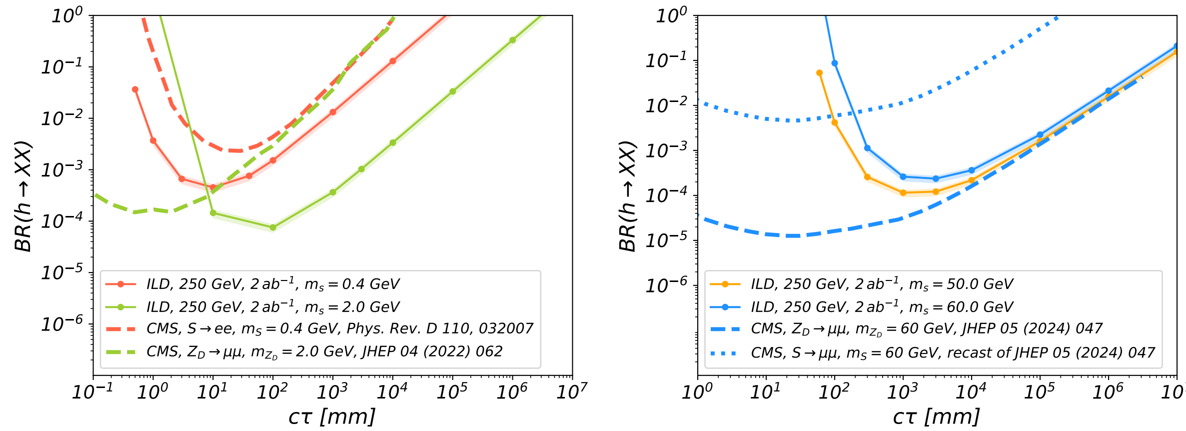


Figure 6.12: Expected 95% C.L. upper limits on the branching ratio of the SM Higgs decays to LLPs for low-mass (left) and high-mass (right) scenarios considered in this study, compared with selected results from the CMS experiment [90, 218, 216, 219]. Different models were considered in the presented studies, and hence X indicates any LLP. See text for details.

terpreted within the framework of the scalar portal model, the same limit becomes more than two orders of magnitude weaker [219]. This is because in HAHM assumed in the original work, the branching ratio of the LLP to muons is of the order of 10%, whereas in the scalar model the LLP decays almost exclusively to $b\bar{b}$ for masses greater than 10 GeV. However, the $b\bar{b}$ channel allowed constraining the Higgs branching ratios only down to the level of 10^{-3} (in the search for decays in the muon spectrometer mentioned above) for the dark scalar mass of 55 GeV [90].

Such a decrease in sensitivity will not occur in the case of the analysis presented in this thesis. The search is designed specifically to detect displaced vertices, irrespective of the (multi-prong) LLP decay mode, in particular allowing for displaced jets in the final state. In this case, it is hard to imagine that decays to b-jets inside the TPC will be missed in the reconstruction or in the analysis if vertices with just two tracks can be detected with high efficiency.

6.4 Search for kinked tracks

6.4.1 Background reduction

In this analysis, the same two types of background — overlay events and high-p_T SM processes — as for the neutral LLP search have been considered. However, in the case of kinked tracks, fully model-independent overlay reduction would be much more challenging. This is because soft $\gamma\gamma$ interactions and incoherent e^+e^- pairs can abundantly produce low-p_T kinked tracks that are indistinguishable from very low-momentum signal processes without using additional assumptions about the final state. Therefore, much stronger cuts are applied on the p_T in this search to fully reject kink vertices in overlay events, as explained in more detail later in this section. On the other hand, this allowed to include decays in the silicon tracker in the analysis. Furthermore, for the high-p_T backgrounds the dilepton production was now fully considered in addition to purely and semi-hadronic processes in order to account for the bremsstrahlung of electrons, as well as for hadronic τ decays, both of which can be a source of kinked tracks. For more details about the high-p_T backgrounds refer to Section 6.2.2 and Table 6.3.

Preliminary cuts

In the case of `KinkFinder`, a set of quality cuts is already implemented as a standard part of the algorithm. Nevertheless, the full set of selection criteria is listed and explained below. The selection procedure described here is the default one as implemented in the processor and used in the reconstruction chain, unless explicitly stated that some modifications have been introduced.

First, a selection related to the quality of the tracks and the position of the kink is applied, as well as a preliminary requirement on their relative momenta, p_1 and p_2, where p_1 is the momentum of the parent track candidate.

- As mentioned above, kink vertices in the inner silicon detector have been considered in addition to decays inside the TPC volume. Thefore, the allowed radii of the vertices are $R_{\text{kink}} > 60\,\text{mm}$, which corresponds to the radius of the last barrel layer of the VTX detector. Consistently, the radii of innermost hits of the parent and daughter track candidates must be greater than 60 mm. The default value of this cut was 100 mm and was changed to increase acceptance.

- In order to reject poor quality tracks, a minimum of five hits on both parent and daughter tracks were required. This is in agreement with the previous cut, as the VTX consists of three double layers of silicon, so the parent particles that decay just after the last layer should produce six hits.

- The preliminary cut on the relative momenta, $|1 - p_2/p_1| > 0.02$, is used to reject the split candidates that are the easiest to identify. This criterion has been introduced by the author of this thesis, since split tracks are considered solely as a background in this analysis.

The requirements are applied on the differences Δr_k between the p.c.a.'s of track helices and $\Delta r_{xy} = r_d^{in} - r_p^{out}$ between the radii of the innermost hit of the daughter track and the outermost hit of the parent track. The values of the corresponding cuts, ΔR_k^{cut} and $\Delta R_{xy}^{\text{cut}}$ respectively, are defined dynamically, based on the kink candidate position in the detector.

- **TPC region**: if only R_{kink} passes the requirement of being outside the VTX, the cut values are assigned first assuming the kink is inside the TPC. The maximum Δr_k is set to $\Delta R_k^{\text{cut}} = 30\,\text{mm}$ and the maximum Δr_{xy} to $\Delta R_{xy}^{\text{cut}} = (\Delta R_{TPC}/N_{row}) \cdot 10$, where $\Delta R_{TPC} = R_{TPC}^{out} - R_{TPC}^{in}$ is the difference between the maximum and minimum radii of the TPC readout and N_{row} is the number of pad rows in TPC. If $\Delta r_k < 5\,(1)\,\text{mm}$, the cut on Δr_{xy} is increased by a factor of 1.5 (2).

- **SIT outer regions**: if $R_{\text{kink}} < R_{TPC}^{in} + \Delta R_{xy}^{\text{cut}}$ the kink is assumed to be in SIT, and ΔR_k^{cut} is set to 10 mm. If the kink candidate is between VTX and SIT, $\Delta R_{xy}^{\text{cut}}$ is set to the difference in radii between the innermost SIT layer and the outermost VTX layer, and if it is between SIT and TPC, $\Delta R_{xy}^{\text{cut}}$ is set to $1.15 \cdot R_{TPC}^{in} - R_{SIT}^{out}$, where R_{SIT}^{out} is the radius of the outermost SIT layer. The factor 1.15 is an additional buffer to take into account that a few hits created by the parent track just after entering the TPC may not have been successfully assigned to the same track.

- **Inside SIT**: if the kink candidate is between the SIT layers, $\Delta R_{xy}^{\mathrm{cut}}$ is set to the difference between the innermost and outermost layers of SIT. However, sometimes hits from the last SIT layer do not get assigned to the corresponding TPC segment. Therefore, if $|r_d^{in} - R_{TPC}^{in}| < 50\,\mathrm{mm}$, $\Delta R_{xy}^{\mathrm{cut}}$ is set to the difference between R_{TPC}^{in} and the radius of the next to last layer of SIT.

- **Additional protection for SIT**: if both r_d^{in} and r_p^{out} are less than 10 mm from one of the other SIT layers, $\Delta R_{xy}^{\mathrm{cut}}$ is set to at least the gap between those other layers (with a minimum of 10 mm). This can happen, for example, if the parent particle decays between the double layers of SIT. In the last step, $\Delta R_{xy}^{\mathrm{cut}}$ is increased by 10% to account for geometry.

When $\Delta R_k^{\mathrm{cut}}$ and $\Delta R_{xy}^{\mathrm{cut}}$ are determined, the kink candidate is accepted for further analysis only if $\Delta r_k < \Delta R_k^{\mathrm{cut}}$ and $\Delta r_{xy} < 2\Delta R_{xy}^{\mathrm{cut}}$. These criteria are maintained for split candidates, while for kinks to be considered as potentially originating from decays of charged particles (both SM and BSM), a tighter cut $\Delta r_{xy} < \Delta R_{xy}^{\mathrm{cut}}$ is used.

To determine whether a kink possibly originates from a split track, requirements are set on the consistency of the momenta of the two tracks.

- The absolute difference between the track momenta, $|p_1 - p_2|$, must be less than 0.2 GeV.

- The fractional difference between the momenta, $2|p_1 - p_2|/(p_1 + p_2)$, must be less than 0.02.

If both of these conditions are met, the kink is further considered as a possible split, and additional criteria on the hit pattern and charge consistency of the tracks are evaluated.

A set of tests is performed to determine whether a kink candidate originates from a decay of an SM hadron. For this purpose, a score function p is used to assess the level of agreement of a kink candidate with an SM particle $X = \mathrm{K}^{\pm}, \pi^{\pm}, \Sigma^{+}, \Xi^{-}$:

$$p = 3.125 \cdot \left(\frac{m_{\mathrm{rec}} - M_X}{\Delta m_X^{\mathrm{cut}}} \right)^2 + t_X. \tag{6.1}$$

m_{rec} is a parent particle mass calculated assuming the masses of particular decay products of a particle X (e.g. $\mu^{\pm}\nu$ for $X = \pi^{\pm}$), M_X is the expected value of the X mass and $\Delta m_X^{\mathrm{cut}}$ is an input parameter. If multiple decay channels of X are possible, the lowest value of p is selected for a given particle. t_X is the ratio of the decay length of a particle to its proper decay length, taking into account its Lorentz boost:

$$t_X = \frac{L_z M_X}{p_z c\tau_X}, \tag{6.2}$$

where L_z is the distance along the beam axis from the start of the track to the vertex, p_z its momentum along z, and $c\tau_X$ the proper decay length of a given particle. A kink candidate is further considered a potential SM decay if $p > 10^{-4}$ for any X and the track charges are consistent. If it also passes the cuts on Δr_k, Δr_{xy} and on charge consistency, the particle type is determined based on the highest value of p and the kink vertex is classified as `KinkVertices`.

Kink candidates that did not get assigned to any of the above categories and passed the criteria $\Delta r_k < \Delta R_k^{\mathrm{cut}}$ and $\Delta r_{xy} < \Delta R_{xy}^{\mathrm{cut}}$, are stored as `ProngVertices`.

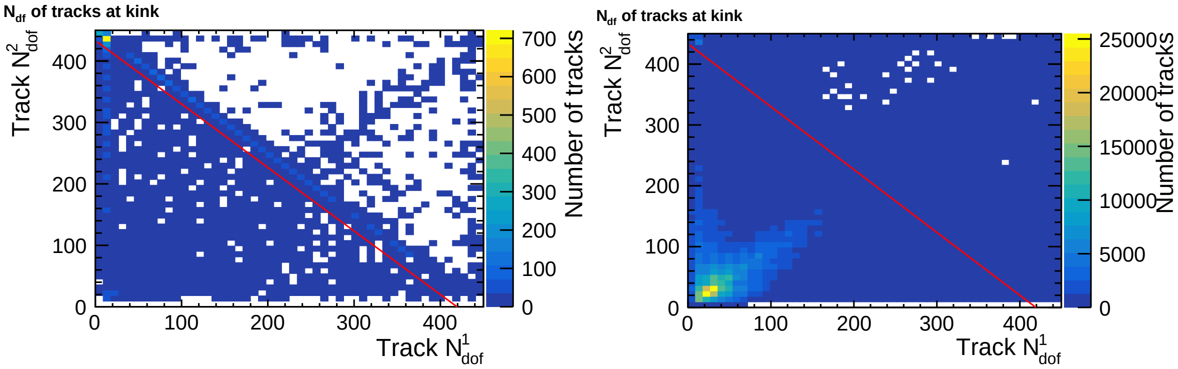


Figure 6.13: Distributions of number of degrees of freedom used to fit tracks forming the identified kink in the signal sample with m_F = 110 GeV, m_s = 109 GeV (left) and for $q\overline{q}$ process (right). The variables are used in the selection and the red line indicates the cut threshold (see text for details). All histograms are normalised to the number of simulated MC events.

Final selection

The preselection used in `KinkFinder` and described above allows for partial rejection of backgrounds from poorly reconstructed (e.g. split) tracks, random coincidences, interactions with detector material, and delayed SM decays. However, more strict selection is needed to further reduce these background sources and increase sensitivity to rare processes. Therefore, preselected `ProngVertices` undergo additional, final selection, aimed at suppressing the SM backgrounds.

In particular, further rejection of fake vertices with unexpectedly short tracks and split segments was necessary. Hence, additional criteria were applied on the cosine of reconstructed kink angle, θ^{rec}_{kink}, and on the number of degrees of freedom, N_{df}, in the track fit (which corresponds to the number of hits used to fit a track).

- To fully eliminate the vertices that originate from split tracks, a strong cut was applied on the kink angle, requiring $|\cos\theta^{rec}_{kink}| < 0.995$. The absolute value of cosine is used because in the split tracks, segments may also have been flipped in the reconstruction, in particular in the low-p_T curling tracks.

- Number of hits on the parent and daughter tracks are expected to be exactly correlated. As visible in Figure 6.13, this is the case for the signal, but not for the $q\overline{q}$ events. A cut using linear dependence between $N^{(1)}_{df}$, $N^{(2)}_{df}$ of the parent and daughter tracks is used to reject kink candidates with unexpectedly short tracks, selecting only kink vertices where $1300/3 - 31N^{(1)}_{df}/30 < N^{(2)}_{df}$. The coefficients were chosen for the most efficient background rejection. ‘

- In order to remove vertices coming from backscattering of tracks with outer walls of the TPC, the radius of the vertex was required to be smaller than 160 cm and its position in z, along the beam axis with respect to IP, to be smaller than 200 cm.

