

Durham E-Theses

Phenomenology of Scalar Particles Assisted by Machine Learning

EDWIN ALI HERRERA CHACON

How to cite:

HERRERA CHACON, EDWIN ALI (2025) Phenomenology of Scalar Particles Assisted by Machine Learning. Doctoral thesis, Durham University.

Use policy



This work is licensed under a [Creative Commons Attribution Non-commercial No Derivatives 3.0 United States \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/3.0/)

Phenomenology of scalar particles assisted by machine learning

Edwin Ali Herrera-Chacon

A Thesis presented for the degree of
Doctor of Philosophy



Institute for Particle Physics Phenomenology
Department of Physics
Durham University
United Kingdom

March 2025

Phenomenology of scalar particles assisted by machine learning

Edwin Ali Herrera-Chacon

Submitted for the degree of Doctor of Philosophy

March 2025

Abstract: In this thesis, we explore the phenomenology of scalar particles within Beyond Standard Model (BSM) frameworks, using Machine Learning techniques to enhance sensitivity and discovery potential at current and future collider experiments, the Large Hadron Collider (LHC) and the High-Luminosity LHC (HL-LHC). Specifically, we study scalar extensions of the Standard Model (SM) such as the Two Higgs Doublet Model Type-III (2HDM-III) and the Froggatt-Nielsen Flavon model.

We perform a detailed collider analysis focusing on charged Higgs boson pair production within the 2HDM-III, examining final states involving muons, neutrinos and quark jets. Our studies identify parameter regions consistent with recent experimental anomalies reported by the ATLAS collaboration, particularly in charged Higgs decays involving charm-bottom quark transitions, and suggest concrete scenarios for achieving statistically significant signals of 5σ at future luminosities.

In the context of the Flavon model, we analyse potential signatures of a new scalar called Flavon decaying into a Higgs boson and a pair of bottom quarks, followed by the channels where the Higgs decays into a pair of bottom quarks or a pair of photons. Additionally, we analyse Lepton-Flavour-Violating (LFV) processes, both

of them achieving discovery level significances of up to 5σ at the HL-LHC.

Using multivariate analysis techniques, specifically Boosted Decision Trees (BDTs), we demonstrate a significant improvement in signal discrimination. Throughout this thesis, Machine Learning methodologies have been integral, notably enhancing the signal from background separation and significantly improving the robustness of phenomenological predictions. The methods and analyses presented here contribute to clarifying the flavour structure mysteries of the SM and offer actionable targets for future experimental searches.

Contents

Abstract	3
List of Figures	9
List of Tables	15
List of Abbreviations	17
1 Introduction	29
1.1 The Standard Model	29
1.2 Research Aims and Key Contributions	31
1.3 Structure of This Thesis	32
2 Theoretical Framework	33
2.1 Gauge Symmetries and the Field Content	33
2.1.1 Local Gauge Group	33
2.1.2 Standard Model Lagrangian	34
2.1.3 Flavour Diagonalisation and CKM	38
2.2 Motivations for Extending the Scalar Sector	39
2.3 Overview of BSM Models	40

3	Beyond Standard Models	43
3.1	Two-Higgs Doublet Model of Type III	43
3.1.1	Scalar Potential	43
3.1.2	Yukawa Lagrangian	47
3.2	Constraints on the 2HDM-III Parameter Space	51
3.2.1	Constraint on $\tan\beta$ and $\cos(\alpha - \beta)$	52
3.2.2	Constraint on M_H , M_A , and M_{H^\pm}	61
3.3	Flavon Model	73
3.3.1	Scalar Potential	74
3.3.2	Yukawa Lagrangian	76
3.4	Constraints on the FNSM Parameter Space	78
3.4.1	Theoretical Constraints	78
3.4.2	Experimental Constraints	81
4	Methodology BDT Data Analysis	85
4.0.1	The LHCO File Format	87
4.0.2	Column Definitions and Structure	87
4.0.3	Example for Reading a LHCO	87
4.0.4	Setup and Usage of Our Python Framework	93
4.0.5	Adding New Observables	94
4.0.6	BDT Training	99
5	Collider Analysis	105
5.1	Collider Analysis for the 2HDM-III	105
5.1.1	Signal and Background	106

5.1.2	Signal Significance	111
5.2	Collider Analysis for the FNSM	115
5.2.1	Multivariate Analysis	125
5.2.2	Lepton-Flavour-Violating Decays $h \rightarrow e\mu$	135
6	Conclusions and Future Work	143
A	Constraints	147
B	2HDM-III Variables and Parameters	153
	Bibliography	157

