TÉCNICO LISBOA

Optimization of the Selection of Hidden Particles in the SHiP Experiment

Guilherme Machado Santos Soares

Thesis to obtain the Master of Science Degree in

Engineering Physics

Supervisors: Dr. Celso Filipe Correia Franco
Prof. Nuno Teotónio Viegas Guerreiro Leonardo

Examination Committee

Chairperson:	Prof. Mário João Martins Pimenta
Supervisor:	Prof. Nuno Teotónio Viegas Guerreiro Leonardo
Member of the Committee:	Prof. Sérgio Eduardo de Campos Costa Ramos
	Prof. Albert De Roeck

January 2021

"*His intentions is tremendous, but will he sway?*"

Daniel Dumile

Acknowledgments

First of all, I would like to thank my supervisors, Celso and Nuno, for accompanying and guiding me through this whole process, and for their never ending availability. I would also like to thank LIP, and more specifically the SHiP department, for providing me with this wonderful experience and letting me work alongside them.

My biggest thanks are for my family, that allowed me to pursue my dreams and were always supportive. Thank you for always being there.

Last but not least, I would like to thank my colleagues and friends, whom I spent endless days and nights with throughout these years, and that always made me want to keep moving forward regardless of the challenge.

Resumo

Apesar do Modelo Padrão (MP) ser uma das maiores conquistas em Física, não consegue explicar alguns fenómenos observáveis. Para isso, novos mecanismos ou partículas são necessários, como Neutrinos Estéreis ou Para-fotões do *Hidden Sector*.

Os Neutrinos Estéreis são neutrinos massivos hipotéticos que não acoplam à força fraca, mas que se misturam com os neutrinos do MP. Os Para-fotões são bosões massivos teoricamente previstos que se misturam com os fotões do MP. A experiência SHiP procurará detectar estas partículas, entre outras, através da observação directa de pelo menos dois decaimentos para produtos do MP. No entanto isto requer sistemas de veto de ruído.

Nesta tese apresentamos vários vetos de ruído optimizados para a experiência SHiP, recorrendo às propriedades cinemáticas das partículas reconstruídas. Os Neutrinos Estéreis (N) são estudados para massas entre 0.7 e 1.4 GeV/c^2, e os Para-fotões (A') para massas entre 0.021 e 4.4 GeV/c^2. Os modos de decaimento mais relevantes são considerados para ambas as partículas. Utilizam-se tanto uma abordagem de cortes cinemáticos, bem como métodos de aprendizagem automática (*Machine Learning*). A utilização de redes neuronais fornece os melhores resultados, com eficiências de seleção superiores a 97% para as amostras de $N \to \mu^{\mp}\pi^{\pm}$ e $A' \to \mu^-\mu^+$. Estudos de seleção são ainda relizados no cenário alternativo em que a câmara de decaimento (*Decay Vessel*) se encontra à pressão atmosférica em lugar de vácuo. Comparam-se as eficiências e níveis de ruído obtidos nos dois cenários.

Palavras-chave: Search for Hidden Particles Experiment; Neutrinos Estéreis; Para-fotões; Eficiências de Seleção; Estudos de Rejeição de Ruído.

Abstract

Although the Standard Model (SM) is one of the biggest achievements in physics, it cannot explain several outstanding phenomena. This requires the introduction of new mechanisms or particles to the SM, such as Heavy Neutral Leptons or Dark Photons of the Hidden Sector.

Heavy Neutral Leptons are hypothetical massive neutrino-like particles that do not couple to any SM forces, but do mix with SM neutrinos. Dark Photons are massive vector-like theorized particles that can mix with SM photons. These Hidden Sector particles are some of the biggest prospects at the SHiP experiment, which will try to discover them through the direct observation of at least two decays to the SM. This requires several background veto systems.

In this thesis we suggest several optimized background veto criteria for the SHiP experiment using the kinematic properties of the reconstructed particles, regarding Heavy Neutral Leptons (N) in the mass range between 0.7 and 1.4 GeV/c^2, and Dark Photons (A') in the mass range between 0.021 and 4.4 GeV/c^2. The most relevant decay modes are considered, and both a cut-based approach and machine learning methods are applied. The utilization of neural networks provided the best results, with selection efficiencies above 97% for the $N \to \mu^{\mp}\pi^{\pm}$ and $A' \to \mu^-\mu^+$ samples. Selection studies are also shown for an alternative scenario where the Decay Vessel is set at atmospheric pressure instead of the currently planned vacuum. HS selection efficiencies and background rejection are compared between both scenarios.

Keywords: Search for Hidden Particles Experiment; Heavy Neutral Leptons; Dark Photons; Selection Efficiencies; Background Rejection Studies.

Contents

List of Tables

List of Figures

List of Abbreviations

SM	Standard Model
SHiP	Search for Hidden Particles
BSM	Beyond the Standard Model
HS	Hidden Sector
HNL	Heavy Neutral Lepton
DP	Dark Photon
BAU	Baryonic Asymmetry of the Universe
QCD	Quantum Chromodynamics
DM	Dark Matter
CC	Charged Current
NC	Neutral Current
SSB	Spontaneous Symmetry Breaking
VEV	Vacuum Expectation Value
ν**MSM**	Neutrino Minimal Standard Model
SND	Scattering and Neutrino Detector
TT	Target Tracker
DIS	Deep Inelastic Scattering
RPC	Resistive Plate Chamber
MRPC	Multi-gap Resistive Plate Chamber
ML	Machine Learning
MLP	Multilayer Perceptron
NN	Neural Network
SHAP	Shapley Additive Explanations

Chapter 1

Introduction

1.1 Framework

The Standard Model [1] (SM) is one of the biggest accomplishments in physics up to this date. It describes all the known elementary particles and their interactions, with very high precision at energies up to the TeV scale.

In spite of this, the SM is not perfect and there are some shortcomings, that range from big theoretical gaps such as the unavailability of a Dark Matter candidate, to more precise problems such as small violations in the constructed leptonic flavour conservation.

The Search for Hidden Particles experiment (SHiP) is a proposed fixed target experiment to be set at the SPS accelerator at CERN, with two distinct main goals: study on an unprecedented level properties of the tau neutrino (ν_τ); and search for physics beyond the Standard Model.

To do this, the SHiP experiment has two very distinct spectrometer apparatuses. The first one, closest to the target, is the Scattering and Neutrino Detector, with the task of performing precision measurements related to the SM neutrinos, as well as indirect searches of Dark Matter, direct observations of pentaquarks [2, 3] and lepton flavour universality violation processes [4]. The second one is the Hidden Sector Spectrometer, and is the namesake of the experiment. It focuses on the direct detection of the decay products of Beyond the Standard Model (BSM) particles that interact very feebly with the SM ones.

From all of the possible BSM particles that SHiP will be able to probe, both the Heavy Neutral Leptons and the Dark Photons are of great interest.

1.2 New Physics at the Hidden Sector

Currently, there are two main ways of searching for new particles in accelerator experiments. Arguably, the most mediatic one is through the energy frontier, at places like the LHC, where the goal is to probe for particles with as big a mass as possible. These experiments search for new physics through theorized particles that are too heavy to have been produced with the technology previously developed, with the

latest successful example being the Higgs Boson.

However, there are other restrictions beyond the collision energy that can inhibit the discovery of new particles. There might be BSM particles that have very feeble effective couplings to the SM, but have currently accessible masses. These are considered to be in the Hidden Sector (HS), or may serve as mediators to the HS. The downside of these particles is that very intense beams are needed to produce them in substantial amounts. As Fig. 1.1 shows, they are in the intensity frontier of unknown physics.

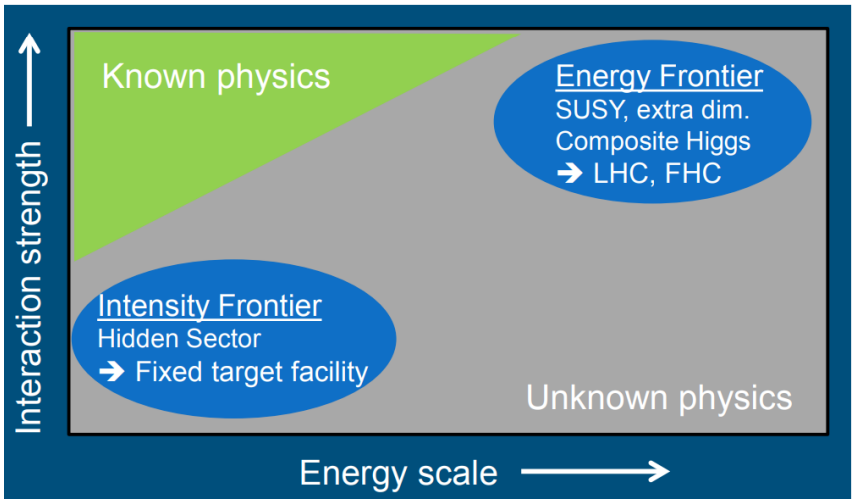

Figure 1.1: New physics that can be explored at intensity frontier experiments and its complementarity with energy frontier experiments, taken from Ref. [5].

Two major constraints arise from performing experiments in the intensity frontier, besides the inherent high beam intensity required. The first one is that since these particles interact very feebly, they may have very long expected lifetimes. As such, these experiments need to be long enough to allow the HS particles to decay into SM ones that we can detect, naturally leading to fixed target facilities. The other one is that with the high intensity of the beam there is a very significant production of background that can drown out the signal of the HS particles. Since the expected signal from the HS is already small, these experiments need to remove the background to a minimum.

The SHiP experiment will probe the intensity frontier, and will be a 0 background experiment, that aims to reduce the background to < 0.1 events for an expected 2×10^{20} proton on target (p.o.t.) interactions during the planned data taking by the experiment.

1.3 Motivation

Two of the HS particles that will be probed at SHiP will be Heavy Neutral Leptons (HNLs) and Dark Photons (DPs). The Heavy Neutral Leptons, are, as the name suggests, neutrinos. However, unlike the SM neutrinos, these are massive right-handed particles that do not couple to the weak force, and consequently to no force in the SM.

The HNLs are one of the biggest prospects at SHiP, since their discovery with a mass $< \mathcal{O}(10)$ GeV/c^2 can solve the SM neutrino mass problem, their oscillations and hierarchy in one go. Additionally some

models also predict one HNL to be a Dark Matter candidate, and under of the existence of an inflaton, can also explain the Baryonic Asymmetry of the Universe (BAU) through neutrino oscillations in what is known as leptogenesis.

Additionally, as Fig. 1.2 shows, SHiP will have access to a privileged parameter space among old and currently planned experiments at ≈ 1 GeV/c^2 mass, most notoriously with the effective coupling to ν_τ.

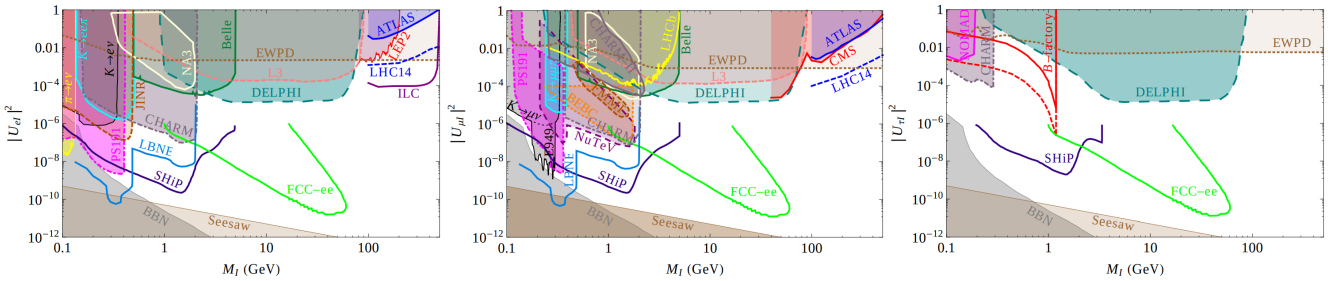

Figure 1.2: Predicted effective coupling to the SM neutrinos $|U_{\alpha I}|^2$ vs mass of the HNL M_I that SHiP will be able to probe at 90% confidence limit, with $\alpha = e, \mu, \tau$, compared against previous and future experiments, at the time of the Technical Proposal [6]. The (dotted, gray) contour labeled BBN corresponds to an HNL lifetime > 1 sec, which is disfavored by Big Bang Nucleosynthesis. The (dashed, brown) line "Seesaw" shows the scale of mixing naively expected in the canonical seesaw [7]. The (dotted, dark brown) contour labeled "EWPD" is the 90% C.L. exclusion limit from electroweak precision data. Figure adapted from Ref. [5].

The Dark Photon is a gauge boson associated with a new $U(1)_D$ symmetry that does not apply to the SM particles, but rather to other postulated fermions from the Hidden Sector, including Dark Matter candidates. Its name comes from the similarity to its SM counterpart, the photon, and due to the fact that it couples to the SM through kinetic mixing with the photon. This allows for indirect searches of Dark Photons in the Scattering and Neutrino Detector, through an excess of Neutral Current interactions.

The DP also allows for Dark Matter to self-interact and annihilate into SM particles. However, unlike the SM photon, the DP is a massive boson, and as such is able to decay into lighter particles, which leads to the possibility of direct searches in the Hidden Sector Spectrometer. Apart from solving precision problems such as the $g-2$ muon magnetic momentum, as of now, models with Dark Photons that do not include any new particles charged under the $U(1)_D$ symmetry are very constrained. This means that in ideal conditions, a direct observation of a Dark Photon strongly suggests the existence of even more BSM physics, from which come Dark Matter candidates.

Not unlike the HNLs, the SHiP experiment will be able to probe a never before reached parameter space of the mixing parameter ϵ vs Dark Photon mass $m_{A'}$, as shown in Fig. 1.3. As with kinetic mixing models, the effective coupling is given by the the mixing parameter, which in this case is denoted by ϵ.

Due to the 0 background requirement that the SHiP experiment has, there is a need to have many redundant ways to remove and distinguish the previously mentioned background from the signals of interest.

The experiment already has planned a use of physical veto systems, such as the timing detectors and the upstream muon detector at the entrance of the decay vessel in order to remove the surviving background that the shields are not able to eliminate. However, performing a kinematic analysis of the

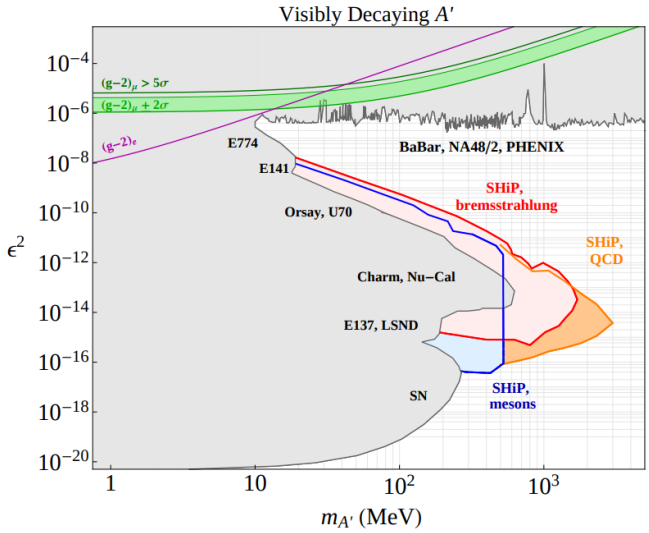

Figure 1.3: Predicted effective coupling to the SM photon ϵ vs mass of the DP $m_{A'}$ that SHiP will be able to probe, compared against previous experiments, at the time of the Technical Proposal. In this plot 3 production modes were considered: The blue line "mesons" is through production from meson decays, the red line with "bremmstrahlung" refers to production from proton bremmstrahlung and the orange line with "QCD" is assigned to Drell-Yan production. In green with the "$(g\text{-}2)_\mu$" label is the region of interest that solves the muon g-2 problem [8, 9] for a model with no new HS particles χ, charged under the newly added $U(1)'$ symmetry, and with masses $m_\chi < m_{A'}/2$. Figure taken from Ref. [5].

reconstructed particles and proceeding to distinguish the possible signal from the background is also a valuable asset to have, since it provides a different avenue to achieve our goals.

Currently, the Decay Vessel is scheduled to be at a pressure of ≈ 1 mbar to minimize the number of interactions inside, and consequently the background for the HS studies. Setting the Decay Vessel at the atmospheric pressure without compromising the HS parameter space that the SHiP experiment is able to probe would be extremely valuable, due to the large volume of the Decay Vessel and the difficulty of maintaining it at a low pressure.

1.4 Objectives

The HNLs and the DPs have had recently published sensitivity studies regarding their parameter space, which can be seen in Ref. [10] and Ref. [11], done for the most up-to-date setup of the SHiP experiment. They were performed with the **FairShip** [12] software, which has both particles implemented with their most relevant decay modes available. These studies considered the pre-selection criteria presented in the most recent Comprehensive Study Design Report [13]. These selection criteria, however, are not optimized, and rely on the veto systems in order to reduce the background to acceptable levels. To our knowledge, currently there are no available sensitivity studies regarding a Decay Vessel filled with air at a pressure of $\simeq 1$ bar.

This thesis has two main goals. First and foremost, we will try to optimize the selection of the HS particles through the kinematic properties of their reconstructions, while maintaining a 0 background environment. The second objective is to perform a similar study on a Decay Vessel with relaxed vacuum

conditions, at a pressure of $\simeq 1$ bar, and assess whether or not it is possible to maintain similar selection efficiencies of the HS particles to those obtained during the first part of this thesis. In order to do this, we will use standard one and two variable kinematic cuts, but we will also employ Machine Learning techniques that can perform non-linear multi-variable cuts, and should be able to improve the results obtained from the most orthodox method.

1.5 Thesis Outline

The thesis is divided into 9 chapters: 1 - Introduction; 2 - Phenomenology behind the Hidden Sector at the SHiP experiment; 3 - The Search for Hidden Particles Experiment; 4 - Samples; 5 - Heavy Neutral Leptons; 6 - Dark Photons; 7 - Machine Learning; 8 - Studies done with Reduced Vacuum; 9 - Concluding remarks.

In chapter 1 we gave a general presentation of the current state of BSM physics, the Hidden Sector, and where the SHiP experiment fits in. We also showcased the thesis objectives and their motivations.

In chapter 2 the theoretical background necessary for this thesis is presented. This includes the relevant parts of the SM, alongside the extensions that include HNLs and DPs that can be probed at SHiP, with their phenomenological implications and expected parameter spaces.

In chapter 3 we give an in-depth description of the SHiP experiment, with details on the currently planned apparatus and conditions in which the simulations were performed.

In chapter 4 all of the samples generated for the studies performed in the current conditions of the SHiP experiment are described. This includes both the background samples, as well as the HS ones. The software used in order to produce these samples is specified, and preliminary cuts are established. For the HS samples, the masses, couplings and decay modes are provided alongside the reasoning behind those choices.

In chapter 5 we present the kinematic properties of the Heavy Neutral Lepton samples generated. Afterwards, the study relying on the standard kinematic cuts is done, and the results of the selection efficiencies obtained are shown. The criteria set for these cuts is also discussed.

In chapter 6 there is a similar analysis to the one done on chapter 5. However, this now relates to the Dark Photon samples instead of the HNLs.

In chapter 7 we have the analysis resorting to Machine Learning methods. First an introduction to Machine Learning and Multilayer Perceptrons is given. Then the analysis is presented, including the selection efficiencies obtained for all the HS samples. In the end, a secondary study with the SHAP software [14] is shown, which gives insight into the results provided by the Neural Networks, and helps validating some of the kinematic cut choices done in chapters 5 and 6.

In chapter 8, the whole study regarding the Decay Vessel set at pressure of ≈ 1 bar is shown. This includes the description of the samples generated, the standard kinematic cuts study and the machine learning approach. The standard study has a comparison of the selection efficiencies obtained on the new samples when the kinematic cuts from chapters 5 and 6 are applied, and a new set of cuts that remove all background events. The Machine Learning approach is also extended, and now includes Neural Networks

that either provide 0 background events selected, or the same signal selection efficiencies obtained as the ones in chapter 7.

Chapter 9 provides some concluding remarks, as well as possible future work that expands on what has been shown here.

Chapter 2

Phenomenology behind the Hidden Sector at the SHiP experiment

2.1 The Standard Model

The Standard Model is one of the biggest accomplishments in science from the 20$^{\text{th}}$ century. It is able to very accurately describe all of the known elementary particles at low energies, and as of 2012, the last particle that it predicted, the Higgs boson, was found at the LHC, by both the CMS [15] and ATLAS [16] experiments.

Even though most predictions made by the Standard Model are incredibly accurate, this does not mean that the theory is complete, and there are several shortcomings that it cannot answer effectively. Some of the problems that still need to be figured out are the existence of Dark Matter (DM), the neutrino masses and consequently their flavour oscillations, the baryonic asymmetry of the universe, the lack of leptonic flavour universality or the anomalous magnetic dipole of the muon.

In this chapter we will look at the basics of the Standard Model and present extensions that can solve some of these shortcomings by introducing new particles, namely the Heavy Neutral Leptons and the Dark Photons. We will also present their expected phenomenology at the SHiP experiment.

2.1.1 Introduction to the Standard Model

The SM is a quantum non-abelian gauge theory that describes the strong, weak and electromagnetic interactions. The theory of electroweak interactions was firstly proposed by Glashow [17], in 1961. Three years later, the Higgs mechanism [18–21] was formulated, and, afterwards, implemented in the Glashow model by Weinberg and Salam [22, 23], in 1967-1968. Some properties of strong interaction were addressed in the quark model by Gell-Mann and Zweig [24, 25], in 1964. Finally, in 1972, 't Hooft and Veltman proved the renormalisability of the theory [26].

The Standard Model Lagrangian is invariant under the local symmetry gauge group

$$G_{SM} = SU(3)_c \times SU(2)_L \times U(1)_Y, \tag{2.1}$$

where c stands for colour charge, L for the left-handedness, and Y for the weak hypercharge. This theory describes the interactions between the fundamental matter constituents (fermions that can either be quarks or leptons), the force carriers (bosons), which mediate the interactions, and the scalar Higgs Boson.

The symmetries seen in the SM gauge group are directly linked to each of the 3 forces depicted in it. The $SU(3)$ symmetry relates to the strong force, while the $SU(2)$ and $U(1)$ relate to the weak and electromagnetic interactions instead, after some transformations.

The number of Gauge Bosons present in the theory comes directly from the generators of each symmetry group. Since the $SU(3)$ group contains 8 generators, there are 8 bosons associated with the strong force called gluons G^a_μ, with $a = 1, \cdots, 8$. The $SU(2)$ gauge group contains 3 generators, and as such there are 3 mediators of the weak force W^i_μ, with $i = 1, \cdots, 3$. At last, the $U(1)$ symmetry only has 1 generator, leading to only 1 boson associated with the electromagnetic interactions B_μ.

All of the interactions between the bosons are a direct consequence of maintaining a gauge invariance when inserting the aforementioned bosons into the SM Lagrangian, which will be shown later.

Unlike the gauge bosons, the amount of fermions is not determined by any underlying conditions in the theory, and as such is somewhat free. There is, however, the need of transforming under the symmetries described in Eq. (2.1).

So far there have only been found 3 generations of pairs for both the quarks and the leptons, and as such 3 generations are featured in the SM. Furthermore, there is also experimental evidence that only 3 types of active neutrinos couple to the weak force [27] for low energies. This was accomplished by measuring the decay width of the Z boson to invisible decays with the one predicted to neutrinos with irrelevant masses in the SM, yielding $N_\nu = \Gamma(invisible)/\Gamma^{theory}(Z \to \nu\overline{\nu}) = 2.992 \pm 0.007$ [28].

The last piece of the Standard Model puzzle is the doublet scalar also known as the Higgs Boson, responsible for generating the masses of the fermions and bosons as well. In Table 2.1 you can see a summary of the SM particle content.

The fermion fields, denoted by ψ, may be represented by the LH and RH chiral eigenfunctions of the matrix γ_5, ψ_L and ψ_R, with eigenvalues -1 and $+1$, respectively. Thus, we may write

$$\psi = \psi_L + \psi_R \quad \text{with} \quad \psi_L = \frac{1-\gamma_5}{2}\psi \quad \text{and} \quad \psi_R = \frac{1+\gamma_5}{2}\psi, \tag{2.2}$$

where $\gamma_5 = \gamma_0\gamma_1\gamma_2\gamma_3$, with γ_i being the Dirac matrices, and ψ is a four-component Dirac spinor and $\psi_{L(R)}$ are two component Weyl spinors. This representation is very helpful, since the SM contains the aforementioned $SU(2)_L$ symmetry, and, as can be seen in Table 2.1, the SM neutrinos only have a left-handed component. This is due to the fact that at the time of the writing of the SM the neutrinos were thought of as being massless particles, and as such they could be represented by a single chiral field. Furthermore, they were known to participate in weak interactions exclusively through their left-handed component. Given this, it was only natural to define them as being a left-handed field only, which was proposed by Landau [29], Lee and Yang [30] and Salam [31].

In this thesis I will be focusing on Heavy Neutral Leptons and Dark Photons. As such I will only

Gauge Fields				
$\mathbf{SU(3)_c}$	$G^a_\mu, a = 1, \ldots, 8$			
$\mathbf{SU(2)_L}$	$W^i_\mu, i = 1, \ldots, 3$			
$\mathbf{U(1)_Y}$	B_μ			
Matter Fields		T	T_3	Q
Quarks	$q_{\alpha L} \equiv \begin{pmatrix} u_{\alpha L} \\ d_{\alpha L} \end{pmatrix} = \left\{ \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix} \right\}$	$1/2$	$+1/2$ $-1/2$	$+2/3$ $-1/3$
	$u_{\alpha R} = \{u_R, c_R, t_R\}$	0	0	$+2/3$
	$d_{\alpha R} = \{d_R, s_R, b_R\}$	0	0	$-1/3$
Leptons	$\ell_{\alpha L} \equiv \begin{pmatrix} \nu_{\alpha L} \\ e_{\alpha L} \end{pmatrix} = \left\{ \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}, \begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}, \begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix} \right\}$	$1/2$	$+1/2$ $-1/2$	0 -1
	$e_{\alpha R} = \{e_R, \mu_R, \tau_R\}$	0	0	-1
Higgs	$\Phi \equiv \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$	$1/2$	$+1/2$ $-1/2$	$+1$ 0

Table 2.1: Standard Model particle content. T, T_3 and Q are the eigenvalues of the weak-isospin operator, its third component, and the eigenvalue of the charge operator, respectively. The hypercharge is given by $Y = 2(Q - T_3)$.

present some parts of the electroweak sector, circunventing anything QCD related, and dive deeper into neutrino physics.

The electroweak sector is invariant under the gauge groups $SU(2)_L \times U(1)_Y$, and we are able to present it separately from quantum chromodynamics since the Higgs mechanism only breaks the electroweak symmetry. The electroweak sector can be represented by the Lagrangian present in Eq. 2.3.

$$\mathcal{L}_{\text{EW}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{matter}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}, \tag{2.3}$$

where each term corresponds to a specific type of interaction. Notice that there are no gauge fixing and Faddeev-Popov terms included here. The first term includes the kinetic terms of the gauge bosons, and consequently, the interactions that involve gauge bosons only, and is given by:

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4} W^i_{\mu\nu} W^{i\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \tag{2.4}$$

where the tensor terms are defined as:

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g\epsilon^{ijk} W^i_\mu W^j_\nu \quad \text{and} \quad B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \tag{2.5}$$

with g the $SU(2)$ gauge coupling constant. The term $\mathcal{L}_{\text{matter}}$ contains the kinetic terms of the fermions, and their interactions with the gauge bosons due to the terms that preserve the gauge invariance, and can be written as:

$$\mathcal{L}_{\text{matter}} = \sum_\alpha i \left(\overline{q_{\alpha L}} \gamma^\mu D_\mu q_{\alpha L} + \overline{u_{\alpha R}} \gamma^\mu D_\mu u_{\alpha R} + \overline{d_{\alpha R}} \gamma^\mu D_\mu d_{\alpha R} + \overline{\ell_{\alpha L}} \gamma^\mu D_\mu \ell_{\alpha L} + \overline{e_{\alpha R}} \gamma^\mu D_\mu e_{\alpha R} \right), \tag{2.6}$$

where we use the notation defined in Table 2.1, and as such α defines the flavour. The term D_μ is the covariant derivative that preserves gauge invariance, and is given by:

$$\begin{aligned} D_\mu \psi_L &= \left[\partial_\mu - igT^k W_\mu^k - ig' \frac{Y(\psi_L)}{2} B_\mu\right] \psi_L, \\ D_\mu \psi_R &= \left[\partial_\mu - ig' \frac{Y(\psi_R)}{2} B_\mu\right] \psi_R. \end{aligned} \tag{2.7}$$

Here, g' is the $U(1)$ coupling constant and $T^k = \tau^k/2$, where τ^k are the SU(2) generators and can be represented by the Pauli matrices, with $k = 1, 2, 3$. The last term $Y(\psi_{L,R})$ is the hypercharge of the fermion that we are considering.

2.1.2 Charged and Neutral Currents

The Lagrangian in Eq. 2.6 can be rewritten in terms of the Charged and Neutral Currents (CC and NC respectively), and consequently of the bosons $W_\mu^\pm$, Z_μ^0 and A_μ. If we do so, we obtain the following Lagrangian:

$$\mathcal{L}_{\text{matter}} = \mathcal{L}_{kinetic} + \mathcal{L}_{CC} + \mathcal{L}_{NC}. \tag{2.8}$$

If we just look at the leptonic terms, and expand the ones that contain the left-handed fields of this Lagrangian, we get:

$$\begin{aligned} \overline{\ell_{\alpha L}} \gamma^\mu D_\mu \ell_{\alpha L} &= \overline{\ell_{\alpha L}} \gamma^\mu \partial_\mu \ell_{\alpha L} + g \overline{\ell_{\alpha L}} \gamma^\mu \frac{\tau^k}{2} W_\mu^k \ell_{\alpha L} + g' \frac{Y(\ell_{\alpha L})}{2} \overline{\ell_{\alpha L}} \gamma^\mu B_\mu \ell_{\alpha L} \\ &= \overline{\ell_{\alpha L}} \gamma^\mu \partial_\mu \ell_{\alpha L} + \frac{1}{2} \begin{pmatrix} \overline{\nu_{\alpha L}} & \overline{e_{\alpha L}} \end{pmatrix} \gamma^\mu \begin{pmatrix} gW_\mu^3 + g'Y(\ell_L)B_\mu & g\left(W_\mu^1 - iW_\mu^2\right) \\ g\left(W_\mu^1 + iW_\mu^2\right) & -gW_\mu^3 + g'Y(\ell_L)B_\mu \end{pmatrix} \begin{pmatrix} \nu_{\alpha L} \\ e_{\alpha L} \end{pmatrix}, \end{aligned} \tag{2.9}$$

while the simpler right-handed terms read:

$$\overline{e_{\alpha R}} \gamma^\mu D_\mu e_{\alpha R} = \overline{e_{\alpha R}} \gamma^\mu \partial_\mu e_{\alpha R} + g' Y(e_R) \overline{e_{\alpha R}} \gamma^\mu B_\mu e_{\alpha R}. \tag{2.10}$$

Defining the SU(2) charged bosons as being $W_\mu^\pm = (W_\mu^1 \mp iW_\mu^2)/\sqrt{2}$, it is easy to see that we can separate the CC interactions from the NC ones, since the quark interactions will be analogous.