The parent particles forming kinked tracks are expected to be prompt and have relatively high momentum. For instance, a dark fermion with a mass as large as 124 GeV, produced at $\sqrt{s}$ = 125 GeV collider, would have a peak in the momentum distribution around 15 GeV.

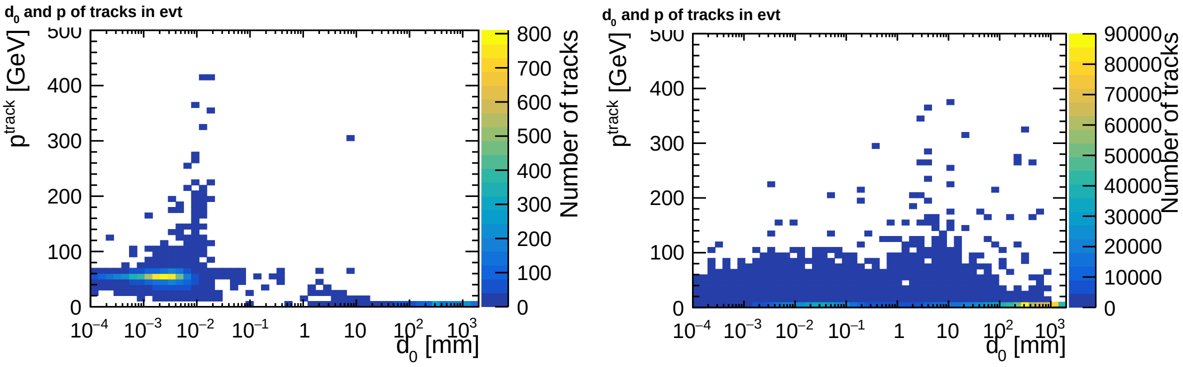


Figure 6.14: Distributions of the momenta of tracks forming the identified kink as a function of their impact parameter d_0, for the signal sample with $m_F = 110$ GeV, $m_s = 109$ GeV (left) and for $q\overline{q}$ process (right). All histograms are normalised to the number of simulated MC events.

Figure 6.14 presents the distributions of the momentum and impact parameter for a selected signal scenario and $q\overline{q}$ events. For the signal, the regions corresponding to the parent and daughter tracks are clearly visible for low and high values of d_0. In the case of background, both peaks occupy low-momentum regions. Unexpectedly high momenta in these distributions result from a much worse momentum resolution for short tracks. Therefore, the following criteria are applied on the momenta of the parent and daughter tracks and on impact parameters.

- If a track that forms a kink has one of its impact parameters, d_0 or z_0, smaller than 1 mm it is required to have a total momentum greater than 10 GeV.
- The parent tracks (the ones with the innermost hit closest to the IP) must have $d_0 \leq 1$ mm and $z_0 \leq 1$ mm.
- The kink candidate is rejected if both the parent and the daughter tracks have momenta smaller than 10 GeV.

These cuts, in particular the last one, also allow to fully neglect the kink vertex finding rates in the low-p_T $\gamma\gamma \rightarrow$ hadrons and incoherent e^+e^- pairs events by explicitly requiring high-momentum parent tracks.

The resulting kink vertex finding rates for the simulated SM processes are summarised in Table 6.8. The backgrounds after selection are an order of magnitude lower than in the case of standard selection and around a factor of 3 higher than for tight selection, as compared to the search for displaced vertices. The greatest improvement with respect to the standard selection for displaced vertices comes from much stronger requirements on the momenta and impact parameters. As described in detail in the following subsections, such a background suppression allows to achieve similar limits to the tight selection for neutral LLPs, partly due to the search acceptance covering also the SIT region.

6.4.2 Kink vertex finding results

The kink finding efficiency following the selection described in the previous section is shown in Figure 6.15, on the left as a function of the LLP decay position in the TPC for selected signal

Table 6.8: The kink finding rates directly obtained for the high-p_T SM backgrounds after the full set of cuts. Large statistical uncertainties for qq and γγ result from small number of MC events remaining after the selection. The total number, N_{bg}, of background events expected after selection, taking into account all channels presented in Table 6.3, is also shown in the last row. N_{bg} is calculated assuming the total luminosity of $2\,\mathrm{ab}^{-1}$ (for details, see Section 6.2.4).

Background channel	Kink finding rate
qq	$(4.09 \pm 1.55) \cdot 10^{-5}$
qqqq	$(1.06 \pm 0.24) \cdot 10^{-4}$
$\ell\bar{\ell}$	$(2.00 \pm 0.21) \cdot 10^{-4}$
$\gamma^{B/W}\gamma^{B/W}$	$(6.73 \pm 4.76) \cdot 10^{-7}$
N_{bg}	10,874

scenarios, and on the right as a function of the radius for all considered benchmarks. The kink vertex is considered "correct" if it is reconstructed within 30 mm of the true decay vertex.

The greatest loss of efficiency for all scenarios is related to the criterion on the $N_{df}^{(1)}$ and $N_{df}^{(2)}$ relation, which is nevertheless necessary to reduce the level of background. The cut results in fluctuations and mostly an irregular dependence of the kink-finding efficiency on the radius. For example, the biggest dip in efficiency, visible for all scenarios around 400-500 mm, can be explained by the removal of many vertices with $N_{df}^{(1)} < 20$ and $N_{df}^{(2)} < 420$. This corresponds to decays just after the LLP entered the TPC, where the hits of parent tracks inside the TPC may not have been picked up by the tracking algorithm.

Despite the fluctuations, the overall efficiency does not change much within the detector, in particular for scenarios with large and moderate mass differences. For these benchmarks, the performance is also very similar, including the relation between efficiency and radius, with slightly better results for the moderate scenarios. For small mass differences, the efficiency increases visibly with distance from the beam axis. This is due to the cut on $N_{df}^{(1,2)}$, which was designed to keep in particular those kinks, for which the total number of hits of the two tracks stays at a constant level. However, at the same time, it was also meant to preserve kinks with low-p_T daughter tracks — because they curl in the detector, they produce a significant number of hits, if only the LLP decayed far enough. Since the cut does not remove the vertices where both tracks produced many hits, for kinks with soft leptons the chance of not being rejected increases with the radius of the vertex.

The total kink-vertex finding efficiencies for all the benchmarks considered in the analysis are presented in Table 6.9. The results are comparable between low- and high-mass scenarios, with the performance slightly better for light LLPs. In particular, the efficiencies for large and moderate mass differences are very similar, while the performance starts to get worse visibly as the final state becomes softer. For all scenarios, the purity, defined as the ratio between the number of "correct" vertices to all found vertices that survive selection, exceeds 99%.

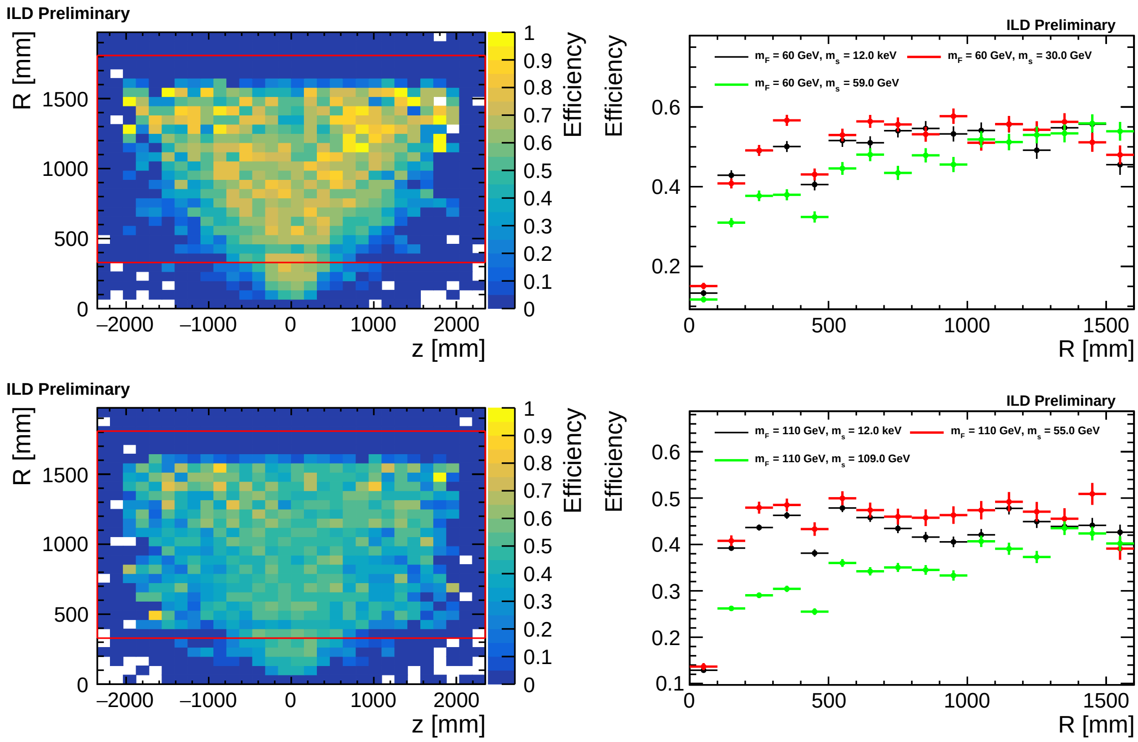


Figure 6.15: Kink finding efficiency after the full selection in different dark fermion production scenarios, with m_F = 60 GeV (top) and m_F = 110 GeV (bottom). Left: the efficiency as a function of the true LLP decay vertex position in the detector for m_s = 12 keV; the TPC volume is shown with the red box. Right: the efficiency as a function of the true LLP decay vertex radius for all scenarios considered in the analysis.

6.4.3 Limits on the production cross section

The obtained expected background levels and signal selection efficiencies were used to calculate 95% C.L. limits on the signal production cross section, in the same way as described in Section 6.2.4. However, in this analysis, overlay events as a stand-alone background could have been neglected following the selection of parent tracks with high momenta (for more details, see Section 6.4.1). The same collider running conditions as in the neutral LLP search were also assumed (the ILC H-20 scenario), resulting in the expected total of 10,874 background events. The same procedure as for the Higgs decays (see Sections 6.2.4 and 6.3.3) was performed to obtain the limits for a range of lifetimes.

The resulting 95% C.L. cross-section limits for all considered benchmarks are presented in Figure 6.16 as a function of proper decay lengths $c\tau$. The errors are purely statistical and originate from the sizes of the MC signal and background samples used in the study. In most cases, the main contribution to the uncertainty comes from the relatively small size of the reprocessed MC background samples. However, for short lifetimes, larger errors can be explained by low signal selection efficiencies for LLP decays at small radii.

For low masses of the dark fermion, the sensitivity below 1 fb was obtained for decay lengths from around 30 mm to almost 20 m. The best results for short lifetimes correspond to moderate mass splitting, with $m_s = m_F/2$, since the kink vertex finding efficiency closer to the beam is

Table 6.9: The kink vertex finding efficiency in the considered search region, for all scenarios of dark fermion pair-production, after the full selection.

m_s	12 keV	$m_F/2$	$m_F - 1$ GeV
	m_F = 60 GeV		
Efficiency [%]	47.2	48.7	40.7
	m_F = 110 GeV		
Efficiency [%]	45.4	43.3	30.2

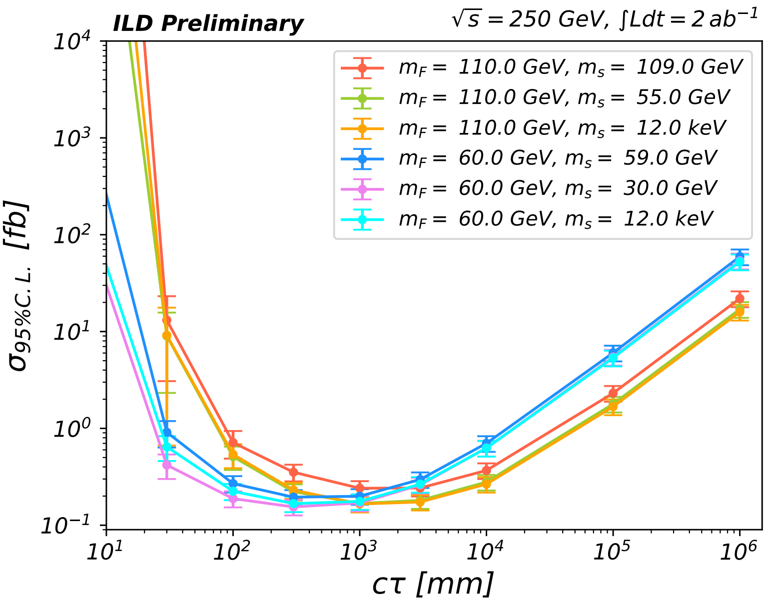


Figure 6.16: Expected 95% C.L. upper limits on the signal production cross section for all benchmarks considered in the charged LLP pair-production at $\sqrt{s}$ = 250 GeV, shown as a function of the LLP proper decay length. The uncertainties are statistical.

the best for this scenario. This is because for a very small mass difference, leptons tend to have a lower p_T, so their reconstruction (particularly in the silicon tracker) is more challenging; on the other hand, for a very large mass difference, a lepton can have a much higher p_T, increasing the chance that its track is merged in the reconstruction with a short parent track. For higher values of $c\tau$, the results start to converge, with the limits for low-mass-difference scenarios being minimally weaker.

In the case of scenarios with m_F = 110 GeV, cross sections below the level of 1 fb can be reached for decay lengths between around 80 mm and 70 m. The results for large and moderate mass differences are nearly identical. The scenario with high dark fermion mass and small mass splitting is the most challenging one of all considered benchmarks, with the limits slightly (but visibly) weaker in this case due to the lowest signal selection efficiency in the majority of the tracker volume.

Chapter 7

Summary

Results presented in this thesis confirm that the future e^+e^- Higgs factories have great potential for LLP searches. Described is the sensitivity to displaced signatures of the ILD experiment operating at the ILC accelerator running at $\sqrt{s} = 250\,\text{GeV}$. The main tracker of the ILD in its baseline design is a TPC, which greatly enhances its ability to detect LLPs. Both neutral and charged LLP production was studied using a diverse set of benchmark scenarios and full `Geant4` simulation of the detector response.