List of Figures

3.1	Generic Feynman diagram for the decays of a neutral meson M (such as B^0 , K^0 , or D^0) into $\mu^+\mu^-$. The black circle indicates a Flavour-changing vertex in the quark sector.	53
3.2	Feynman diagrams contributing to a_μ . Here ϕ represents a CP-even scalar, CP-odd scalar, or the SM-like Higgs boson. H^\pm denotes charged scalar bosons. In these diagrams $\ell_i = \mu$	56
3.3	Feynman diagram for process $\ell_i \rightarrow \ell_j \gamma$ at one-loop level induced by $\phi = h, H, A$	58
3.4	Feynman diagrams contributing to the process $\ell_i \rightarrow \ell_j \ell_k \bar{\ell}_k$. Panel (a) shows the tree-level and (b) one-loop level, where the black circle denotes a loop of the type as Feynman diagram of Fig. 3.3.	59
3.5	Feynman diagram for the tree-level decay $h \rightarrow \ell_i \ell_j$	59
3.6	In the $\cos(\alpha - \beta) - \tan \beta$ plane, the blue points represent parameter values allowed by all signal strength modifiers μ_X , whereas the orange points indicate those consistent with the LFV processes. The dataset was generated using the SpaceMath framework.	61
3.7	Feynman diagram for the production of ϕ in association with a bottom quark at the LHC, followed by its decay into a $\tau\tau$ pair.	62

- 3.8 Observed and expected 95% CL upper limits on the production cross-section multiplied by the di-tau BR for a scalar boson produced in association with a bottom quark. The results are shown as a function of (a) M_A and (b) M_H , where we consider $\tan\beta = 5, 10, 20$ and $\cos\alpha_\beta = 0.01$ 63
- 3.9 Feynman diagram illustrating the production process of H and its subsequent decay into a pair of Higgs bosons hh 65
- 3.10 The 95% CL upper limits on the production cross-section multiplied by the di-Higgs BR are presented for a scalar boson produced in pp collisions as a function of M_H . We observed and expected limits are shown, considering $\tan\beta = 5, 10, 20$ and $\cos(\alpha - \beta) = 0.01$ 66
- 3.11 Radiative decay $b \rightarrow s\gamma$ mediated by charged Higgs bosons. The loop involves up-type quarks (u, c, t), with the dominant contribution coming from the top quark. The photon is emitted from the charged scalar line. 67
- 3.12 Scatter plot in the $M_{H^\pm} - \tan\beta$ plane. (a) Red points stand for these regions allowed by R_{quark} for the mass interval $114 \leq M_{H^\pm} \leq 140$ GeV. (b) The same as in (a) but for the mass range $m_t + m_b \leq M_{H^\pm} \leq 1000$. The R_{quark} is defined in Eq. (3.2.27). 68
- 3.13 Scatter plot in the $\chi_{bb} - \chi_{tt}$ plane. Red points represent parameter combinations allowed by allowed by R_{quark} as defined in Eq. (3.2.27). 69
- 3.14 $\chi_{cb} - \chi_{tb}$ plane. The coloured points indicate the values of χ parameters that accommodate the current excess of events reported by the ATLAS collaboration. (a) Blue points represent parameter values that explain the excess for $M_{H^\pm} = 120$ GeV. (b) and (c) show the same as in (a) but for $M_{H^\pm} = 130$ GeV and $M_{H^\pm} = 140$ GeV, respectively. 71
- 3.15 Allowed points (simultaneously) for the difference of scalar masses by oblique parameters. 73