As such the interactions Lagrangian $\mathcal{L}_{CC}$ can be written as:

$$\mathcal{L}_{CC} = \mathcal{L}_{CC}^\ell + \mathcal{L}_{CC}^q, \tag{2.11}$$

where we have

$$\mathcal{L}_{CC}^\ell = \frac{g}{\sqrt{2}} \left(\overline{e_{\alpha L}} \gamma^\mu W_\mu^- \nu_{\alpha L} + \overline{\nu_{\alpha l}} \gamma^\mu W_\mu^+ e_{\alpha L}\right), \tag{2.12}$$

$$\mathcal{L}_{CC}^q = \frac{g}{\sqrt{2}} \left(\overline{d_{\alpha L}} \gamma^\mu W_\mu^- u_{\alpha L} + \overline{u_{\alpha l}} \gamma^\mu W_\mu^+ d_{\alpha L}\right). \tag{2.13}$$

On the other hand, the NC interactions only involve the W_μ^3 and B_μ bosons. We can perform a transformation through the weak mixing angle θ_W into the mass eigenstates, and represent the interactions

through the A_μ and Z^0_μ bosons instead, with:

$$\begin{pmatrix} Z^0_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W^3_\mu \\ B_\mu \end{pmatrix} \quad \text{with} \quad \sin\theta_W = \frac{g'}{\sqrt{g'^2+g^2}} = \frac{g'}{g}\cos\theta_W. \tag{2.14}$$

This leads to the following Lagrangian of neutral currents:

$$\mathcal{L}_{NC} = \mathcal{L}^\ell_{NC} + \mathcal{L}^q_{NC}, \tag{2.15}$$

where now each term is given by

$$\mathcal{L}^\ell_{NC} = -e\overline{e_\alpha}\gamma^\mu A_\mu e_\alpha + \frac{g}{2\cos\theta_W}\left[\overline{\nu_\alpha}\gamma^\mu\left(g_V^{\nu_\alpha} - g_A^{\nu_\alpha}\gamma_5\right)\nu_\alpha + \overline{e_\alpha}\gamma^\mu\left(g_V^{e_\alpha} - g_A^{e_\alpha}\gamma_5\right)e_\alpha\right]Z^0_\mu, \tag{2.16}$$

$$\begin{aligned}\mathcal{L}^q_{NC} = &-e[Q(u)\overline{u_\alpha}\gamma^\mu u_\alpha + Q(d)\overline{d_\alpha}\gamma^\mu d_\alpha]A_\mu + \\ &+ \frac{g}{2\cos\theta_W}\left[\overline{u_\alpha}\gamma^\mu\left(g_V^{u_\alpha} - g_A^{u_\alpha}\gamma_5\right)u_\alpha + \overline{d_\alpha}\gamma^\mu\left(g_V^{d_\alpha} - g_A^{d_\alpha}\gamma_5\right)d_\alpha\right]Z^0_\mu.\end{aligned} \tag{2.17}$$

Here, e is the elementary electric charge, and g_V^ψ and g_A^ψ are the vector and axial couplings of the boson Z^0_μ to the fermion fields ψ, and are defined as

$$e \equiv g\sin\theta_W \ , \quad g_V^\psi \equiv T_3(\psi) - 2Q(\psi)\sin^2\theta_W \quad \text{and} \quad g_A \equiv T^3(\psi). \tag{2.18}$$

Note that the values of $T^3(\psi)$ and $Q(\psi)$ are present in Table 2.1.

2.1.3 Fermion Masses and Mixings

In the SM it is impossible to assign Dirac mass terms to fermions by construction, such as $-m_\psi\overline{\psi}\psi = -m_\psi\left(\overline{\psi_L}\psi_R + \overline{\psi_R}\psi_R\right)$. This happens because LH and RH particles differ in hypercharge, and cannot form a $SU(2)$ singlet. There is the need of a different mechanism that is able to provide these particles their masses. Spontaneous Symmetry Breaking (SSB) of the electroweak sector through the Higgs boson is such a mechanism, when we take into account the Yukawa couplings.

SSB happens due to perturbations around the vacuum expectation value (VEV) of the Higgs field. The Higgs Lagrangian in Eq. 2.3 reads as:

$$\mathcal{L}_{\text{Higgs}} = (D_\mu\Phi)^\dagger(D_\mu\Phi) - V(\Phi^\dagger\Phi) \ , \text{ where } \quad V(\Phi^\dagger\Phi) = \mu^2\Phi^\dagger\Phi + \lambda\left(\Phi^\dagger\Phi\right)^2, \tag{2.19}$$

where $V\left(\Phi^\dagger\Phi\right)$ is the Higgs potential, and D_μ is the covariant derivative in the fermion Lagrangian. $(D_\mu\Phi)^\dagger(D_\mu\Phi)$ is the term responsible for giving the $W^\pm_\mu$ and Z^0_μ bosons their masses.

The non-trivial vacuum of the theory can be computed by determining the field values that minimize the Higgs potential for $\lambda > 0$ and $\mu^2 < 0$. Only the neutral component of the Higgs field acquires a VEV, v, since the charge neutrality of the vacuum must be preserved. Thus, the expectation value of the Higgs

field is given by

$$\langle\Phi\rangle_0 \equiv \begin{pmatrix} 0 \\ v \end{pmatrix} \text{ , with } \quad v^2 = -\frac{\mu^2}{2\lambda}. \tag{2.20}$$

We can parameterize the Higgs scalar as a function of the perturbations around the VEV. Thus, we can write

$$\Phi(x) = e^{\frac{i\xi^k(x)\cdot\tau^k}{2v}} \begin{pmatrix} 0 \\ v + \frac{H(x)}{\sqrt{2}} \end{pmatrix}, \tag{2.21}$$

where $\xi^k(x)$ are the 3 Nambu-Goldstone bosons that are absorbed by the $SU(2)$ bosons, and the real $H(x)$ field is the Higgs boson.

When the Higgs field acquires a vaccum expectation value, just like in Eq. 2.20, the $SU(2)_L \times U(1)_Y$ symmetry is spontaneously broken, and hence the name of the mechanism arises. Nevertheless, a $U(1)_Q$ symmetry is preserved, which guarantees the existence of a massless gauge boson (the photon) associated to the conservation of the electric charge Q.

With the parameterization of Eq. 2.21, we can now create $SU(2)$ singlets with hypercharge 0. Thus, the Yukawa Lagrangian reads:

$$\mathcal{L}_{\text{Yukawa}} = -\mathbf{Y}^{\ell}_{\alpha\beta}\overline{\ell_{\alpha L}}\Phi e_{\beta R} - \mathbf{Y}^{u}_{\alpha\beta}\overline{q_{\alpha L}}\tilde{\Phi}u_{\beta R} - \mathbf{Y}^{d}_{\alpha\beta}\overline{q_{\alpha L}}\Phi d_{\beta R} + \text{H.c.}, \tag{2.22}$$

where the matrices $\mathbf{Y}^{\ell}_{\alpha\beta}$, $\mathbf{Y}^{u}_{\alpha\beta}$ and $\mathbf{Y}^{d}_{\alpha\beta}$ are the Yukawa coupling matrices for charged leptons, up and down quarks respectively. Here, we also use the definition

$$\tilde{\Phi} = i\tau_2\Phi^*(x) = \begin{pmatrix} \phi^{0*}(x) \\ -\phi^-(x) \end{pmatrix} \xrightarrow{SSB} \begin{pmatrix} v + \frac{H(x)}{\sqrt{2}} \\ 0 \end{pmatrix}. \tag{2.23}$$

When the Higgs field acquires a non-zero VEV, the fermions become massive and the Yukawa Lagrangian can then be rewritten as

$$\mathcal{L}_{\text{Mass}} = -\overline{e_{\alpha L}}\mathbf{M}^{\ell}_{\alpha\beta}e_{\beta R} - \overline{u_{\alpha L}}\mathbf{M}^{u}_{\alpha\beta}u_{\beta R} - \overline{d_{\alpha L}}\mathbf{M}^{d}_{\alpha\beta}d_{\beta R} + \text{H.c.}, \tag{2.24}$$

with $\mathbf{M}^{\ell}_{\alpha\beta} = v\mathbf{Y}^{\ell}_{\alpha\beta}$, $\mathbf{M}^{u}_{\alpha\beta} = v\mathbf{Y}^{u}_{\alpha\beta}$ and $\mathbf{M}^{d}_{\alpha\beta} = v\mathbf{Y}^{d}_{\alpha\beta}$. Since there are no RH neutrinos present in the SM, they are strictly massless.

The mass matrices $\mathbf{M}^{\ell ud}$ are arbitrary 3×3 complex matrices, and as such might not be in the physical (diagonal) basis. To diagonalize these matrices we can multiply them by two 3×3 unitary matrices $\mathbf{U_L}$ and $\mathbf{U_R}$, rotating the LH and RH fields respectively:

$$\begin{aligned} \mathbf{U_L}^{\ell\dagger}\mathbf{M}^{\ell}\mathbf{U_R}^{\ell} &= \text{diag}(m_e, m_\mu, m_\tau) \equiv \mathbf{d}^{\ell}, \\ \mathbf{U_L}^{u\dagger}\mathbf{M}^{u}\mathbf{U_R}^{u} &= \text{diag}(m_u, m_c, m_t) \equiv \mathbf{d}^{u}, \\ \mathbf{U_L}^{d\dagger}\mathbf{M}^{d}\mathbf{U_R}^{d} &= \text{diag}(m_d, m_s, m_b) \equiv \mathbf{d}^{d}, \end{aligned} \tag{2.25}$$

where the masses, m_ψ, are real and positive. Notice that m_ψ may not necessarily be the eigenvalues of the matrix M^ψ, but correspond instead to the square root of the positive and real eigenvalues of the

matrix $\mathbf{H} = M^{\psi\dagger}M^{\psi}$, diagonalized by ${\mathbf{U_R}}^{\psi}$ or ${\mathbf{U_L}}^{\psi}$.

We can now write the Lagrangian in Eq. 2.24 in the mass basis, by performing the following transformations:

$$\begin{aligned} e_{\alpha L} &\to \left({\mathbf{U_L}}^{\ell}\right)_{\alpha\beta} e'_{\beta L}, \quad e_{\alpha R} \to \left({\mathbf{U_R}}^{\ell}\right)_{\alpha\beta} e'_{\beta R}, \\ u_{\alpha L} &\to \left({\mathbf{U_L}}^{u}\right)_{\alpha\beta} u'_{\beta L}, \quad u_{\alpha R} \to \left({\mathbf{U_R}}^{u}\right)_{\alpha\beta} u'_{\beta R}, \\ d_{\alpha L} &\to \left({\mathbf{U_L}}^{d}\right)_{\alpha\beta} d'_{\beta L}, \quad d_{\alpha R} \to \left({\mathbf{U_R}}^{d}\right)_{\alpha\beta} d'_{\beta R}, \end{aligned} \tag{2.26}$$

where the primed fields now correspond to fermion fields with definite mass and definite flavours. The Lagrangian in the physical basis then reads as

$$\begin{aligned} \mathcal{L}_{\text{Mass}} &= -\overline{e'_{\alpha L}}\mathbf{d}^{\ell}_{\alpha\beta}e'_{\beta R} - \overline{u'_{\alpha L}}\mathbf{d}^{u}_{\alpha\beta}u'_{\beta R} - \overline{d'_{\alpha L}}\mathbf{d}^{d}_{\alpha\beta}d'_{\beta R} + \text{H.c.} \\ &= -m_{e\alpha}\overline{e'_{\alpha L}}e'_{\alpha R} - m_{u\alpha}\overline{u'_{\alpha L}}u'_{\alpha R} - m_{d\alpha}\overline{d'_{\alpha L}}d'_{\alpha R} + \text{H.c.}. \end{aligned} \tag{2.27}$$

In the new basis, the NC Lagrangian in Eq. 2.15 remains unchanged, since it relates fermion states with the same weak isospin. However, the quark CC Lagrangian in Eq. 2.13 gets modified and the interaction with the charged gauge bosons becomes non-diagonal, being written as

$$\mathcal{L}^{q}_{CC} = \frac{g}{\sqrt{2}}\overline{u'_{\alpha L}}\gamma^{\mu}\left(\mathbf{V}_{\text{CKM}}\right)_{\alpha\beta} d'_{\beta L}W^{+}_{\mu} + \text{H.c.}, \tag{2.28}$$

where $\mathbf{V}_{\text{CKM}} = {\mathbf{U_L}}^{u\dagger}{\mathbf{U_L}}^{d}$ is the unitary 3×3 Cabbibbo-Kobayashi-Maskawa (CKM) matrix [32, 33].

In contrast with the quark sector, the leptonic sector has no mixing in the CC interactions. We can show this by applying the transformation in Eq. 2.26 to Eq. 2.12, leading to:

$$\mathcal{L}^{\ell}_{CC} = \frac{g}{\sqrt{2}}\overline{e'_{\alpha L}}\gamma^{\mu}\left({\mathbf{U_L}}^{\ell}\right)^{\dagger}_{\alpha\beta}\nu'_{\beta L}W^{-}_{\mu} + \text{H.c.}. \tag{2.29}$$

Due to the massless character of the neutrinos in the SM, we can freely transform neutrino fields without spoiling any of the mass terms in the Lagrangian. Therefore, the transformation $\nu_{\beta L} \to \left({\mathbf{U_L}}^{\ell}\right)_{\alpha\beta}\nu_{\alpha L}$ may be applied, giving

$$\mathcal{L}^{\ell}_{CC} = \frac{g}{\sqrt{2}}\overline{e'_{\alpha L}}\gamma^{\mu}\nu'_{\alpha L}W^{-}_{\mu} + \text{H.c.}. \tag{2.30}$$

Hence, the CC current interactions cannot induce flavour changes in the leptonic sector, in the SM, leading to leptonic flavour symmetry.

With this framework, the electroweak sector of the SM can be described by six quark masses, four parameters that define the CKM matrix, three charged lepton masses, the Higgs mass, its VEV, and the gauge couplings g and g'. All of these physical parameters are experimental input, but when measured are able to constrain each other and verify the model.

Although the Standard Model makes very accurate predictions, it does not provide a complete picture of our universe. It is now known that neutrinos have very small masses, since they perform flavour oscillations [34, 35], and the SM not only forbids leptonic flavour change, but also explicitly states that neutrinos are massless particles. Adding to this, one still needs to find a mechanism that explains the

smallness of the neutrino masses.

Other unanswered questions by the SM include the anomalous magnetic moment of the muon [8, 36], also known as the $g-2$ problem, or the existence of Dark Matter, proposed as far back as 1933 by Zwicky [37] and in 1936 by Smith [38], due to missing mass in cluster observations, with reignited interest since the 1970s [39], to which the SM has no candidate available.

2.2 Heavy Neutral Leptons

The Heavy Neutral Leptons (HNLs), are, as the name suggests, hypothetical neutrino-like leptons, but with considerable masses. They were introduced in 1977 by Minkowsky [40] in order to treat the Standard Model as a limit case of a system with a $SU(2)_L \times SU_R \times U(2)^V$ symmetry. Due to the electroweak precision measurement constraints related to the number of light neutrino species that couple to the Z boson presented in Section 2.1.1, if these have masses of $M_N < \mathcal{O}(10^2)$ GeV/c^2, which will be in the range of masses that the SHiP experiment will be able to probe, these cannot couple to the weak force, and as such are also called Sterile Neutrinos.

In order to search for these particles at intensity frontier experiments like SHiP they must couple to the Standard Model through neutrino oscillations.

Due to the left-handness of the SM neutrinos, right-handed Heavy Neutral Leptons are an attractive idea that can solve the small masses of the SM neutrinos, through seesaw mechanisms [7, 41], while having the added bonus of extending the symmetries of the SM and providing very promising Dark Matter candidates [42–44]. Additionally, some extensions that include HNLs are able to solve the Baryonic Assymetry of the Universe through leptogenesis [43, 45, 46], making these very desirable particles.

Extensions to the SM that include 3 HNLs with masses below the electroweak scale are known as Neutrino Minimal Standard Models, often abreviated as νMSM. These theoretical proposals were first presented in 2005 by Shaposhnikov and Asaka [42], and can encompass all the aforementioned solutions, and as such are of the most interest at SHiP.

2.2.1 Neutrino Minimal Standard Model (νMSM)

The simplest renormalisable extension of the standard model, consistent with neutrino experiments, contains $\mathcal{N}$ right-handed $SU(2)\times U(1)$ singlet neutrinos N_I $(I = 1, \cdots, \mathcal{N})$ with the most general gauge-invariant interactions described by the Lagrangian extension:

$$\mathcal{L}_N = \overline{N}_I i\partial_\mu \gamma^\mu N_I - \left[F_{\alpha I}\overline{\ell_\alpha} N_I \tilde{\Phi} + \frac{M_I}{2}\overline{N}^c_I N_I + \text{H.c.}\right], \tag{2.31}$$

where Φ is the Higgs doublet as defined in 2.21, ℓ_α are the lepton doublets with $\alpha = (e, \mu, \tau)$, and $F_{\alpha I}$ is a $3 \times \mathcal{N}$ complex matrix that provides Yukawa-like couplings. Notice that it is implied that the SM lepton doublets are left-handed, while the new HNLs are right-handed. It is also worthy to mention that we are considering the Sterile Neutrinos to be Majorana particles [47], with a Majorana mass term M_I, since they do not have the need to conserve the SM lepton number. Furthermore, we are creating the

portal term

$$\mathcal{L}_{\text{portal}} = -F_{\alpha I}\overline{\ell_\alpha} N_I \tilde{\Phi}, \tag{2.32}$$

which provides not only a Dirac mass term for the SM neutrinos, $M_D = F\langle\Phi\rangle_0$, but is also responsible for a small symmetry break of the leptonic flavour conservation. This term also means that phenomenologically HNLs can couple to the SM through any process that involves a SM neutrino, provided it is kinematically accepted.

Since we are working with the νMSM and want to fulfill the conditions mentioned in the introduction of Chapter 2.2, we can assume several things, with the first being that there are only 3 HNLs ($\mathcal{N} = 3$). We can also assume without loss of generality that we are using a basis in which the mass matrices of the charged leptons and right-handed neutrinos are real and diagonal.

In these conditions, we are adding 18 new free parameters in comparison with the SM. From these, 3 are physical parameters, related to the masses of the HNLs, while the remaining 15 are owed to the Yukawa matrix, that can be parameterized by 3 diagonal couplings, 6 mixing angles and 6 CP-violating phases in the following way:

$$F = \tilde{K}_L f_d \tilde{K}_R^\dagger, \tag{2.33}$$

where

$$f_d = \text{diag}(f_1, f_2, f_3)\ , \quad \tilde{K}_L = K_L P_\alpha \quad \text{and} \quad \tilde{K}_R^\dagger = K_R^\dagger P_\beta. \tag{2.34}$$

The diagonal matrices for the Majorana phases are

$$P_\alpha = \text{diag}\left(e^{i\alpha_1}, e^{i\alpha_2}, 1\right)\ , \quad \textit{and} \quad P_\beta = \text{diag}\left(e^{i\beta_1}, e^{i\beta_2}, 1\right), \tag{2.35}$$

and the CKM-like matrix K_L is

$$K_L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{L23} & s_{L23} \\ 0 & -s_{L23} & c_{L23} \end{pmatrix} \begin{pmatrix} c_{L13} & 0 & s_{L13}e^{-i\delta_L} \\ 0 & 1 & 0 \\ -s_{L13}e^{-i\delta_L} & 0 & c_{L13} \end{pmatrix} \begin{pmatrix} c_{L12} & s_{L12} & 0 \\ -s_{L12} & c_{L12} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{2.36}$$

where $c_{Lij} = \cos\theta_{Lij}$ and $s_{Lij} = \sin\theta_{Lij}$, and with the construction of the other CKM-like matrix K_R being completely analogous with the replacement of L by R. Here we fix the indices so that 1 corresponds to the lightest HNL, and 3 to the heaviest.

This model can explain the SM neutrino masses and their oscillations, provided it obeys several cosmology and astrophysics constraints such as the active neutrino mass upper bound [48], that limits the strength of the Yukawa couplings.

Since the Majorana masses of the Sterile Neutrinos are assumed to be of the order of the electroweak scale or below, but still much greater than those of the SM neutrinos, seesaw mechanisms can still be applied, and from there a diagonalization of the active neutrino mass matrix provides the neutrino mixing matrix [49–52] (Pontecorvo-Maki-Nakagawa-Sakata matrix), responsible for the neutrino oscillations.

However, the νMSM is very attractive because it is able to solve other problems at the same time as well. One of those solutions is the proposal of a Dark Matter candidate alongside a leptogenesis generated

baryonic asymmetry method, that applies more restrictions [53–55] to the model.

2.2.2 Constructing a νMSM with HNLs that we can search

The easiest way to construct a νMSM that provides a light Dark Matter candidate, while solving the masses and oscillations of the neutrinos, is to start from simple symmetries, and then add an *ad hoc* term that breaks those symmetries. Let us consider that the lightest neutrino is the Dark Matter candidate [56]. Thus in a first approximation we can state that it does not couple to the SM neutrinos, and that it is also exactly massless.

This leaves us with the 2 heaviest HNLs to explain the neutrino phenomena and offer a CP-violating effect strong enough to generate the baryonic asymmetry. While the former condition does not impose big restrictions, the latter requires the Sterile Neutrinos to have as similar masses as possible [45]. Therefore we will consider them to be exactly degenerate in mass.

We shall denote these new fields by $\tilde{N}_I$, in order to differentiate from the HNLs presented in Section 2.2.1, that were in the basis with a diagonal mass matrix. By imposing these conditions we imply that $\tilde{N}_2$ and $\tilde{N}_3$ are under the same $U(1)$ symmetry, and transform according to

$$\tilde{N}_2 \to e^{-i\alpha_L}\tilde{N}_2 \quad \text{and} \quad \tilde{N}_3 \to e^{i\alpha_L}\tilde{N}_3, \tag{2.37}$$

where for $\tilde{N}_1$ we have a simple chiral symmetry

$$\tilde{N}_1 \to e^{i\beta_L}\tilde{N}_1. \tag{2.38}$$

Effectively, we are considering that $\tilde{N}_2$ and $\tilde{N}_3$ are particle and anti-particle, and this leads to the free Lagrangian of the HNLs to be

$$\mathcal{L}_{\text{free}} = \overline{\tilde{N}}_I i\partial_\mu\gamma^\mu\tilde{N}_I - M\overline{\tilde{N}_2^c}\tilde{N}_3. \tag{2.39}$$

In this model we now have the symmetry $U(1)_{23} \times U(1)_1$. However, we can also state that there is a single $U(1)_L$ symmetry, with the quantum numbers $(q, -1, 1)$, with $q \neq 0$, so that $\tilde{N}_1$ cannot have a Majorana mass term. Notice that for $q = \pm 1$ our theory still has one massless state in the sterile neutrino sector, and as such, it is allowed.

Since we want to have the active neutrinos interact with the sterile ones, these must also be charged under the newly proposed $U(1)_L$ symmetry. As such, there are 4 possible different charge assignments to the SM neutrinos: $(1, 1, 1)$, $(1, 1, -1)$, $(1, -1, 1)$ and $(-1, 1, 1)$. This makes it so that $\tilde{N}_1$ does not interact with any neutrino for $q \neq \pm 1$, sterile or otherwise.

Due to charge conservation, this leads to 4 different F_0 Yukawa matrices, that read as following:

$$\begin{aligned} \text{Model I: } F_0 &= \begin{pmatrix} 0 & h_{12} & 0 \\ 0 & h_{22} & 0 \\ 0 & h_{32} & 0 \end{pmatrix}, \quad \text{Model II: } F_0 = \begin{pmatrix} 0 & 0 & h_{13} \\ 0 & 0 & h_{23} \\ 0 & h_{32} & 0 \end{pmatrix}, \\ \text{Model III: } F_0 &= \begin{pmatrix} 0 & 0 & h_{13} \\ 0 & h_{22} & 0 \\ 0 & 0 & h_{33} \end{pmatrix}, \quad \text{Model IV: } F_0 = \begin{pmatrix} 0 & h_{12} & 0 \\ 0 & 0 & h_{23} \\ 0 & 0 & h_{33} \end{pmatrix}, \end{aligned} \tag{2.40}$$

where the parameters h_{ij} can be taken to be real.

By developing this simple model we have now achieved a way to create a new massless state $\tilde{N}_1$ that does not interact with the SM, while also creating 2 HNLs that are degenerate in mass and can interact with the active neutrinos. If we define the mass matrix of all the neutrinos, active and sterile, to be $\mathbf{M}^\nu$, it is now a non-zero matrix, which is given by

$$\mathbf{M}^\nu = \begin{pmatrix} \mathbf{0} & vF_0 \\ vF_0^\intercal & \mathbf{M}^{\tilde{N}} \end{pmatrix}, \text{ with } \quad \mathbf{M}^{\tilde{N}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & M \\ 0 & M & 0 \end{pmatrix}, \tag{2.41}$$

and $\mathbf{0}$ the SM neutrino null mass matrix. If we compute the eigenvalues of $\mathbf{M}^\nu$ for all of the 4 models, we can see that for models II through IV the active neutrinos have one massless state, and two others with degenerate masses. This leads us to assume that these models only apply to our reality if the masses are in the inverted hierarchy, where $m_3 < m_1 < m_2$.

Therefore, we can compare the predicted mixing angles between the active neutrinos to the measurements available, as was done in Ref. [43], through the elements of the active neutrino mass matrix M_ν. This matrix can be parameterized following Ref. [57]:

$$M_\nu = V^* \cdot \text{diag}\left(m_1, m_2 e^{2i\delta_1}, m_3 e^{2i\delta_2}\right) \cdot V^\dagger, \tag{2.42}$$

with $V = R(\theta_{23})\text{diag}\left(1, e^{i\delta_3}, 1\right) R(\theta_{13}) R(\theta_{12})$ the active neutrino mixing matrix.

This rules out both models II and III. However, model IV is still possible, since it corresponds to a $(L_e - L_\mu - L_\tau)$ symmetry [58, 59], even though it needs a not so trivial symmetry breaking mechanism [60].

As for model I, all three active neutrinos are massless, and as such all masses and mixing angles are defined by the symmetry breaking terms.

No matter the model, we are still reliant on a small symmetry breaking term in order to explain the observed properties of the active neutrinos. At the same time, there is still no Dark Matter candidate since $\tilde{N}_1$ is massless and the heavier HNLs are unstable.

If we assume that the symmetry breaking is generated through both the Majorana mass and Yukawa coupling terms, the symmetry breaking Lagrangian is given by:

$$\mathcal{L}_{\text{breaking}} = \Delta\mathcal{L}_{\text{Mass}} + \Delta\mathcal{L}_{\text{Yukawa}}, \tag{2.43}$$

where

$$\Delta\mathcal{L}_{\text{Mass}} = -\frac{\Delta M_{IJ}}{2}\overline{\tilde{N}_I^c}\tilde{N}_J, \tag{2.44}$$

and

$$\Delta\mathcal{L}_{\text{Yukawa}} = -\Delta F_{\alpha I}\bar{\ell}_\alpha \tilde{N}_I \tilde{\Phi}. \tag{2.45}$$

Here we omitted the H.c. term for simplicity, and explicitly we have

$$\Delta M = \begin{pmatrix} m_{11}e^{i\alpha} & m_{12} & m_{13} \\ m_{12} & m_{22}e^{i\beta} & 0 \\ m_{13} & 0 & m_{33}e^{i\gamma} \end{pmatrix}, \tag{2.46}$$

where all $m_{ij} \ll M$ can be taken to be real, and α, β and γ are 3 CP breaking Majorana phases. On the other hand, the Yukawa component is taken to be

$$\Delta F = \begin{pmatrix} h_{11} & 0 & h_{13} \\ h_{21} & 0 & h_{23} \\ h_{31} & 0 & h_{33} \end{pmatrix}, \tag{2.47}$$

where $h_{i1} \ll h_{k2}$ and $h_{i3} \ll h_{k2}$ are in general complex, containing 3 physical CP-breaking phases. Notice that we are assuming that the SM neutrinos couple very feebly to both $\tilde{N}_1$ and the heavier $\tilde{N}_3$, when compared to $\tilde{N}_2$. While this distinction is not necessary for the degenerate pair [44], and functionally similar models can still be built with analogous conclusions, it is used here for the sake of simplicity. On the other hand, the smallness of the h_{i1} couplings is required to ensure the viability of $\tilde{N}_1$ as a Dark Matter candidate.

The mass eigenstates ($N_{2,3}$) are related to $\tilde{N}_{2,3}$ by the unitary transformation [43]

$$\tilde{N} = U_R N_I. \tag{2.48}$$

The matrix U_R can be written in the form

$$U_R \simeq \frac{e^{i\phi^0}}{\sqrt{2}} \begin{pmatrix} e^{i\phi_1} & e^{i\phi_2} \\ -e^{-i\phi_2} & e^{-i\phi_1} \end{pmatrix}, \tag{2.49}$$

where the phases ϕ_i can be written as functions of the elements of ΔM, with an explicit form that is irrelevant to us. If we define the quantity $\epsilon = F_3/F_2$, where $F_i^2 = [h^\dagger h]_{ii}$ and $h = F_0 + \Delta F$, and take the aforementioned *ansatz* of $h_{i3} \ll h_{k2} \implies \epsilon \ll 1$, then the couplings of the mass eigenstates N_2 and N_3 to the standard model neutrinos are almost similar, and in first order depend on h_{i2}.

It is clear that at SHiP the particles that will be probed will be the heavier Sterile Neutrinos N_2 and N_3. We are interested in their effective couplings to the active neutrinos, $|U_{\alpha 2}|^2 \simeq |U_{\alpha 3}|^2 \equiv |U_\alpha|^2$. In the mass eigenstate basis M_2 and M_3 are not exactly degenerate, and as such they are produced and decay independently, showing a similar behaviour for different masses. With this, phenomenologically the model can be simplified to have only 4 free parameters: the possible masses of $N_{2,3}$ (M_N) and the

effective couplings to each of the 3 flavours of SM neutrinos ($|U_\alpha|^2$).