The main idea behind this study was to use fully model-independent, experiment-orientated approach, mainly because at the time of writing this thesis there was no clear indication what kind of BSM phenomena are realised by Nature. In addition, LLPs, due to the exotic character of their signatures, are a good case study to test the robustness of a detector design and its software ecosystem, in particular the reconstruction algorithms. Therefore, instead of selecting a specific BSM model and a set of points in its parameter space, benchmark scenarios for this study were chosen in a way allowing to test signatures with the most challenging and interesting kinematics. This approach was also tested for neutral LLP production on scenarios with exotic Higgs boson decays to dark long-lived scalars, where the search was further optimised to improve the sensitivity.

In the case of neutral LLPs, two classes of scenarios were first analysed. One involved production of heavy scalar LLP, which decayed into SM gauge bosons and DM. Scenarios with small mass difference between LLP and DM were selected, which provided non-pointing low-p_T final states, difficult to distinguish from beam-induced backgrounds in a linear collider. The second class of benchmarks involved the production of a very light pseudoscalar, which can be produced in e^+e^- collisions in association with a photon. Small LLP mass results in a large boost and high collimation of its decay products, bringing challenges for precise vertex reconstruction.

The analysis was carried out using a vertex-finding algorithm designed specifically for this study. The search was performed considering only decays inside the TPC volume, with as few assumptions about the final state as possible – in particular, final state objects other than the displaced vertex were ignored in the analysis. Soft, low-p_T beam-induced (overlay) events were considered as a standalone background, and a dedicated selection procedure was proposed to suppress their contribution. Hard, high-p_T SM processes with hadronic jets and di-lepton events were also taken into account as background sources. Two working points of the selection were considered: standard and tight.

Expected 95% C.L. limits on the LLP production cross section were calculated for a range of lifetimes, using the obtained background levels and signal selection efficiencies. The resulting sensitivity allows to probe cross sections down to the level of femtobarns using standard selection for most of considered scenarios, and the tight selection provides an improvement by an order of magnitude. The most challenging benchmarks, production of the heavy scalar with the mass difference $\Delta m_{AH} = 1\,\text{GeV}$ between LLP and DM, and the light long-lived pseudoscalar production with the mass $m_a = 0.3\,\text{GeV}$, could be accessed down to the level of 10 fb using the standard model-independent selection.

However, this sensitivity can be improved when optimising the search for a specific BSM scenario. This has been demonstrated in the search for exotic Higgs decays to LLPs, using the Higgsstrahlung production channel with the Z boson decays to neutrinos. The expected signature was at least one displaced vertex with no other activity inside the detector. Two scenarios with low LLP masses and two with high masses were considered. Additionally, requiring no high-p_T prompt tracks in the event and using cuts on the total p_T of tracks forming the vertex allowed to improve the reach by two orders of magnitude with respect to the previous analysis mentioned above. When translated to branching ratios of the Higgs boson to dark scalars, the results indicate that the ILD could provide an improvement with respect to current limits from the LHC in the regime of long lifetimes, depending on the LLP mass and BSM model, already at the first stage of ILC running.

Charged LLP production was studied profiting from the experience gained in the search for displaced vertices. Model predicting pair-production of long-lived dark fermions that decay into DM and SM leptons was used to generate six benchmark scenarios, with low and high masses of LLPs, and different mass differences between LLP and DM. The kinked track signature was considered in the analysis, and the search was carried out by optimising a tool designed for kink-finding which is available in the ILC software stack. A similar approach was taken to the model-independent analysis for displaced vertices, with the same background sources considered, but this time the search was performed in the entire detector volume. For all scenarios cross sections below 1 fb could be probed in the range of decay lengths from around 100 mm to 10 m, with the strongest limits reaching down to the level of 0.1 fb. The results weakly depend on the mass difference between LLP and DM.

The results presented in this thesis demonstrate the great potential of the ILD experiment for LLP searches. The proposed analysis framework allows both to probe wide range of scenarios in a model-independent manner obtaining good sensitivity, and test the detector capabilities. Optimisation of selection criteria for a specific model provides further background reduction down to the level of single events resulting in orders of magnitude better sensitivity. The TPC plays a crucial role in the ILD discovery potential, which is confirmed not only by the results obtained in all presented analyses, but also by a direct comparison of the tracking acceptance with an alternative all-silicon ILD design. The results obtained in this thesis could be even further improved by collecting more data at other stages of a linear collider operation.

Bibliography

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

[2] CMS collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, Phys. Lett. B **716** (2012) 30 [arXiv:1207.7235].

[3] P. Bambade et al., *The International Linear Collider: A Global Project*, arXiv:1903.01629.

[4] M. Aicheler, P. Burrows, M. Draper, T. Garvey, P. Lebrun, K. Peach et al., *A Multi-TeV Linear Collider Based on CLIC Technology*, CERN Yellow Reports: Monographs, CERN Scientific Information Service, Geneva (2012), 10.5170/CERN-2012-007.

[5] E. Adli et al., *The Compact Linear e^+e^- Collider (CLIC)*, arXiv:2503.24168.

[6] M. Benedikt, W. Bartmann, J.-P. Burnet, C. Carli, A. Chance, P. Craievich et al., *Future circular collider feasibility study report volume 2: Accelerators, technical infrastructure and safety*, Tech. Rep. CERN-FCC-ACC-2025-0004, CERN, Geneva (2025), https://doi.org/10.17181/CERN.EBAY.7W4X.

[7] CEPC Study Group, *CEPC Conceptual Design Report: Volume 1 - Accelerator*, arXiv:1809.00285.

[8] Linear Collider collaboration, *The Linear Collider Facility (LCF) at CERN*, arXiv:2503.24049.

[9] European Strategy Group, *2020 Update of the European Strategy for Particle Physics*, CERN Council, Geneva (2020), 10.17181/ESU2020.

[10] J. de Blas, M. Dunford, E. Bagnaschi, A. Freitas, P.P. Giardino, C. Grefe et al., *Physics Briefing Book: Input for the 2026 update of the European Strategy for Particle Physics*, Tech. Rep. CERN-ESU-2025-001, Geneva (2025).

[11] J. Altmann et al., *ECFA Higgs, electroweak, and top Factory Study*, vol. 5/2025 of *CERN Yellow Reports: Monographs*, CERN Scientific Information Service (6, 2025), 10.23731/CYRM-2025-005, [arXiv:2506.15390].

[12] ILD Concept Group, *International Large Detector: Interim Design Report*, arXiv:2003.01116.

[13] J. Klamka and A.F. Zarnecki, *Searching for displaced vertices with a gaseous tracker for a future e^+e^- Higgs factory*, *JHEP* **02** (2025) 112 [arXiv:2409.13492].

[14] J. Klamka, *Reconstructing long-lived particles with the ILD detector*, *PoS* **EPS-HEP2023** (2024) 458 [arXiv:2310.20534].

[15] J.F. Klamka, A.F. Żarnecki and D. Jeans, *Long-lived particle searches with the ILD experiment*, *EPJ Web Conf.* **315** (2024) 01033.

[16] D. Jeans, J. Klamka and A.F. Zarnecki, *Long-lived particle searches with the ILD experiment*, *PoS* **ICHEP2024** (2025) 320.

[17] Super-Kamiokande collaboration, *Evidence for oscillation of atmospheric neutrinos*, *Phys. Rev. Lett.* **81** (1998) 1562 [arXiv:hep-ex/9807003].

[18] LSND collaboration, *Evidence for neutrino oscillations from the observation of $\bar{\nu}_e$ appearance in a $\bar{\nu}_\mu$ beam*, *Phys. Rev. D* **64** (2001) 112007 [arXiv:hep-ex/0104049].

[19] KamLAND collaboration, *First results from KamLAND: Evidence for reactor anti-neutrino disappearance*, *Phys. Rev. Lett.* **90** (2003) 021802 [arXiv:hep-ex/0212021].

[20] RENO collaboration, *Observation of Reactor Electron Antineutrino Disappearance in the RENO Experiment*, *Phys. Rev. Lett.* **108** (2012) 191802 [arXiv:1204.0626].

[21] A.D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, *Pisma Zh. Eksp. Teor. Fiz.* **5** (1967) 32.

[22] F.R. Klinkhamer and N.S. Manton, *A Saddle Point Solution in the Weinberg-Salam Theory*, *Phys. Rev. D* **30** (1984) 2212.

[23] V.A. Rubakov and M.E. Shaposhnikov, *Electroweak baryon number nonconservation in the early universe and in high-energy collisions*, *Usp. Fiz. Nauk* **166** (1996) 493 [arXiv:hep-ph/9603208].

[24] J.H. Christenson, J.W. Cronin, V.L. Fitch and R. Turlay, *Evidence for the 2π Decay of the K_2^0 Meson*, *Phys. Rev. Lett.* **13** (1964) 138.

[25] V.A. Kuzmin, V.A. Rubakov and M.E. Shaposhnikov, *On the Anomalous Electroweak Baryon Number Nonconservation in the Early Universe*, *Phys. Lett. B* **155** (1985) 36.

[26] D.E. Morrissey and M.J. Ramsey-Musolf, *Electroweak baryogenesis*, *New J. Phys.* **14** (2012) 125003 [arXiv:1206.2942].

[27] F. Zwicky, *Die Rotverschiebung von extragalaktischen Nebeln*, *Helv. Phys. Acta* **6** (1933) 110.

[28] P. Salucci, *The distribution of dark matter in galaxies*, *Astron. Astrophys. Rev.* **27** (2019) 2 [arXiv:1811.08843].

[29] Planck collaboration, *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [arXiv:1807.06209].

[30] P. Dayal and A. Ferrara, *Early galaxy formation and its large-scale effects*, *Phys. Rept.* **780-782** (2018) 1 [arXiv:1809.09136].

[31] S. Tulin and H.-B. Yu, *Dark Matter Self-interactions and Small Scale Structure*, *Phys. Rept.* **730** (2018) 1 [arXiv:1705.02358].

[32] W. Tucker, P. Blanco, S. Rappoport, L. David, D. Fabricant, E.E. Falco et al., *1e0657-56: a contender for the hottest known cluster of galaxies*, *Astrophys. J. Lett.* **496** (1998) L5 [arXiv:astro-ph/9801120].

[33] M. Markevitch, A.H. Gonzalez, L. David, A. Vikhlinin, S. Murray, W. Forman et al., *A Textbook example of a bow shock in the merging galaxy cluster 1E0657-56*, *Astrophys. J. Lett.* **567** (2002) L27 [arXiv:astro-ph/0110468].

[34] D. Clowe, A. Gonzalez and M. Markevitch, *Weak lensing mass reconstruction of the interacting cluster 1E0657-558: Direct evidence for the existence of dark matter*, *Astrophys. J.* **604** (2004) 596 [arXiv:astro-ph/0312273].

[35] M. Milgrom, *MOND theory*, *Can. J. Phys.* **93** (2015) 107 [arXiv:1404.7661].

[36] T. Clifton, P.G. Ferreira, A. Padilla and C. Skordis, *Modified Gravity and Cosmology*, *Phys. Rept.* **513** (2012) 1 [arXiv:1106.2476].

[37] S.S. McGaugh, *A tale of two paradigms: the mutual incommensurability of* Λ*CDM and MOND*, *Can. J. Phys.* **93** (2015) 250 [arXiv:1404.7525].

[38] M. Cirelli, A. Strumia and J. Zupan, *Dark Matter*, arXiv:2406.01705.

[39] R. Jackiw and C. Rebbi, *Vacuum Periodicity in a Yang-Mills Quantum Theory*, *Phys. Rev. Lett.* **37** (1976) 172.

[40] C.-Y. Seng, *Reexamination of The Standard Model Nucleon Electric Dipole Moment*, *Phys. Rev. C* **91** (2015) 025502 [arXiv:1411.1476].

[41] C. Abel et al., *Measurement of the Permanent Electric Dipole Moment of the Neutron*, *Phys. Rev. Lett.* **124** (2020) 081803 [arXiv:2001.11966].

[42] Particle Data Group, *Review of particle physics*, *Phys. Rev. D* **110** (2024) 030001.

[43] H. Georgi and S.L. Glashow, *Unity of All Elementary Particle Forces*, *Phys. Rev. Lett.* **32** (1974) 438.

[44] M.E. Peskin and D.V. Schroeder, *An Introduction to quantum field theory*, Addison-Wesley, Reading, USA (1995), 10.1201/9780429503559.

[45] M. Thomson, *Modern particle physics*, Cambridge University Press, New York (2013), 10.1017/CBO9781139525367.

[46] CMS collaboration, *CMS technical design report, volume II: Physics performance*, *J. Phys. G* **34** (2007) 995.

[47] ATLAS collaboration, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, arXiv:0901.0512.

[48] ATLAS collaboration, *Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector*, Phys. Lett. B **786** (2018) 59 [arXiv:1808.08238].

[49] CMS collaboration, *Observation of Higgs boson decay to bottom quarks*, Phys. Rev. Lett. **121** (2018) 121801 [arXiv:1808.08242].

[50] CMS collaboration, *Observation of the Higgs boson decay to a pair of τ leptons with the CMS detector*, Phys. Lett. B **779** (2018) 283 [arXiv:1708.00373].

[51] ATLAS collaboration, *Cross-section measurements of the Higgs boson decaying into a pair of τ-leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **99** (2019) 072001 [arXiv:1811.08856].

[52] CMS collaboration, *Evidence for Higgs boson decay to a pair of muons*, JHEP **01** (2021) 148 [arXiv:2009.04363].

[53] ALEPH, DELPHI, L3, OPAL, LEP Working Group for Higgs Boson Searches, *Search for neutral MSSM Higgs bosons at LEP*, Eur. Phys. J. C **47** (2006) 547 [arXiv:hep-ex/0602042].

[54] ALEPH, DELPHI, L3, OPAL collaboration, *Searches for Higgs bosons beyond the Standard Model at LEP (Part B): Fermiophobia, anomalous couplings, invisible and charged Higgs*, PoS **HEP2001** (2001) 145.

[55] Belle-II collaboration, *Search for an Invisibly Decaying Z' Boson at Belle II in $e^+e^- \to \mu^+\mu^-(e^\pm\mu^\mp)$ Plus Missing Energy Final States*, Phys. Rev. Lett. **124** (2020) 141801 [arXiv:1912.11276].