3.16	The VEV of the FN singlet v_s as a function of the cosine of the mixing angle α . The red points denote the region of parameter space that is consistent with all the theoretical constraints discussed in the text.	80
3.17	This plot shows the ultraviolet mass scale Λ as a function of the VEV of the complex singlet v_s . The area enclosed by the solid black lines corresponds to the intersection of all constraints, while the different points represent the individual \mathcal{R}_X measurements. The green rectangle indicates the parameter space allowed by the LFVp constraints.	83
5.1	Branching Ratios (a) $BR(H^+ \rightarrow cb)$, (b) $BR(H^+ \rightarrow \mu\nu)$, as a function of the charged scalar mass M_{H^+}	107
5.2	Feynman diagrams of the production cross-section of the signal $pp \rightarrow H^+ H^- \rightarrow \mu\nu_\mu cb$	108
5.3	Production cross-section of a charged Higgs pair with their subsequent decays into $\mu\nu_\mu cb$ for $M_H = 500$ GeV.	109
5.4	Production cross-section for a pair of charged Higgs bosons followed by their subsequent decays into $\mu\nu_\mu cb$, for $M_H = 800, 1000$ GeV in scenario S2	110
5.5	Plot of the discriminant for signal and background data.	111
5.7	Kinematic distributions for the signal and SM background processes are shown for the following observables: (a) transverse momentum of the b -jet, (b) missing transverse energy \cancel{E}_T , (c) transverse momentum of the c -jet, (d) transverse mass of the muon $M_T[\mu]$ and (e) invariant mass of one of the charged Higgs bosons decaying into a bc -jet pair, $M_{\text{inv}}[bc]$. In all plots the signal is labeled as SGN , while the total background is labeled as Full_BGD	114
5.8	<i>Signal Significance</i> as a function of the charged Higgs boson mass M_{H^\pm} and the integrated luminosity for the following scenarios: (a) S2 , (b) S3 with $M_H = 500$ GeV, and (c) S2 with $M_H = 800$ GeV.	116

5.9	Feynman diagrams inducing the $H_F \rightarrow \bar{f}fh$ decay in the FNSM. Diagram (a) shows an effective vertex corresponding to a dimension-5 operator suppressed by $1/\Lambda$. Such interactions are non renormalisable and are treated within the effective field theory framework, where higher dimensional operators are systematically incorporated and suppressed by powers of the new physics scale Λ , as shown in Eq. (3.3.15). This ensures the consistency of the model at energies below Λ	117
5.10	Production cross-section of the signal $pp \rightarrow H_F \rightarrow h\bar{b}b$. The centre-of-mass energy was set to 14 TeV.	120
5.12	b -jet transverse momentum normalised distributions for the signal and total background following the application of the acceptance cuts.	123
5.13	Normalised distribution of the reconstructed invariant mass $M_{\text{inv}}(b_1b_2b_3b_4)$ for the signal and background processes.	124
5.15	Normalised p_T distributions for photons (γ) and b -jets, comparing signal and total background after acceptance cuts.	127
5.16	Normalised distribution of the reconstructed invariant mass $M_{\text{inv}}(\gamma_1\gamma_2b_1b_2)$ for the signal and background processes.	128
5.17	Density plot showing the <i>Signal Significance</i> for the $h \rightarrow \bar{b}b$ channel as a function of the integrated luminosity and the Flavon mass M_{H_F} , the subplots correspond to the following benchmark scenarios: (a) S1 , (b) S2 , and (c) S3	130
5.18	ROC curve for $h \rightarrow \bar{b}b$ channel showing the model's ability to predict training data (used to fit the model) and independent test data (evaluating performance on new, unseen data). Each plot corresponds to a different Flavon mass as indicated.	132

5.19	1/Background Acceptance against the Signal Acceptance for the $h \rightarrow b\bar{b}$ channel. Each plot corresponds to a different Flavon mass as indicated.	133
5.20	Density plot displaying the <i>Signal Significance</i> as a function of the integrated luminosity and the Flavon mass M_{HF} for the $h \rightarrow \gamma\gamma$ decay channel, the subplots correspond to the following benchmark scenarios: (a) S1 , (b) S2 , and (c) S3	134
5.21	Feynman diagram of the LFV process $h \rightarrow e\mu$	136
5.22	Feature importances from the XGBoost classifier, showing which variables contribute most strongly to separate signal from background for $h \rightarrow e\mu$. (Correlated features were removed prior to training, as discussed in the dataset preparation section of the previous chapter.)	137
5.23	Parameter space in the plane $(v_s, Z_{e\mu})$ for $h \rightarrow e\mu$. Since this decay has not been observed, no signal strength is measured for this channel. Instead, only upper limits on the BR are available. In this analysis, we impose the experimental constraint $\mathcal{B}(h \rightarrow e\mu) < 6.2 \times 10^{-5}$. . .	138
5.24	Boosted Decision Tree (BDT) classifier distribution for S6 , comparing signal (blue) and background (red) events in both training and test samples. The legend shows the KS p-values for each distribution. . .	139
5.25	ROC curve for S6 , illustrates purity vs. efficiency for both, the training (blue) and test (red) datasets. The AUC values are ~ 0.996 (train) and ~ 0.996 (test), reflecting strong discrimination.	139
5.26	<i>Signal Significance</i> as a function of the integrated luminosity and the XGB cut for the decay $h \rightarrow e\mu$ in the scenarios: (a) S4 , (b) S5 , and (c) S6	141
A.1	Signal strengths: (a) μ_b , (b) μ_τ , (c) μ_W , (d) μ_Z , (e) μ_γ ($114 \leq M_{H^\pm} \leq 140$), and (f) μ_γ ($m_t + m_b \leq M_{H^\pm} \leq 1000$). In all cases, 50K random points are generated.	148