It is also trivial that in the mass eigenstate basis of both the active and sterile neutrinos $|U_\alpha|^2$ can be seen as a small mixing parameter, and as such the HNLs can now couple to the weak force SM bosons through an active neutrino-like suppressed interaction term given by

$$\begin{aligned}\mathcal{L}_{\text{int}} = \frac{g}{\sqrt{2}}\left(\overline{e_{\alpha L}}\gamma^\mu W_\mu^- \sum_\alpha U_\alpha N + \overline{N}^c \sum_\alpha U_\alpha^* \gamma^\mu W_\mu^+ e_{\alpha L}\right) + \\ + \frac{g}{4\cos\theta_W}\left(\overline{\nu_{\alpha L}}\gamma^\mu Z_\mu^0 \sum_\alpha U_\alpha N + \overline{N}^c \sum_\alpha U_\alpha^* \gamma^\mu Z_\mu^0 \nu_{\alpha L}\right).\end{aligned} \tag{2.50}$$

As was shown in Ref. [43], we can express the ratios of the effective couplings $f_\alpha \equiv |h_{i2}|$ through elements of the active neutrino mixing matrix. A particularly simple expression can be derived for a case where $\theta_{13} = 0$ and $\theta_{23} = \pi/4$, which is in agreement with experimental data and can be constructed through model I. For normal hierarchy, where $m_1 < m_2 < m_3$, there are possibilities:

$$f_e^2 : f_\mu^2 : f_\tau^2 \approx \frac{m_2}{m_3}\sin^2(\theta_{12})|1 \pm x^2| : \frac{1}{2}|1 - x^2|^2 : \frac{1}{2}|1 \pm x|^4, \tag{2.51}$$

where $x = ie^{i(\delta_1-\delta_2-\delta_3)}\sqrt{\frac{m_2}{m_3}}\cos(\theta_{12})$, and all combinations of signs are admitted. For a numerical estimate we follow Ref. [57] and take $\sin^2(\theta_{12}) \simeq 0.3$, leading to $x \simeq 0.35ie^{i(\delta_1-\delta_2-\delta_3)}$ and to $f_e^2/(f_\mu^2+f_\tau^2) \sim 0.05$.

As for inverted hierarchy, the solutions are:

$$f_e^2 : f_\mu^2 : f_\tau^2 \simeq \frac{1+p}{1-p} : \frac{1}{2} : \frac{1}{2}, \tag{2.52}$$

where $p = \pm\sin(\delta_1)\sin(2\theta_{12})$. Taking the same value of θ_{12}, we obtain $f_e^2/(f_\mu^2+f_\tau^2) \sim 0.04-25$, depending on the CP-violating phase δ_1.

The corrections to Eqs. (2.51) and (2.52) are of the order $\mathcal{O}(\epsilon)$. Taking the previously mentioned *ansatz* of $\epsilon \ll 1$, these equations are very good approximations of the ratios.

With Eqs. (2.51) and (2.52) we have the basis for our analysis of experimental signatures of HNLs, if we take into account that our experimental apparatus is able to probe viable regions of the parameter space[1].

2.2.3 Sensitivity Studies done on the HNLs

In order to predict a parameter space that can be probed at experiments like SHiP, and perform sensitivity studies, there is a need to establish the effective couplings of the theorized particles to the SM ones. As was mentioned in the previous section 2.2.2, due to the similarity between N_2 and N_3, in the case where there is a small breaking of the leptonic symmetry, there is only the need to consider 4 new parameters in the νMSM: the coupling of $N_{2,3}$ to each active neutrino $|U_\alpha|$, and the HNL masses M_N.

Following [44], from the coupling ratios for the normal and inverted hierarchies, in Eqs. (2.51) and (2.52) respectively, we can define 3 different sets of Yukawa couplings related to "extreme hier-

[1] For further details on constraints related to the effective couplings see Refs. [43, 44]

archies", where the couplings f_α and f_β are minimized in comparison with f_γ, with $\gamma \neq \alpha, \beta$, since U_α is determined through f_α in a first approximation. These 3 sets are considered to be the benchmark models used in all HNL related studies at SHiP.

$$\begin{aligned} \text{model I:} \quad & f_e^2 : f_\mu^2 : f_\tau^2 \approx 52 : 1 : 1, \quad \kappa = 2, \\ \text{model II:} \quad & f_e^2 : f_\mu^2 : f_\tau^2 \approx 1 : 16 : 3.8, \quad \kappa = 1, \\ \text{model III:} \quad & f_e^2 : f_\mu^2 : f_\tau^2 \approx 0.061 : 1 : 4.3, \quad \kappa = 1. \end{aligned}$$

Model I is obtained by increasing the ratio $\frac{1+p}{1-p}$ presented in Eq. (2.52), as much as possible, taking into account the appropriate combination of signs, and the center value of the neutrino mixing matrix of $\theta_{12} \simeq 34^\circ$ [57].

Only a normal ordering of the active neutrino masses is able to provide significantly different couplings to the muon and tau flavours. Thus, in models II and III we look at Eq. (2.51). It is clear that for model III, if we take x to be real and positive, we have

$$\frac{|f_\tau|}{|f_\mu|} \simeq \left(\frac{1+p}{1-p}\right)^2. \tag{2.53}$$

Once again, taking the center values of the neutrino mixing matrix from [57], we get $x \simeq 0.35$ and consequently $\frac{|f_\tau|}{|f_\mu|} \simeq 4.3$. Analogously, the ratio between the tau and electron couplings is given by

$$\frac{|f_\tau|}{|f_e|} \simeq \left(\frac{m_2}{2m_3}\sin^2\theta_{12} \cdot \left(\frac{1-x}{|1+x^2|}\right)^2\right)^{-1} \simeq 71. \tag{2.54}$$

We get the couplings in model II through similar considerations.

From Eq. (2.50), we know that the Sterile Neutrinos only interact with the SM through leptonic processes. At fixed target experiments, such as SHiP, GeV-scale Heavy Neutral Leptons are mostly produced through the (semi)leptonic decays of c- and b- mesons [61], favouring the lightest ones.[2] In Fig. 2.1 are the two basic diagrams of these decays. Notice that charged conjugate channels are possible due to the Majorana character of the HNLs, but were omitted for simplicity. In these diagrams, the pair (U,D) can be any pair of up and down quarks (u_i, d_j), where, for the SHiP case, at least one of U or D needs to be a charm or bottom quark, respectively.

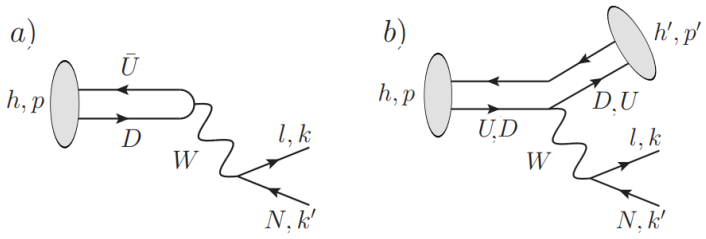

Figure 2.1: Diagram for the favoured channels of HNL production mediated by leptonic a) and semi-leptonic b) meson decays. Figure adapted from [61].

[2] These are not limited to open flavour mesons, but also include quarkonia states such as J/ψ.

The decays of the HNLs are always mediated by the weak current, and consequently always include at least one lepton in the final state. The two basic diagrams that include all decays can be seen in Fig. 2.2. Once again, charge conjugate channels were omitted for simplicity. For the charged current-mediated decays (Fig.2.2(a)), the pair (U, D) can either be leptonic $(\nu_\alpha, \ell_\alpha)$ or a pair of up and down quarks (u_i, d_j). As for f in the neutral current decays, it can be any fermion. Obviously, both currents demand that the decay is kinematically acceptable.

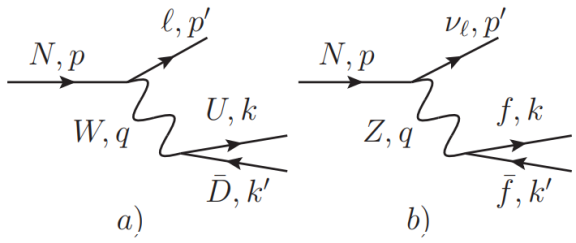

Figure 2.2: Diagram for the HNL decays mediated by charged a) and neutral b) currents. Figure taken from [61].

Several sensitivity studies have been done regarding the HNL probing performance at a SHiP-like experiment, from the simpler ones in the embryonic *Proposal to Search for Heavy Neutral Leptons at the SPS* [62] and the SHiP Technical Proposal [6], to more concrete studies regarding a better established SHiP experiment [10, 63]. Due to its recency and more complete character, this thesis followed the parameter space established in [10].

The sensitivity to HNLs can be estimated through a combination of calculations and simulations of HNL events in the SHiP apparatus in order to determine specific efficiencies, done with the **FairShip** software, based on the **FairRoot** framework [64].[3] The amount of estimated detected HNL events (N_{events}) can be written as

$$N_{events} = N_{\text{prod}} \times P_{\text{det}}, \tag{2.55}$$

where N_{prod} is the number of produced HNLs that fly in the direction of the fiducial volume, and P_{det} is the probability of detection of these HNLs in the Hidden Sector Spectrometer[4].

The number of produced HNLs is given by

$$N_{prod} = \sum_{q\in(c,b)} N_q \times \sum_h f(q \to h) \times \text{BR}(h \to N + X) \times \epsilon_{\text{decay}}, \tag{2.56}$$

where $f(q \to h)$ is the h meson fragmentation fraction at SHiP, $\text{BR}(h \to N + X)$ is the mass dependent inclusive branching ratios for h mesons decay with HNLs in their final state, and ϵ_{decay} is the geometrical acceptance of the detector. As for N_q, it is the total number of produced quarks and anti-quarks with flavour q, taking into account the quark-antiquark production fraction from the initial collisions $X_{q\bar{q}}$ and

[3] The full framework of **FairShip** is mentioned in section 4.1.
[4] For more detail on the Hidden Sector or Decay Spectrometer see section 3.2.5.

the cascade enhancement factor $f_{cascade}$ from secondary p.o.t. interactions.

$$N_q = 2 \times X_{q\bar{q}} \times f_{cascade} \times N_{POT}. \tag{2.57}$$

As for the probability of detection P_{det}, it only depends on the probability that the HNL decays within the allowed volume (P_{decay}), the branching ratio of the specific decay to visible channels (BR($N \to$ visible)), and the detection efficiency ϵ_{det}, being given by:

$$P_{\text{det}} = P_{\text{decay}} \times \text{BR}(N \to \text{visible}) \times \epsilon_{\text{det}}, \tag{2.58}$$

with

$$P_{\text{decay}} = \exp\left(-\frac{\ell_{\text{ini}}}{\ell_{\text{decay}}}\right) - \exp\left(-\frac{\ell_{\text{fin}}}{\ell_{\text{decay}}}\right). \tag{2.59}$$

Notice that ℓ_{ini} is the distance travelled before the HNL enters the Decay Vessel, ℓ_{fin} is the distance travelled until the end of the Decay Vessel, and $\ell_{\text{decay}} = c\gamma\tau_N$ is the usual decay length.

Alongside the three benchmark models described previously, we can produce the estimated sensitivity of SHiP to HNLs as a function of their masses and the strongest coupling to the HNL (U_α^2), presented in Fig.2.3. Notice that the limits of the sensitivity curves were estimated for a total of 2.4 events detected over the whole running time of the experiment.

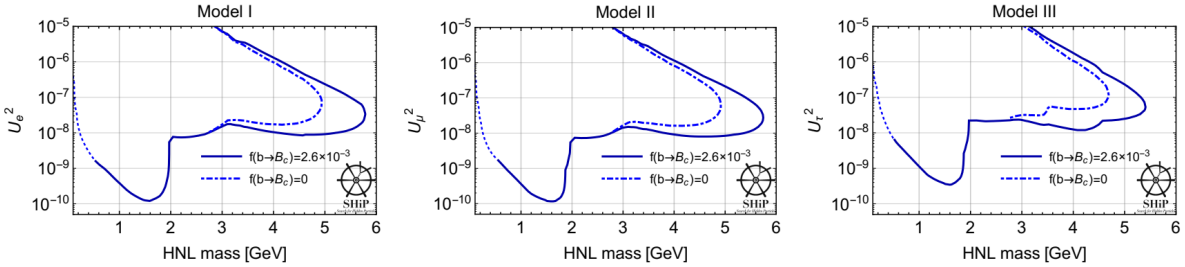

Figure 2.3: Sensitivity curves for the 3 benchmark models I-III (90% CL). To indicate the uncertainty related to the unknown production fraction of B_c mesons at SHiP, there are two types of curves for each flavour. Solid curves show the sensitivity contours when their production fraction equals to that at LHC energies: $f(b \to B_c) = 2.6 \times 10^{-3}$. Dashed-dotted lines do not include contributions from B_c. Below 0.5 GeV only production from D and B mesons is included (dotted lines). Figure taken from Ref. [10].

2.3 Dark Photons

The Dark Photons (DPs), or para-photons as they were initially proposed [9], are hypothetical gauge bosons associated with a new $U(1)_D$ gauge group, similar to the one present in the SM, and as such, they are similar to the SM photons and Z^0 boson. However, the existing particles in the SM are not charged under this new gauge group, unlike for the hypercharge of the existing $U(1)_Y$ symmetry. Instead, $U(1)_D$ can only charge particles either from the Hidden Sector, χ, which can include Dark Matter candidates, or from higher energy scales.

Since the DP has an analogous behaviour to the neutral bosons in the SM, it cannot couple to particles

that do not have a charge under $U(1)_D$, and as such should not be able to interact with SM particles. There is, however, an analogous behaviour to the one seen with the HNLs, where the DP kinnetically mixes with the neutral bosons of the SM. Consequently, it has an interaction Lagrangian similar to these, albeit highly suppressed[5].

The addition of Dark Photons to the SM is a very attractive idea since the extensions to the SM can range from very simple, with just one new particle (the DP), to being able to contain many Hidden Sector particles that only interact with the SM through the DP itself.

If the DPs are massive and able to decay to SM particles, thus being observable, they can be Dark Matter candidates. In the case that they do not fulfill the role of Dark Matter, they still provide very solid hints of the existence of new particles χ, since the simplest model is highly constrained [65, 66]. Stemming from the fact that, due to their nature, the DPs provide a way for the χ particles to self-annihilate into SM ones, these can be good Dark Matter candidates. Additionally, the fact that there is kinetic mixing with the SM photons also gives an avenue to explain the $g-2$ muon magnetic momentum discrepancy [8, 9], while in return being constrained by it.

2.3.1 Dark Photon mediated Hidden Sector

There are several ways of extending the Standard Model through the addition of a Dark Photon [67, 68]. [6] Regardless of the method, lies the fact that any theory with a visible Dark Photon is an effective theory for energies below the electroweak scale. This always leads to the emergence of a kinetic mixing between the new Gauge Boson A'_μ and the already existing B_μ. The effective extension to the SM can be given by the Hidden Sector Lagrangian, $\mathcal{L}_{\text{HS}}$, which is given by

$$\mathcal{L}_{\text{HS}} = \mathcal{L}_{A'} + \mathcal{L}_{\chi}, \tag{2.60}$$

where we have $\mathcal{L}_{A'}$ the Lagrangian that only contains the DP (A'), and $\mathcal{L}_{\chi}$ is the Lagrangian of the particles charged under $U(1)_D$. The easiest way of adding a Dark Photon to the currently existing paradigm does not include direct couplings to the SM particles, and as such is usually denoted as a *fermiophobic* model. Instead, the Dark Photon interacts with the Standard Model through the kinetic mixing factor [68], leading to very feeble couplings and allowing for very low masses of the new gauge boson. Given that, the simplest extension can just be given by the former Lagrangian, that can be written as

$$\mathcal{L}_{A'} = -\frac{1}{4}\left(F'_{\mu\nu}\right)^2 - \frac{\epsilon}{2}F'_{\mu\nu}F^{\mu\nu} - \frac{1}{2}m^2_{\gamma_D}\left(A'_\mu\right)^2. \tag{2.61}$$

Here $F^{\mu\nu}$ is the SM $U(1)_e$ field strength, given by $F_{\mu\nu} = \partial_\mu F_\nu - \partial_\nu F_\mu$, $F'_{\mu\nu}$ is field strength of the Dark Photon with an analogous definition, ϵ is the mixing factor and m_{γ_D} is the mass of the Dark Photon. This mixing parameter provides the interaction terms of the DP with the SM, since through the equations of motion we can define $\partial_\mu F^\nu = eJ^{EM}_\nu$, where J^{EM}_ν is the electromagnetic current, and rewrite the

[5] While it is not necessary that the Lagrangian is suppressed at all, due to the nature of SHiP only feebly interacting particles are of interest.

[6] For a good breakdown of the most popular extensions, and their consequences, see Ref. [69].

interaction term as

$$\frac{\epsilon}{2} F'_{\mu\nu} F^{\mu\nu} = A'_{\mu} \times (e\epsilon) J^{EM}_{\mu}. \tag{2.62}$$

Eq. (2.62) means that the DP couples to the eletromagnetic current, but with the interactions suppressed by a factor of ϵ, giving it a behaviour akin to that of the SM photon. One can generalize the Lagrangian in Eq. 2.61 to the SM $U(1)_Y$ by replacing $F^{\mu\nu}$ by $B^{\mu\nu}$, defined in Eq. 2.5. With this generalization, the DP now also mixes with the weak neutral current in an analogous way[7].

From this model, we quickly realize that there are only 2 effective parameters needed to describe an experimental signature: the mixing parameter ϵ, and the mass of the Dark Photon m_{γ_D}. Additionally, if we look at the simpler model where the mixing occurs between $U(1)_D$ and $U(1)_e$, it also becomes clear that there is now a new one loop interaction that provides an upward correction to the muon magnetic moment, as can be seen in Fig. 2.4.

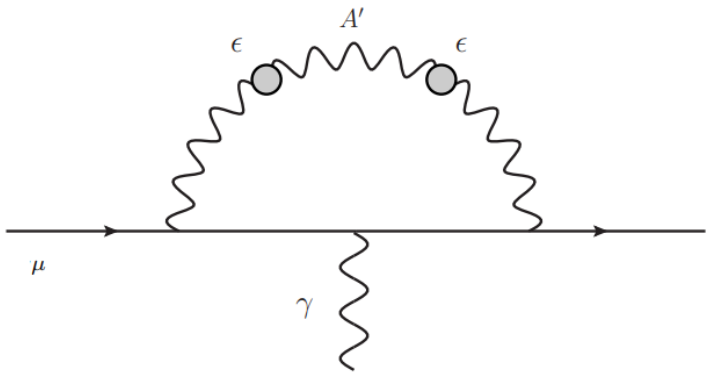

Figure 2.4: One loop correction to the muon magnetic moment from the newly added interaction involving the Dark Photon. Figure taken from Ref. [5].

2.3.2 Sensitivity Studies done on the Dark Photons

As was previously mentioned, the Dark Photons only need 2 effective parameters to project the parameter space that will be probed at SHiP. These are the mixing parameter ϵ and the mass of the Dark Photon m_{γ^D}. The difficulty in estimating the sensitivity at SHiP comes from the production methods available to the DP. If we are to take the simpler model, where the portal to the SM comes from the electromagnetic current J^{EM}_{μ}, the Dark Photons can be produced through all the mechanisms that also produce SM photons. There are three mechanisms of production for such particles at fixed-target experiments [11]: meson decays, proton bremmstrahlung and Drell-Yan like processes in Quantum Chromodynamics (QCD).

In the meson decays, we are looking for decays that involve SM photons, since they kinnetically mix with the DPs, and due to branching ratios this is only relevant for masses below 0.9 GeV/c^2. Bremmstrahlung production is relevant for masses until $\simeq$ 2 GeV/c^2, given that the process is highly supressed when the mass of the DP surpasses that of the proton. Production from QCD mechanisms can only be estimated for masses of 1.4 GeV/c^2 onwards, due to the validity of the parton model. Figures 2.5

[7] While the construction of the kinetic mixing should be done through $U(1)_Y$ and not $U(1)_e$, this model provides a good approximation, since A' must be massive and the mixing with the Z^0 boson is suppressed by $1/m^2_Z$ [70].

and 2.6 show simplified Feynman diagrams of these processes. In Fig. 2.5(a), m stands for the π^0, η and η' mesons, and q can be an up or down quark for all cases, but also a strange one for η and η'. Notice that for the decays involving neutral pions, only the ω meson is considered, since it is the only one with a significant branching fraction.

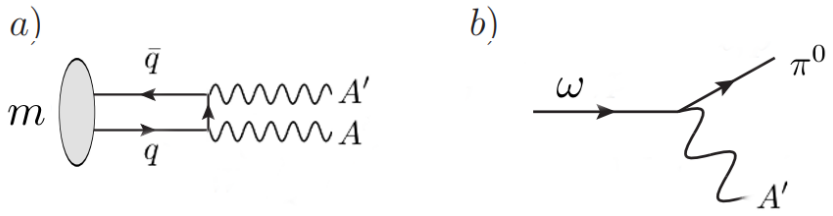

Figure 2.5: Production through meson decays for Dark Photons at SHiP. a) Decays into a pair of DP and SM photon. b) Decays into a DP alongside a neutral pion.

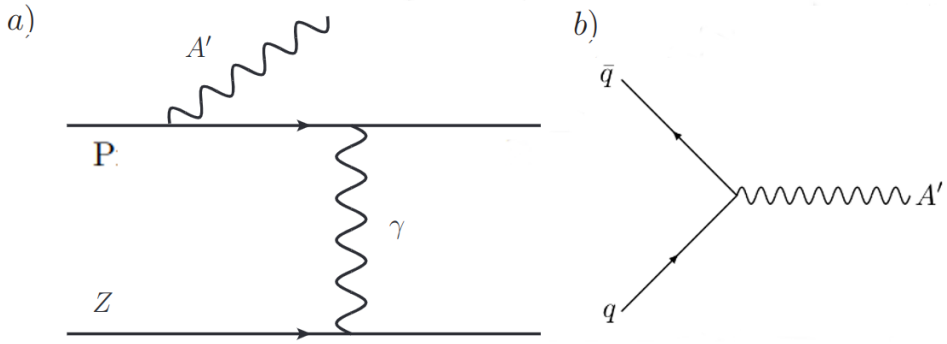

Figure 2.6: Production mechanisms for Dark Photons at SHiP. a) Bremmstrahlung off an incoming proton scattering off protons in a target with atomic number Z. b) QCD Drell-Yan-like production.

The decay modes of the Dark Photons to SM particles are quite straightforward. Since they behave like a massive SM photon, the decays are always to particle/anti-particle pairs of charged fermions, as Fig. 2.7 illustrates, where f can be any electrically charged fermion.

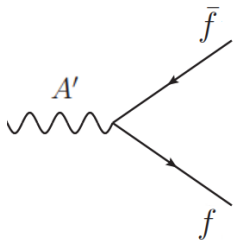

Figure 2.7: Diagram for SM decays of the Dark Photon.

With this, we are able to estimate the parameter space for the Dark Photons. The most recent sensitivity study can be seen in Ref. [11]. The study was done in an analogous way to that of the HNLs, but in this case the number of produced DPs needs to consider all of three production modes. There are two special details about these calculations. The first is the fact that two models were considered for the

bremmstrahlung production, when the mass of the DP surpasses the mass of the proton. The first one considers that for this range of masses, there are no QCD contributions, and the bremmstrahlung process does not depend on the partons. This model was denominated as Dipole, since it used the standard form dipole. The second model tried to take these contributions into account, enhancing the cross-section of the process due to nuclear resonances in the so-called vector meson dominance (VMD) model. In this case, the form factor was altered, leading to enhancements around the ρ and ω meson masses. The second detail is that, cascade production of mesons of target is taken into account, no cascade production is used for the DPs. While Ref. [71] states that this does not affect significantly electromagnetic processes, no studies have been done on the QCD processes yet, and as such the expected sensitivity for high masses may be improved further.

Fig. 2.8 shows the estimated sensitivity of the SHiP experiment to Dark Photons, already with detection efficiencies, and with the caveat that the sensitivity is hampered by hadronic decay modes with long lived neutral hadrons, since the SHiP experiment only considers particles reconstructed from at least 2 charged tracks to be HS candidates.

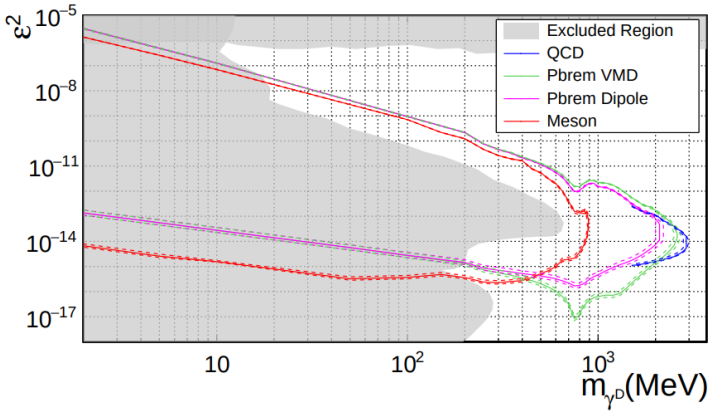

Figure 2.8: Estimated sensitivity curve for the Dark Photons (90% CL). Notice that the two different estimates for production through proton bremsstrahlung are presented. The dashed lines highlight the $1-\sigma$ uncertainty due to systematics. The excluded region in grey is taken from [72]. Figure taken from Ref. [11].

Chapter 3

The Search for Hidden Particles Experiment

3.1 Introduction

The SHiP experiment [6, 73, 74] is planned to be a fixed target multi-purpose experiment, located in a new beam dump at the north area of the CERN SPS accelerator. This experiment focuses on the study of neutrino physics, with more relevance to the tau neutrino (ν_τ), and most importantly, the search for new physics.

The search for new physics will be done in several ways, such as the study of possible lepton flavour violation and indirect detection of dark matter through its scattering with matter, but primarily through the detection of several long lived Beyond the Standard Model (BSM) particles that should decay within 50-120 m from the production vertex. Such BSM particles may have very feeble couplings to the SM and masses below the Fermi scale, which are currently inaccessible to the experiments located at the Large Hadron Collider (LHC). A complementary overall search strategy for BSM will thus be attained.

Due to this, the SHiP experiment will be able to search for hidden particles with effective couplings to the Standard Model (SM) as low as $O(10^{-10})$. The provided beam will be of 400 GeV/c momentum protons, and it is expected to yield a total amount of 2×10^{20} proton on target interactions along the estimated 5 years of operational time of the experiment. In this way, the experiment will explore the intensity frontier of new physics, complementary to the energy frontier studied in the LHC experiments, that search for high-mass short-lived BSM particles.

These hidden particles are postulated in some of the simpler extensions that can be done to the SM by only being able to interact through portals. These portals can be scalars (e.g. ALPs and Dark Higgs Boson), vectors (e.g. Dark Photons) or fermions (e.g. Heavy Neutral Leptons).

The goal of the on-target collisions is to generate charm and beauty hadrons, which decay quickly, either through leptonic decay channels, creating neutrinos, or through the several channels that can produce the desired HS particles. In cases such as the Dark Photon, direct production of the particles upon the scattering of beam onto the target is also allowed.

There will also be a large production of pions and kaons that do not have big enough masses to produce the particles of interest. They can, however, decay into undesired muons and neutrinos that will become the main background sources, especially to the Hidden Sector searches, and as such must be removed as early as possible.

In order to distinguish the feeble HS signal from the background, while exploring as big of a parameter space as is allowed, SHiP will be a 0 background experiment, and as such, the goal is to have beam-induced background reduced to less than 0.1 expected events over the lifetime of the experiment.

With these factors in mind, SHiP will have an experimental setup akin to that shown in figure 3.1, which can be roughly divided into 5 parts: the target; the long-lived hadron absorber alongside the magnetic muon shield intended to deflect the undesired muons produced at the target; the Scattering and Neutrino Detector (SND), that as the name implies, will perform all the neutrino and precision related studies; the Decay Vessel, or Decay Volume, where the Hidden Sector particles are supposed to decay into Standard Model ones; and the Decay Spectrometer, that is in charge of detecting the decay products of the HS particles.

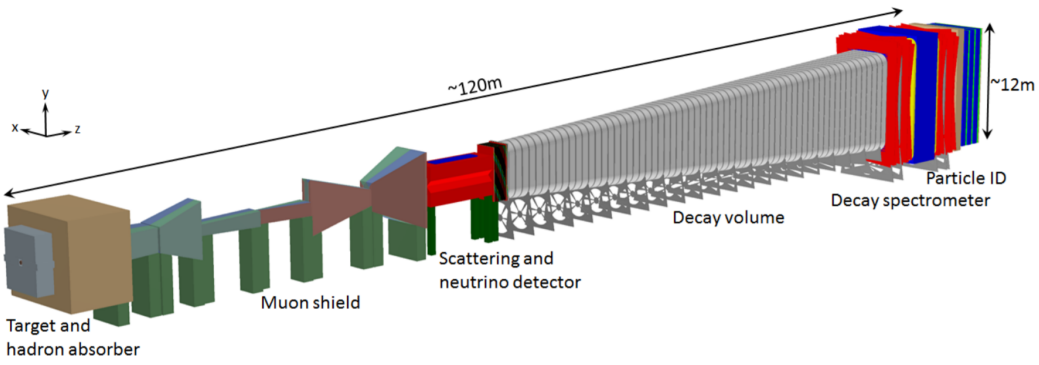

Figure 3.1: Overview of the SHiP experiment taken from Ref. [74].

3.2 SHiP Experimental Setup as of 2019

As of now, there is no definitive version of the experimental setup. However, several features are necessary, and as such, detector studies have already been performed in order to zero in onto the final setup build.

All of the updates done upon the original design [6] of the experiment as of the beginning of the studies presented in this thesis can be found in the progress report from 2019 [74], as well as in the more recent Comprehensive Design Study Report [13].

3.2.1 Target

As was said in section 3.1, the target that doubles as an hadron shield needs to be able to stop as efficiently as possible the produced pions and kaons before they decay, as well as being able to intercept virtually all the protons from the incident beam.

The target consists of segmented blocks of a titanium-zirconium doped molybdenum alloy, followed by blocks of pure tungsten, all interleaved with 5 mm wide slots for water cooling, in a cylindrical shape with a diameter of 250 mm and a length of 150 cm.

This design encompasses 12 interaction lenghts for the protons in the beam.

3.2.2 Hadron and Muon Shields

In order to be able to deal with the muon background, the flux through the detectors must be at least 6 orders of magnitude smaller (10^{-6}) than at the target. An hadron absorber of 5 meters in length follows the target. As of now the absorber is expected to have 4 meters magnetized with a field of ≈ 1.6 T along the beam axis with the help of an integrated coil.

The free standing muon shield is placed immediately afterwards. It has 35 m in length, and is made of grain-oriented steel that is estimated to produce an average magnetic field of 1.7 T. It has been subjected to many shape optimization studies and has ended with the very eye-catching shape that can be seen in Figure 3.1.

3.2.3 Scattering and Neutrino Detector

The SND is the first of the two spectrometers in the SHiP experiment. In this first stage, the aim of the experiment is to detect feeble interactions. While this term encompasses the possible scattering of dark matter with SM particles, and the detection of pentaquarks, the main goal is to study an unprecedented amount of neutrino interactions and, most importantly, detection of both tau neutrinos (ν_τ) and anti-neutrinos ($\overline{\nu}_\tau$).

This goal leads this spectrometer to have a very unorthodox build, since it will use an emulsion tracking apparatus, that follows in the footsteps of the DONUT and OPERA experiments, responsible for the first detection of the tau neutrino [75] and first detection of an oscillation from a muon neutrino to a tau neutrino [76] respectively.