[56] Belle-II collaboration, *Search for Axion-Like Particles produced in e^+e^- collisions at Belle II*, Phys. Rev. Lett. **125** (2020) 161806 [arXiv:2007.13071].

[57] NA62 collaboration, *Search for heavy neutral lepton production in $K+$ decays to positrons*, Phys. Lett. B **807** (2020) 135599 [arXiv:2005.09575].

[58] NA64 collaboration, *Search for invisible decays of sub-GeV dark photons in missing-energy events at the CERN SPS*, Phys. Rev. Lett. **118** (2017) 011802 [arXiv:1610.02988].

[59] S. Alekhin et al., *A facility to Search for Hidden Particles at the CERN SPS: the SHiP physics case*, Rept. Prog. Phys. **79** (2016) 124201 [arXiv:1504.04855].

[60] XENON collaboration, *The XENONnT dark matter experiment*, Eur. Phys. J. C **84** (2024) 784 [arXiv:2402.10446].

[61] PandaX collaboration, *PandaX-xT—A deep underground multi-ten-tonne liquid xenon observatory*, Sci. China Phys. Mech. Astron. **68** (2025) 221011 [arXiv:2402.03596].

[62] LUX-ZEPLIN collaboration, *The LUX-ZEPLIN dark matter experiment*, *Astron. Nachr.* **342** (2021) 423.

[63] DEAP-3600 collaboration, *Design and Construction of the DEAP-3600 Dark Matter Detector*, *Astropart. Phys.* **108** (2019) 1 [arXiv:1712.01982].

[64] DarkSide collaboration, *First Results from the DarkSide-50 Dark Matter Experiment at Laboratori Nazionali del Gran Sasso*, *Phys. Lett. B* **743** (2015) 456 [arXiv:1410.0653].

[65] DarkSide-20k collaboration, *DarkSide-20k: A 20 tonne two-phase LAr TPC for direct dark matter detection at LNGS*, *Eur. Phys. J. Plus* **133** (2018) 131 [arXiv:1707.08145].

[66] CAST collaboration, *New CAST Limit on the Axion-Photon Interaction*, *Nature Phys.* **13** (2017) 584 [arXiv:1705.02290].

[67] ALPS collaboration, *Towards the first axion search results of the Any Light Particle Search II experiment*, *PoS* **TAUP2023** (2024) 049 [arXiv:2401.14773].

[68] Fermi-LAT collaboration, *Searching for Dark Matter Annihilation from Milky Way Dwarf Spheroidal Galaxies with Six Years of Fermi Large Area Telescope Data*, *Phys. Rev. Lett.* **115** (2015) 231301 [arXiv:1503.02641].

[69] AMS collaboration, *The Alpha Magnetic Spectrometer (AMS) on the international space station: Part II — Results from the first seven years*, *Phys. Rept.* **894** (2021) 1.

[70] H.E.S.S. collaboration, *Search for dark matter annihilations towards the inner Galactic halo from 10 years of observations with H.E.S.S*, *Phys. Rev. Lett.* **117** (2016) 111301 [arXiv:1607.08142].

[71] D0 collaboration, *Observation of the top quark*, *Phys. Rev. Lett.* **74** (1995) 2632 [arXiv:hep-ex/9503003].

[72] ALEPH, CDF, D0, DELPHI, L3, OPAL, SLD, LEP Electroweak Working Group, Tevatron Electroweak Working Group, SLD Electroweak, Heavy Flavour Groups, *Precision Electroweak Measurements and Constraints on the Standard Model*, arXiv:1012.2367.

[73] W. Buchmuller and D. Wyler, *Effective Lagrangian Analysis of New Interactions and Flavor Conservation*, *Nucl. Phys. B* **268** (1986) 621.

[74] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085 [arXiv:1008.4884].

[75] J. Ellis, C.W. Murphy, V. Sanz and T. You, *Updated Global SMEFT Fit to Higgs, Diboson and Electroweak Data*, *JHEP* **06** (2018) 146 [arXiv:1803.03252].

[76] J. Ellis, M. Madigan, K. Mimasu, V. Sanz and T. You, *Top, Higgs, Diboson and Electroweak Fit to the Standard Model Effective Field Theory*, *JHEP* **04** (2021) 279 [arXiv:2012.02779].

[77] SMEFiT collaboration, *Combined SMEFT interpretation of Higgs, diboson, and top quark data from the LHC*, *JHEP* **11** (2021) 089 [arXiv:2105.00006].

[78] J. de Blas, A. Goncalves, V. Miralles, L. Reina, L. Silvestrini and M. Valli, *Constraining new physics effective interactions via a global fit of electroweak, Drell-Yan, Higgs, top, and flavour observables*, arXiv:2507.06191.

[79] J. de Blas et al., *Higgs Boson Studies at Future Particle Colliders*, *JHEP* **01** (2020) 139 [arXiv:1905.03764].

[80] J. de Blas, Y. Du, C. Grojean, J. Gu, V. Miralles, M.E. Peskin et al., *Global SMEFT Fits at Future Colliders*, in *Snowmass 2021*, 6, 2022 [arXiv:2206.08326].

[81] I. Neutelings, *Particle mass versus lifetime & decay length*, 12, 2022, https://tikz.net/sm_particles_masses/. Accessed: 2025-03-28. Licensed under Creative Commons CC BY 4.0.

[82] R.D. Peccei and H.R. Quinn, *CP Conservation in the Presence of Instantons*, *Phys. Rev. Lett.* **38** (1977) 1440.

[83] F. Wilczek, *Problem of Strong P and T Invariance in the Presence of Instantons*, *Phys. Rev. Lett.* **40** (1978) 279.

[84] M. Bauer, M. Neubert and A. Thamm, *Collider Probes of Axion-Like Particles*, *JHEP* **12** (2017) 044 [arXiv:1708.00443].

[85] S.D. Thomas and J.D. Wells, *Phenomenology of Massive Vectorlike Doublet Leptons*, *Phys. Rev. Lett.* **81** (1998) 34 [arXiv:hep-ph/9804359].

[86] R.T. Co, F. D'Eramo, L.J. Hall and D. Pappadopulo, *Freeze-In Dark Matter with Displaced Signatures at Colliders*, *JCAP* **12** (2015) 024 [arXiv:1506.07532].

[87] J. Alimena et al., *Searching for long-lived particles beyond the Standard Model at the Large Hadron Collider*, *J. Phys. G* **47** (2020) 090501 [arXiv:1903.04497].

[88] I. Neutelings, *Signatures of long-lived particles in the CMS detector*, 2, 2023, https://tikz.net/bsm_longlived/. Accessed: 2025-03-28. Licensed under Creative Commons CC BY 4.0.

[89] L. Calibbi, F. D'Eramo, S. Junius, L. Lopez-Honorez and A. Mariotti, *Displaced new physics at colliders and the early universe before its first second*, *JHEP* **05** (2021) 234 [arXiv:2102.06221].

[90] CMS collaboration, *Search for long-lived particles decaying in the CMS muon detectors in proton-proton collisions at $\sqrt{s} = 13$ TeV*, arXiv:2402.01898.

[91] ATLAS collaboration, *Search for events with one displaced vertex from long-lived neutral particles decaying into hadronic jets in the ATLAS muon spectrometer in pp collisions at $\sqrt{s} = 13$ TeV*, arXiv:2503.20445.

[92] CMS collaboration, *Search for decays of stopped exotic long-lived particles produced in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, *JHEP* **05** (2018) 127 [arXiv:1801.00359].

[93] CMS collaboration, *Search for long-lived particles using displaced jets in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, *Phys. Rev. D* **104** (2021) 012015 [arXiv:2012.01581].

[94] CMS collaboration, *Search for long-lived particles using displaced vertices and missing transverse momentum in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, arXiv:2402.15804.

[95] ATLAS collaboration, *Search for long-lived, massive particles in events with displaced vertices and missing transverse momentum in* $\sqrt{s} = 13$ *TeV pp collisions with the ATLAS detector*, *Phys. Rev. D* **97** (2018) 052012 [arXiv:1710.04901].

[96] ATLAS collaboration, *Search for light long-lived neutral particles from Higgs boson decays via vector-boson-fusion production from pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, arXiv:2311.18298.

[97] CMS collaboration, *Search for long-lived particles decaying to final states with a pair of muons in proton-proton collisions at* $\sqrt{s} = 13.6$ *TeV*, arXiv:2402.14491.

[98] CMS collaboration, *Search for supersymmetry in final states with disappearing tracks in proton-proton collisions at s=13 TeV*, *Phys. Rev. D* **109** (2024) 072007 [arXiv:2309.16823].

[99] ATLAS collaboration, *Search for long-lived charginos based on a disappearing-track signature using 136* fb^{-1} *of pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, *Eur. Phys. J. C* **82** (2022) 606 [arXiv:2201.02472].

[100] ATLAS collaboration, *Search for displaced leptons in* $\sqrt{s} = 13$ *TeV and* 13.6 *TeV pp collisions with the ATLAS detector*, *Phys. Rev. D* **112** (2025) 012016 [arXiv:2410.16835].

[101] ATLAS collaboration, *Search for Nearly Mass-Degenerate Higgsinos Using Low-Momentum Mildly Displaced Tracks in pp Collisions at s=13 TeV with the ATLAS Detector*, *Phys. Rev. Lett.* **132** (2024) 221801 [arXiv:2401.14046].

[102] CMS collaboration, *Search for compressed electroweakinos with low-momentum isolated tracks*, Tech. Rep. CMS-PAS-SUS-24-012, CERN, Geneva (2025).

[103] ATLAS collaboration, *Search for long-lived charged particles using large specific ionisation loss and time of flight in 140* fb^{-1} *of pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, *JHEP* **07** (2025) 140 [arXiv:2502.06694].

[104] CMS collaboration, *Search for heavy long-lived charged particles with large ionization energy loss in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, *JHEP* **04** (2025) 109 [arXiv:2410.09164].

[105] J.L. Feng, I. Galon, F. Kling and S. Trojanowski, *ForwArd Search ExpeRiment at the LHC*, *Phys. Rev. D* **97** (2018) 035001 [arXiv:1708.09389].

[106] FASER collaboration, *The FASER detector*, *JINST* **19** (2024) P05066 [arXiv:2207.11427].

[107] FASER collaboration, *First neutrino interaction candidates at the LHC*, *Phys. Rev. D* **104** (2021) L091101 [arXiv:2105.06197].

[108] FASER collaboration, *Search for dark photons with the FASER detector at the LHC*, *Phys. Lett. B* **848** (2024) 138378 [arXiv:2308.05587].

[109] B. Aitken et al., *Conceptual Design Report for the MATHUSLA Long-Lived Particle Detector near CMS*, arXiv:2503.20893.

[110] M. Bauer, O. Brandt, L. Lee and C. Ohm, *ANUBIS: Proposal to search for long-lived neutral particles in CERN service shafts*, arXiv:1909.13022.

[111] M. Revering, G. Aielli, O. Brandt, L. Pizzimento, A. Shah and P. Swallow, *Searches for long-lived particles with ANUBIS: first commissioning results from proANUBIS*, *PoS* **ICHEP2024** (2025) 983.

[112] G. Aielli et al., *Expression of interest for the CODEX-b detector*, *Eur. Phys. J. C* **80** (2020) 1177 [arXiv:1911.00481].

[113] CODEX-B collaboration, *COmpact DEtector for EXotics at LHCb: CODEX-b*, *PoS* **ICHEP2024** (2025) 1077.

[114] FPF Working Groups, *The Forward Physics Facility at the Large Hadron Collider*, 3, 2025 [arXiv:2503.19010].

[115] Belle-II collaboration, *Search for a long-lived spin-0 mediator in $b \to s$ transitions at the Belle II experiment*, *Phys. Rev. D* **108** (2023) L111104 [arXiv:2306.02830].

[116] Belle collaboration, *Search for a heavy neutral lepton that mixes predominantly with the tau neutrino*, *Phys. Rev. D* **109** (2024) L111102 [arXiv:2402.02580].

[117] T. Ferber, C. Garcia-Cely and K. Schmidt-Hoberg, *BelleII sensitivity to long–lived dark photons*, *Phys. Lett. B* **833** (2022) 137373 [arXiv:2202.03452].

[118] S. Dreyer et al., *Physics reach of a long-lived particle detector at Belle II*, arXiv:2105.12962.

[119] R. Schäfer et al., *Near or far detectors? A case study for long-lived particle searches at electron-positron colliders*, *Phys. Rev. D* **107** (2023) 076022.

[120] SHiP, HI-ECN3 Project Team, *SHiP experiment at the SPS Beam Dump Facility*, arXiv:2504.06692.

[121] K. Asai, S. Iwamoto, Y. Sakaki and D. Ueda, *New physics searches at the ILC positron and electron beam dumps*, *JHEP* **09** (2021) 183 [arXiv:2105.13768].

[122] Y. Sakaki and D. Ueda, *Searching for new light particles at the international linear collider main beam dump*, *Phys. Rev. D* **103** (2021) 035024 [arXiv:2009.13790].

[123] M.M. Nojiri, Y. Sakaki, K. Tobioka and D. Ueda, *First evaluation of meson and τ lepton spectra and search for heavy neutral leptons at ILC beam dump*, *JHEP* **12** (2022) 145 [arXiv:2206.13523].

[124] K. Asai, A. Das, J. Li, T. Nomura and O. Seto, *Chiral Z' in FASER, FASER2, DUNE, and ILC beam dump experiments*, *Phys. Rev. D* **106** (2022) 095033 [arXiv:2206.12676].

[125] ATLAS collaboration, *ATLAS Run 2 searches for electroweak production of supersymmetric particles interpreted within the pMSSM*, *JHEP* **05** (2024) 106 [arXiv:2402.01392].

[126] S. Dawson et al., *Report of the Topical Group on Higgs Physics for Snowmass 2021: The Case for Precision Higgs Physics*, in *Snowmass 2021*, 9, 2022 [arXiv:2209.07510].

[127] T. Robens and R. Santos, *BSM: Extended Scalar Sectors*, 7, 2025 [arXiv:2507.21910].

[128] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, arXiv:1307.1347.

[129] ATLAS collaboration, *A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery*, *Nature* **607** (2022) 52 [arXiv:2207.00092].

[130] CMS collaboration, *A portrait of the Higgs boson by the CMS experiment ten years after the discovery.*, *Nature* **607** (2022) 60 [arXiv:2207.00043].

[131] ATLAS collaboration, *Projected sensitivity of measurements of Higgs boson pair production with the ATLAS experiment at the HL-LHC*, Tech. Rep. ATL-PHYS-PUB-2025-006, CERN, Geneva (2025).

[132] CMS collaboration, *Combination of searches for nonresonant Higgs boson pair production in proton-proton collisions at sqrt(s) = 13 TeV*, Tech. Rep. CMS-PAS-HIG-20-011, CERN, Geneva (2024).

[133] ILD Design Study Group, *HZ Recoil Mass and Cross Section Analysis in ILD*, arXiv:1202.1439.

[134] M. Narain et al., *The Future of US Particle Physics - The Snowmass 2021 Energy Frontier Report*, arXiv:2211.11084.

[135] Linear Collider Vision collaboration, *A Linear Collider Vision for the Future of Particle Physics*, arXiv:2503.19983.

[136] CLICdp collaboration, *Double Higgs boson production and Higgs self-coupling extraction at CLIC*, *Eur. Phys. J. C* **80** (2020) 1010 [arXiv:1901.05897].

[137] P. Stylianou and G. Weiglein, *Constraints on the trilinear and quartic Higgs couplings from triple Higgs production at the LHC and beyond*, *Eur. Phys. J. C* **84** (2024) 366 [arXiv:2312.04646].

[138] M. Berggren et al., *Tackling light higgsinos at the ILC*, *Eur. Phys. J. C* **73** (2013) 2660.

[139] M. Berggren, M.T. Núñez Pardo de Vera and J. List, *$\tilde{\tau}$ searches at future e+e− colliders*, *EPJ Web Conf.* **315** (2024) 01030 [arXiv:2410.14622].

[140] CEPC Physics Study Group, *The Physics potential of the CEPC. Prepared for the US Snowmass Community Planning Exercise (Snowmass 2021)*, in *Snowmass 2021*, 5, 2022 [arXiv:2205.08553].

[141] M. Benedikt, W. Bartmann, J.-P. Burnet, C. Carli, A. Chance, P. Craievich et al., *Future Circular Collider Feasibility Study Report Volume 1: Physics and Experiments*, Tech. Rep. CERN-FCC-PHYS-2025-0002, CERN, Geneva (2025), https://doi.org/10.17181/CERN.9DKX.TDH9.

[142] CEPC Study Group, *CEPC Conceptual Design Report: Volume 2 - Physics & Detector*, arXiv:1811.10545.

[143] N. Bacchetta et al., *CLD – A Detector Concept for the FCC-ee*, arXiv:1911.12230.

[144] ILD Concept Group, *The ILD Detector: A Versatile Detector for an Electron-Positron Collider at Energies up to 1 TeV*, arXiv:2506.06030.

[145] T. Behnke, J.E. Brau, B. Foster, J. Fuster, M. Harrison, J.M. Paterson et al., eds., *The International Linear Collider Technical Design Report - Volume 1: Executive Summary*, arXiv:1306.6327.

[146] M. Altarelli et al., eds., *XFEL: The European X-Ray Free-Electron Laser. Technical design report*, DESY-06-097, (2006).

[147] T. Barklow, J. Brau, K. Fujii, J. Gao, J. List, N. Walker et al., *ILC Operating Scenarios*, arXiv:1506.07830.

[148] CLICdp, CLIC collaboration, *The Compact Linear Collider (CLIC) - 2018 Summary Report*, arXiv:1812.06018.

[149] ILC International Development Team, *The International Linear Collider: Report to Snowmass 2021*, arXiv:2203.07622.

[150] H. Abramowicz et al., *The International Linear Collider Technical Design Report - Volume 4: Detectors*, arXiv:1306.6329.

[151] M. Thomson, *Particle flow calorimetry and the PandoraPFA algorithm*, *Nucl. Instrum. Meth. A* **611** (2009) 25–40.

[152] K. Fujii et al., *Physics Case for the 250 GeV Stage of the International Linear Collider*, arXiv:1710.07621.

[153] R.K. Ellis et al., *Physics Briefing Book: Input for the European Strategy for Particle Physics Update 2020*, arXiv:1910.11775.

[154] N. Alipour Tehrani et al., *CLICdet: The post-CDR CLIC detector model*, CLICdp-Note-2017-001, (2017).

[155] M. Bai et al., *C^3: A "Cool" Route to the Higgs Boson and Beyond*, in *Snowmass 2021*, 10, 2021 [arXiv:2110.15800].

[156] B. Foster, R. D'Arcy and C.A. Lindstrom, *A hybrid, asymmetric, linear Higgs factory based on plasma-wakefield and radio-frequency acceleration*, *New J. Phys.* **25** (2023) 093037 [arXiv:2303.10150].

[157] S. Gessner et al., *Design Initiative for a 10 TeV pCM Wakefield Collider*, arXiv:2503.20214.

[158] V.N. Litvinenko, N. Bachhawat, M. Chamizo-Llatas, Y. Jing, F. Méot, I. Petrushina et al., *The ReLiC: Recycling Linear e^+e^- Collider*, arXiv:2203.06476.

[159] V.I. Telnov, *A high-luminosity superconducting twin e^+e^- linear collider with energy recovery*, *JINST* **16** (2021) P12025 [arXiv:2105.11015].

[160] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens et al., *DELPHES 3, A modular framework for fast simulation of a generic collider experiment*, *JHEP* **02** (2014) 057.

[161] S. Antusch, E. Cazzato and O. Fischer, *Displaced vertex searches for sterile neutrinos at future lepton colliders*, *JHEP* **12** (2016) 007 [arXiv:1604.02420].

[162] M. Bauer, M. Heiles, M. Neubert and A. Thamm, *Axion-Like Particles at Future Colliders*, *Eur. Phys. J. C* **79** (2019) 74 [arXiv:1808.10323].

[163] G. Burdman, Z. Chacko, H.-S. Goh and R. Harnik, *Folded supersymmetry and the LEP paradox*, *JHEP* **02** (2007) 009 [arXiv:hep-ph/0609152].

[164] M.J. Strassler and K.M. Zurek, *Echoes of a hidden valley at hadron colliders*, *Phys. Lett. B* **651** (2007) 374 [arXiv:hep-ph/0604261].

[165] K. Cheung and Z.S. Wang, *Probing Long-lived Particles at Higgs Factories*, *Phys. Rev. D* **101** (2020) 035003 [arXiv:1911.08721].

[166] S. Alipour-Fard and others, *Long Live the Higgs Factory: Higgs Decays to Long-Lived Particles at Future Lepton Colliders*, *Chin. Phys. C* **43** (2019) 053101.

[167] G. Polesello, *Sensitivity of the FCC-ee to decay of an axion-like particle into two photons*, arXiv:2502.08411.

[168] L. Bellagamba, G. Polesello and N. Valle, *Searches for Heavy Neutral Leptons at FCC-ee in final states including a muon*, arXiv:2503.19464.

[169] G. Ripellino, M.V. Voorde, A. Gallén and R. Gonzalez Suarez, *Searching for long-lived dark scalars at the FCC-ee*, arXiv:2412.10141.

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

[171] M. Kucharczyk and M. Goncerz, *Search for exotic decays of the Higgs boson into long-lived particles with jet pairs in the final state at CLIC*, *JHEP* **03** (2023) 131 [arXiv:2212.04147].

[172] Y. Zhang, C. Mo, X. Chen, B. Li, H. Chen, J. Hu et al., *Search for Long-lived Particles at Future Lepton Colliders Using Deep Learning Techniques*, arXiv:2401.05094.

[173] L. Jeanty, L. Nosler and C. Potter, *Sensitivity to decays of long-lived dark photons at the ILC*, arXiv:2203.08347.

[174] ILD Concept Group, *The International Large Detector: Letter of Intent*, arXiv:1006.3396.

[175] D. Jeans, *Beamstrahlung backgrounds in ILD at linear (ILC) and circular (FCCee) colliders*, *EPJ Web Conf.* **315** (2024) 02017.

[176] C. Ligtenberg, *A GridPix TPC readout for the ILD experiment at the future International Linear Collider*, Ph.D. thesis, Vrije U., Amsterdam, Vrije U., Amsterdam, 2021.

[177] M. van Beuzekom et al., *Towards a Pixel TPC, Part II: Particle identification with a 32-chip GridPix detector*, *Nucl. Instrum. Meth. A* **1081** (2026) 170849.

[178] W. Kilian et al., *WHIZARD: Simulating Multi-Particle Processes at LHC and ILC*, *Eur. Phys. J. C* **71** (2011) 1742.

[179] M. Moretti, T. Ohl and J. Reuter, *O'Mega: An Optimizing matrix element generator*, arXiv:hep-ph/0102195.

[180] T. Sjostrand, S. Mrenna and P.Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **05** (2006) 026 [arXiv:hep-ph/0603175].

[181] T. Sjöstrand, S. Ask, J.R. Christiansen, R. Corke, N. Desai, P. Ilten et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [arXiv:1410.3012].

[182] T. Ohl, *CIRCE version 1.0: Beam spectra for simulating linear collider physics*, *Comput. Phys. Commun.* **101** (1997) 269 [arXiv:hep-ph/9607454].

[183] F. Staub, *SARAH 4 : A tool for (not only SUSY) model builders*, *Comput. Phys. Commun.* **185** (2014) 1773 [arXiv:1309.7223].

[184] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer and T. Reiter, *UFO - The Universal FeynRules Output*, *Comput. Phys. Commun.* **183** (2012) 1201 [arXiv:1108.2040].

[185] A. Alloul, N.D. Christensen, C. Degrande, C. Duhr and B. Fuks, *FeynRules 2.0 - A complete toolbox for tree-level phenomenology*, *Comput. Phys. Commun.* **185** (2014) 2250 [arXiv:1310.1921].

[186] S. Agostinelli et al., *Geant4 — a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250 .

[187] M. Frank, F. Gaede, C. Grefe and P. Mato, *DD4hep: A detector description toolkit for high energy physics experiments*, *Journal of Physics: Conference Series* **513** (2014) 022010.

[188] J.M. Carceller et al., *Five years of Key4hep - Towards production readiness and beyond*, *PoS* **ICHEP2024** (2025) 1029.

[189] *iLCSoft – Linear Collider Software*, GitHub repository, `https://github.com/iLCSoft`.

[190] F. Gaede, *Marlin and LCCD: Software tools for the ILC*, *Nucl. Instrum. Meth. A* **559** (2006) 177.

[191] *ILDConfig*, GitHub repository, `https://github.com/iLCSoft/ILDConfig/tree/master/StandardConfig/production`.

[192] F. Gaede, S. Aplin, R. Glattauer, C. Rosemann and G. Voutsinas, *Track reconstruction at the ILC: the ILD tracking software*, *Journal of Physics: Conference Series* **513** (2014) 022011.

[193] T. Suehara and T. Tanabe, *LCFIPlus: A Framework for Jet Analysis in Linear Collider Studies*, *Nucl. Instrum. Meth. A* **808** (2016) 109 [`arXiv:1506.08371`].

[194] T. Kramer, *Track parameters in LCIO*, LC-DET-2006-004, (2006).

[195] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, *Nucl. Instrum. Meth. A* **262** (1987) 444.

[196] K. Fujii, *Extended Kalman Filter*. ACFA-Sim-J Group. `https://www-jlc.kek.jp/subg/offl/kaltest/doc/ReferenceManual.pdf`.

[197] B. Li, K. Fujii and Y. Gao, *Kalman-filter-based track fitting in non-uniform magnetic field with segment-wise helical track model*, *Comput. Phys. Commun.* **185** (2014) 754 [`arXiv:1305.7300`].

[198] R. Glattauer, R. Fruhwirth, J. Lettenbichler and W. Mitaroff, *Forward Tracking in the ILD Detector*, in *International Workshop on Future Linear Colliders (LCWS11)*, 2, 2012 [`arXiv:1202.2761`].

[199] Y. Shrivastava, S. Dasgupta and S. Reddy, *Guaranteed convergence in a class of hopfield networks*, *IEEE Transactions on Neural Networks* **3** (1992) 951.

[200] F. Gaede, T. Behnke, N. Graf and T. Johnson, *LCIO: A Persistency framework for linear collider simulation studies*, *eConf* **C0303241** (2003) TUKT001 [`arXiv:physics/0306114`].

[201] S. Aplin, J. Engels, F. Gaede, N.A. Graf, T. Johnson and J. McCormick, *LCIO: A Persistency Framework and Event Data Model for HEP*, in *2012 IEEE Nuclear Science Symposium and Medical Imaging Conference and 19th Workshop on Room-Temperature Semiconductor X-ray and Gamma-ray Detectors*, pp. 2075–2079, 11, 2012, https://doi.org/10.1109/NSSMIC.2012.6551478.

[202] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, *Nucl. Instrum. Meth. A* **389** (1997) 81.

[203] K.S. Sasikumar, *Photon-Photon processes at the International Linear Collider and BSM signatures with small mass differences*, Ph.D. thesis, Hamburg U., Hamburg U., Hamburg, 2020. 10.3204/PUBDB-2021-01241.

[204] D. Schulte, *Beam-beam simulations with GUINEA-PIG*, CERN-PS-99-014-LP, CERN-PS-99-14-LP, CLIC-NOTE-387, CERN-CLIC-NOTE-387, (1999).

[205] C. Grefe, S. Poss, A. Sailer and A. Tsaregorodtsev, *ILCDIRAC, a DIRAC extension for the Linear Collider community*, *J. Phys.: Conf. Ser.* **513** (2013) 032077. 5.

[206] J. Kalinowski et al., *Benchmarking the Inert Doublet Model for e^+e^- colliders*, *JHEP* **12** (2018) 081.

[207] G. Bélanger et al., *LHC-friendly minimal freeze-in models*, *JHEP* **02** (2019) 186 [arXiv:1811.05478].

[208] CLICdp collaboration, *Conformal tracking for all-silicon trackers at future electron–positron colliders*, *Nucl. Instrum. Meth. A* **956** (2020) 163304.

[209] ALEPH collaboration, *Search for gauge mediated SUSY breaking topologies in e^+e^- collisions at center-of-mass energies up to 209-GeV*, *Eur. Phys. J. C* **25** (2002) 339 [arXiv:hep-ex/0203024].