A schematic of the spectrometer can be seen in Figure 3.2, and we can divide it into 4 separate components.

Magnet

The magnet used in this spectrometer has a rather unique configuration, since it accommodates the Emulsion Target, Target Trackers and Downstream Trackers, and has to produce a very low stray field, in order to not interfere with the diverted muons, lending itself to be aligned with the beam axis, with the approximate dimensions of $\approx 1 \times 1.6 \times 6.4$ m^3

It is a warm magnet, with coils made of copper and aluminum and with a 0.6 m thick yolk of iron. This allows for a uniform magnetic field in its volume of 1.25 T, with $\Delta B/B < 1\%$. The coils also have a cooling system in order to be compatible with the required temperature in the emulsion target.

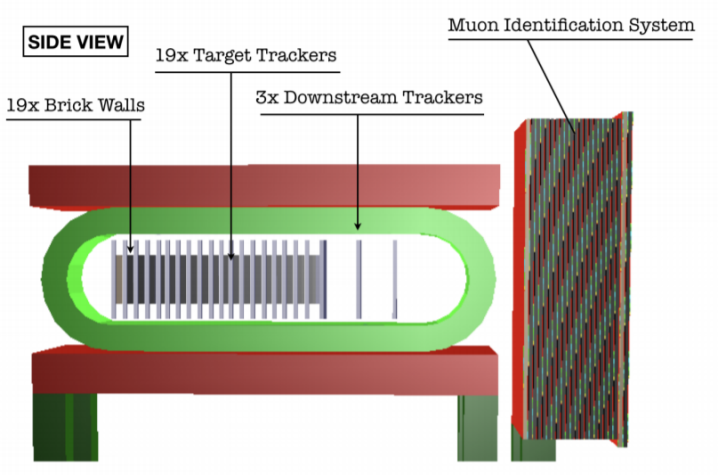

Figure 3.2: Overview of the Scattering and Neutrino Detector taken from Ref. [74].

Emulsion Target

The Emulsion Target is made of 19 bricks, interleaved with 19 Target Trackers (TT), followed by the 3 Downstream Trackers. They have a transverse dimension of 80×80 cm^2 and a length of approximately 3 m, with a functioning temperature of 18° C.

Each brick is divided into 4 cells, distributed 2×2, with transverse dimensions of 40×40 cm^2 each, that contain an Emulsion Cloud Chamber (ECC) and a Compact Emulsion Spectrometer (CES).

The ECC acts as a tracking device with sub-micrometric position and milliradian angular resolution. This allows the disentanglement of the production and decay vertexes of the τ leptons and charmed hadrons. In order to do so, the ECC uses 57 nuclear emulsion films, interleaved with 1 mm thick lead absorber layers, to total an approximate thickness of ≈ 8 cm, which translates to roughly $\approx 10X_0$.

The CES aims to measure the charge and momentum of the hadrons produced in τ lepton decays, allowing for a distinction between tau neutrino and tau anti-neutrino DIS interactions, for the first time. Additionally, it also provides good measurements of low momenta muons. They are made of 3 emulsion films, interleaved by 2 layers of low density material. As of now, the spacers are not defined. However, the use of air gaps introduced by either a typical 175 μm PMMA base, more commonly known as acrylic, or an alternative 500 μm glass base are promising.

Just as with the aforementioned DONUT and OPERA experiments, the Emulsion Target bricks are left with the tracks of the charged particles that go through them. This leads to frequent replacements, and the bricks are analyzed through automated scanning systems after being replaced.

Target Trackers and Downstream Tracker

Due to the nature of the Emulsion Target, there is a need to use Target Trackers that time stamp the reconstructed events in the bricks, and then connect them with the ones seen in the Downstream Trackers and in the Muon Identification System.

There are 19 TT planes with transverse dimensions of 930×1500 mm^2, where the first one, located

before the first brick, functions as a veto system. After all the Emulsion Target bricks, there are 3 final TTs, spaced by 50 cm without any material in between, whose purpose is to measure the momentum and charge of long tracks, such as muons.

These are Scintilating Fibre Trackers (SciFi) [77] developed in LHCb, and have high granularity tracking with a spatial resolution of 50 μm over a surface of ≈ 1 m^2, single plane time resolution of 400 ps, high detection efficiency of $> 99.5\%$, and are insensitive to the magnetic field.

Muon Identification System

The SND Muon Identification System has the primary and very straightforward objective of accurately measuring the muons produced in neutrino interactions, as well as in τ decays. However, it also doubles as a Decay Vessel veto system, or Upstream Background Tagger, charged with the detection of possible neutrino and muon DIS interactions in its upstream portion, that might generate background events in the Hidden Sector part of the experiment.

This leads to having a first section of 8 iron filters, where the first 4 are 15 cm thick, and the last 4 only 10 cm thick. They are interleaved with tracking planes, instrumented with Resistive Plate Chambers (RPC). These RPC planes will be made with 3 gas gaps, with an active area of 1900×1200 **mm**^{2}, and will be arranged in a way such that the total area of the first 8 layers has transverse dimensions of approximately $\approx 200 \times 471$ cm^2.

Due to the necessity of having it double as the Upstream Background Tagger, there is a second part, with two layers that needs to cover the whole Decay Vessel entrance, and they have the slightly bigger transverse dimensions of $\approx 2000 \times 4910$ cm^2.

While the iron planes maintain the 10 cm thickness, the tracking planes are slightly different, and are now Multigap Resistive Plating Chambers (MRPCs) instead, in order to better the timing resolution to about ≈ 300 ps. These MRPCs are currently being developed by our colleagues at LIP-Coimbra.

This model is in Sealed Glass Stack [78] (SGS) MRPCs, where both the gas volume and the High Voltage insulation are confined inside a permanently sealed plastic box. The trackers have 2 gaps, and the RPC structure is defined by three 2 mm thick float glass electrodes of about 2070×1020 mm^2, separated by 1 mm nylon mono-filaments. The electrodes are made of a resistive layer applied to the outer surface of the outermost glasses, and the whole structure is sealed inside a PMMA gas tight box with a 1 mm lid.

Since each MRPC will have an active area of 1.5×1.2 m^2, each of these planes is made by 5 MRPCs staggered by 20 cm.

All of this leads the whole configuration to total about 2 m in length.

3.2.4 Decay Vessel

The Decay Vessel is effectively where the Hidden Sector experiment starts, and is where we are expected to see the HS particles decay. As such we want to minimize the amount of events that might generate particles with a behaviour similar to those of interest, such as K_L^0. This leads to the necessity of having a

complete control of all the charged particles that enter its fiducial volume so that we are able to distinguish DIS events from HS decays.

In order to do this, the Decay Vessel has a Surrounding Background Tagger (SBT), and has a pyramidal shape with a length of approximately 50 m, with upstream outer/inner dimensions of 2.2×5.0 m^2/1.5×4.3 m^2, and downstream dimensions of 5.9×11.9 m^2/5×11 m^2.

However, charged particles entering the volume is not the only background source, and neutrinos can also interact with the material inside the Decay Vessel. To minimize these interactions, it is at a constant pressure of 1 mbar.

Since the Decay Spectrometer has a Straw Tracker, it also needs to be placed inside vacuum. The best way to do this is by extending the Vessel past its pyramidal section just enough to accommodate the Straw Tracker, leading to a total vacuum volume of ≈ 2400 m^3.

The vessel is made of steel and the SBT uses a liquid scintillator made of linear alkylbenzene (LAB) together with 2.0 g/l diphenyl-oxazole (PPO) as the fluorescent. This is enough to keep an inside pressure of 1 mbar.

3.2.5 Decay Spectrometer

The Decay Spectrometer is more orthodox than the SND, and is composed by the aforementioned Straw Tracker, a Timing Detector, the Calorimeters, and at the end a Muon Identification System, as can be seen in Fig. 3.3.

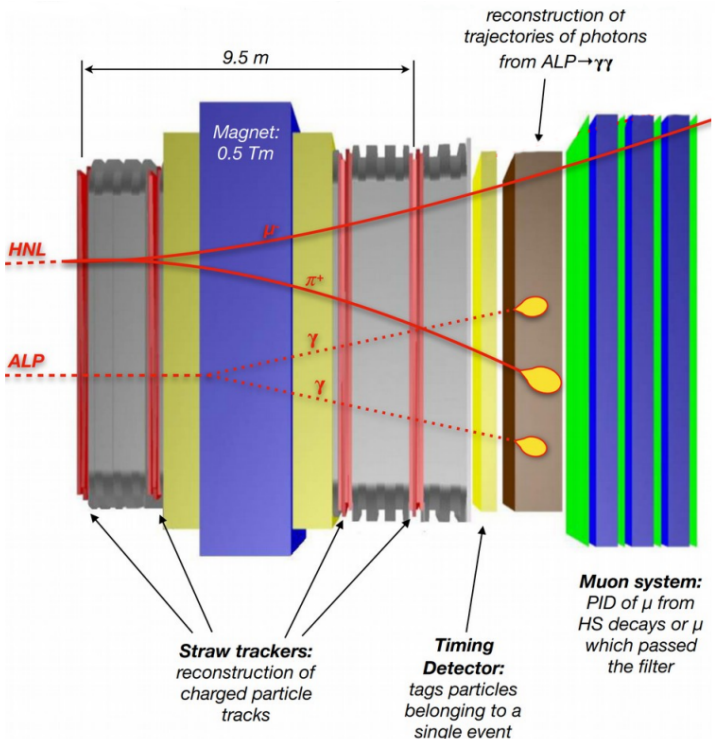

Figure 3.3: Schematic of the Hidden Sector or Decay Spectrometer adapted from Ref. [79].

Spectrometer Straw Tracker

The first part is the Spectrometer Straw Tracker, that consists of 2 tracking telescopes, made of 2 tracking stations each, with a large dipole magnet in between.

This magnet produces a magnetic field in the x direction of 0.14 T at its maximum and about 0.08 T at the location of the closest TTs. total integral field of about 0.5 Tm. On the longitudinal axis the field integral between the second and third station is approximately 0.65 Tm.

The tracking stations are ultra-thin straw drift tubes oriented horizontally, with 4 views in a Y-U-V-Y arrangement. They all have a nominal acceptance of 5 meters in X and 10 meters in Y.

Timing Detector

The Timing Detector will be located right after the Straw Tracker, and just outside the cap of the Decay Vessel. This is still in development, but two options have already been found that provide a time resolution lower than then 100 ps required: the first one based on scintillators with silicon photomultipliers (SiPM), and a second one based on SGS MRPCs [80] analogous to those in 3.2.3. While only a brief mention of the SiPM version will be made, we will explore the MRPCs in more depth, since they are produced by LIP.

The SiPM version is made of EJ-200 plastic scintillator with dimensions $168 \times 6 \times 1$ cm^3, arranged in three columns and 182 rows with 0.5 cm overlap between bars, for a total area of 5×10 m^2, and has demonstrated a resolution of the mean time of around ≈ 85 ps along the whole length of the bar [81].

The SHiP Timing Detector based on MRPC

Multigap Resistive Plating Chambers [82] are, in essence, ionization chambers. When a charged particle goes through these chambers, the gas ionizes. By applying a high voltage electric field, we are able to create Townsend avalanches, and collect a significant amount of charges in the electrodes. Naturally, the time response of these detectors depends on the average drift speed of the ionized particles, and the distance that they have to travel to reach the electrodes. However, just increasing the electrical field in regular RPCs would increase the avalanches to a point where it would start creating delayed streamer signals, which are usually unwanted. MRPCs solve this problem by dividing a large gas chamber into smaller ones, stopping the avalanches before they get too big, allowing for an increase in the electric field, consequent drift velocity and subsequently providing a better timing resolution.

The MRPC option is currently being developed by our colleagues at LIP-Coimbra, and just as with the Upstream Background Tagger it is based on the SGS MRPCs from [78]. As such, it is confined inside of a completely sealed plastic box, and has a multigap RPC structure with six gas gaps defined by seven 1 mm thick float glass electrodes of about 1600×1200 mm^2, separated by 0.3 mm nylon monofilaments. The chamber is composed of a sandwich of two identical sensitive modules with a plane of pick-up electrodes, consisting of 1600×30 mm^2 copper strips, in the middle, as can be seen in Fig 3.4.

A prototype of this Timing Detector, with dimensions 1500×1200 mm^2, was tested with a beam of 8 GeV negative pions, at CERN, in October 2018. The main objective was to test the detection efficiency,

the timing accuracy, and their dependence on the position of the detector. Fig. 3.5 summarizes the results obtained. Due to mechanical restrictions in the setup, the whole prototype could not be scanned. However, there seems to be no significant dependence on the position along the detector, inside its fiducial volume. For an electrical field above 2450 V/gap, the average detection efficiency is of about 98%, and the average timing accuracy is of 54 ps, reaching a minimum of 40 ps in some cases, far surpassing the requirements for the SHiP timing detector, all within the reasonable budgets and production times required by these type of detectors.

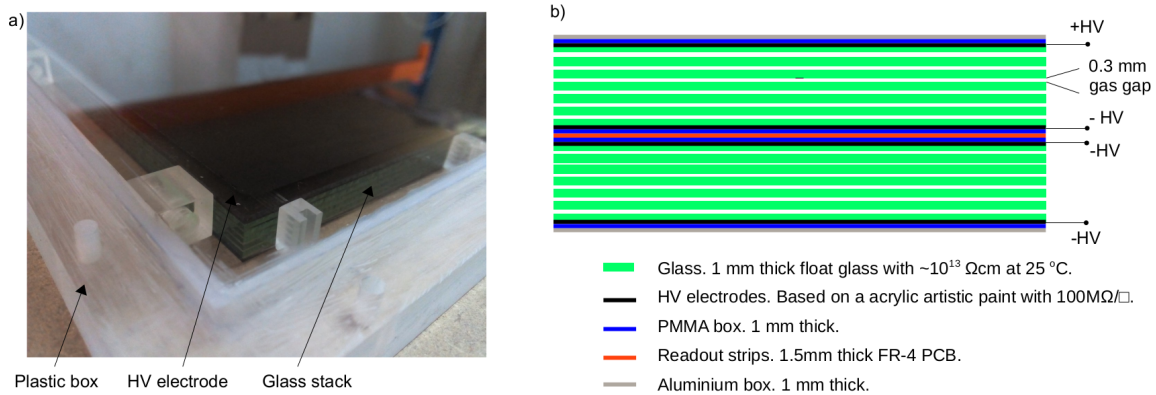

Figure 3.4: Prototype of the MRPC-based Timing Detector. a) Sealed Glass Stack consisting of a plastic tight box, which encloses and confines the gas gaps and the glass and High Voltage electrodes. b) Sketch of the inner structure of the chamber. Figure taken from Ref. [80].

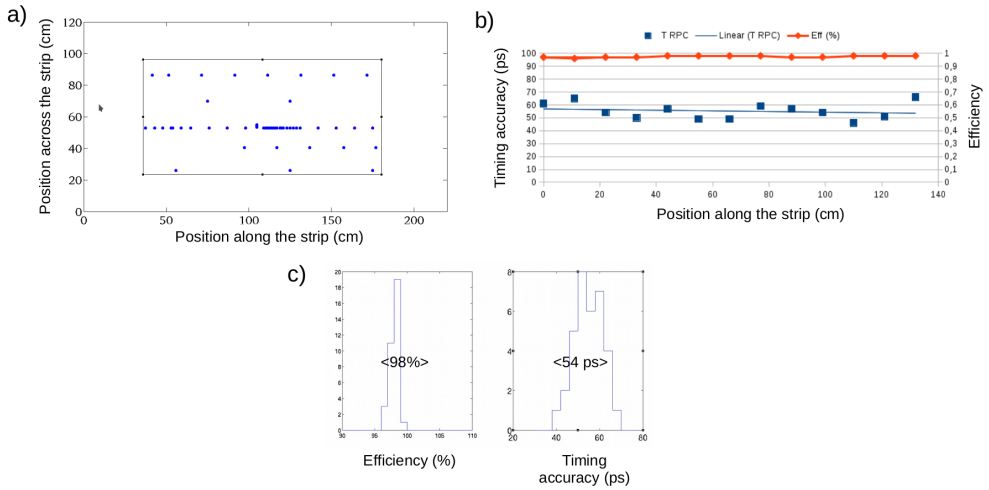

Figure 3.5: Summary of the results obtained during the tests of the MRPC-based Timing Detector prototype performed at CERN. a) Location of the different measured positions. b) Efficiency and timing accuracy along one of the strips of the module. c) Average values of efficiency and timing accuracy in different locations. Figure adapted from Ref. [80].

To cover the whole downstream area of the Decay Vessel, a 7×5 matrix of modules will be arranged. These will overlap the active areas by 10 cm (vertically and horizontally), in order to create a layout without dead regions.

Calorimeters

As of 2019, the SHiP calorimeter section of the HSS has been established to use the so called *SplitCal* configuration, which can measure $ALP \to \gamma\gamma$ decays and functions both as an electromagnetic as well as a hadronic calorimeter, and provides better electron/hadronic separation then the Technical Proposal (TP) one. However, due to the unavailability of a finished implementation of this module in the **FairShip** software as of the time of the study, the *caloDesign 3* configuration will also be described.

The *SplitCal* is a 20 X_0 long lead sampling calorimeter, with lead absorber plates orthogonal to the proton beam direction, and with two kinds of active layers. Most sampling layers have scintillator bars readout.

As for the *caloDesign 3* option, the calorimeters are separated. The first one is a shashlik electromagnetic calorimeter made of 140 layers of 1 mm lead plates, interleaved with 2 mm scintillator planes. This configuration gives the calorimeter a length of 420 cm, which is equivalent to the same 25 X_0 seen in the most recent configuration, with a transverse area of 530×1060 cm^2. The hadronic calorimeter follows, with a length of ≈ 2 m, which totals about 8.5λ. It is also based on shashlik technology made of 66 layers of 1.5 cm steel absorbers, interspersed with scintillators.

Muon Identification System

The last component of the full apparatus is the Muon System. It has four stations of active layers interleaved by the three muon filters. The stations have a width, height and thickness of $600 \times 1200 \times 1$ cm^3 respectively, and will be made of scintillating tiles with a direct SiPM readout, with each tile having dimensions of $10 \times 20 \times 1$ cm^3.

As for the muon filters, they are 60 cm thick iron walls, corresponding to 3.4 λ_l. If we couple this with the amount of material of the calorimeters, which corresponds to 6.7 λ_l, then only muons with an energy of at least 2.6 GeV reach the first station, and only those with 5.3 GeV make it to the last one.

Chapter 4

Samples

This thesis focuses on the study of the Dark Photon and Heavy Neutral Lepton, both of which are Hidden Sector particles, and if they are to be directly detected it will be on the Hidden Sector Spectrometer. Due to its veto systems, the only meaningful background events are the ones produced either in the vicinity of the Decay Vessel, with SM products with similar topologies to the HS ones, or the ones produced inside it by particles that slip through the veto systems. This becomes apparent when we consider that the abridged goal of the HS search is to detect particles that had a production vertex located inside the Decay Vessel, without any sign of being produced by SM particles that entered the volume.

The particles that SHiP will search for are expected to be predominantly accessed through meson decays. SHiP will be a charm and beauty meson factory, with an unprecedented production of charm mesons. The compromise to having a large enough amount of these mesons is the associated high amount of light mesons, muons and SM neutrinos produced.

While the light mesons are innocuous due to the hadronic shield, both the muons and neutrinos cannot be totally deflected, thus creating the main background sources for the studies related to the detection of HS particles in the Hidden Sector Spectrometer.

In the case of muons, performing inelastic scattering interactions with the apparatus can produce long-lived neutral hadrons just before the decay vessel. As for neutrinos, apart from a similar problem to the muons, DIS interactions can also create meaningful background inside the vessel, since unlike the muons we cannot detect them when they enter the volume, and veto the event.

There can also be what is known as random muon combinatorial background, where unrelated muons are detected, and the reconstruction algorithm artificially recreates a vertex inside the volume of interest.

One last possible background source would be that of cosmic muons. However, this has already been dismissed since the Technical Proposal, where it was shown not to be relevant [6].

We will present all the samples generated for the studies in the current SHiP setup, and their reasoning in this chapter. Analysis of the data will follow in the subsequent chapters.

4.1 Sample Generation

The samples employed in this work were generated with the **FairShip** software, based on the **FairRoot** framework [64]. In **FairShip** simulations, the initial proton on target interactions are done with **Pythia 8** [83], and the subsequent propagation and interaction of the produced particles is done with **GEANT4** [84]. Neutrino interactions are simulated with **GENIE** [85], and as for heavy flavour production and inelastic muon interactions a mixture of **Pythia 6** [86] and **GEANT4** are used.

As expected, the HSS response resorts to **GEANT4**. The pattern recognition of the Spectrometer Straw Tracker can be seen in Ref. [87], and the algorithms for particle identification are in Ref. [88].

All events were generated using the configurations described in Chapter 3, with a special nod to the use of the *caloDesign 3* configuration for the calorimeters, which is now outdated, but does not affect our analysis, since it revolves around the kinematic properties of the decay products, and none of the studied cases is impacted by the change. After the generation of the events, they are reconstructed with the algorithms provided in the software.

Even though the samples generated for the background have different selection criteria according to the signal that is being studied, a few preliminary cuts are always applied. The first one is to ensure that the particle was reconstructed inside the Decay Vessel, with a cutoff of the fiducial volume at 5 cm from any walls. The second one was to make sure that the events were only reconstructed from a pair of tracks, since all decays studied only involve two particles in the final state. The only exception to this rule are decays that involve a third neutral pion, which was not automatically reconstructed by the algorithms at the time of the simulations, and as such is seen by these cuts as having 2 daughter particles.

Background Type	Events Generated	Particles Reconstructed	Reconstructions within the Decay Vessel	Events Reconstructed from 2 Tracks
ν_e DIS	1.0×10^6	16804	7741	5211
$\overline{\nu_e}$ DIS	1.0×10^6	11653	5959	3918
ν_μ DIS	1.0×10^6	4913	1807	1340
$\overline{\nu_\mu}$ DIS	1.0×10^6	2986	1046	788
μ DIS	1.0325×10^6	10511	3015	2144
μ Comb.	2.18×10^9	4	3	2

Table 4.1: Summary of the background samples generated with the **FairShip** software. Unless stated in otherwise the reconstruction takes into account a combination of at least 2 tracks. Notice that μ Comb. stands for the muon combinatorial background.

4.1.1 Neutrino DIS

Neutrino DIS events should be the hardest ones to differentiate from the desirable signal, since they cannot be discarded as effectively by the physical veto systems, and as such they were given the most attention throughout this thesis.

While an unparalleled amount of both τ flavoured neutrinos and anti-neutrinos will be produced at SHiP, naturally the amount is much inferior to that of electron and muon flavoured ones. Thus, the

neutrino DIS background studies only considered the latter flavours.

All neutrino simulations done with the **FairShip** software require the input of a file containing all the neutrino events previously generated in the target by the proton beam. The original neutrino files used in the simulations can be found at the */eos/experiment/ship/data/GenieEvents*. For each flavour of neutrinos and anti-neutrinos considered, 1×10^6 DIS events were forced to interact from the beginning of the Muon Identification System in the SND, until the second Target Tracker of the Spectrometer Straw Tracker.

A summary of the preliminary neutrino DIS events can be found in Table 4.1, alongside the rest of the background.

4.1.2 Muon DIS

From the preliminary studies presented since the Technical Proposal [6], the Muon DIS background should not pose as big of a problem as the neutrino one. However, we can never be too complete in our studies, and as such this background was included as well.

In a similar vein to the neutrino background, muon DIS simulations need muons produced at the target as an input, and the input files can be found at */eos/experiment/ship/data/muonDIS/*. Three different files were used this time, which were named under *muonDIS_1.root*, *muonDIS_3.root* and *muonDIS.root*. While the first 2 files were used to generate 2.575×10^5 events from each, the latter was used in order to generate 5.175×10^5 events, leading to a total of 1.0325×10^6 events.

Once again, the muons were forced to interact between the beginning of the Muon ID System, in the SND, and the second Target Tracker of the Spectrometer Straw Tracker.

The preliminary muon DIS events, alongside the rest of the background, can be found in Table 4.1.

4.1.3 Muon Combinatorial Background

The last relevant background source is the Muon Combinatorial Background. In this case, all of the files used as inputs for the simulations can be found in */eos/experiment/ship/data/Mbias/background-prod-2018/*. The events came from 67 different files, which were named under *pythia8_Geant4_10.0_withCharmandBeautyXXX_mu.root*, where *XXX* numerates the files.

From the file *pythia8_Geant4_10.0_withCharmandBeauty0_mu.root*, 13.6×10^6 events were generated, while for the remainder 66 ones, 6.4×10^6 events were generated per file.

This process was repeated 4 additional times, always using the command *–phiRandom*, allowing for a total of 2.18×10^9 events generated. However, despite the much bigger amount of events available, there were only 4 reconstructed particles, from which 3 were inside the fiducial volume of the Decay Vessel, rendering any possible conclusions from a study done regarding this topic not statistically significant. As such, this background was discarded and will not be considered any further in this thesis.

Since the production of the samples used in this thesis, the study on this background type has advanced, mostly regarding a more efficient production of meaningful events by using Generative Adversarial Networks (GANs) [89], which should make this type of study possible. Nevertheless, just like for

the previous background types, a summarized version of this sample can be seen in Table 4.1.

4.1.4 Heavy Neutral Leptons

In this thesis several points of the parameter space for the HNLs were considered, as well as several relevant decay modes. Here we shall explain the thought process behind the choices, as well as present preliminary data on the samples generated.

It was shown in Fig. 2.3 that SHiP will have its best sensitivity to HNLs around a mass of 1 GeV/c^2. As stated in studies relating to the νMSM [43], the masses preferred in order to create a Dark Matter candidate while also providing a viable mechanism to explain BAU are of the order of $\mathcal{O}(1)$ GeV/c^2. As such, the masses of the HNL samples should stay around the aforementioned mark. Accordingly samples were generated for masses that range from 0.7 to 1.4 GeV/c^2, in 0.1 GeV/c^2 intervals.

For masses smaller than $m_D \approx 1.9$ GeV/c^2, the biggest source of production of HNLs is the decay of charmed mesons. This lead to the decision of producing HNL samples solely through the decay of open-charm mesons, foregoing production from open-beauty mesons. Fig. 4.1 provides better insight into this decision. It shows the product of the branching ratio of hadrons h into HNLs by the hadron fragmentation fraction, with h being charmed (left) and beauty (right) mesons. This plot took into consideration interactions of protons on a molybdenium target, and HNL mixing angles with the SM neutrinos $U_e = 1$ and $U_\mu = U_\tau = 0$.

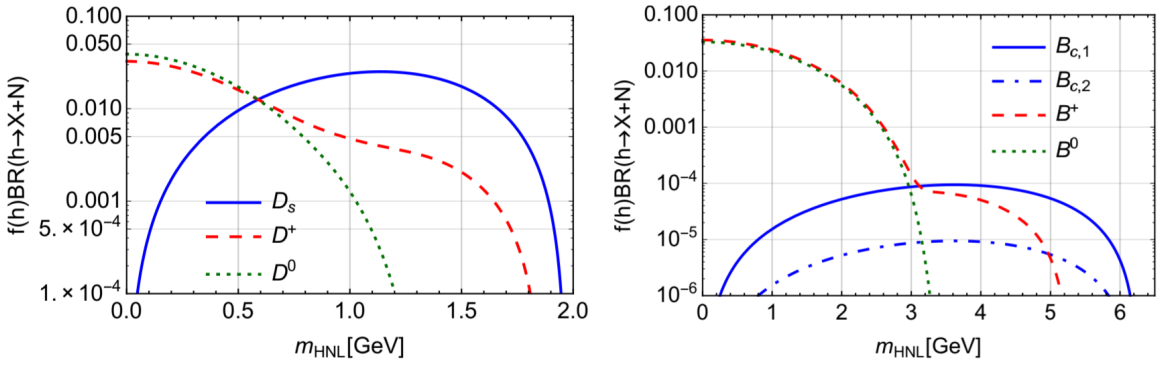

Figure 4.1: Fraction of mesons h that decay into HNLs multiplied by their branching ratios to HNLs, for charm (left) and beauty (right) mesons. The HNL mixing angles were set as $U_e = 1$ and $U_\mu = U_\tau = 0$, and two estimates for B_c production are shown, with $f(b \to B_c) = 2 \times 10^{-3}$ ($B_{c,1}$, line) and $f(b \to B_c) = 2 \times 10^{-4}$ ($B_{c,2}$, dash-dotted line). Figure taken from Ref. [10].

Just like in the most recent sensitivity study [10], heavy flavour cascade production originated by the initial proton collisions [90] was taken into account, since it enhances the production rate of HNLs, and is a default feature of the **FairShip** software. The file used in order to provide the initial charmed particles was the default one in the **FairShip** software.

In this thesis, six different decay modes were considered. The first criterion to select the relevant ones was very basic, and only required that the decay modes had to have at least two products with opposite electrical charges. The second criterion was to consider the modes with the biggest branching ratios

possible for the range of masses studied. The third criterion was that there would need to be at least one decay mode with missing energy in the final state. This means that not all decay products can be detected, which is equivalent to saying that at least one decay mode had to involve a SM neutrino. Therefore, the decay modes chosen were: $N \to e^-\pi^+$, $N \to \mu^-\pi^+$, $N \to e^-\rho^+$, $N \to \mu^-\rho^+$, $N \to \nu_\alpha\mu^+\mu^-$ and $N \to \nu_\alpha\rho^0$. Notice that conjugate charge channels are also included. Decays involving ρ mesons were only generated for HNLs with masses bigger than 0.9 GeV/c^2.

While the fully reconstructible decays all fit the criteria set in a very straightforward fashion, the missing energy ones do not necessarily have the highest branching ratios[1], and as such we will elaborate further on our choices.

From the two decays with missing energy, the decay that involves ρ^0 is the simplest one, since it does have the maximum branching ratio of the semi-leptonic decay channels that involve SM neutrinos. As for the decay into a muon pair, while it does not necessarily have the biggest branching ratio, it is still very close to it, and can lead to more interesting studies that compare the kinematic properties of similar decay products generated from different types of signal from the Hidden Sector (e.g. $DP \to \mu^+\mu^-$).

The last thing to take into account is the couplings $|U_\alpha|^2$ provided to the samples. Since optimal detection is of bigger importance when the predicted amount of particles detected over the course of the experiment is minimal, the couplings chosen were done near the limits established in Fig. 2.3, taking into account the three benchmark models presented in Section 2.2.3. Since the sensitivity limits to models I and II were so similar, the same maximum couplings were applied for simulations with similar masses. Fig. 4.2 shows a simplified version of the selected parameter space points, and the detailed maximum couplings, defined for each model as a function of the HNL mass, are highlighted in Table 4.2.

For every point defined in the parameter space, 5000 events were generated, to ensure that the samples a much larger number of events than the expected ≈ 2.4 detected particles from the HS, established as the limit criterion in the projected parameter space in Ref. [10].

We present in Table 4.3 a summary of the preliminary cuts of the HNL events by decay mode.

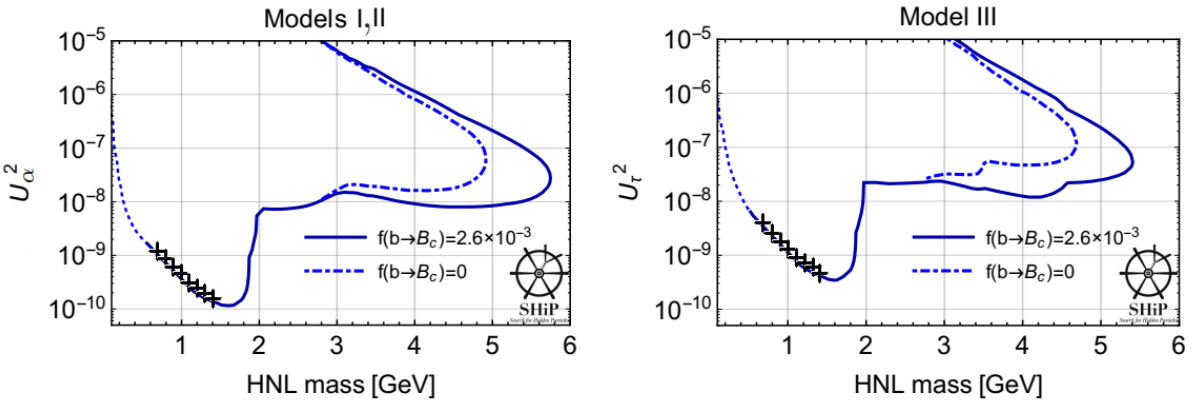

Figure 4.2: Approximate sensitivity curves for the 3 benchmark models I-III, with the chosen couplings for each benchmark model represented. Since models I and II share the same couplings, they were condensed in one graphic. In the left $\alpha = e, \mu$ for models I and II, respectively. Figure adapted from Ref. [10].

[1]For further information about the branching ratios see the appendix A of Ref. [10].

Mass [GeV/c^2]	Maximum Coupling $\|U_\alpha\|^2$	
	Models I, II	Model III
0.7*	1.2×10^{-9}	4×10^{-9}
0.8*	9×10^{-10}	2.4×10^{-9}
0.9	6×10^{-10}	1.8×10^{-9}
1.0	5×10^{-10}	1.2×10^{-9}
1.1	3×10^{-10}	9×10^{-10}
1.2	2.5×10^{-10}	7×10^{-10}
1.3	2.0×10^{-10}	6×10^{-10}
1.4	1.8×10^{-10}	4×10^{-10}

Table 4.2: Maximum couplings selected for each HNL mass and benchmark model. *No simulations involving decays with ρ particles were done for masses lower than 0.9 GeV/c^2.

Decay Mode	Events Generated	Particles Reconstructed	Reconstructions within the Decay Vessel	Events Reconstructed from 2 Tracks
$N \to e^-\pi^+$	1.2×10^5	26307	25116	25058
$N \to \mu^-\pi^+$	1.2×10^5	28338	26867	26808
$N \to e^-\rho^+$	9.0×10^4	23226	21923	21193
$N \to \mu^-\rho^+$	9.0×10^4	25994	24235	23458
$N \to \nu_\alpha\mu^+\mu^-$	1.2×10^5	30800	29001	29001
$N \to \nu_\alpha\rho^0$	9.0×10^4	22781	21373	21312

Table 4.3: Summary of the HNL samples generated with the **FairShip** software. Decay modes involving ρ particles were only simulated for HNLs with masses of 0.9 GeV/c^2 or higher. Notice that the charge conjugate channels are also included.

4.1.5 Dark Photons

At SHiP, the Dark Photons have a very different sensitivity curve to that of the HNLs. As we have seen, the probing capability of SHiP for DPs has a cigar shape, and both the upper and lower bounds, coupling-wise, are of interest.

Since effective theories with Dark Photons only have two free parameters, as opposed to the four in HNL extensions, doing a full sweep of the parameter space that will be probed is simpler. Unlike the HNLs, however, the amount of events generated was not evenly spaced, and the couplings did not always follow the edge of the parameter space.

One of the big implications of a new vector-like particle that kinetically mixes with the SM currents is the shift in the theoretical magnetic dipole of the muon. Consequently, the muon g-2 anomaly becomes a constraint. It has been shown that solutions to this problem that involve DPs are heavily disfavoured for masses $m_{\gamma_D} \geq 400$ MeV/c^2 [91, 92]. This lead to selecting masses under 400 MeV/c^2 within small intervals, alongside the sensitivity bounds shown in Fig. 2.8. As for bigger masses, intervals of ≈ 1 GeV/c^2 were chosen in order to complete the sweep, up until 4.4 GeV/c^2 [2].

While Fig. 4.3 visually shows the parameter space points selected for the simulations, Table 4.4

[2] At the time of the sample generation the most recent sensitivity study was not available yet, reducing the probing range to $\mathcal{O}(3)$GeV/c^2.

includes the detailed masses used when generating the samples alongside the couplings provided. For some masses two couplings were provided, since there is both a lower and upper bound based on the minimum 2.4 expected events. Due to the reduced physical relevance for masses that surpass 1 GeV, and also due to the naturally shrinking range of couplings for such masses, only one "central" coupling was provided for 3.6 and 4.4 GeV/c^2.

The DPs are produced through different methods, for different masses, as was referenced in Section 2.3. While all the DPs simulated with masses below 0.5 GeV/c^2 were produced through meson decays, for the higher masses direct QCD production was used[3].

At low enough masses of $m_{\gamma_D} \leq 2m_\mu$, the only available decay mode into SM particles is $A' \to e^- e^+$ and, as such, this decay channel is a must study. As for masses in the range $2m_\mu \leq m_{\gamma_D} \leq 2m_\pi \approx 400$ MeV/c^2, which encompasses the remaining area of very high interest, the only other SM decay channel that opens up is $A' \to \mu^- \mu^+$, making it into a channel of interest in our study. These two fully leptonic decay modes are the most relevant ones for low masses, and have very simple and distinguishable signatures, and as such were the ones considered in this thesis. They were used in the whole range of masses chosen since they are always relevant[4].

Just as with the HNLs, 5000 events were generated for each parameter space point and decay channel selected. A summary of the DP samples generated can be found in Table 4.5.

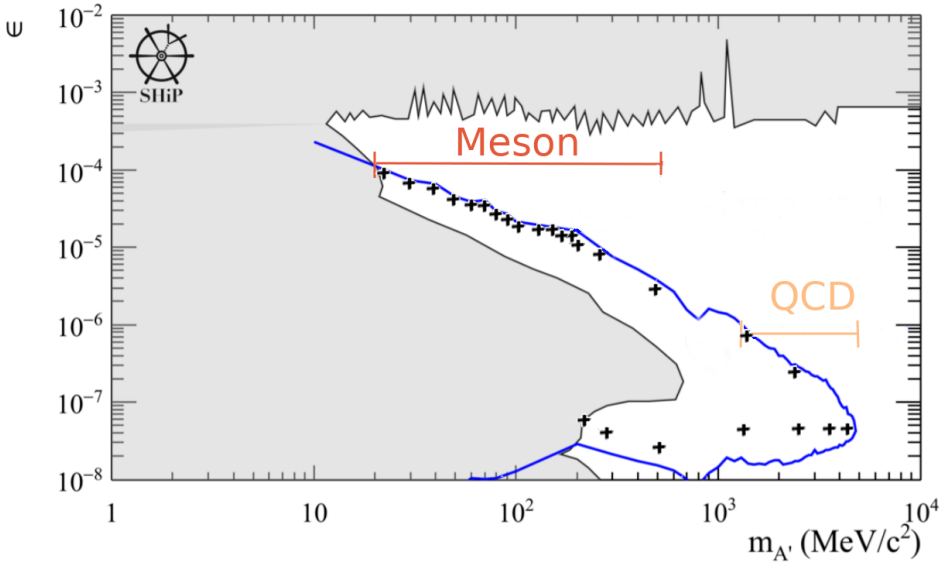

Figure 4.3: Excluded region (gray) and estimated sensitivity curve for the Dark Photons, with the couplings for each mass represented. An updated sensitivity curve is shown in Fig. 2.8. Figure adapted from Ref. [73].

[3] For further details on the relevance of each production mode see Ref. [11].

[4] For a more in-depth breakdown of the branching ratios, once again, see the most recent sensitivity study done on DPs at SHiP [11].

Mass [MeV/c^2]	Effective Coupling ε	Mass [MeV/c^2]	Effective Coupling ε	Mass [MeV/c^2]	Effective Coupling ε
21	1.0×10^{-4}	120	2.0×10^{-5}	500*	4.0×10^{-8}
30	7×10^{-5}	140	2.0×10^{-5}	1400*	1.0×10^{-6}
40	6×10^{-5}	160	1.4×10^{-5}		5×10^{-8}
50	4.4×10^{-5}	180	1.4×10^{-5}	2500*	3.0×10^{-7}
60	4.0×10^{-5}	200	1.0×10^{-5}		5×10^{-8}
70	4.0×10^{-5}		8×10^{-8}	3600*	5×10^{-8}
80	3.0×10^{-5}	250*	1.0×10^{-5}	4400*	5×10^{-8}
90	2.4×10^{-5}		5×10^{-8}	-	-
100	2.0×10^{-5}	500*	3.0×10^{-6}	-	-

Table 4.4: Detailed masses chosen for the Dark Photon simulations, and respective couplings selected for each. Notice that for a mass of 500 MeV/c^2, there are two different couplings. *The channel $A' \to \mu^+\mu^-$ is only open for simulations with masses of 250 MeV/c^2 or higher.

Decay Mode	Events Generated	Particles Reconstructed	Reconstructions within the Decay Vessel	Events Reconstructed from 2 Tracks
$A' \to e^-e^+$	1.25×10^5	34810	33324	33277
$A' \to \mu^-\mu^+$	5.0×10^4	17667	16954	16933

Table 4.5: Summary of the DP samples generated with the **FairShip** software. Unless stated in contrary, the reconstructions take into account a combination of at least 2 tracks. The decay to a pair of muons was only simulated for DPs with masses of 250 MeV/c^2 or higher.

Chapter 5

Heavy Neutral Leptons

As has been mentioned ever since the introduction to this thesis, the Search for Hidden Particles experiment has as a primary objective to find particles from the Hidden Sector, based on as few as 2 or 3 events detected over the whole operational time. This lead to the 0 background condition, which translates to having an expected 0.1 background events detected throughout the 5 years of running in the Hidden Sector Spectrometer.

Even though preliminary studies were done on the background and its sources, as can be seen as far back as the Technical Proposal [6], that provide a standard set of cuts able to reduce the background to the accepted levels, taking into account the use of physical veto systems, it was done for the preliminary configuration of the SHiP apparatus, which as since been heavily altered. More recent studies have been done, and a summary can be seen in the latest Comprehensive Design Study Report [13]. However this latter study also takes into account the use of physical veto systems, and as such runs parallel to our own.

Unlike for the first study, we do not have access to precise numbers on the amount of background events expected throughout the whole experiment, and as such we will not be able to define cuts based on a clear cut-off of 0.1 events expected over the whole experiment. This lead to the decision of making unorthodox cuts that remove every single background event, regardless of their significance. With this said, we shall proceed to explain the cuts applied, as well as their reasoning, and the results obtained.

All of the decay modes selected for the HNLs effectively have as final products combinations of pions and light leptons, indeed the ρ particles decay fairly quickly into pairs of pions, be it $\rho^{\pm} \to \pi^{\pm}\pi^{0}$ or $\rho^{0} \to \pi^{-}\pi^{+}$. The first selection past the preliminary cuts presented in Chapter 4 is to only keep the events that decay into the following combinations:

$$e^{\mp}\pi^{\pm},\ \mu^{\mp}\pi^{\pm},\ \pi^{\mp}\pi^{\pm},\ e^{\mp}\mu^{\pm} \text{ or } \mu^{\mp}\mu^{\pm}, \tag{5.1}$$

where the last two combinations of products take into account misidentifications of charged pions.

Recall that the reconstruction algorithm in **FairShip** at the time of the analysis did not reconstruct neutral pions automatically from photons. As such, after every single reconstructed event that fit all the previous selection criteria, a search for pairs of photons that can come from neutral pions was done.

In an ideal world, if the SHiP apparatus were to detect two charged tracks, we would like to be able to state whether it is an event from the Hidden Sector or just background immediately, from their kinematic properties alone. Practically, this is near impossible. However, we can try to do the next best thing, which is to distinguish signal from background by just looking at the kinematic properties of the two tracks that can be decay products of HS particles. This lead to the decision of comparing every sample of HNLs against the same background, which contained all the possible combinations[1].

After the particles were reconstructed, and the neutral pions were found, in case there were any, the kinematic properties shown in Table 5.1 were taken into account, for the reconstructed HS candidate particle and for the particles that interacted in the detectors, when applicable. From here on out, we shall refer to the reconstructed HS candidate particle as Mother and the particles that interacted in the detectors as Tracks, interchangeably. The quantities related to the Total Momentum of the HS candidate took into consideration the addition of the neutral pion, if one was found. Notice that the Impact Parameter (IP) is calculated by propagating the reconstructed HS candidate particle to the point where it is expected to have been created, at the SHiP target, and measuring the closest distance.

Feature	Symbol
Total Momentum	TotMom
Transverse Momentum	TransMom
Fraction of Transverse Momentum	FracMom
Opening Angle between the charged tracks	OA
Impact Parameter	IP
Coordinates of the Decay Vertex	Decay X, Y and Z

Table 5.1: Kinematic features extracted from the reconstructed particles, and respective symbols that will be used throughout this thesis. When considering the coordinates of the decay vertex Z relates to the initial proton beam axis, while X and Y are the transverse coordinates.

Figure 5.1 and 5.2 show the distributions of the aforementioned properties for the combined HNL samples (blue), overlapped with the the distributions for the background (orange), normalized as densities, where the integral over the whole distribution is approximately $\simeq 1$.

5.1 Kinematic Cuts

As can be seen from Figs. 5.1 and 5.2, the biggest distinction between distributions is in the Impact Parameter, where the peaks of the distributions are clearly separated. This is expected, since we are contrasting the HS particles that propagate without interacting throughout the apparatus, with the background created from violent interactions, which is very likely to have missing energy.

Three other distributions that also have visible discrepancies are the Decay Position of the particle along the Decay Vessel, the Total Momentum of the HS candidate and the Total Momentum of the Tracks. Despite the similar behaviour, there are different reasons as to why this happens for the decay coordinate, then for the latter two.

[1] This choice also allows for a cross-evaluation by Neural Networks, which we will explain in Chapter 7.

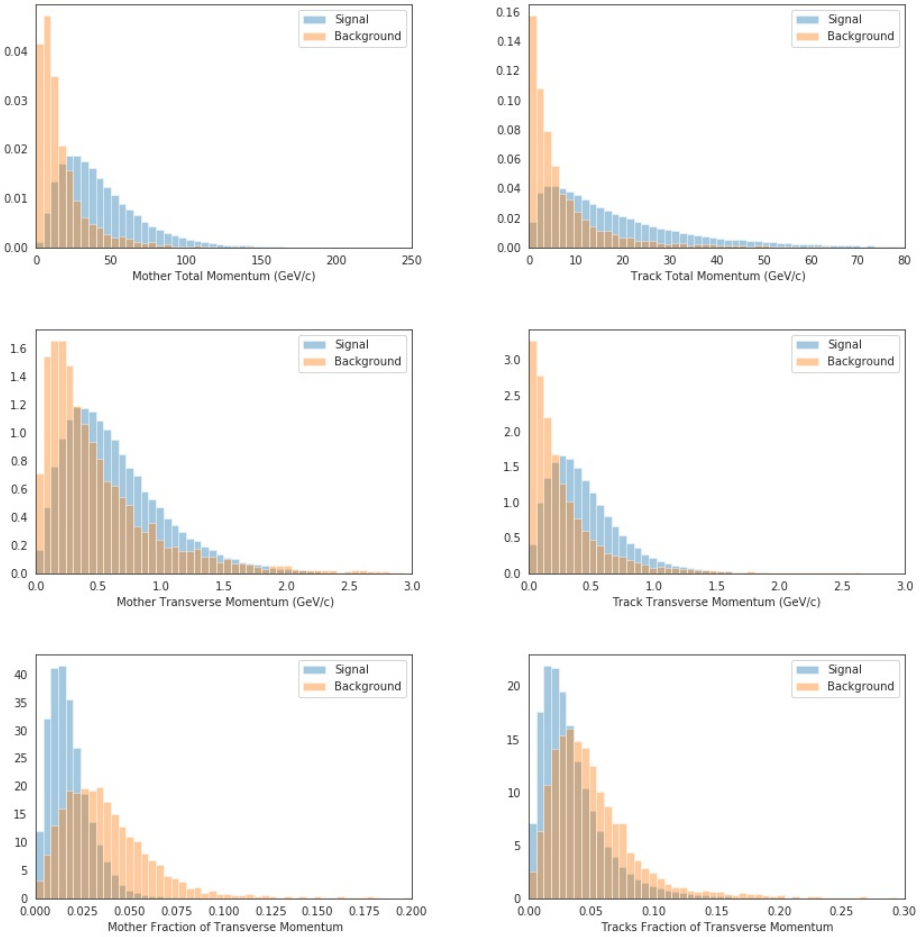

Figure 5.1: First set of distributions containing the momentum related properties of the HS candidate (Mother), as well as of the Tracks.

The decay coordinate along the vessel (Decay Z) is easily explained by the limits we imposed when generating the neutrino and muon interactions. Since the vessel is in vacuum, most of the interactions of the neutrinos and muons will occur with the walls of the Vessel, or in the Muon Detector of the SND, leading to the decreasing shape of that distribution, while the HNLs have a low chance of decaying rapidly in the beginning of the vessel due to their long expected mean life-time.

As for the Total Momentum features, the background shows a big amount of events with very low momenta, when compared to the HNLs, which is also to be expected, since the HNLs have a much bigger mass than SM neutrinos. This leads to a smaller minimum momentum carried by the neutrinos since they are produced through the same meson decays, but in the three-body decays the minimum fraction of momentum of each product is equal to its mass.

No figures of merit were established in order to determine the cuts that removed all the background in the most efficient way. Instead, some preliminary, and fairly loose, cuts were done involving these three variables. Afterwards all the combinations of two variables available were compared in order to find cuts that could remove every single background event. This leads to some signal selection inefficiencies. As such, after all events were removed, every single cut was individually removed, and new cut thresholds

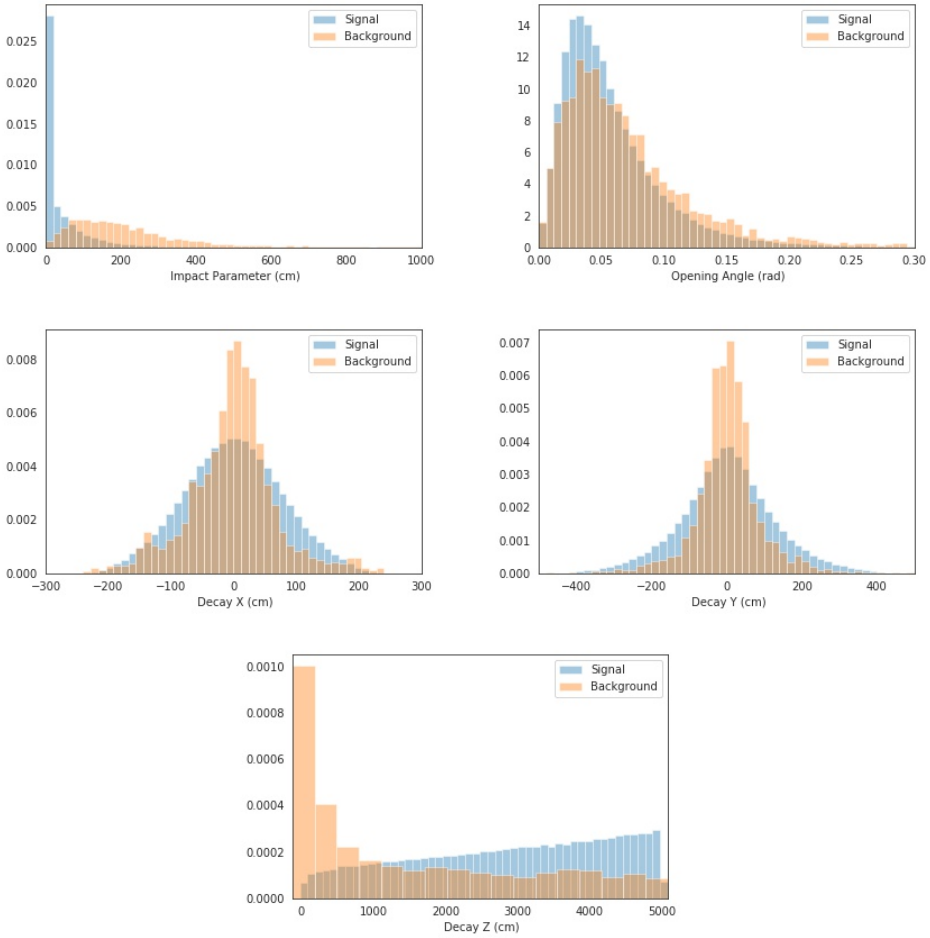

Figure 5.2: Second set of distributions containing the remaining features listed in Table 5.1. While for the decay positions in Y and X the origin represents the incident beam trajectory, for Z the origin represents the beginning of the Decay Vessel.

were sought for the cut, such that the selection efficiency of the HNLs was as high as possible, while maintaining the 0 background events selected. In this case, 7 different cuts were necessary in order to reach the background condition. These cuts are described in Table 5.2.

The selection efficiencies obtained from applying the cuts presented in Table 5.2, are shown in Table 5.3, and the corresponding visual representations can be seen from Figs 5.3 to 5.9. Since the sensitivity estimates presented in Fig 2.3 already take into account geometric efficiencies, the percentages shown are only as a reference against the HS candidates reconstructed within the Decay Vessel from Table 4.3.

A final summary of the selection efficiencies by decay mode is shown in Table 5.4. From this breakdown two things stand out: while the simpler decays $N \to \ell^{\mp}\pi^{\pm}$ have very acceptable selection efficiencies, the decays with missing energy in the form of SM neutrinos have low ones. This very high contrast stems from the fact that most of the cuts were based around the Impact Parameter, which does seem to be a very good distinguishing feature[2].

[2]In Chapter 7, we will use the **SHAP** software in order to try and extrapolate the weights that the neural networks attribute to each kinematic variable provided, and verify that indeed the Impact Parameter is highly regarded even for the machine learning algorithms.

Cut	Thresholds for Selection
1	IP < 33.0 cm.
2	TotMom of the Tracks > 1.0 GeV/c.
3	IP > 6.0 or > 10.0 cm requires TotMom of the Mother < 58.0 or 38.0 GeV/c, respectively.
4	IP > 12.0 or 16.0 cm requires TransMom of the Track < 0.25 GeV/c or > 1.45 GeV/c, respectively.
5	IP < 12.0 cm requires IP $< 412 \times (0.05 - \text{TransMom})^2 + 2.0$ cm, where TransMom is from the Tracks, in GeV/c.
6	IP $< 47043 \times (0.0005 - \text{FracMom})^2 + 2.0$ cm, where FracMom is from the Mother in GeV/c.
7	Decay Z > 35.3 cm.

Table 5.2: Kinematic cuts applied to the HNL samples, and their respective background.

The drawback to this approach is that most of its strength lies on the fact that neutrino and muon DIS events are usually reconstructed with either missing energy or off the beam axis, since the Impact Parameter is calculated by propagating the reconstructed particle to the point where it is expected to have been created, at the SHiP target, and measuring the closest distance. This leads to very low selection efficiencies for decays that cannot be fully reconstructed due to their undetectable decay products.

A good testament to this are the decays involving charged ρ mesons, $N \to \ell^{\mp}\rho^{\pm}$, where sometimes the neutral pion stemming from the almost instantaneous $\rho^{\pm} \to \pi^{\pm}\pi^0$ decay cannot be reconstructed[3], leading to missing energy, and as such a selection efficiency that can never be as good as for the simpler decays, but that is better than for the cases where there is always missing energy.

Machine learning tools should be especially effective at separating signal from background in these types of decays, since they do not necessarily score kinematic variables as step functions (unlike the kinematic cuts presented previously), and as such the Impact Parameter should not be as big of an impairment.

Cut	Background Events Selected (%)	HNL Events Selected (%)
Decay products in Eq. (5.1)	4828 (24.7)	143742 (96.8)
1	171 (0.9)	90574 (61.0)
2	137 (0.7)	89456 (60.2)
3	81 (0.4)	76219 (51.3)
4	26 (0.1)	73234 (49.3)
5	18 (0.1)	69898 (47.1)
6	5 (0.0)	64796 (43.6)
7	0 (0.0)	64633 (43.5)

Table 5.3: Selection efficiency of the kinematic cuts presented in this section. The numbered cuts refer to the ones presented on Table 5.2. All percentages are calculated against the total number of HS candidate particles reconstructed within the Decay Vessel, presented in Table 4.3, which gives a total of 19568 background events, and 148515 HNL events.

[3] From the 21923 $N \to e^{\mp}\rho^{\pm}$ events with decay vertices inside the Decay Vessel, only 10767 have a detected π^0, and for $N \to \mu^{\mp}\rho^{\pm}$ it is only 12287 out of 24235.

Decay Mode	Mother reconstructed within the Decay Vessel	Events with the desired Decay Products (%)	Surviving Events after the Cuts (%)
$N \to e^- \pi^+$	25116	24053 (95.8)	20417 (81.3)
$N \to \mu^- \pi^+$	26867	26732 (99.5)	24524 (91.3)
$N \to e^- \rho^+$	21923	19798 (93.4)	7729 (36.5)
$N \to \mu^- \rho^+$	24235	22991 (94.9)	9494 (39.2)
$N \to \nu_\alpha \mu^+ \mu^-$	29001	28983 (99.9)	1054 (3.6)
$N \to \nu_\alpha \rho^0$	21373	21185 (99.1)	1415 (6.6)

Table 5.4: Summary of the selection efficiency obtained by applying the kinematic cuts, detailed by Decay Mode. Efficiencies are based on the HS candidate particles reconstructed within the Decay Vessel, presented in Table 4.3, which are reiterated here.

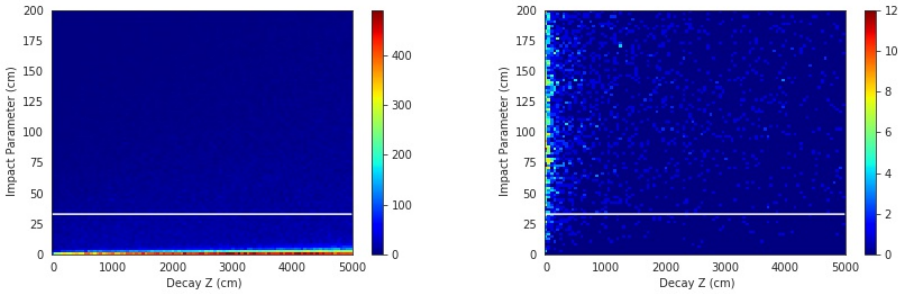

Figure 5.3: First cut applied to the samples. The image on the left corresponds to the signal, while the image on the right corresponds to the background.

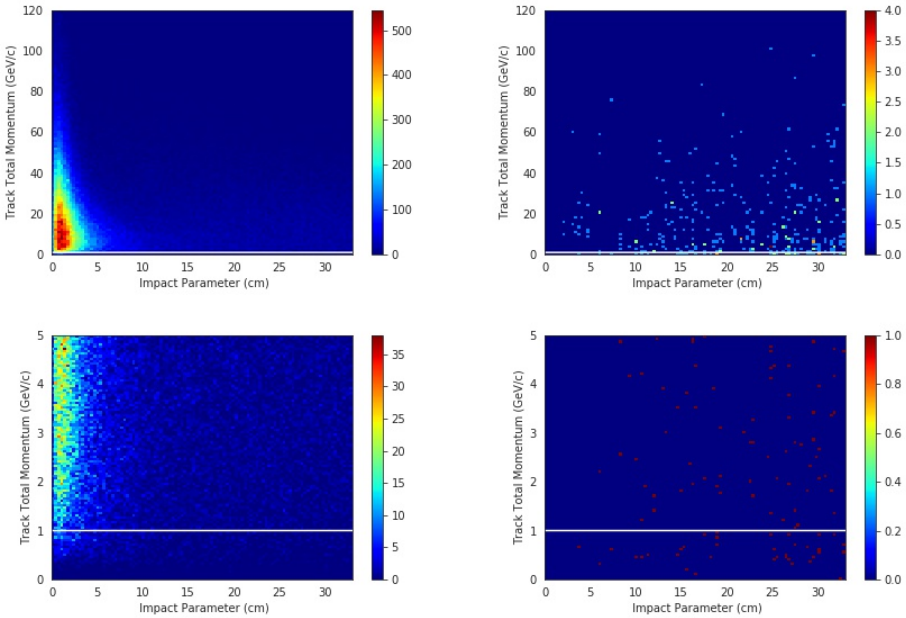

Figure 5.4: Second cut applied to the samples. The images on the left correspond to the signal, while the images on the right correspond to the background. The images on the bottom are zoom-ins of the Total Momentum of the Track around the region of most interest.

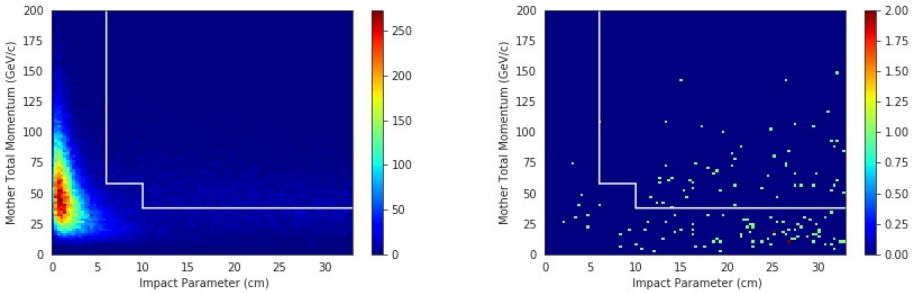

Figure 5.5: Third cut applied to the samples. The images on the left correspond to the signal, while the images on the right correspond to the background.

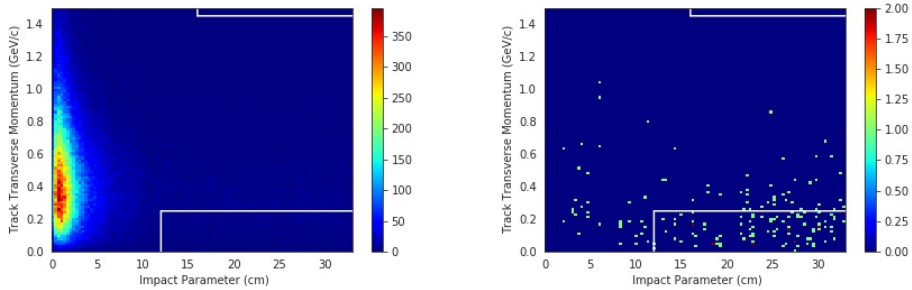

Figure 5.6: Forth cut applied to the samples. The images on the left correspond to the signal, while the images on the right correspond to the background.