[210] OPAL collaboration, *Searches for gauge-mediated supersymmetry breaking topologies in e+ e- collisions at LEP2*, *Eur. Phys. J. C* **46** (2006) 307 [arXiv:hep-ex/0507048].

[211] M. Berggren private communication.

[212] G.J. Feldman and R.D. Cousins, *A Unified approach to the classical statistical analysis of small signals*, *Phys. Rev. D* **57** (1998) 3873 [arXiv:physics/9711021].

[213] T. Robens, T. Stefaniak and J. Wittbrodt, *Two-real-scalar-singlet extension of the SM: LHC phenomenology and benchmark scenarios*, *Eur. Phys. J. C* **80** (2020) 151 [arXiv:1908.08554].

[214] CMS collaboration, *Search for light scalar particles from Higgs boson decays in exclusive final states with two muons and two hadrons*, CMS-PAS-EXO-24-034, (2025).

[215] ATLAS collaboration, *Search for Light Long-Lived Particles in pp Collisions at $\sqrt{s}$ = 13 TeV Using Displaced Vertices in the ATLAS Inner Detector*, *Phys. Rev. Lett.* **133** (2024) 161803 [arXiv:2403.15332].

[216] CMS collaboration, *Dark sector searches with the CMS experiment*, arXiv:2405.13778.

[217] CMS collaboration, *Search for long-lived particles decaying to leptons with large impact parameter in proton–proton collisions at* $\sqrt{s}$ = 13 *TeV*, *Eur. Phys. J. C* **82** (2022) 153 [arXiv:2110.04809].

[218] CMS collaboration, *Search for long-lived particles decaying into muon pairs in proton-proton collisions at* $\sqrt{s}$ = *13 TeV collected with a dedicated high-rate data stream*, *JHEP* **04** (2022) 062 [arXiv:2112.13769].

[219] A.E. del Valle, *Dark scalar reinterpretation of searches for Higgs boson decays into long-lived particles at the LHC*, arXiv:2509.02564.

[220] S. Weinberg, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (1967) 1264.

[221] S.L. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl. Phys.* **22** (1961) 579.

[222] A. Salam and J.C. Ward, *Weak and electromagnetic interactions*, *Nuovo Cim.* **11** (1959) 568.

[223] S. Weinberg, *Nonabelian Gauge Theories of the Strong Interactions*, *Phys. Rev. Lett.* **31** (1973) 494.

[224] T.-P. Cheng and L.-F. Li, *Gauge Theory of Elementary Particle Physics*, Oxford University Press, Oxford, UK (1984).

[225] E.D. Bloom et al., *High-Energy Inelastic e p Scattering at 6-Degrees and 10-Degrees*, *Phys. Rev. Lett.* **23** (1969) 930.

[226] M. Breidenbach, J.I. Friedman, H.W. Kendall, E.D. Bloom, D.H. Coward, H.C. DeStaebler et al., *Observed behavior of highly inelastic electron-proton scattering*, *Phys. Rev. Lett.* **23** (1969) 935.

[227] CDF collaboration, *Observation of top quark production in* $\bar{p}p$ *collisions*, *Phys. Rev. Lett.* **74** (1995) 2626 [arXiv:hep-ex/9503002].

[228] Pluto collaboration, *Observation of a Narrow Resonance Formed in e+ e- Annihilation at 9.46-GeV*, *Phys. Lett. B* **76** (1978) 243.

[229] PLUTO collaboration, *Evidence for Gluon Bremsstrahlung in e+ e- Annihilations at High-Energies*, *Phys. Lett. B* **86** (1979) 418.

[230] TASSO collaboration, *Evidence for Planar Events in e+ e- Annihilation at High-Energies*, *Phys. Lett. B* **86** (1979) 243.

[231] D.P. Barber et al., *Discovery of Three Jet Events and a Test of Quantum Chromodynamics at PETRA Energies*, *Phys. Rev. Lett.* **43** (1979) 830.

[232] JADE collaboration, *Observation of Planar Three Jet Events in e+ e- Annihilation and Evidence for Gluon Bremsstrahlung*, *Phys. Lett. B* **91** (1980) 142.

[233] UA1 collaboration, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV*, *Phys. Lett. B* **122** (1983) 103.

[234] UA2 collaboration, *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider*, *Phys. Lett. B* **122** (1983) 476.

[235] UA1 collaboration, *Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider*, *Phys. Lett. B* **126** (1983) 398.

[236] UA2 collaboration, *Evidence for $Z^0 \to e^+e^-$ at the CERN $\bar{p}p$ Collider*, *Phys. Lett. B* **129** (1983) 130.

[237] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321.

[238] P.W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508.

[239] G.S. Guralnik, C.R. Hagen and T.W.B. Kibble, *Global Conservation Laws and Massless Particles*, *Phys. Rev. Lett.* **13** (1964) 585.

[240] E. Fermi, *Tentativo di una teoria dell'emissione dei raggi beta*, *Ric. Sci.* **4** (1933) 491.

[241] A.V. Manohar, *Introduction to Effective Field Theories*, `arXiv:1804.05863`.

[242] K. Fujii et al., *Physics Case for the International Linear Collider*, `arXiv:1506.05992`.

[243] H. Hölbe and F. Gaede, *C Event Display (CED) – User Manual.* DESY. `https://ilcsoft.desy.de/CED/current/doc/manual.pdf`.

Appendix A

The Standard Model of particle physics

A.1 Formulation of the Standard Model

The Standard Model (SM) of particle physics is a theory that combines the whole knowledge about the constituents of matter and interactions between them gathered over the last century. The SM introduces twelve fermionic fields, which correspond to particles of matter, vector bosons mediating the three fundamental forces (electromagnetic, weak and strong), and one fundamental scalar. Each particle in the SM is associated with an antiparticle, which carries an opposite charge. The mediators are represented by gauge fields—or their combinations—introduced to preserve a local symmetry of the theory.

A.1.1 Fermions and gauge bosons

The following Lagrangian describes electroweak interactions between fermions:

$$\mathcal{L}_{\mathrm{EW}} = \overline{Q}_j i\gamma^\mu D_\mu Q_j + \overline{u}_{Rj} i\gamma^\mu D_\mu u_{Rj} + \overline{d}_{Rj} i\gamma^\mu D_\mu d_{Rj} + \overline{L}_j i\gamma^\mu D_\mu L_j + \overline{\ell}_{Rj} i\gamma^\mu D_\mu \ell_{Rj}, \quad \text{(A.1)}$$

where the j index sums over three generations of quarks and leptons [220, 45]. Q_j and L_j represent left-handed doublets in the fundamental representation of the SU(2) group, where $Q_j = (u_L, d_L)_j$ contains left-handed up- and down-type quark fields, and $L_j = (\nu_L, \ell_L)_j$ the left-handed neutrinos and leptons. u_{Rj}, d_{Rj} and e_{Rj} are right-handed SU(2) singlets corresponding to right-handed quarks and leptons, and γ^μ are the Dirac matrices. The covariant derivative

$$D_\mu \equiv \partial_\mu - ig'\frac{1}{2}Y_W B_\mu - ig\frac{1}{2}T^a W^a_\mu \quad \text{(A.2)}$$

is used in place of a regular partial derivative, ∂_μ, to assure a local gauge symmetry. It introduces U(1) and SU(2) gauge fields, B_μ and W^a_μ, where a is the group index. Three components of the vector T are the SU(2) group generators (the vector is also called the weak isospin), Y_W is the weak hypercharge, and g and g' are coupling constants.

The physical fields A_μ and Z_μ, corresponding to the photon and the Z boson respectively, are linear combinations of the B_μ field and the third component of W_μ:

$$A_\mu = +B_\mu \cos\theta_W + W^{(3)}_\mu \sin\theta_W, \quad \text{(A.3)}$$

$$Z_\mu = -B_\mu \sin\theta_W + W^{(3)}_\mu \cos\theta_W, \quad \text{(A.4)}$$

where θ_W is the so-called weak mixing angle [221, 222, 220].

The equivalent Lagrangian for strong interaction between quarks can be simply written as

$$\mathcal{L}_{\text{QCD}} = \overline{\psi}_k i\gamma^\mu D'_\mu \psi_k, \tag{A.5}$$

$$D'_\mu \equiv \partial_\mu - i\alpha_s \frac{1}{2}\lambda^a \mathcal{A}^a_\mu, \tag{A.6}$$

where ψ_k is a triplet containing quark Dirac spinors for each colour charge of a quark [223, 224]. Index k runs over the the quark flavours and α_s is a strong coupling constant. $\mathcal{A}_\mu$ is an SU(3) gauge field with eight components corresponding to gluons carrying different colour charges, λ^a (the SU(3) group generators).

The kinetic term for gauge bosons in the SM Lagrangian is following:

$$\mathcal{L}_{kinetic} = -\frac{1}{4}W^a_{\mu\nu}W^{\mu\nu}_a - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a. \tag{A.7}$$

$W^a_{\mu\nu}$, $B^{\mu\nu}$, and $G^a_{\mu\nu}$ are gauge field strength tensors of SU(2), U(1) and SU(3) groups, respectively [224]. For instance, $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$, as in classical electrodynamics.

The full particle content of the Standard Model is summarised in Table A.1. The parton model, first to suggest that baryons and mesons are not elementary, was successfully used to explain data from the Stanford Linear Accelerator (SLAC) in 1969 [225, 226]. However, the complete set of quarks was not directly measured until the top quark discovery in Tevatron at Fermilab in 1995 [227, 71]. The observation of a gluon was first indicated, and later confirmed, in the measurements of the Ypsilon resonance at DESY in 1978 and 1979 [228, 229, 230, 231, 232]. The existence of the W and Z bosons was experimentally confirmed at the CERN Super Proton Synchrotron in 1983 [233, 234, 235, 236]. The Higgs boson and the origin of the mass terms and particle masses in the SM are discussed in the next section.

A.1.2 The Higgs boson

By default, the gauge bosons introduced to provide the local gauge symmetry of the theory are massless. In the case of mediators of the weak interaction, this scenario has been strongly disfavored by experimental observations that showed a very short range of this force, which suggests its carriers should be massive. However, inserting the mass terms by hand in the Lagrangian of the SM would break the local gauge symmetry and create problems related to renormalisability of the theory [224, 45]. Similarly, fermion (e.g. lepton) mass terms written simply as $m_L \bar{L}\ell_R$ would not preserve the symmetry, because right-handed field components are singlets under the SU(2) transformations. Following the logic in Ref. [45], this issue can be resolved by introducing a complex scalar SU(2) doublet

$$\phi = \frac{1}{\sqrt{2}}\begin{pmatrix}\phi^+ \\ \phi^0\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}\phi_1 + i\phi_2 \\ \phi_3 + i\phi_4\end{pmatrix}. \tag{A.8}$$

The simplest renormalizable Lagrangian for this field can be written as

$$\mathcal{L}_{\text{Higgs}} = (D_\mu\phi)^\dagger(D^\mu\phi) - V(\phi), \tag{A.9}$$

$$V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2, \tag{A.10}$$

Table A.1: Summary of fermions and gauge bosons in the Standard Model [42]. Limits on neutrino masses are given for measurements associated with a corresponding charged lepton; since the neutrino mass basis is different than the flavour basis, the limits do not refer to the mass of a given flavour. The extraction and interpretation of quark masses is highly non-trivial, and this topic is beyond the scope of this dissertation; for details, see Refs. [42] and references therein.

Category	Particle	Symbol	Electric charge [e]	Mass [MeV]
Gauge Bosons				
Electromagnetic	Photon	γ	0	0
Weak	W boson	$W^{\pm}$	± 1	80,369
Weak	Z boson	Z	0	91,188
Strong	Gluon	g	0	0
Quarks (Fermions)				
Generation 1	Up quark	u	$+\frac{2}{3}$	2.16
	Down quark	d	$-\frac{1}{3}$	4.7
Generation 2	Charm quark	c	$+\frac{2}{3}$	1,273
	Strange quark	s	$-\frac{1}{3}$	93.5
Generation 3	Top quark	t	$+\frac{2}{3}$	172,570
	Bottom quark	b	$-\frac{1}{3}$	4,183
Leptons (Fermions)				
Generation 1	Electron	e^-	-1	0.511
	Electron neutrino	ν_e	0	$< 0.8 \cdot 10^{-6}$
Generation 2	Muon	μ^-	-1	105.658
	Muon neutrino	ν_μ	0	< 0.19
Generation 3	Tau	τ^-	-1	1,776.93
	Tau neutrino	ν_τ	0	< 18.2

where μ and λ are parameters. For the potential to have a finite minimum, λ must be positive. For $\mu^2 < 0$ the vacuum expectation value (vev) is non-zero and satisfies

$$\phi^\dagger \phi = \frac{-\mu^2}{2\lambda} \equiv \frac{v^2}{2}, \tag{A.11}$$

which corresponds to an infinite set of degenerate vacuum states. The choice of a particular state is called spontaneous symmetry breaking (SSB), since the Lagrangian symmetry now breaks from SU(2)×U(1) to U(1). Without the loss of generality, the minimum can be chosen in the real direction of the neutral scalar field, ϕ^0, at $(\phi^+, \phi^0) = (0, v)$. Then, by expanding ϕ around that minimum

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ v + h(x) + i\phi_4 \end{pmatrix}, \tag{A.12}$$

and substituting into Equation (A.10), one gets a scalar, $h(x)$, with a mass

$$m_h = \sqrt{-2\mu^2} \tag{A.13}$$

and three massless Goldstone bosons. However, writing the Lagrangian A.9 in the so-called *unitary* gauge allows one to eliminate the unphysical massless bosons and automatically obtain the mass terms for the $W^\pm$ and Z^0 bosons from the kinetic terms. Therefore, the degrees of freedom associated with the Goldstone bosons become longitudinal polarisations of the $W^\pm$ and Z^0 bosons. This is the so-called Brout-Englert-Higgs-Guralnik-Hagen-Kibble (or, shortly, BEH) mechanism [237, 238, 239].