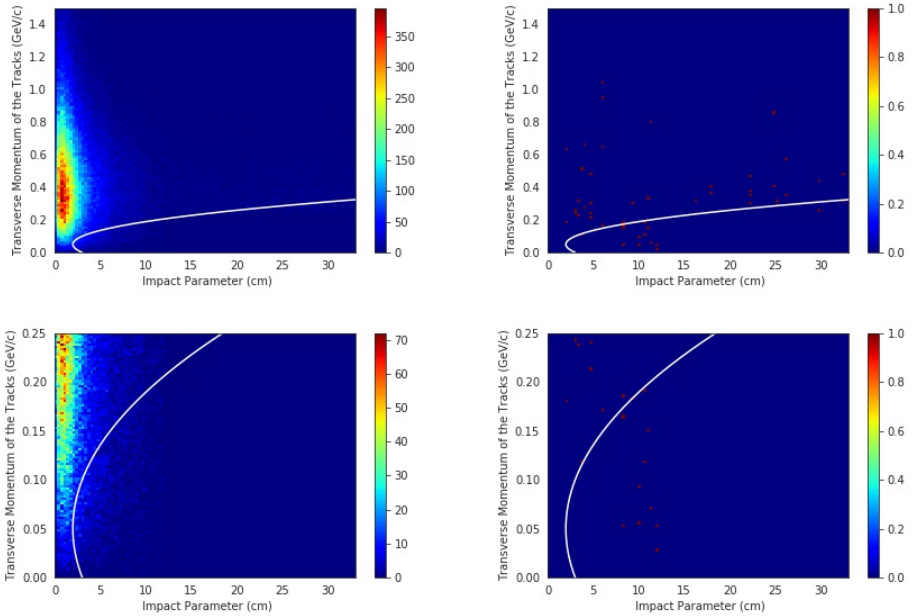

Figure 5.7: Fifth cut applied to the samples. The images on the left correspond to the signal, while the images on the right correspond to the background. The images on the bottom are zoom-ins of the Transverse Momentum of the Track around the region of most interest.

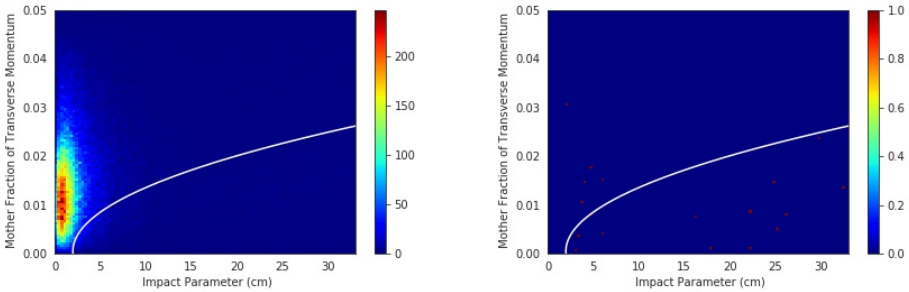

Figure 5.8: Sixth cut applied to the samples. The images on the left correspond to the signal, while the images on the right correspond to the background.

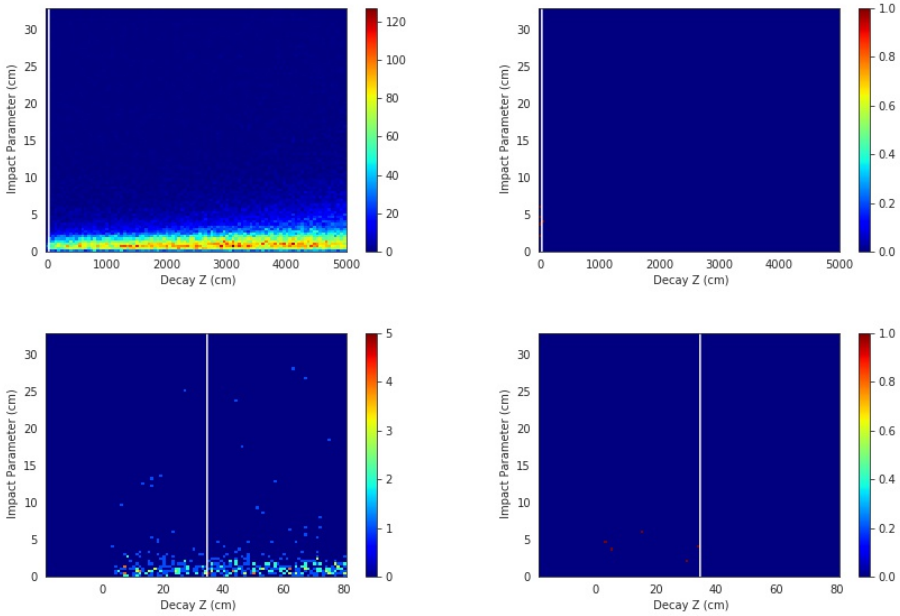

Figure 5.9: Seventh cut applied to the samples. The images on the left correspond to the signal, while the images on the right correspond to the background. The images on the bottom are zoom-ins of the Z coordinate of the decay vertex of the HS candidate particle around the region of most interest.

Chapter 6

Dark Photons

The analysis of the Dark Photon samples was done in a very similar fashion to that of the Heavy Neutral Leptons, presented in the previous chapter. We will first reiterate the variables considered in the study, followed by the cuts applied. As was stated in Section 4.1.5, the sensitivity limits were updated recently, and as such this study included some DP samples with masses that most likely will not be able to be probed at SHiP. We will therefore also present the selection efficiencies of the cuts applied on a sample of Dark Photons that does not take into account events with masses of 3.6 and 4.4 GeV/c^2.

The only decay modes studied for the Dark Photons were the decays into pairs of charged leptons. So, the first cut applied after the preliminary ones was to only consider combinations of muons and electrons, which correspond to the next three combinations:

$$e^-e^+,\ \mu^-\mu^+ \text{ and } e^\mp\mu^\pm. \tag{6.1}$$

Even though there are no decays involving neutral pions, we still searched for them, since they should not appear in our signal samples this time around, providing a good veto for background events. Once again, every decay mode studied with the Dark Photons was compared against the full background obtained from these cuts.

After this preliminary analysis, the same 11 kinematic variables that were used for the HNLs, and presented in Table 5.1 were taken into account here. The "Mother" and "Track" denominations defined in Chapter 5 are also used here. Recall that the Momentum related features apply to both the Mother and Tracks.

Figs. 6.1 and 6.2 show the distributions of the all the aforementioned properties for the combined DP samples (blue), overlapped with the the distributions for the background (orange), normalized as densities, where the integral of the distributions is $\simeq 1$.

6.1 Kinematic Cuts

Just like for the Heavy Neutral Leptons, Figs. 6.1 and 6.2 show that there are clear distinctions between the distributions of some kinematic properties. The most flagrant case is, once again, the Impact Pa-

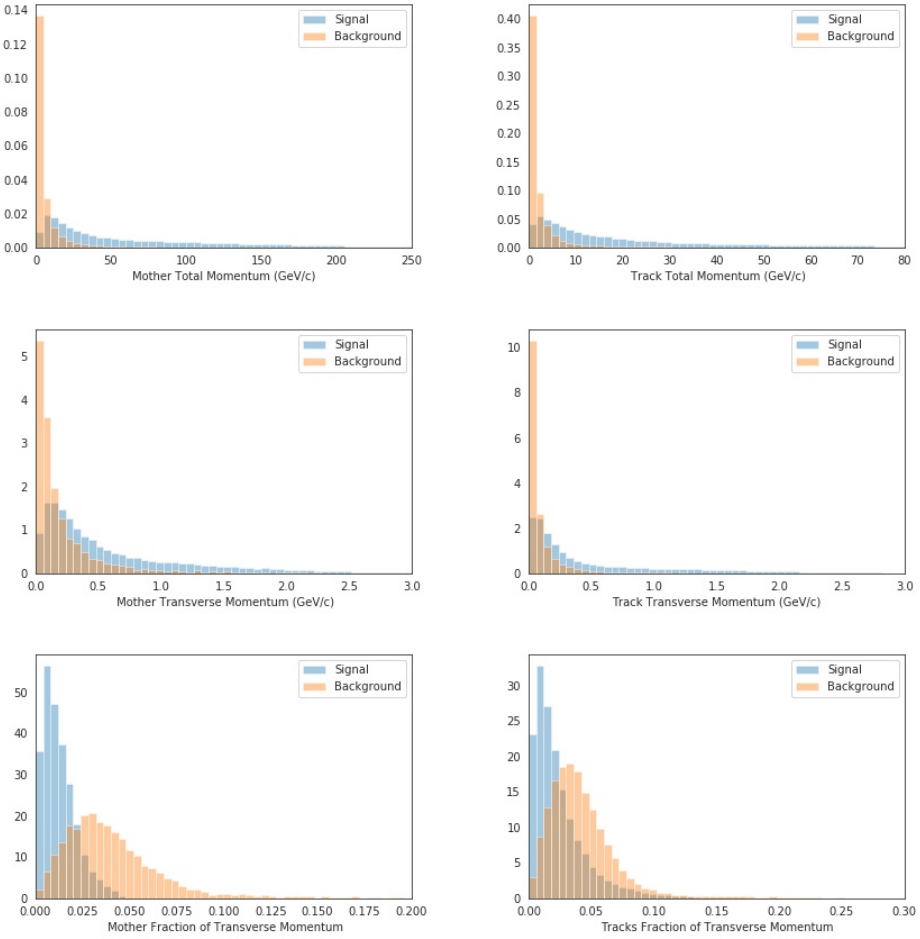

Figure 6.1: First set of distributions containing the momentum related properties of the final reconstructions (Mother), as well as of the tracks.

rameter, which is even more pronounced this time around. This is to be expected, since in the case of the Dark Photons we did not generate any events that had innate missing energy in the reconstructions. On the other hand, events that produce pairs of charged leptons from muon or neutrino DIS interactions should also lead to the creation of additional neutrinos and consequently missing energy in the final reconstructions, in order to conserve the leptonic numbers.

The biggest difference between the previous case and this one is that the clear tendency for the background events to be generated before the Decay Vessel is not as sharp this time around. This is most likely due to the fact that pairs of muons are coming from misidentifications of the decays of particles like neutral kaons ($K_L^0 \to \mu^{\mp}\pi^{\pm}\nu_{\mu}$). Nevertheless, this leads to missing energy, which should only exacerbate the difference in the Impact Parameter distributions, which we can verify when compare the Impact Parameter distributions for the DPs and their relevant background, from Fig. 6.2, and the ones for the HNLs in Fig. 5.2.

Regardless, since we still expect the distinction of the HS signal from the background to be mostly related to the Impact Parameter, we approached the cuts in an analogous fashion to the ones done on Chapter 5. We established some preliminary cuts, then searched for combinations of variables that would

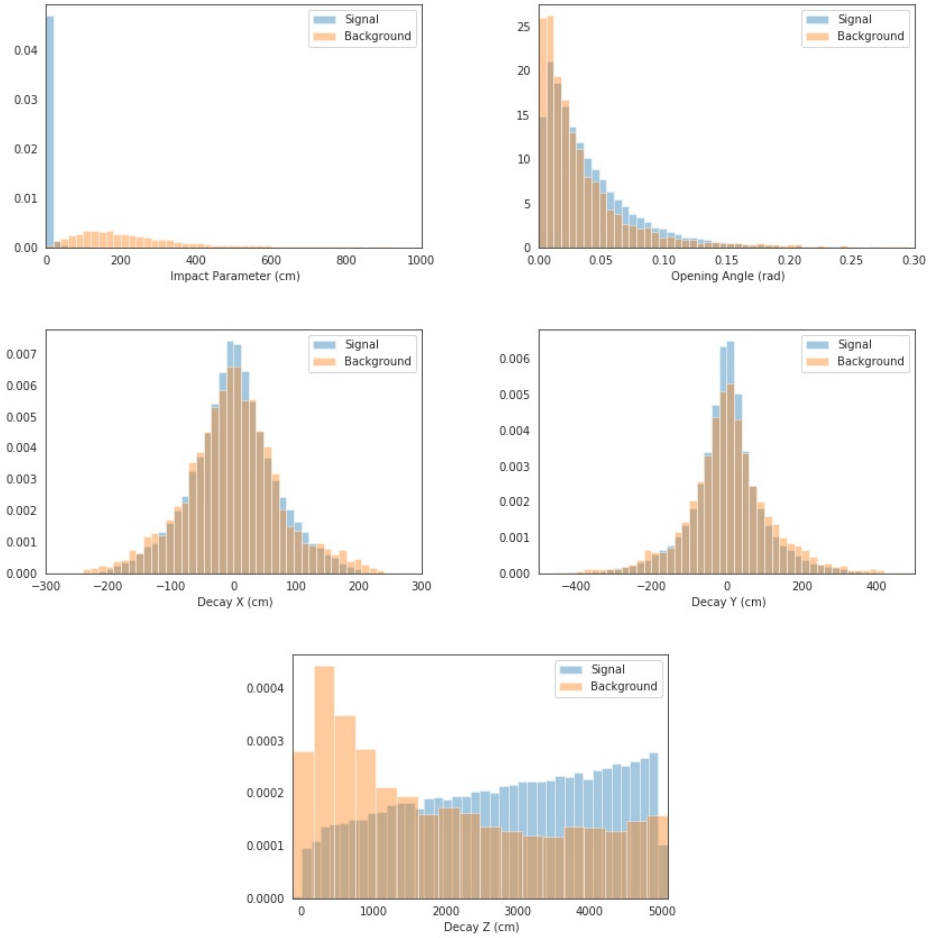

Figure 6.2: Second set of distributions containing the alongside the Impact Parameter and Opening Angle, alongside the decay positions of the final reconstructions (Mother). Similarly to Fig. 5.2, for the decay position in Y and X the origin represents the incident beam trajectory, and for Z the origin represents the beginning of the Decay Vessel.

be able to remove all background events while maintaining a high selection efficiency for the DPs, and in the end revised every single cut one by one against the rest to reduce as much as possible the vetoing of the signal. In this case, however, we already had the set of cuts that were applied to the HNLs, and as such we took those as the preliminary ones. Here we did not try to relax as much as possible the Impact Parameter cut, since it should not be as harmful to the signal samples, due to the lack of decay modes with missing energy.

Albeit different from the Heavy Neutral Leptons, there were also 7 cuts applied to the data in this case, which are presented in Table 6.1.

The aforementioned cuts, and respective selection efficiencies, are shown in Table 6.2, and the corresponding visual representations can be seen from Figs 6.3 to 6.9. Since the sensitivity estimates presented in Fig 2.8 already take into account geometric efficiencies, the percentages shown are only as a reference against the HS candidate particles reconstructed within the Decay Vessel from Table 4.5.

A final summary of the selection efficiencies by decay mode is shown in Table 6.3. As was expected, the selection efficiencies are very high for the Dark Photons, even though the Impact Parameter cut was

Cut	Thresholds for Selection
1	IP < 27.0 cm.
2	IP $< 8.7 \times (0.466 - \text{TotMom})^2 + 6.5$ cm, where TotMom is from the Tracks in GeV/c.
3	IP > 14.8 or 23.7 cm requires TotMom of the Mother < 66.0 or 47.0 GeV/c, respectively.
4	IP > 10.15 cm requires IP $< 168.87 \times$ TransMom $- 3.41$ cm, where TransMom is from the Tracks in GeV/c.
5	IP $< 50484 \times (0.0012 - \text{TransMom})^2 + 1.79$ cm, where TransMom is from the Tracks, in GeV/c.
6	IP > 6.15 cm requires OA > 0.0026 rad.
7	Decay Z > 156.0 cm.

Table 6.1: Kinematic cuts applied to the DP samples, and their respective background.

tightened. Once again, decay channels involving muons are favoured in comparison with their electron counterparts, which is natural, since the muons tend to interact much less with the environment, and usually produce clearer experimental signatures.

Even though the criteria established are quite satisfactory, notoriously the muonic decay channel with $\approx 95\%$ selection efficiency within the most recent sensitivity estimates, there is still room for improvement with the electronic channel, which is the only one available for a majority of the mass region of interest. As such, we are expecting the Machine Learning algorithms to improve substantially upon this result.

Cut	Background Events Selected (%)	DP Events Selected (%)	DP Events Selected within new Sensitivity (%)
Decay products in Eq. (6.1)	6806 (34.8)	50209 (99.9)	44109 (99.9)
1	120 (0.6)	47982 (95.4)	42447 (96.1)
2	25 (0.1)	47048 (93.6)	41514 (94.0)
3	22 (0.1)	46515 (92.5)	41218 (93.3)
4	7 (0.0)	45742 (91.0)	40445 (91.6)
5	4 (0.0)	44803 (89.1)	39506 (89.5)
6	3 (0.0)	44773 (89.1)	39476 (89.4)
7	0 (0.00)	44082 (87.7)	28861 (88.0)

Table 6.2: Selection efficiency of the kinematic cuts presented in this section. The numbered cuts refer to the ones presented on Table 6.1. All percentages are calculated against the total number of HS candidate particles reconstructed within the Decay Vessel, presented in Table 4.5, which gives a total of 19568 background events, 50278 DP events, and 44161 DP events within the most recent sensitivity estimates.

Decay Mode	Reconstructions within the Decay Vessel	Events with the desired Decay Products (%)	Surviving Events after the Cuts (%)
$A' \to e^- e^+$	33324	33276 (99.9)	27981 (84.0)
$A' \to \mu^- \mu^+$	16954	16933 (99.9)	16101 (95.0)
$^*A' \to e^- e^+$	30218	30183 (99.9)	25608 (84.7)
$^*A' \to \mu^- \mu^+$	13943	13926 (99.9)	13253 (95.1)

Table 6.3: Summary of the selection efficiency obtained by applying the kinematic cuts, detailed by Decay Mode. Once again notice that all the efficiencies are based on the Reconstructions within the Decay Vessel, presented in Table 4.5, which are reiterated here. *Dark Photon events without the 3.6 and 4.4 GeV/c^2 mass samples, that are not within the most recent sensitivity estimates.

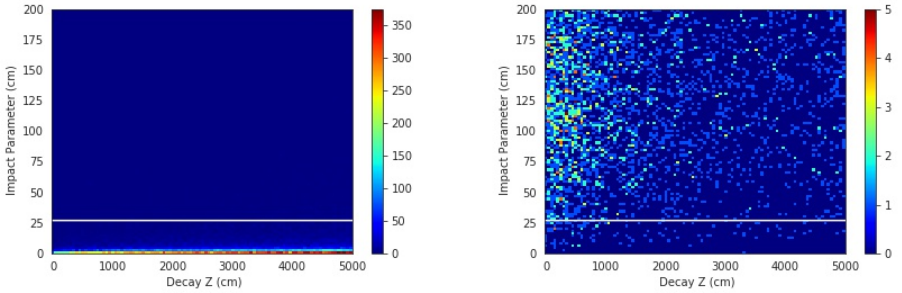

Figure 6.3: First cut applied to the samples. The image on the left corresponds to the signal, while the image on the right corresponds to the background.

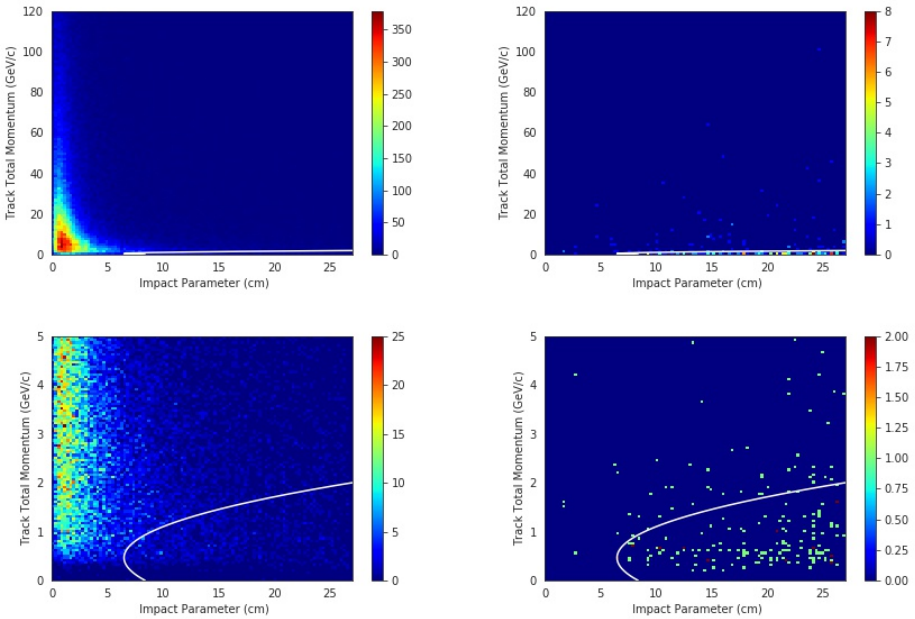

Figure 6.4: Second cut applied to the samples. The images on the left correspond to the signal, while the images on the right correspond to the background. The images on the bottom are zoom-ins of the Total Momentum of the Tracks around the region of most interest.

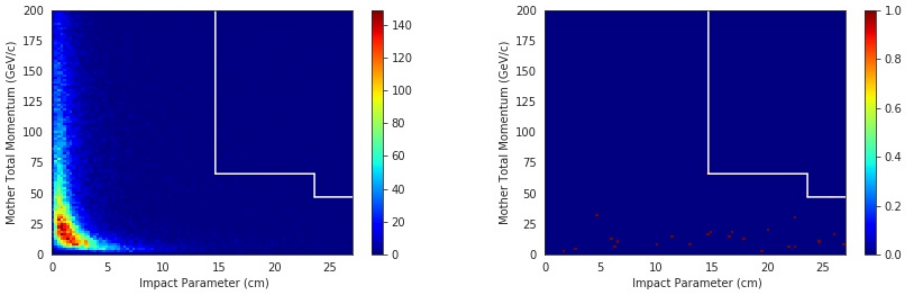

Figure 6.5: Third cut applied to the samples. The images on the left correspond to the signal, while the images on the right correspond to the background.

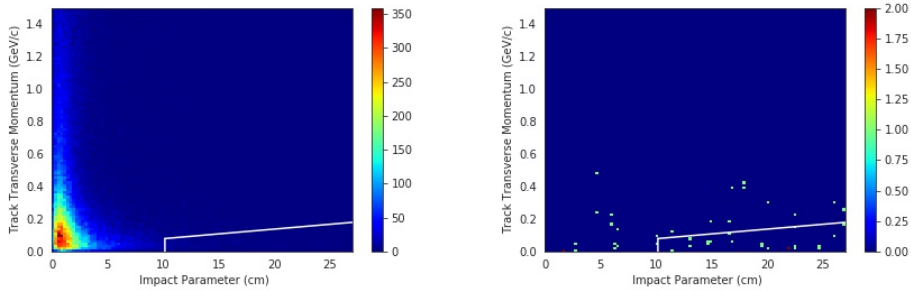

Figure 6.6: Forth cut applied to the samples. The images on the left correspond to the signal, while the images on the right correspond to the background.

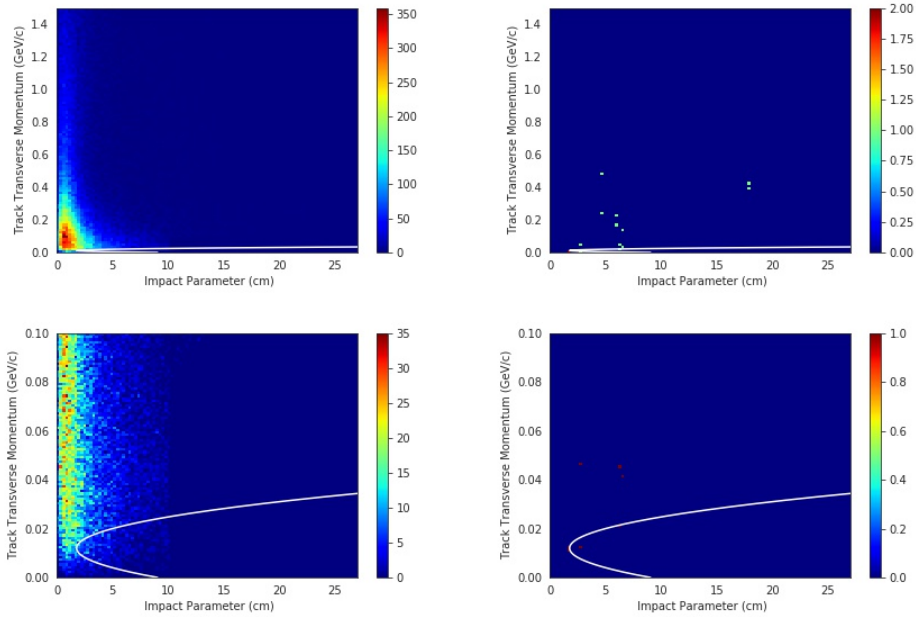

Figure 6.7: Fifth cut applied to the samples. The images on the left correspond to the signal, while the images on the right correspond to the background. The images on the bottom are zoom-ins of the Transverse Momentum of the Tracks around the region of most interest.

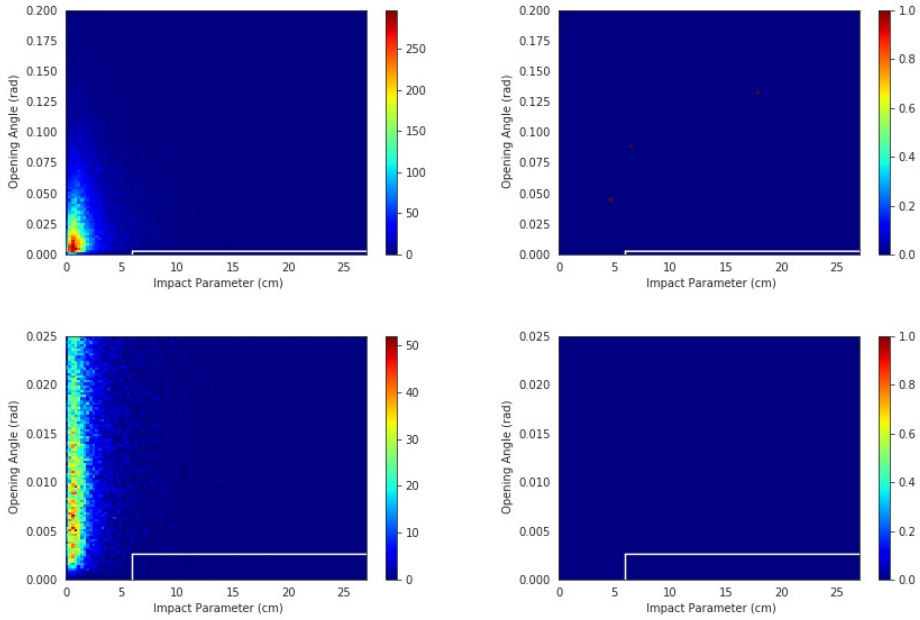

Figure 6.8: Sixth cut applied to the samples. The images on the left correspond to the signal, while the images on the right correspond to the background. The images on the bottom are zoom-ins of the Opening Angle around the region of most interest.

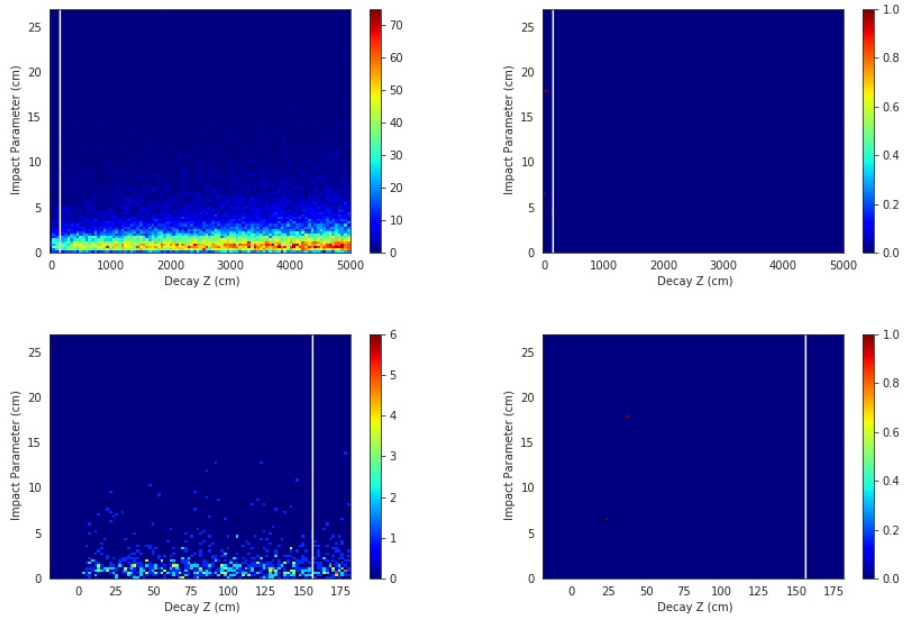

Figure 6.9: Seventh cut applied to the samples. The images on the left correspond to the signal, while the images on the right correspond to the background. The images on the bottom are zoom-ins of the Z coordinate of the decay vertex of the HS candidate particle around the region of most interest.

Chapter 7

Machine Learning

Machine Learning (ML) algorithms are very bright prospects in analysing data sets, since they can compare the information taking into account multi-variable correlations that are non-trivial. Their use has become quite frequent nowadays, including for what regards selection problems such as the namesake of this thesis [93]. Such an avenue was employed in this study.

One very popular set of algorithms is the Multilayer Perceptron (MLP), which encompasses feedforward artificial neural networks. The ML analysis was done resorting to these Neural Networks.

In this chapter we will present a small introduction to MLPs, alongside the basic structure used for the networks that were trained. Afterwards, we will present the best results obtained, and lastly provide the final selection efficiencies obtained. Additionally, the **SHAP** software is used to estimate the relevance given to every variable fed to the networks. This further allows to verify whether or not prioritizing the Impact Parameter when applying the kinematic cuts in Chapters 5 and 6 was a sensible choice.

7.1 Multilayer Perceptrons

A typical Neural Network (NN) consists of an interconnected group of nodes, or neurons, hence its name, arranged in layers, where each node processes information and then passes the results to the next layer of nodes. A generic diagram of a Neural Network with 4 layers can be seen in Fig. 7.1. The first layer is called the input layer, as it receives the data features, and last one is called the output layer, since it provides the NN response. There needs to be one node for each input feature on the input layer, and there is usually one node for each set that we are trying to classify. For example a NN trying to distinguish Dark Photons, Heavy Neutral Leptons and background at once should have three output nodes, attributing the three final scores to each event that can be interpreted as the chances of being from each set. The common exception to this protocol is for binary classifications, where there is only the need for one score, since the chance of not being from the set is just the maximum score attainable, minus the score provided by the network.