The weak boson masses arise from the kinetic terms of Equation (A.9):

$$m_W = \frac{1}{2} g v, \tag{A.14}$$

$$m_Z = \frac{1}{2} v \sqrt{g^2 + g'^2}. \tag{A.15}$$

By measuring m_W and g, it can be found that [42]

$$v \simeq 246\,\text{GeV}. \tag{A.16}$$

Now the fermion masses can also be written with the help of the Higgs doublet, by introducing the so-called Yukawa interactions:

$$\mathcal{L}_{\text{Yukawa}} = -y_{\ell j} \overline{L}_j \phi \ell_{Rj} - y_{dj} \overline{Q}_j \phi d_{Rj} - y_{uj} \overline{Q}_j \phi_c u_{Rj} + \text{h.c.} \tag{A.17}$$

The index j runs over generations and h.c. stands for hermitian conjugate of each term. $y_f = m_f \sqrt{2}/v$ are Yukawa couplings, and $\phi_c \equiv -i\sigma_2 \phi^*$, with σ_2 being the second Pauli matrix, is a conjugate Higgs field defined to allow writing the mass terms also for up-type quarks. ϕ_c transforms under SU(2) in the same way as ϕ.

The Higgs boson was discovered in 2012 by the ATLAS and CMS experiments in the Large Hadron Collider (LHC) at CERN [1, 2]. Its first measurements were performed using decay channels: $h \to \gamma\gamma$ (through a triangle loop) and $h \to ZZ^* \to 4\ell$, where $\ell = e^\pm, \mu^\pm$. The current Particle Data Group (2024) average of its mass measurements is $m_h = 125.2\,\text{GeV}$ [42]. Together with the vev estimate, one can use Equation (A.11) to determine the SM prediction for the parameter λ. However, independent measurements of the trilinear or quartic Higgs couplings are therefore necessary to test the shape of the SM Higgs potential.

A.2 Observables in the Quantum Field Theory

The Quantum Field Theory (QFT) provides the most comprehensive and well-tested known framework allowing to make predictions about processes that involve the elementary constituents of matter and the mediators of fundamental forces. Within this theory, the particles are represented by excitations of quantum fields that can have certain energies and carry momenta. In order to confront the predictions of a model written within this framework with an outcome of an experiment, one has to define some observables that could be measured. The standard quantities predicted by QFTs for that purpose are the cross sections and the particle decay rates.

A.2.1 Cross section

A typical particle physics experiment involves scattering of a beam of particles on a target or another beam going in an opposite direction. For quantitative description, one needs a measure of probability that an arbitrary process can occur. Following classical definition of probability, it can be then related to a number of events of that process observed in the experiment. The cross section is a quantity that provides such a measure of probability. Following the logic and notation in Ref. [44], given that we are colliding bunches $\mathcal{A}$ and $\mathcal{B}$, with lengths $l_{\mathcal{A}}$ and $l_{\mathcal{B}}$, and densities $\rho_{\mathcal{A}}$ and $\rho_{\mathcal{B}}$, the cross section σ is the ratio

$$\sigma = \frac{\text{Number of scattering events}}{l_{\mathcal{A}} l_{\mathcal{B}} \rho_{\mathcal{A}} \rho_{\mathcal{B}} A}, \tag{A.18}$$

where A is the area of overlap of the two bunches in the plane perpendicular to the motion of particles[1]. The "scattering events" refer to events of any arbitrary process, and the denominator here corresponds to the total number of scattering events over the common area of the bunches. Therefore, the cross section can be a measure of probability for the process to occur, even though σ has a dimension of area. It is also possible to define a differential cross section in order to obtain the probability density for some exact, specific values of final-state momenta. Due to four-momentum conservation, some of the momentum components are constrained. For the simplest case of only two particles in the final state, usually taken as unconstrained are the two angles of one of the particles. After integrating out the constrained components, that gives the differential cross section with respect to the solid angle, $\mathrm{d}\sigma/\mathrm{d}\Omega$.

In reality, the particle density in a bunch is not constant. To obtain the exact, useful expression for calculating the cross section, one must also take into account the wavepacket representation of the initial- and final-state particles, including the time evolution of initial states in the language of QFT. This gives the probability amplitude that can later be translated into a cross section. The initial and final states are related by a unitary operator called the S-matrix, for which a momentum-conserving matrix element $\mathcal{M}$ is defined. The expression that allows to calculate the cross section for two particles in the final state is:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{2E_{\mathcal{A}} 2E_{\mathcal{B}} |v_{\mathcal{A}} - v_{\mathcal{B}}|} \frac{|\mathbf{p}_1|}{(2\pi)^2 4E_{\text{cm}}} |\mathcal{M}(p_{\mathcal{A}}, p_{\mathcal{B}} \to p_1, p_2)|^2. \tag{A.19}$$

[1] This formula is valid only in the limit of small cross sections, when multiple scattering can be neglected, assuming the range of interaction and the size of wavepackets are small compared to the diameter of the bunches.

$p_{\mathcal{A}}, p_{\mathcal{B}}$ and p_1, p_2 are momenta of the initial- and final-state particles, respectively, expressed in the centre-of-mass frame. $E_{\mathcal{A}}$ and $E_{\mathcal{B}}$ are the energies of the initial-state particles, and $|v_{\mathcal{A}} - v_{\mathcal{B}}|$ their relative velocity in the laboratory frame[2], while E_{cm} is the total initial energy. The matrix element $\mathcal{M}$ is process-specific and depends on all possible Feynman diagrams that contribute to the process in perturbation theory. The details and exact derivation of Equation (A.19) are beyond the scope of this dissertation, and can be found in Ref. [44]. However, the presence of $\mathcal{M}$ in the cross section formula is a bridge between theory and experiment, and is the reason why the measurements of (differential) cross sections provide the perfect test for theory predictions.

A.2.2 Decay width and lifetime

Another observable that can be typically measured in the experiment and is also important in the context of this work is the decay rate. It is defined as a number of particles of a specific type in some ensemble that decay in the unit of time, divided by a number of particles in this ensemble. The decay rate is equal to the width of the Breit-Wigner function describing the shape of a resonance [44].

Similarly to the cross section, a differential decay rate, $\mathrm{d}\Gamma/\mathrm{d}\Omega$, can be defined. The formula is constructed in analogy to the expression for cross section, but featuring only one particle in the initial state. For two-body final state

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}\Omega} = \frac{1}{2m_{\mathcal{A}}} \frac{|\mathbf{p}_1|}{(2\pi)^2 4m_{\mathcal{A}}} |\mathcal{M}(p_{\mathcal{A}} \to p_1, p_2)|^2, \tag{A.20}$$

since in this case $E_{\text{cm}} = E_{\mathcal{A}} = m_{\mathcal{A}}$, where $m_{\mathcal{A}}$ is the mass of a decaying particle.

The decay rate is an important quantity in the context of this thesis, as it is related to the mean lifetime τ of a particle by $\Gamma \cdot \tau = \hbar$. Since it depends on the matrix element $\mathcal{M}$, the space of all possible final states (both in terms of allowed kinematics and decay modes) is responsible for the value of this parameter.

A.3 Effective field theories

The effective description of particles and interactions has also received a lot of attention in recent years. In general, effective theories are based on describing a particular regime in which they are valid, and many of them have been very successful throughout history. Good examples not related strictly to particle physics are Newtonian gravity, valid in the regime of non-relativistic velocities, or low-energy quantum mechanics, which is a limit of Quantum Electrodynamics for energies much smaller than the electron mass.

The most famous effective description in particle physics is the Fermi theory of weak interaction [240]. Although it is now well known that the muon decay is mediated by the heavy W boson, $\mu^- \to W^- \nu_\mu \to e^- \nu_e \nu_\mu$ (as follows from Equation (A.1)), the W particle was only discovered almost 30 years after Fermi's death. Despite that, this process was successfully described by Fermi using contact interaction, $\mu^- \to e^- \nu_e \nu_\mu$, with a single vertex connecting four nodes. In

[2] The expression $1/E_{\mathcal{A}} E_{\mathcal{B}} |v_{\mathcal{A}} - v_{\mathcal{B}}|$ is invariant under boosts along the beam axis.

the limit of momenta much smaller than the W mass the muon decay Lagrangian takes the form

$$\mathcal{L} = -\frac{G_F}{\sqrt{2}}\left[\bar{\nu}_\mu \gamma^\mu (1-\gamma^5)\mu\right]\left[\bar{e}\gamma_\mu (1-\gamma^5)\nu_e\right], \tag{A.21}$$

where

$$\frac{G_F}{\sqrt{2}} \equiv \frac{g^2}{8M_W^2} \tag{A.22}$$

is the historical Fermi constant [241]. This dimension-6 four-fermion operator can be used to make predictions on scattering amplitudes. By measuring the muon lifetime, one can determine the value of G_F and learn the energy scale Λ_{EW} of the weak interaction, which cannot be directly observed in this process. Since G_F is suppressed by the square of W mass, the relation between G_F and Λ_{EW} should be

$$G_F = \frac{1}{\Lambda_{EW}^2} \approx 1.17 \cdot 10^{-5}\,\frac{1}{\mathrm{GeV}^2}, \tag{A.23}$$

where the value of Fermi constant is derived from the formula for muon lifetime [42]. This leads to an estimation of the electroweak energy scale that matches the Higgs vev after the EWSB: $\Lambda_{EW} \approx 245\,\mathrm{GeV} \approx v$.

In a similar way, there may be some new physics phenomena mediated at higher energy scales to which there is currently no direct access, and the SM is only a successful effective description. However, the scale of this high-energy theory could be estimated by precisely measuring the properties of the SM and using a description that features all possible operators of dimension greater than four. While several approaches exist for constructing such an effective theory, the most popular example is the Standard Model Effective Field Theory (SMEFT) [73, 74]. In this model, in addition to the SM Lagrangian (before choosing the particular Higgs vacuum state), all possible higher-dimensional operators invariant under the SM gauge group are included:

$$\mathcal{L}_{\mathrm{SMEFT}} \supset \mathcal{L}_{\mathrm{SM}} + \sum_i \frac{c_i}{\Lambda_i^2}\mathcal{O}_i, \tag{A.24}$$

where only dimension-6 operators $\mathcal{O}_i$ are written out. c_i are the so-called Wilson coefficients that are typically treated as free parameters, and Λ_i is the energy scale of the new physics corresponding to a given operator.

One can try to fit this model to a set of experimental variables in order to search for deviations from the SM and establish the scale of new physics. In such a fit, the χ^2 function is minimised that depends on the difference between the prediction of an EFT (e.g. SMEFT) and the corresponding values given by the experimental data. In this way, the optimal values of the Wilson coefficients can be determined that match the observations. The experimental uncertainties are also included in these kinds of fits, allowing to calculate the errors on the coefficients.

Since SMEFT is constructed as the SM plus the contribution from higher-dimension operators, its predictions $\boldsymbol{\mu}(\boldsymbol{c})$ corresponding to a set of observables $\boldsymbol{y}$ can be defined as $\boldsymbol{\mu} = \boldsymbol{\mu}_{SM} + \boldsymbol{H}\cdot\boldsymbol{c}$, where $\boldsymbol{\mu}_{SM}$ is the prediction of the SM, $\boldsymbol{c}$ are the SMEFT Wilson coefficients, and $\boldsymbol{H}$ is a matrix used to parametrise corrections to the SM from SMEFT [75]. Consequently, if the SM is exact, all of the obtained coefficients should be consistent with zero. However, if any non-SM higher-energy processes contribute to the observations (and the measurements are sufficiently precise), one should see a significant deviation from zero in some of the coefficients. This can be achieved without making assumptions about the theory beyond the SM, and therefore, using EFTs such as SMEFT is a very powerful way to search for new physics phenomena in an almost model-independent manner.

Furthermore, a result suggesting the existence of new physics can be interpreted and used to provide some information about the underlying theory. A clear deviation would point to particular operators and processes, helping to identify the source of the anomaly. A particular pattern in the deviations could even directly point to a specific model [242], and the values of coefficients to a certain energy scale of new physics. In the case where the results are consistent with zero, a lower bound can be set on this scale. Depending on the precision of the measurements, this often allows one to probe energies much higher than those at which the data were collected.

Appendix B

Event displays

Some examples of event displays are presented in this appendix to give a better understanding and intuition of how the signal and the background can look like in the detector. In all figures only the TPC cage is shown, without other parts of the detector, and their content is described in the respective captions.

The event display in ILCSoft is handled by C Event Display (CED) [243], which allows drawing dynamic 3D pictures and saving screenshots. Different layers of detector and data can be displayed in CED. For the detector, one can show or hide any component. In case of data, the available layers correspond to

- True MC particles – the "perfect" trajectory of a particle assuming the information at the generator level and the magnetic field in the detector.
- Simulated hits – raw hits generated by Geant4 in trackers, calorimeters, and muon detectors.
- Reconstructed hits – hits in trackers, calorimeters, and muon detectors after digitisation.
- Tracks – digitised hits that were used in the track fit, together with a corresponding helix (for the parameters corresponding to a reference point in the IP).
- PandoraPFOs – particle-flow objects with all the assigned hits.

Charged MC particles are always shown with green lines. Simulated hits are represented by small dots, with different colours corresponding to MC particles that created them, whereas reconstructed hits are shown with slightly larger red squares. For tracks, separate segments are also presented with different colours of corresponding hits.

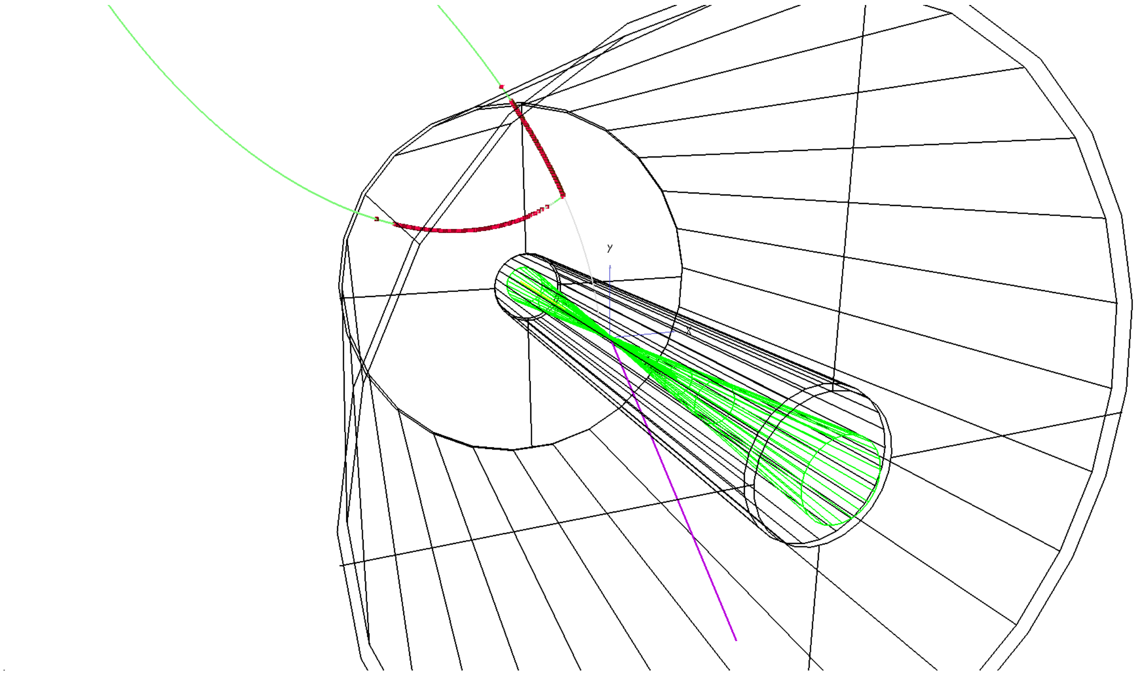


Figure B.1: Trajectories of the true MC particles (green lines) and digitised hits (red points) in the heavy scalar LLP production in the $\Delta m_{AH} = 5\,\text{GeV}$ scenario. Grey line represents reconstructed helix of one of the tracks. The more displaced particle (non-pointing to the IP) did not generate any hits close to the LLP decay vertex due to the missing hits problem.

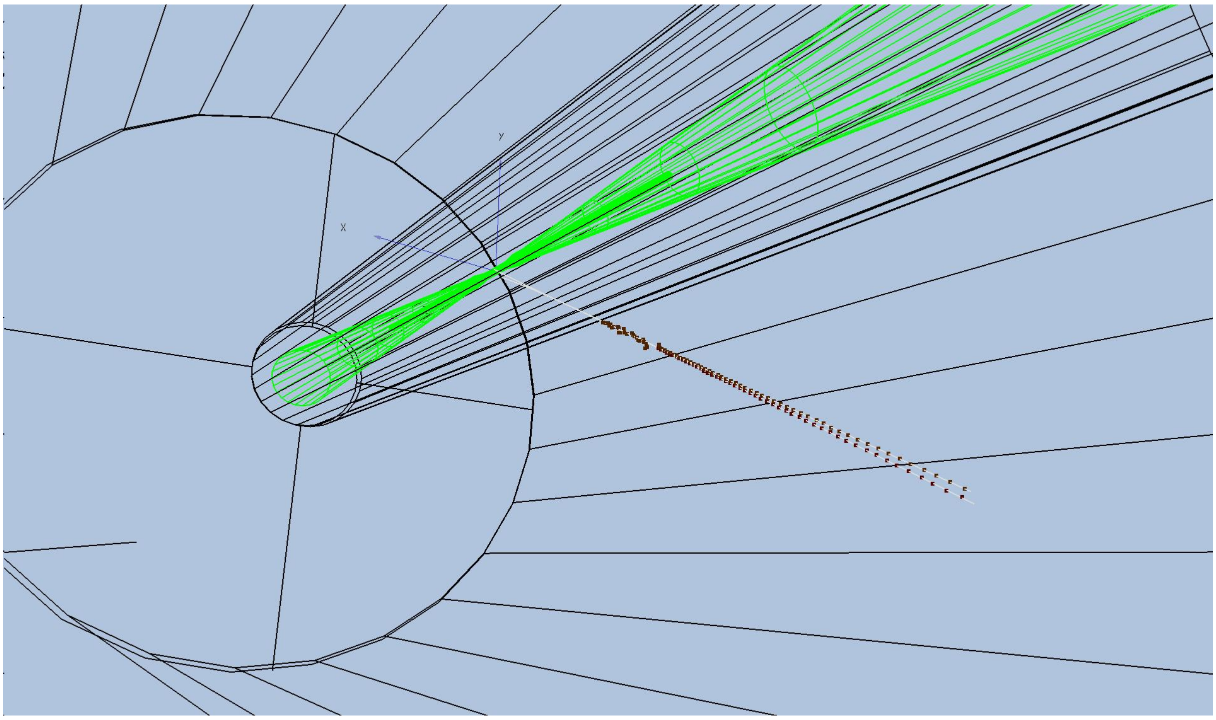


Figure B.2: Reconstructed track helices (grey lines) and hits used in the track fits (points) in the light pseudoscalar production with $m_a = 0.3\,\mathrm{GeV}$. The hits merged close to the LLP decay point making the momentum and vertex reconstruction more challenging and less precise.

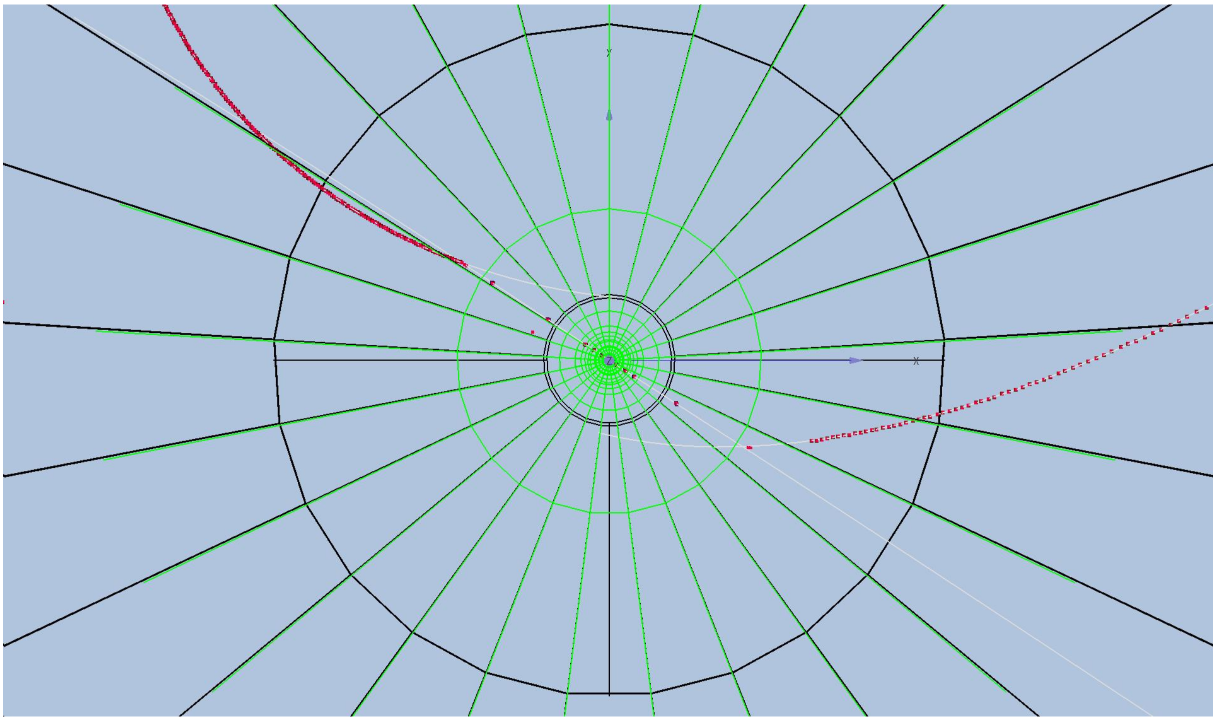


Figure B.3: Reconstructed track helices (grey lines) and digitised reconstructed hits (red points) in the dark fermion pair-production. Both decays occurred just behind the inner TPC wall, so the parent particles generated only a few TPC hits, which did not get assigned to any track.

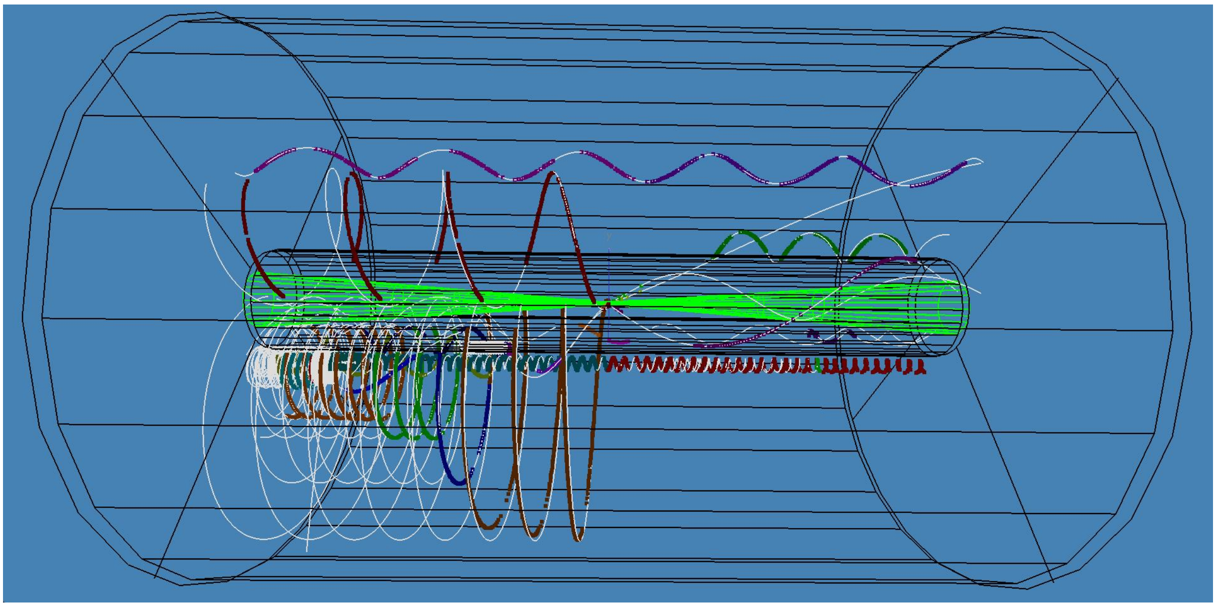


Figure B.4: Reconstructed track helices (grey lines) and hits used in the track fits (coloured points) in a low-p_T $\gamma\gamma \rightarrow$ hadrons event. Multiple split curlers can be seen, with a particularly good example of a track split in two in the top part (purple and dark-blue).

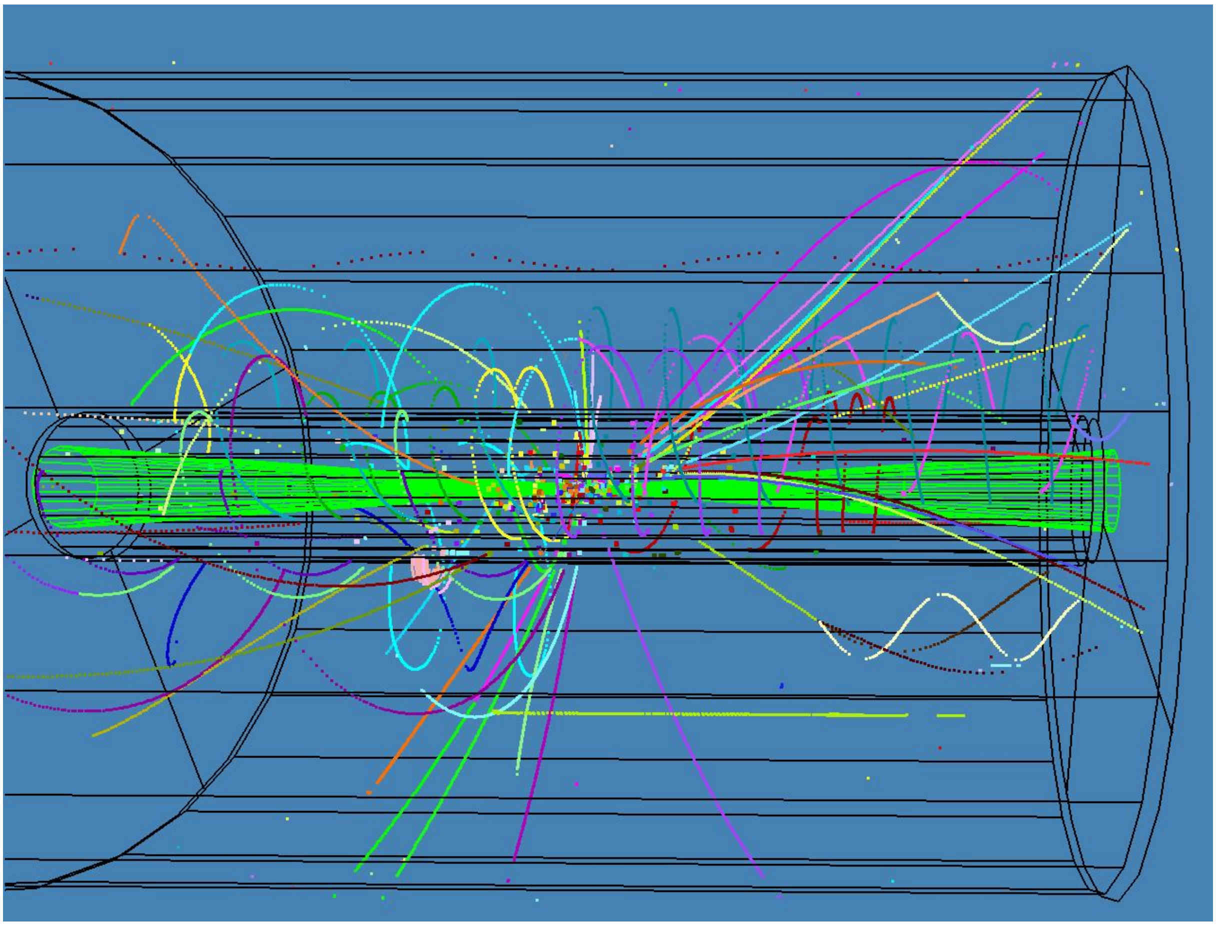


Figure B.5: Simulated hits in a W^+W^- production at $\sqrt{s} = 250\,\text{GeV}$ with the fully hadronic final state. Different colours correspond to different MC particles that generated the hits. A displaced vertex produced by a charged particle (green) is visible in the bottom-right corner. In the top-right corner a kinked track (orange) can be seen.