The connections between the nodes are defined by weights, that are adapted during the training phase, and each of the nodes might have a threshold or bias. Once the information that was fed to a node is

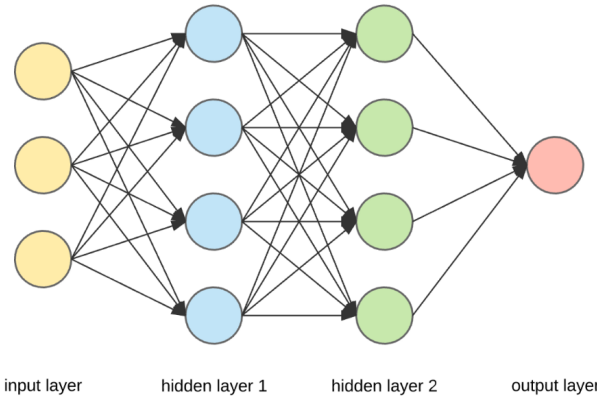

Figure 7.1: Typical structure of a Neural Network where each circular node represents a neuron, and the arrows represent the connection of the output of a neuron to the input of another on the next layer.

processed, it is sent through an activation function before being transmitted as output. The property that distinguishes MLPs from simple NNs is that the activation functions are non-linear for all layers besides the first, allowing for backpropagation[1]. This makes it so that MLPs need to have three or more layers, since if all activation functions are linear we can reduce them to two layers through algebra. The layers between the input and output are called hidden layers. Due to the high number of neurons, and a need for fast and easy processing of information, activation functions need to be simple to calculate, and need to maintain the output within specific ranges. This leads to a few types of activation functions. A simple function that combines a step-like behaviour with a linear one is the Rectified Linear Unit (ReLU), which is 0 until a threshold and linear afterwords, and is very easy to compute, but does not provide upper bounds for its output. On the other hand, there are functions with smooth gradients, which are able to easily discern outliers and provide outputs within very clear boundaries, such as the sigmoid and hyperbolic tangent (TanH) functions, which are functionally the same but where the sigmoid provides an output between 0 and 1, the TanH ranges from -1 to 1. However, they do carry an increased complexity. As such, ReLU is usually used in inner layers of deep neural networks, while the other two are commonly found in the output layer.

MLPs are NNs with supervised learning, which means they learn by processing examples with clearly defined inputs and results, forming a probability-weighted association between the two. The NN updates these weights according to a *loss-function*, which is commonly a *cross-entropy* function in classification problems. Since it is impossible to find the optimal weights for a given network, due to its complexity, optimization algorithms involving gradient descent are used in order to reach near-optimal solutions[2]. The rate at which they are updated is known as the *learning rate*. As mentioned previously, MLPs have supervised learning, and as such a separate data set that is not used to train it is used to validate the results, resorting to the loss-function. This procedure attempts to stop the network from overfitting the training data, becoming obsolete when faced with new data.

[1] Backpropagation is a training algorithm based on the gradient of the loss-function. For further information on this topic see Chapter 6 of Ref. [94].

[2] Although optimizers can influence the speed at which a near-optimal solution is found, they do not change drastically that solution, and as such it is not a critical concept. However, for further information on common optimizing functions see Ref. [95].

Each time the algorithm sees the whole data is called an *epoch*. The weights however do not need to be updated once each epoch. The number of events that the network sees before making an update is called a *batch*, and both of these parameters can influence the time it takes for the algorithm to find the optimal solution. Having a low number of epochs provides very little time and information for the network, which might lead to underfitting. As for the batches, having small batches might lead to an increase in training time without any visible results, due to a high volatility in the events being used for training. However, having batches that are too big can, once again, delay the process of finding a near-optimal weight solution.

Overfitting and underfitting are ever-present problems in ML, and can also be influenced by the size and complexity of the network. If the NN is too complex for the problem, overfitting is unavoidable. On the other hand, too simple of a network will hardly lead to satisfactory results. In general several iterations are required until the balance is struck.

7.2 Analysis employing MLPs

The ML learning algorithms used in order to differentiate the data provided were Multilayer Perceptrons, as was stated previously. For the implementation of our algorithms we used Keras [96], an open-source library that provides a Python interface for artificial neural networks. Keras acts as an interface for the TensorFlow [97] library.

The first step taken in order to apply Neural Networks to the data was to normalize all features to the same scale, in order to prevent unwanted biases during the validation of the network, since the variables do not have the same units, and some have much larger ranges than others (e.g. the Total Momentum of the Mother particle can easily reach 100 GeV/c, while the Fraction of Transverse Momentum mathematically can never be larger than 1). This was done by applying a standard scaler to all features fed onto the network, which transforms the whole set so that the mean value is approximately 0, with a standard deviation of 1.

In this project, several configurations were tried, including different combinations of activation functions, both for the hidden layers and final layer, loss functions and optimizers. Training networks using all of the decay modes from the HS samples was also compared to training individual networks for each decay mode, and the latter option proved to be better. We followed an under sampling approach, where due to the asymmetric amounts of data of signal versus background, a smaller subset of HS events was selected for training, with the remaining being used for validation alone. A standard of 70% of the background events was selected for training the NN, alongside a comparable signal sample.

Several test networks were fed all the variables available, but the results were not satisfactory, and as such we settled on only including the variables that had already been used in the standard kinematic cuts from Tables 5.2 and 6.1, during the training, which showed much more promise. Several rates of information *dropout* along the layers were also tested, in order to reduce overfitting effects that might show up during training. The best solution found was to only deny information from hidden layers, varying the rate at which we did so depending on the data set.

Due to the inherent randomness of ML algorithms, several configurations were used in order to differentiate the data set. Additionally, networks with the exact same characteristics were also trained several times. From all of these tests, the NN that showed the best selection efficiency for the HS signal, while removing all of the background events, was chosen as the best one. In order to remove all of the background events, we determined the selection threshold to be equal to the maximum score obtained by a background event, maintaining the 0 background criterion set in the kinematic cuts.

In this study, we set from the get go that the background would be the same for each HS particle, regardless of the decay mode studied. A direct consequence of this is that the NNs were not trained to distinguish specific types of signals, but rather what was not a background event. As such, all the NNs related to the same background samples can be used to separate the data sets, regardless of their decay mode.

This allows for the final selection efficiency for every decay mode to be obtained by passing the events through all the NNs that relate to same background, and selecting as true positives any events that were given as signal by at least one of these NNs (e.g. If we are selecting an event of $N \to e^{\mp}\pi^{\pm}$, all NNs trained on HNL data sets score the event, and if at least one score is above the threshold given for the respective network, the event is considered a positive.).

In the earliest tests involving neural networks, no special configuration stood out from the remaining. This lead to the use of networks that provided a binary separation, with the ReLU activation functions for the initial and hidden layers, followed by a final single node layer, using a sigmoid activation function, which is pretty standard configuration due to its relatively fast training speed, allowing for a testing of more networks. The loss function defined was the binary cross-entropy, and the optimization algorithm selected was Adam, with a learning rate of 0.001.

The only network that initially provided better efficiencies for some decay modes, when compared to the efficiencies obtained by using a single binary NN on the data from a specific decay mode, was a multi-classifier network, used on the full HNL data set. The only difference to the previously described standard configuration was the logical transition from a binary cross-entropy loss function to categorical cross-entropy, and changing the optimizer to a stochastic gradient descent. Obviously, since we are now attributing 7 scores to each event, one for each decay mode and a background score, the final layer had 7 nodes, and it was activated by a *softmax* function. However, these types of networks had much slower training times in order to provide effective results, and once all the individual binary NNs were applied to the HNL data, any efficiency that was initially higher was easily surpassed, validating our choice.

The basic properties of all the selected networks can be found in Table 7.1, defined by their training sample and respective background, and the results obtained from applying such networks in the samples can be seen in Table 7.2.

Comparing the selection efficiencies obtained with the standard kinematic cuts, presented in Tables 5.4 and 6.3 for the HNLs and DPs, respectively, with the ones obtained by using machine learning algorithms in Table 7.2, we can clearly see that there is a improvement when NNs are involved in the analysis. All samples show higher efficiencies, regardless of whether there is missing energy or not in the decay.

The samples that receive a smaller boost in their selection efficiency are the ones where usually there

Training Sample	Training Epochs	Batch Size	Number of Hidden Layers	Nodes on first Hidden Layer	Step between Hidden Layers	Information Dropout [%]
$N \to e^-\pi^+$	600	100	3	25	5	5
$N \to \mu^-\pi^+$	600	200	3	20	5	5
$N \to e^-\rho^+$	800	100	3	25	5	8
$N \to \mu^-\rho^+$	800	100	3	25	5	10
$N \to \nu_\alpha\mu^+\mu^-$	1000	200	2	30	10	5
$N \to \nu_\alpha\rho^0$	1000	200	2	30	10	10
$A' \to e^-e^+$	800	100	3	20	5	15
$A' \to \mu^-\mu^+$	800	100	3	20	5	10

Table 7.1: Summary of every NN trained, containing their basic properties, alongside some training choices. No NN is shown for DP samples that do not include masses outside the most recent sensitivity studies, since no NN was specifically trained on such a data set. The Step between hidden layers refers to the number of nodes that are being removed from layer to layer. The information dropout also applies from hidden layer to hidden layer.

is no energy missing in their reconstructions, such as the DPs and the HNL decays that involve charged leptons accompanied by charged pions ($N \to \ell^\mp\pi^\pm$), which range from an extra 3.2 % to 10.7% for the selection of $A' \to \mu^-\mu^+$, within the most recent sensitivity limits, and $N \to e^-\pi^+$ events respectively. On the other hand, HNL samples involving decays with missing energy had surprisingly good selection efficiencies, even more so when compared to the standard kinematic cuts. The selection of HNL events that decay to a SM neutrino and a pair of muons received an astounding increase of 47.6%, and the selection of events that decay to a neutral rho meson alongside a SM neutrino got an even more impressive increase of 64.6%, clearly showcasing the benefit of using machine learning methods in studies such as these.

7.3 SHAP Analysis

The best explanation for any given model is the model itself, since it represents itself, and if it is intuitive, there is no need to find better explanations. However not all models are simple enough, and with the advent of black-box-like machine learning models it is clearer than ever that in some cases the inherent complexity does not allow for the model to explain itself in a fully understandable way. As such, we try to find simpler explanation models that provide better insight.

Previously we mentioned the use of a software that analyses Neural Networks, and tries to provide insight into the relative importance that a specific network attributes to each variable fed onto it. One type of software that aims at solving this is **SHapley Additive exPlanations (SHAP)** [14]. As the name suggests, it relies on additive feature attribution methods, where the explanation model is a linear function of binary variables:

$$g(x') = \varphi_0 + \sum_{i=1}^{M} \varphi_i x' i, \tag{7.1}$$

where x' are the simplified features, $x' \in \{0,1\}^M$, with M the number of simplified input features, and φ_i are real numbers. Notice that the explanation method $g(x')$ tries to approximate the output of the

Sample	Score Threshold	Signal Events Selected (%)	Signal Events Selected Applying all adequate NNs (%)
$N \to e^- \pi^+$	0.9875	22467 (89.5)	23098 (92.0)
$N \to \mu^- \pi^+$	0.9770	25825 (96.1)	26077 (97.1)
$N \to e^- \rho^+$	0.9994	9847 (44.9)	13558 (61.8)
$N \to \mu^- \rho^+$	0.9993	10817 (44.6)	15528 (64.1)
$N \to \nu_\alpha \mu^+ \mu^-$	0.9981	11875 (41.0)	14855 (51.2)
$N \to \nu_\alpha \rho^0$	0.9950	14303 (66.9)	15225 (71.2)
$A' \to e^- e^+$	0.9989	29061 (87.2)	30216 (90.7)
$A' \to \mu^- \mu^+$	0.9872	16608 (98.0)	16698 (98.5)
$*A' \to e^- e^+$	N.A.	N.A.	27274 (90.3)
$*A' \to \mu^- \mu^+$	N.A.	N.A.	13704 (98.3)

Table 7.2: Summary of the results obtained by the chosen Neural Networks, discriminated by the samples they were trained with. The final selection efficiencies are obtained by applying all the NNs that share the same background as the sample. *Samples within the most recent sensitivity estimates. For these only the final selection efficiency is shown, since these data sets were not used to train any NN and as such N.A. stands for Not Applicable. All selection efficiencies are relative to the number of HS candidate particles reconstructed within the Decay Vessel, presented in Tables 4.3 and 4.5 and reiterated in Tables 5.4 and 6.3.

actual model $f(h_x(x'))$, where $x = h_x(x')$ is an arbitrary mapping function commonly used in models of this type.

Its output, the SHAP number, relies on the Shapley value, which is a solution concept in cooperative game theory [98]. Games such as these assume that there are a number n of players that cooperate in order to obtain a certain overall gain. One common problem seen within these games is how to find an optimal way of distributing the gain between all players, so that each individual player is rewarded according to their contribution. The Shapley value is a very enticing approach, since it attributes the individual gain to each player (ϕ_i) as a function of its marginal contributions, and only depends on a real-valued function $v(S)$, which relates to the gain obtained by a subset S of players without outside help. Its formal definition can be seen in Eq. 7.2, where N is the full set of players and n the aforementioned total number of players,

$$\phi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n - S - 1)!}{n!} \left(v(S \cup \{i\}) - v(S)\right). \tag{7.2}$$

If we break down Eq. 7.2, the Shapley value attributed to a player i depends on the value that the player adds when he is inserted in a specific group S that does not include him ($v(S \cup \{i\}) - v(S)$). This is then repeated for all possible combinations of groups, hence the sum over all subsets of N that do not include the player i himself. Since neural networks essentially take into account a set N of initial features, which are then evaluated in order to provide a score $v(N)$, it is clear as to why the Shapley number is a good approach when trying to estimate the contributions of these features. To sweeten the deal, it brings along some very helpful properties, which include Efficiency: $\sum_i \phi_i$ = total gain; Symmetry: if $v(S \cup \{i\}) = v(S \cup \{j\})$ applies to all subsets S of N that do not include either i or j, then $\phi_i = \phi_j$;

Linearity: $\phi(v+w)_i = \phi_i(v) + \phi_i(w)$ and $\phi_i(a.v) = a\phi_i(v)$ for any real a; and the Null Player: $\phi_i = 0$ for a null player, or if a player has no marginal contribution ($v(S \cup \{i\}) = v(S)$).

The only drawback to the traditional Shapley value is that it is not so trivial as to how one estimates the value functions for all the different subsets of attributes. Nevertheless, delving deeper into these themes is not the focus of this project, and the explanation of the Shapley value should suffice to understand the results of the SHAP analysis ran on the used Neural Networks. In order to gain insight into the NNs described in Section 7.2, the specific Deep SHAP model was used. It is important to notice that a SHAP number can be defined for every event, and as such what we are interested in is the mean value for each variable. In Table 7.3, we present the three most important features according to their SHAP numbers, for each Network used when we separated the HS signal from the background.

Training Sample	Feature Importance		
	First	Second	Third
$N \to e^-\pi^+$	Mother TotMom	Impact Parameter	Mother FracMom
$N \to \mu^-\pi^+$	Mother TotMom	Impact Parameter	Mother FracMom
$N \to e^-\rho^+$	Mother TotMom	Impact Parameter	Mother FracMom
$N \to \mu^-\rho^+$	Mother TotMom	Mother FracMom	Impact Parameter
$N \to \nu_\alpha\mu^+\mu^-$	Mother TotMom	Mother FracMom	Impact Parameter
$N \to \nu_\alpha\rho^0$	Mother TotMom	Impact Parameter	Mother FracMom
$A' \to e^-e^+$	Mother TotMom	Mother FracMom	Opening Angle
$A' \to \mu^-\mu^+$	Mother TotMom	Mother FracMom	Opening Angle

Table 7.3: Features with the highest mean-SHAP values for the Neural Networks used in Section 7.2. The momentum related features refer to the HS candidate particle, denoted as Mother throughout this analysis. Notice that the Training Sample also includes the respective background.

The SHAP analysis, presented in Table 7.3, brought a mixture of expected results alongside some unexpected outcomes. The most resounding feature is the clear predominance of the Total Momentum of the Mother Particle as the most important feature for every NN. Although this feature was mentioned as clearly showing a discrepancy between the distributions of the signal vs background, it did not have a clear cut separation. Another surprise is that the machine learning algorithms seem to heavily favour the kinematic properties of the Mother, while not paying nearly as much attention to the properties of the Tracks. The only outlier to this trend is the importance given to the Opening Angle formed between the Tracks of the decay products of Dark Photons.

The expected result that came from this analysis is the predominance of the Impact Parameter as the second or third highest rated feature by SHAP. The only Neural Networks that fled this trend were the ones that were trained on the Dark Photon samples, where surprisingly the Opening Angle was held in high regard. With this in mind, we can very safely state that approaching the standard kinematic cuts mostly through the Impact Parameter, and two feature combinations including it, was definitely a sensible choice, at the very least for the Heavy Neutral Lepton samples.

However, one can make the case that applying the same strategy on the DP samples is not optimal, and that we should search for other features to base most of the cuts around. To this, we can only state

that despite the relatively low importance attributed to the Impact Parameter by these networks, it still stands that the most distinct distributions between the signal and the background were still from this feature, and that the NNs do not apply binary cuts, where the evaluation is either pass or veto. Instead, as Fig. 7.2 shows, even though there might be clear tendencies as to what the network highly regards as signal or background, scoring a higher or lower SHAP number respectively, the specific value of a feature does by no means determine it.

Fig 7.2 shows that despite there being a tendency to highly value a slightly below average Total Momentum of the Mother particle, with values below 0[3], its actual influence in the final score fluctuates from being very relevant, with SHAP values near 3, to being nearly irrelevant, with SHAP values near 0. At the same time, the classification is very sensitive, and the score of events with a Total Momentum slightly above average is heavily penalized. However, once this feature starts becoming unusually high, it ceases to influence the decision of the network. Clearly, this type of behaviour cannot be emulated through kinematic cuts, and as such using the behaviour of a NN as a guideline for the more conventional studies should always be done with a grain of salt.

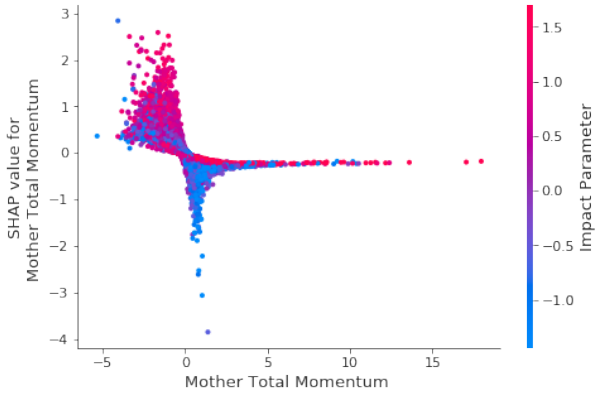

Figure 7.2: Distribution of the SHAP value for the Total Momentum of the Mother particle as a function of the Total Momentum of the Mother particle, for the Neural Network trained on the $N \to e^{\mp}\pi^{\pm}$ sample. Each point corresponds to an event from that sample, or from the corresponding background, that was not used to train the Deep SHAP model. Notice that the Total Momentum of the Mother particle was normalized so that the whole sample had a mean value of 0 and a standard deviation of 1. The colour coding refers to the feature that obtained the second highest mean SHAP value, which in this case is the Impact Parameter.

Recapitulating, by using ML methods, not only were we able to verify that the Impact Parameter is indeed a good distinguishing feature of the HS events, but we were also able to improve the selection efficiencies of the HS particles, while maintaining the 0 background requirement. Applying NNs allows us to obtain selection efficiencies of over 90% for the $N \to \ell^{\mp}\pi^{\pm}$ and $A' \to \ell^{-}\ell^{+}$ samples, which are the most relevant decay modes for both particles, and the lowest selection efficiency obtained was greatly enhanced to $\approx 51\%$ for the $N \to \nu_{\alpha}\mu^{-}\mu^{+}$ sample.

[3] As previously stated, all features are transformed before being fed into the NNs so that the mean value is approximately 0 and the standard deviation 1.

Chapter 8

Studies done with Reduced Vacuum

In this chapter, we go a step further and carry out a complementary study to probe an alternative scenario where there is a potential increase of the air pressure inside the Decay Vessel. To do this new samples were generated for a SHiP setup with exactly the same conditions as the ones done previously, but with the caveat that the Decay Vessel would now stand at an air pressure of 1 bar instead of the currently planned 1 mbar.

After the selection studies are done, selection efficiencies will be compared in order to determine whether the kinematic selection is effective enough to allow an increase in the pressure of the vessel without compromising severely the sensitivity of the SHiP experiment to the Hidden Sector.

While the Machine Learning analysis has shown to provide better selection efficiencies than the conventional analysis, a set of standard cuts was still applied to the data sets studied, in order to provide a good comparison point. In the cut-based study, we will also apply the cuts that were used in the samples generated for the current SHiP configuration, to study any possible degradation in the reconstruction of the HS events from increasing the Decay Vessel pressure alone. As for the ML approach, we will search for Neural Networks that either provide the best selection efficiency for 0 background events selected, or provide the least amount of background events selected, while maintaining the HS signal selection efficiency obtained for the vacuum setup.

8.1 Sample Generation

During the whole ML analysis, which provided the best selection efficiencies, we found out that the neutrino DIS background proved to be the hardest to distinguish from the HS signals. As such, unlike in the main study, the muon DIS background was disregarded. Needless to say, no study regarding the muon combinatorial background is provided since the same problems that were seen previously still apply.

New samples were generated in similar conditions to the ones presented in Chapter 4, apart from the fact that the Decay Vessel now has a pressure of 1 bar instead of the one in the currently planned setup, of 1 mbar. While this change does not alter the behaviour of the HS particles, it does increase the number of neutrino and muon DIS interactions inside the Decay Vessel, increasing the amount of

background events that we must remove, and also increases the amount of average interactions that the decay products from the HS particles suffer until they reach the Spectrometer Straw Tracker at the Hidden Sector Spectrometer.

Once again, 1.0×10^6 neutrino DIS interactions for electron and muon flavoured neutrinos and anti-neutrinos were simulated, using the same input files as previously. The range of forced interaction was also maintained between the beginning of the Muon ID system in the SND and the second Target Tracker of the Hidden Sector Spectrometer. The Hidden Sector samples also retained the parameter space points shown in Tables 4.2 and 4.4 for the HNLs and DPs respectively. A summary of the new samples, alongside the preliminary cuts, can be found in Table 8.1.

Sample	Events Generated	Particles Reconstructed	Reconstructions within the Decay Vessel	Events with allowed Decay Products	
$N \to e^- \pi^+$	1.2×10^5	26709	26181	23965	
$N \to \mu^- \pi^+$	1.2×10^5	27006	26676	26487	
$N \to e^- \rho^+$	9.0×10^4	25286	22203	19455	
$N \to \mu^- \rho^+$	9.0×10^4	27674	24526	22594	
$N \to \nu_\alpha \mu^+ \mu^-$	1.2×10^5	28671	28660	28624	
$N \to \nu_\alpha \rho^0$	9.0×10^4	21428	21074	20843	
$A' \to e^- e^+$	1.25×10^5	37024	34982	34169	
$A' \to \mu^- \mu^+$	5.0×10^4	17557	16746	16656	
$^*A' \to e^- e^+$	1.15×10^5	33508	31632	30978	
$^*A' \to \mu^- \mu^+$	4.0×10^4	14496	13781	13715	
ν_e DIS	1.0×10^6	129329	96975	36032	15624
$\overline{\nu_e}$ DIS	1.0×10^6	74777	54523	18633	14922
ν_μ DIS	1.0×10^6	31376	24068	10786	2168
$\overline{\nu_\mu}$ DIS	1.0×10^6	20988	16202	8718	1830

Table 8.1: Summary of all the samples generated with the **FairShip** software with a Decay Vessel set at a pressure of 1 bar. For the "Events within allowed Decay Modes" entries relative to the background samples, the left column relates to the HNL decay modes, and the right one to the DP modes. *Dark Photon samples without the events with masses of 3.6 and 4.4 GeV/c^2, which are not within the most recent sensitivity studies.

8.2 Kinematic Cuts

As stated previously, the first step in analysing the new batch of samples is to apply the cuts that were applied to the corresponding HS samples in Chapters 5 and 6. As such, in Table 8.2, we show the surviving events for each sample of the Hidden Sector, alongside the surviving background events for both sets of cuts.

Comparing the data from Table 8.2 with the one from Tables 5.4 and 6.3, we can clearly see that all the HS samples with decays involving electrons have lost significant amounts of their selection efficiency. From these, the HNL samples have suffered the most, with both decays involving electrons losing about one third of their selection efficiency, with a drop of 81% to 52% and 36% to 20%, for the pion and

Sample	Surviving Events after the Cuts (%)	Sample	Surviving Events after the Cuts (%)
$N \to e^- \pi^+$	13516 (51.6)	$A' \to e^- e^+$	27981 (80.0)
$N \to \mu^- \pi^+$	23231 (87.1)	$A' \to \mu^- \mu^+$	16101 (96.2)
$N \to e^- \rho^+$	5002 (22.5)	$^*A' \to e^- e^+$	25608 (81.0)
$N \to \mu^- \rho^+$	7418 (30.3)	$^*A' \to \mu^- \mu^+$	13253 (96.2)
$N \to \nu_\alpha \mu^+ \mu^-$	1092 (5.2)	HNL Background	107 (0.1)
$N \to \nu_\alpha \rho^0$	1351 (4.7)	DP Background	55 (0.0)

Table 8.2: Summary of the selection efficiency obtained by applying the kinematic cuts defined in Chapters 5 and 6 on the corresponding samples. All the efficiencies are relative to the HS candidate particles reconstructed within the Decay Vessel, presented in Table 8.1. *Dark Photon events without the 3.6 and 4.4 GeV/c^2 mass samples, that are not within the most recent sensitivity estimates.

rho decays respectively. On the other hand, samples that involve decays to muons do not show a clear tendency, because even though the HNL samples with semi-leptonic decays have slight decreases in selection efficiency, the HNL and DP samples that lead to the detection of muon pairs actually show an increase, even though it is only a net gain of around 1%.

The first phenomenon was anticipated, since now that the Decay Vessel is filled with air, the electrons resulting from decays of the HS particles will interact much more. While this does not necessarily lead to a sizeable decrease in the number of reconstructed particles, as can be seen from Table 8.1, the increasing number of interactions leads to a much bigger energy loss by the electrons, immediately skewing the Impact Parameter of the reconstructed particle. Since pions do not radiate as much as the electrons, due to their mass, an asymmetric momentum loss appears, increasing the Impact Parameter, which is the main selection criterion for our analysis. This is further corroborated by the much smaller decrease in selection efficiency of the DP to an electron pair sample of just 4%, since the momentum loss is now more symmetrical.

As for the increase in the selection efficiency of some samples, we cannot see a fundamental reason as to why this should happen, besides inherent statistical fluctuations.

The last noteworthy takeaway from Table 8.2 is that be it for the HNL or the DP background, the previous kinematic cuts are no longer reasonable. Even though most of the interactions happen in the Muon ID System of the Scattering and Neutrino Detector, the mere presence of air in the Decay Vessel leads to more neutrino DIS events, and, since our goal is to check whether or not the selection efficiencies can be maintained with this change, if our background sample remains at 4.0×10^6 neutrino interactions we need to keep the background rejection the same as in the previous case, at the very least. As such, following a procedure similar to the one in the analysis of the current SHiP setup, we applied a whole new set of kinematic cuts to these samples. Naturally, since there are much more reconstructed background events inside the Decay Vessel, it is expected that a lot more fringe ones will appear. As such, the amount of kinematic cuts with low relevance will increase in order to keep the 0 background events selected policy.

Due to the redundancy of the distributions, no plots of the used kinematic properties will be shown in this chapter. The same approach of taking preliminary loose cuts, finding combinations of features

that could remove the most fringe events, and then revisiting all cuts one by one in order to optimize the selection efficiency of the HS signal was taken.

8.2.1 Heavy Neutral Leptons

The loose preliminary cuts taken for the HNL samples was the set established during the first part of this thesis. After some tinkering, we reached at the following final 16 cuts:

Cut	Thresholds for Selection
1	IP $<$ 0.00138 $\times$ Decay Z cm, with Decay Z in cm.
2	IP $>$ 1.88, 2.89, 3.37, 6.0 or 10.0 cm requires TotMom of the Mother $<$ 163.2, 75.7, 63.8, 58.0 or 38.0 GeV/c, respectively. IP $<$ 0.61 cm requires TotMom of the Mother $>$ 57.6 GeV/c.
3	IP $>$ 12.0 cm requires TransMom of the Tracks $>$ 0.25 GeV/c.
4	IP $<$ 412 $\times$ (0.05 $-$ TransMom)^2 + 2.0 cm, where TransMom is from the Tracks in GeV/c.
5	IP $<$ 22182 $\times$ (0.0005 $-$ FracMom)^2 + 2.0 cm, where FracMom is from the Mother in GeV/c.
6	IP $>$ 7.4 cm requires TotMom of the Tracks $<$ 39.1 GeV/c.
7	Decay Z $>$ 183 cm.
8	OA $<$ 0.5 rad.
9	TransMom of the Tracks $<$ 46.8 $\times$ (0.23 $-$ OA)^2 + 0.50 GeV/c, for OA in rad.
10	OA $>$ 0.0195 rad requires TransMom of the Tracks $<$ 1.60 GeV/c.
11	OA $>$ 0.08 rad requires TotMom of the Tracks $<$ 65.0 GeV/c.
12	OA $>$ $-$14.13 $\times$ (0.022 $-$ FracMom)^2 + 0.0064 rad, where FracMom is from the Tracks.
13	FracMom of the Mother $<$ 0.0044 requires TransMom of the Tracks $<$ 1.00 GeV/c.
14	FracMom of the Mother $>$ 0.0185 or 0.0266 requires FracMom of the Tracks $>$ 0.0147 or 0.0180, respectively. FracMom of the Mother $<$ 0.0065, 0.0128 or 0.0183 requires FracMom of the Tracks $<$ 0.0451, 0.0597 or 0.0747, respectively.
15	TotMom of the Mother $<$ 29.0 GeV/c requires FracMom of the Mother $>$ 0.0080.
16	TotMom of the Tracks $<$ 4.51, 7.23 or 11.77 GeV/c requires a FracMom of the Tracks $>$ 0.0461, 0.0357 or 0.0147, respectively.

Table 8.3: Kinematic cuts applied to the HNL samples, and their respective background. Recall that Decay Z = 0 stands for the beginning of the Decay Vessel.

Due to the sheer amount of cuts, we do not present any Figures illustrating the kinematic cuts in this Section.

In the spirit of Chapter 5, Tables 8.4 and 8.5 show the event selection for all the cuts in Table 8.3, and a final selection efficiency broken down by decay mode of the samples, respectively. In the latter Table, we also present the selection efficiencies of the HNL samples when no cuts that remove 5 or fewer background events are applied. We denominated those entries as "Surviving Events without irrelevant Cuts" This makes it so that cuts 6, 8, 10, 11, 12, 13, 15 and 16 are not applied, leading to 11 background events that survive the cuts.

No matter how we look at it, requiring 0 background events to go through the kinematic filters leads to very substantial losses of the selection efficiency of HNLs. The only decay mode that still has a reasonable percentage of events selected is $N \to \mu^- \pi^+$. Comparing Table 8.2 with Table 8.5, it also becomes clear

that the extra cuts, and the tightening of the ones from Table 5.2, sap away at the selection efficiencies that were already fragilized from the increased number of interactions of the decay products with the air in the Decay Vessel. Other than that, there are no outlandish surprises, since the selection efficiency hierarchy remains similar to what was expected.

Cut	Background Events Selected (%)	HNL Events Selected (%)
1	328 (0.2)	65464 (43.8)
2	191 (0.1)	54318 (36.4)
3	185 (0.1)	53974 (36.2)
4	135 (0.1)	50049 (33.5)
5	52 (0.0)	42137 (28.2)
6	51 (0.0)	42063 (28.2)
7	27 (0.0)	41452 (27.8)
8	26 (0.0)	41449 (27.8)
9	19 (0.0)	40137 (26.9)
10	18 (0.0)	40035 (26.8)
11	16 (0.0)	39955 (26.8)
12	12 (0.0)	39194 (26.3)
13	11 (0.0)	39192 (26.3)
14	4 (0.0)	33242 (22.3)
15	3 (0.0)	33145 (22.2)
16	0 (0.0)	30889 (20.7)

Table 8.4: Selection efficiency of the kinematic cuts presented in this section, applied on the Heavy Neutral Lepton samples. All percentages are calculated against the total number of events reconstructed inside the Decay Vessel, presented in Table 8.1, which gives a total of 191768 background events, and 149320 HNL events.

Decay Mode	Surviving Events after the Cuts (%)	Surviving Events without irrelevant Cuts (%)
$N \to e^- \pi^+$	7046 (26.9)	7522 (28.7)
$N \to \mu^- \pi^+$	15710 (58.9)	16603 (62.2)
$N \to e^- \rho^+$	2676 (12.1)	3343 (15.1)
$N \to \mu^- \rho^+$	4441 (18.1)	5522 (22.5)
$N \to \nu_\alpha \mu^+ \mu^-$	396 (1.4)	448 (1.6)
$N \to \nu_\alpha \rho^0$	620 (2.9)	665 (3.2)

Table 8.5: Summary of the selection efficiency obtained by applying the kinematic cuts on the Heavy Neutral Lepton samples, detailed by Decay Mode. The Surviving Events without irrelevant Cuts are from samples where cuts that did not remove more than 5 background events are not applied. All the efficiencies are based on the Reconstructions within the Decay Vessel, presented in Table 8.1.

8.2.2 Dark Photons

In the same fashion as the previous subsection, the preliminary cuts established for the new DP samples were the ones done for the current SHiP setup. In order to remove all background events, the following

10 cuts were needed:

Cut	Thresholds for Selection
1	IP $<$ 16.0 cm.
2	IP $<$ 12.2 cm requires TotMom from the Mother $<$ 38.0 GeV/c.
3	TotMom of the Mother $>$ 1.7 GeV/c.
4	IP $< 9.93 \times (0.556 - \text{TotMom})^2 + 4.2$ cm, where TotMom is from the Tracks in GeV/c.
5	IP $< 50484 \times (0.0012 - \text{TransMom})^2 + 1.79$ cm, where TransMom is from the Tracks, in GeV/c. IP $>$ 4.2, 5.2, 6.3 or 6.4 cm requires TransMom of the Tracks $>$ 0.021, 0.048, 0.059 or 0.095 GeV/c, respectively. IP $>$ 8.9 cm requires TransMom of the Tracks $<$ 1.19 GeV/c. IP $> 63.86 \times \text{TransMom} + 1.8$ cm requires IP $< 291.97 \times \text{TransMom} - 1.3$ cm.
6	IP $>$ 2.6 or 7.9 cm requires OA $>$ 0.0021 or 0.0082 rad, respectively. IP $>$ 8.6 cm requires OA $<$ 0.112 rad.
7	IP $<$ 0.6 cm requires FracMom of the Mother $<$ 0.011. IP $>$ 9.5 cm requires FracMom of the Mother $<$ 0.042.
8	OA $<$ 0.0065 rad requires FracMom of the Mother $<$ 0.032. OA $>$ 0.1090 rad requires FracMom of the Mother $>$ 0.003.
9	Decay Z $>$ 329 cm.
10	TotMom of the Tracks $<$ 2.19 or 3.66 GeV/c requires FracMom of the Tracks $<$ 0.052 or 0.068, respectively.

Table 8.6: Kinematic cuts applied to the DP samples, and their respective background.

The resulting selection efficiencies from applying the aforementioned cuts to the Dark Photon samples can be found in Table 8.7, and the breakdown by decay mode, including the samples with masses exclusively within the most recent sensitivity study, can be seen in Table 8.8.

Although the selection efficiencies obtained for the Dark Photon samples are substantially higher than those obtained for the Sterile Neutrinos, the events that decay into pairs of electrons do not have particularly satisfying efficiencies. In this case no selection efficiency for low significance cuts is presented since it would only encompass the last 4 cuts. Additionally, notice that the ladder like selection from cut 5 has a lot of tailored cuts to remove the remaining background events. If we simply remove the criterion where an Impact Parameter bigger than 6.4 cm requires the Transverse Momentum of the Tracks to be bigger than 0.095 GeV/c, there is only 1 background event that passes through the filters, but the selection efficiencies are raised to 55.4% and 81.1%, for the full samples. Even though this is the most fringe background event, it is a clear reminder that our approach should only be taken as a term of comparison against the kinematic cuts applied in a Vacuum filled Decay Vessel, as we cannot perform a better study without a clear estimate of the background levels in the newly proposed setup.

8.3 Machine Learning

In order to analyze the samples with some machine learning algorithms, the approach taken in Chapter 7 was followed, where several configurations of MLPs were trained, in order to obtain the NNs that provided the best results. However, unlike in the previous study, we are no longer just looking for the best selection

Cut	Background Events Selected (%)	DP Events Selected (%)	DP Events Selected within new Sensitivity (%)
1	228 (0.1)	42688 (82.5)	37910 (83.5)
2	202 (0.1)	42073 (81.3)	37472 (82.5)
3	186 (0.1)	42062 (81.3)	37461 (82.5)
4	101 (0.1)	40813 (78.9)	36212 (79.7)
5	26 (0.0)	36103 (69.8)	31730 (69.9)
6	13 (0.0)	35832 (69.3)	31467 (69.3)
7	10 (0.0)	34819 (67.3)	30615 (67.4)
8	8 (0.0)	34746 (67.2)	30566 (67.3)
9	3 (0.0)	33632 (65.0)	29587 (65.2)
10	0 (0.0)	32487 (62.8)	28503 (62.8)

Table 8.7: Selection efficiency of the kinematic cuts presented in this section applied on the Dark Photon samples. All percentages are calculated against the total number of events reconstructed inside the Decay Vessel, presented in Table 8.1, which gives a total of 191768 background events and 51728 DP events, of which 45413 are within the most recent sensitivity estimates.

Decay Mode	Surviving Events after the Cuts (%)
$A' \to e^- e^+$	18983 (54.3)
$A' \to \mu^- \mu^+$	13504 (80.6)
$^*A' \to e^- e^+$	17433 (55.1)
$^*A' \to \mu^- \mu^+$	11070 (80.3)

Table 8.8: Summary of the selection efficiency obtained by applying the kinematic cuts in the Dark Photon samples, detailed by Decay Mode. Once again notice that all the efficiencies are based on the Reconstructions within the Decay Vessel, presented in Table 8.1. *Dark Photon events without the 3.6 and 4.4 GeV/c^2 mass samples, that are not within the most recent sensitivity estimates.

efficiencies for 0 background events selected, we are also looking at the NNs that provide the lowest number of background events, for selection efficiencies equal to those obtained by NNs during the study presented in Chapter 7. The comparison point was the selection efficiency relative to the number of events selected after the Decay Mode preliminary cuts.

For some samples, the best network for one goal was not the same as for the second, and this lead to the selection of a total 14 networks. The only cases where one NN was able to achieve both objectives were the ones for the $A' \to \mu^- \mu^+$ and $N \to \mu^- \pi^+$ samples. Additionally, we trained NNs with either just the 9 features used previously, or the 12 features used in the kinematic cuts presented in this section.

The basic network features that change depending on the NN can be seen in Tables 8.9 and 8.10, defined by training sample and respective background. A summary of the results obtained by the selected networks can be seen in Tables 8.11 and 8.12, for the 0 background, and similar selection efficiencies, respectively. Notice that the selection efficiencies presented in Table 8.12 are not exactly equal to those in Table 7.2 since these efficiencies are relative to the HS candidate particles reconstructed within the Decay Vessel, and not the particles that are also within the selected decay modes, which was used as the criterion.

Training Sample	Training Epochs	Batch Size	Number of Hidden Layers	Nodes on first Hidden Layer	Step between Hidden Layers	Information Dropout [%]
$N \to e^- \pi^+$	500	200	3	25	5	0
$^*N \to \mu^- \pi^+$	300	1000	3	25	5	0
$N \to e^- \rho^+$	500	200	3	25	5	5
$^*N \to \mu^- \rho^+$	600	200	3	25	5	5
$^*N \to \nu_\alpha \mu^+ \mu^-$	500	500	3	25	5	5
$N \to \nu_\alpha \rho^0$	600	200	2	30	10	8
$^*A' \to e^- e^+$	500	500	3	20	5	10
$^*A' \to \mu^- \mu^+$	800	200	3	20	5	10

Table 8.9: Summary of every NN trained to obtain the best selection efficiency for 0 background events selected, containing their basic properties, alongside some training choices. *Neural Networks trained on the same 9 features that were used for the current SHiP setup.

Training Sample	Training Epochs	Batch Size	Number of Hidden Layers	Nodes on first Hidden Layer	Step between Hidden Layers	Information Dropout [%]
$N \to e^- \pi^+$	600	200	3	25	5	8
$^*N \to \mu^- \pi^+$	300	1000	3	25	5	0
$^*N \to e^- \rho^+$	500	500	3	25	5	5
$N \to \mu^- \rho^+$	500	200	3	25	5	5
$^*N \to \nu_\alpha \mu^+ \mu^-$	500	500	3	25	5	5
$N \to \nu_\alpha \rho^0$	500	200	3	25	5	0
$^*A' \to e^- e^+$	800	500	3	25	5	10
$^*A' \to \mu^- \mu^+$	800	200	3	20	5	10

Table 8.10: Summary of every NN trained to obtain the same selection efficiencies as in the current SHiP setup, containing their basic properties, alongside some training choices. *Neural Networks trained on the same 9 features that were used for the current SHiP setup.

Comparing the results shown in Table 8.11 with those in Table 8.5, most of the selection efficiencies were improved. As expected, this was more substantial for the decay modes where a significant fraction of the reconstructions have missing energy. The sample with the biggest net gain was $N \to e^- \rho^+$, where the selection efficiency went from 12.1% to 23.1%, where as the sample with the biggest relative gain was $N \to \nu_\alpha \mu^- \mu^+$, with an improved selection efficiency of 7.6% from the underwhelming 1.4% events selected from the standard cuts. However, the $N \to \mu^- \pi^+$ sample is a clear outlier in this trend. The selection efficiency obtained through the Neural Network selection processes decreased to $\approx 75\%$ of the one obtained in Section 8.2, with a net 13.5% less HS events selected. While surprising, this might be an artifact from the criterion set for the network. As stated in Chapter 7, NNs are trained according to loss functions, and while the goal is to find the global minimum, this is almost impossible, and near-optimal solutions are used from the local minimums instead. An increase in the background events leads to an increase in the number of the ones almost indistinguishable from the HS, and makes finding the global minimum much harder. Alas, the score threshold defined for the 0 background criterion, and consequently the selection efficiency of the signal, becomes highly volatile. If we we compare the selection

efficiency for this sample when we allow for 11 background events to pass through the filter (Surviving Events without irrelevant Cuts), we obtain a total selection efficiency of 67.4%, that surpasses the 62.2% obtained through the standard kinematic cuts, which is more in line with the expected improvement of the NNs over the previous method.

Comparing the DP efficiencies from Table 8.8 with the NN ones (Table 8.9), we find a similar behaviour to that of the HNL samples. While the decay into a pair of muons shows selection efficiencies that increased by $\approx 5\%$ and $\approx 3\%$ for the full sample and the one within the most recent sensitivity limits, respectively, the decay to a pair of electrons shows a decrease of $\approx 3\%$ and $\approx 6\%$. When we allow for 10 false positives[1], the NN selected for 0 background events yields a selection efficiency of 74.0%, which is much bigger than the 60.5% obtained from the standard kinematic cuts, for the full sample, further corroborating the previously established idea.

Regarding the data in Table 8.12, we can safely state that the amount of false positives for the HNL samples is too big if we want to maintain the HNL selection efficiencies similar to those obtained in Chapter 5. As analysing the HNL data in Table 8.2 led us to believe, the decays involving electrons provided substantially more background than their muon counterparts. Additionally, the decays with missing energy also come with a lot of background events attached. The only surprise comes from the relative low background in the $N \to \ell^- \rho^+$ samples, but this is most likely due to the already lower selection efficiencies, when compared to the $N \to \ell^- \pi^+$ samples. As for the DP samples, the background is much more manageable with just 82 events that pass through the filters. However, this is not satisfactory unless the estimated background levels are around the 10^4 neutrino interactions for a Decay Vessel at a pressure of 1 bar, which is not likely since the estimates for the current setup are of the order of $\sim 10^6$.

Sample	Score Threshold	Signal Events Selected (%)	Signal Events Selected Applying all adequate NNs (%)
$N \to e^- \pi^+$	0.9990	6249 (23.9)	7849 (30.0)
$N \to \mu^- \pi^+$	0.9832	8047 (30.2)	12113 (45.4)
$N \to e^- \rho^+$	0.9979	3123 (14.1)	5120 (23.1)
$N \to \mu^- \rho^+$	0.9995	5356 (21.8)	6663 (27.2)
$N \to \nu_\alpha \mu^+ \mu^-$	0.9956	757 (2.6)	1497 (7.6)
$N \to \nu_\alpha \rho^0$	0.9970	927 (4.4)	1611 (5.2)
hline $A' \to e^- e^+$	0.9997	14694 (42.0)	18062 (51.6)
$A' \to \mu^- \mu^+$	0.9988	14010 (79.8)	14342 (85.6)
$*A' \to e^- e^+$	N.A.	N.A.	15656 (49.5)
$*A' \to \mu^- \mu^+$	N.A.	N.A.	11501 (83.5)

Table 8.11: Summary of the results obtained by the NNs that provided the best selection efficiencies for 0 background events selected. All selection efficiencies are relative to the number of HS candidate particles reconstructed within the Decay Vessel, presented in Table 8.1. *Samples within the most recent sensitivity estimates. N.A. stands for Not Applicable, since no NNs were trained on these samples.

[1]The number of background events that survive if we remove the last four cuts from the DP cuts in Section 8.2 is 10.

Sample	Score Threshold	Signal Events Selected (%)	Background Events Selected (%)
$N \to e^- \pi^+$	0.563	22590 (86.28)	720 (0.38)
$N \to \mu^- \pi^+$	0.829	24966 (93.59)	102 (0.05)
$N \to e^- \rho^+$	0.964	9689 (43.64)	57 (0.03)
$N \to \mu^- \rho^+$	0.980	10647 (43.41)	18 (0.01)
$N \to \nu_\alpha \mu^+ \mu^-$	0.947	11778 (41.10)	230 (0.12)
$N \to \nu_\alpha \rho^0$	0.883	14092 (66.87)	450 (0.23)
Total HNL Background	-	-	1001 (0.52)
$A' \to e^- e^+$	0.955	29857 (85.35)	70 (0.04)
$A' \to \mu^- \mu^+$	0.734	16337 (97.56)	34 (0.02)
Total DP Background	-	-	82 (0.04)

Table 8.12: Summary of the results obtained by the NNs that provided the least amount of background events selected in order to obtain the same HS selection efficiencies obtained in the studies presented in Chapter 7. All selection efficiencies are relative to the number of HS candidate particles reconstructed within the Decay Vessel, presented in Table 8.1.

Chapter 9

Concluding remarks

Although the Standard Model is very successful, it is not a complete theory and BSM physics are required to explain several oustanding phenomena. SHiP is a cutting edge, intensity frontier experiment that will try to provide experimental evidence for, and discover, theorized particles from the Hidden Sector, such as the Heavy Neutral Leptons, Dark Photons or Axion Like Particles, that solve several SM shortcomings. The goal is to discover these particles based on the direct observation of at least two of their decays into SM products. This requires a 0 background environment with several redundant veto systems. If these are poorly optimized, the parameter space that can be probed will be compromised.

In the present thesis, we suggested several optimized background veto criteria for the search of HNLs and DPs based on the kinematic properties of the reconstructed particles. We started by providing a theoretical introduction to the SM and Hidden Sector extensions containing HNLs and DPs, alongside their phenomenological implications and experimental signatures at SHiP, in Chapters 2 and 3. With the parameter spaces well defined, we generated HNL samples around the 1 GeV/c^2 region of interest, and DP samples that provided a full sweep of the parameter space, with special focus for masses below 0.4 GeV/c^2. The most relevant decay modes for both HS particles were considered and, in the case of the HNLs, decays with missing products (energy) were also studied. All the relevant background sources at SHiP were considered. In Chapters 5 and 6 we presented a kinematic cut-based approach exploring in particular the discriminating power of the Impact Parameter to remove all the background events generated, while maintaining the selection efficiencies of the HS particles as high as possible. Starting with signal and background samples of comparable size, we managed to provide selection efficiencies above 80% for both the $N \to \ell^{\mp}\pi^{\pm}$ and $A' \to \ell^{-}\ell^{+}$ samples, while removing all of the generated background events. The samples where the reconstructions had missing energy, however, had poor selection efficiencies, of less than 7%. In Chapter 7, we used a Machine Learning approach in order to separate the HS signal from the background. This approach showed improvements upon the previous method, leading to selection efficiencies of over 97% for the $N \to \mu^{\mp}\pi^{\pm}$ and $A' \to \mu^{-}\mu^{+}$ samples. The missing energy samples were the most improved with new selection efficiencies of over 51%.

Having established satisfying selection efficiencies for standard detector conditions, we proceeded to verify in Chapter 8 whether the Decay Vessel vacuum conditions could be relaxed. We generated HS

samples with the exact same specifications as previously, apart from imposing an air pressure inside the Decay Vessel of 1 bar. Only neutrino DIS interactions were considered as background this time. We proceeded to perform the cut-based analysis on the new samples, followed by the use of Machine Learning techniques. It became apparent from both analyses that the new conditions heavily degrade the selection efficiencies of the HS particles, even if we assume that the expected background events stay the same. The presence of more particles inside the Decay Vessel enhance the number of interactions particularly of the electrons coming from HS particle decays, leading to higher values of the Impact Parameter, which was one of the main distinguishing factors. We conclude accordingly that the extreme scenario where the SHiP experiment would be conducted at atmospheric pressure (in place of the standard vacuum conditions foreseen) would result in drastic losses in sensitivity, and thus should not be viable.

An accurate estimation of the background yields expected, for realistic beam and detector conditions, would allow to further tune the selection criteria. This could allow to remove statistically irrelevant cuts and features, or to relax the thresholds established for the Neural Network classifiers. Enhanced selection efficiencies for the HS particles could thus potentially be attained, in particular for HNLs that decay into ρ mesons or SM neutrinos. The Distance of Closest Approach between decay products, which has not been used, while partially correlated with other features, may provide additional discriminating power. Further studies, therefore, may use it, as it might be able to improve the selection efficiencies obtained through both of the used methods.

Even though obtaining near 100% selection efficiencies for particles coming from the Hidden Sector is crucial at SHiP, it is not the end of the journey. There are a lot of HS particles that share decay modes into the SM, and as such just detecting events that pass all the veto systems does not guarantee the discovery of any specific particle, but rather the discovery of a Hidden Sector (e.g. $N \to \nu_\alpha \mu^- \mu^+$ and $A' \to \mu^- \mu^+$ have the same detectable products). In order to claim the discovery of a specific particle, when we see these decay modes, we need to establish criteria capable of distinguishing different types of signals. Alas, this kind of separation studies can be done by continuing the work provided in this thesis.

Bibliography

[1] R. Oerter, B. Holstein, *Physics Today - PHYS TODAY* **59** (2006).

[2] C. Gignoux, B. Silvestre-Brac, J. Richard, *Physics Letters B* **193**, 323 (1987).

[3] H. J. Lipkin, *Physics Letters B* **195**, 484 (1987).

[4] A. Buonaura, *Journal of Physics: Conference Series* **1586**, 012031 (2020).

[5] SHiP Collaboration, *Reports on Progress in Physics* **79**, 124201 (2016).

[6] SHiP Collaboration, A facility to search for hidden particles (ship) at the cern sps (2015).

[7] Y. Cai, T. Han, T. Li, R. Ruiz, Lepton number violation: Seesaw models and their collider tests (2018).

[8] T. Aoyama, *et al.*, The anomalous magnetic moment of the muon in the standard model (2020).

[9] L. Okun, *Sov. Phys. JETP* **56**, 502 (1982).

[10] SHiP Collaboration, Sensitivity of the ship experiment to heavy neutral leptons (2019).

[11] SHiP Collaboration, Sensitivity of the ship experiment to dark photons decaying to a pair of charged particles (2020).

[12] D. B. M. Al-Turany et al., *J.Phys.Conf.Ser.* **396**, 022001 (2012).

[13] SHiP Collaboration, SHiP Experiment - Comprehensive Design Study report, *Tech. Rep. CERN-SPSC-2019-049. SPSC-SR-263*, CERN, Geneva (2019).

[14] S. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions (2017).

[15] S. Chatrchyan, *et al.*, *Physics Letters B* **716**, 30 (2012).

[16] G. Aad, *et al.*, *Physics Letters B* **716**, 1 (2012).

[17] S. Glashow, *Nucl. Phys.* **22**, 579 (1961).

[18] F. Englert, R. Brout, *Phys. Rev. Lett.* **13**, 321 (1964).

[19] P. W. Higgs, *Phys. Rev. Lett.* **13**, 508 (1964).

[20] G. S. Guralnik, C. R. Hagen, T. W. B. Kibble, *Phys. Rev. Lett.* **13**, 585 (1964).

[21] P. W. Higgs, *Phys. Rev.* **145**, 1156 (1966).

[22] S. Weinberg, *Phys. Rev. Lett.* **19**, 1264 (1967).

[23] A. Salam, *Conf. Proc. C* **680519**, 367 (1968).

[24] M. Gell-Mann, *Phys. Lett.* **8**, 214 (1964).

[25] G. Zweig, *An SU(3) model for strong interaction symmetry and its breaking. Version 2* (1964), pp. 22–101.

[26] G. 't Hooft, M. Veltman, *Nucl. Phys. B* **44**, 189 (1972).

[27] *Physics Reports* **427**, 257–454 (2006).

[28] R. Aleksan, S. Jadach, *Physics Letters B* **799**, 135034 (2019).

[29] L. Landau, *Nucl. Phys.* **3**, 127 (1957).

[30] T. D. Lee, C. N. Yang, *Phys. Rev.* **105**, 1671 (1957).

[31] A. Salam, *Nuovo Cim.* **5**, 299 (1957).

[32] N. Cabibbo, *Phys. Rev. Lett.* **10**, 531 (1963).

[33] M. Kobayashi, T. Maskawa, *Progress of Theoretical Physics* **49**, 652 (1973).

[34] Super-Kamiokande Collaboration, *Phys. Rev. Lett.* **81**, 1562 (1998).

[35] SNO Collaboration, *Phys. Rev. Lett.* **87**, 071301 (2001).

[36] G. W. Bennett, *et al.*, *Physical Review D* **73** (2006).

[37] F. Zwicky, *Helvetica Physica Acta* **6**, 110 (1933).

[38] S. Smith, *Astrophys. J.* **83**, 23 (1936).

[39] S. M. Faber, J. S. Gallagher, *Annual Review of Astronomy and Astrophysics* **17**, 135 (1979).

[40] P. Minkowski, *Physics Letters B* **67**, 421 (1977).

[41] S. Correia, R. Felipe, F. Joaquim, *Physical Review D* **100** (2019).

[42] T. Asaka, M. Shaposhnikov, *Physics Letters B* **620**, 17 (2005).

[43] M. Shaposhnikov, *Nuclear Physics B* **763**, 49–59 (2007).

[44] D. Gorbunov, M. Shaposhnikov, *Journal of High Energy Physics* **2007**, 015–015 (2007).

[45] E. K. Akhmedov, V. A. Rubakov, A. Y. Smirnov, *Physical Review Letters* **81**, 1359–1362 (1998).

[46] D. M. Barreiros, A minimal seesaw model for neutrino masses and the origin of matter, Master's thesis, Instituto Superior Técnico (2017).

[47] E. Majorana, *Nuovo Cim.* **14**, 171 (1937).

[48] M. Aker, *et al.*, *Physical Review Letters* **123** (2019).

[49] B. Pontecorvo, *Sov. Phys. JETP* **6**, 429 (1957).

[50] B. Pontecorvo, *Sov. Phys. JETP* **7**, 172 (1958).

[51] Z. Maki, M. Nakagawa, S. Sakata, *Progress of Theoretical Physics* **28**, 870 (1962).

[52] B. Pontecorvo, *Sov. Phys. JETP* **26**, 984 (1968).

[53] A. Boyarsky, A. Neronov, O. Ruchayskiy, M. Shaposhnikov, I. Tkachev, *Physical Review Letters* **97** (2006).

[54] A. Boyarsky, A. Neronov, O. Ruchayskiy, M. Shaposhnikov, *Journal of Experimental and Theoretical Physics Letters* **83**, 133–135 (2006).

[55] T. Asaka, M. Shaposhnikov, M. Laine, *Journal of High Energy Physics* **2007**, 091–091 (2007).

[56] T. Asaka, S. Blanchet, M. Shaposhnikov, *Physics Letters B* **631**, 151–156 (2005).

[57] A. Strumia, F. Vissani, *Nuclear Physics B* **726**, 294–316 (2005).

[58] S. Petcov, *Physics Letters B* **110**, 245 (1982).

[59] *Nuclear Physics B - Proceedings Supplements* **118**, 267 (2003).

[60] G. Altarelli, R. Franceschini, *Journal of High Energy Physics* **2006**, 047 (2006).

[61] K. Bondarenko, A. Boyarsky, D. Gorbunov, O. Ruchayskiy, *Journal of High Energy Physics* **2018** (2018).

[62] W. Bonivento, *et al.*, Proposal to search for heavy neutral leptons at the sps (2013).

[63] E. Graverini, SHiP sensitivity to Heavy Neutral Leptons (2016).

[64] M. Al-Turany, *et al.*, *Journal of Physics: Conference Series* **396**, 022001 (2012).

[65] J. Lees, *et al.*, *Physical Review Letters* **113** (2014).

[66] E. Goudzovski, *EPJ Web of Conferences* **96**, 01017 (2015).

[67] P. Galison, A. Manohar, *Physics Letters B* **136**, 279 (1984).

[68] B. Holdom, *Physics Letters B* **166**, 196 (1986).

[69] P. Langacker, *Rev. Mod. Phys.* **81**, 1199 (2009).

[70] J. D. Bjorken, R. Essig, P. Schuster, N. Toro, *Physical Review D* **80** (2009).

[71] D. Gorbunov, A. Makarov, I. Timiryasov, *Phys. Rev. D* **91**, 035027 (2015).

[72] J. Beacham, *et al.*, *Journal of Physics G: Nuclear and Particle Physics* **47**, 010501 (2019).

[73] SHiP Collaboration, A facility to search for hidden particles at the cern sps: the ship physics case (2015).

[74] SHiP Collaboration, SHiP Experiment - Progress Report, *Tech. Rep. CERN-SPSC-2019-010. SPSC-SR-248*, CERN, Geneva (2019).

[75] DONUT Collaboration, *Physics Letters B* **504**, 218–224 (2001).

[76] OPERA Collaboration, *Physics Letters B* **691**, 138–145 (2010).

[77] LHCb Collaboration, LHCb Tracker Upgrade Technical Design Report, *Tech. Rep. CERN-LHCC-2014-001. LHCB-TDR-015* (2014).

[78] L. Lopes, *et al.*, *Journal of Instrumentation* **9**, C10023 (2014).

[79] C. Franco, The SHiP experiment at CERN and the role of the LIP group within the collaboration [Conference Session], Jornadas LIP - 2020.

[80] A. Blanco, *et al.*, *Journal of Instrumentation* **15**, C10017 (2020).

[81] A. Korzenev, *et al.*, *Proceedings of the 5th International Workshop on New Photon-Detectors (PD18)* (2019).

[82] E. Zeballos, *et al.*, *Nuclear Instruments Methods in Physics Research Section A-accelerators Spectrometers Detectors and Associated Equipment* **374**, 132 (1996).

[83] T. Sjöstrand, S. Mrenna, P. Skands, *Computer Physics Communications* **178**, 852–867 (2008).

[84] S. Agostinelli, *et al.*, *Nucl. Instrum. Meth. A* **506**, 250 (2003).

[85] C. Andreopoulos, *et al.*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **614**, 87–104 (2010).

[86] T. Sjöstrand, S. Mrenna, P. Skands, *Journal of High Energy Physics* **2006**, 026–026 (2006).

[87] E. Van Herwijnen, H. Dijkstra, M. Ferro-Luzzi, T. Ruf (2015).

[88] B. Hosseini, W. M. Bonivento, Particle Identification tools and performance in the SHiP Experiment (2017).

[89] C. Ahdida, *et al.*, *Journal of Instrumentation* **14**, P11028 (2019).

[90] SHiP Collaboration, Heavy Flavour Cascade Production in a Beam Dump (2015).

[91] W. Altmannshofer, S. Gori, M. Pospelov, I. Yavin, *Physical Review D* **89** (2014).

[92] W. Altmannshofer, S. Gori, M. Pospelov, I. Yavin, *Physical Review Letters* **113** (2014).

[93] P. C. Bhat, *Submitted to Ann.Rev.Nucl.Part.Sci.* **61** (2010).

[94] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning* (MIT Press, 2016). `http://www.deeplearningbook.org`.

[95] S. Ruder, An overview of gradient descent optimization algorithms (2017).

[96] F. Chollet, *et al.*, Keras, `https://github.com/fchollet/keras` (2015).

[97] M. Abadi, *et al.*, TensorFlow: Large-scale machine learning on heterogeneous systems (2015). Software available from tensorflow.org.

[98] L. S. Shapley, *Notes on the n-Person Game —II: The Value of an n-Person Game* (RAND Corporation, Santa Monica, CA, 1951).