A.2	Allowed values by the measurement on $\mathcal{BR}(B_d^0 \rightarrow \mu^+ \mu^-)$ (black triangles) and by the upper limit on $\mathcal{BR}(B_s^0 \rightarrow \mu^+ \mu^-)$ (light red circles). For both cases we generate 50K random points.	149
A.3	Radiative decays $\ell_i \rightarrow \ell_j \gamma$: (a) $\mu \rightarrow e \gamma$, (b) $\tau \rightarrow \mu \gamma$, (c) $\tau \rightarrow e \gamma$. For all cases, we generate 100K random points.	150
A.4	Decays $\ell_i \rightarrow \ell_j \ell_k \bar{\ell}_k$: (a) $\mu \rightarrow 3e$, (b) $\tau \rightarrow 3\mu$. In all cases we generate 100K random points.	151
A.5	Allowed values by the upper limit on $\mathcal{BR}(h \rightarrow \tau \mu)$. We generate 50K random points.	152

List of Tables

3.1	The fact that the up-type quark couples to the doublet ϕ_2 is by convention. Type III exhibits a flavour violation.	49
3.2	The coefficients for the ϕ -Fermion couplings in the 2HDM-III with a \mathcal{CP} -conserving Higgs potential.	50
3.3	Scanned range of the parameters. The values of χ_{ij} are set to 1 for fermions i and j not included in the Table (in general $i \neq j$).	69
3.4	Scanned range of the parameters. The values of χ_{ij} are set to 1 for fermions i and j not included in the Table (in general $i \neq j$).	70
3.5	Benchmark scenarios to be used in the subsequent calculations.	73
3.6	Diagonal SXX interactions, ($S = h, H_F, A_F$).	77
3.7	Scanned parameters for theoretical viability in the FNSM.	80
3.8	Scanned parameters for \mathcal{R}_X evaluation in the FNSM.	82
4.1	LHCO File Column Definitions.	88
4.2	Detailed example of LHCO event data formatted for multiple event blocks.	88
5.1	Leading contributions to the cross-section and BRs for scenario S2	108
5.2	Cross-section of the dominant SM background processes.	108
5.3	Scenarios (S1 , S2 , S3) used in the calculations.	119

5.4	Cross-section of the signal for scenarios S1 , S2 , S3	120
5.5	Cross-section of the SM background processes for the $h \rightarrow b\bar{b}$ channel.	121
5.6	Cross-section of the SM background processes for the $h \rightarrow \gamma\gamma$ channel.	125
5.7	The <i>Signal Significance</i> for the decay channel $h \rightarrow b\bar{b}$ evaluated at $\mathcal{L}_{\text{int}}=3000 \text{ fb}^{-1}$	131
5.8	Parameters for scenarios S4 , S5 and S6	138
A.1	Range of scanned parameters for the signal strengths μ_X	147
A.2	Range of scanned parameters. On the left: $B_d^0 \rightarrow \mu^+\mu^-$, on the right: $B_s^0 \rightarrow \mu^+\mu^-$	149
A.3	Range of scanned parameters. On the left: $\mu \rightarrow e\gamma$, on the right: $\tau \rightarrow \mu\gamma$, and below: $\tau \rightarrow e\gamma$	151
A.4	Range of scanned parameters. On the left: $\mu \rightarrow 3e$, on the right: $\tau \rightarrow 3\mu$	152
A.5	Range of scanned parameters associated to the decay $h \rightarrow \tau\mu$. . .	152
B.1	Variables used to train and test the signal and background events list. The transverse mass is defined as $M_T[\ell] = \sqrt{2P_T^\ell E_T(1 - \cos \phi_{\ell E_T})}$. .	153

List of Abbreviations

2HDM-III	Two-Higgs Doublet Model of type III	31
ATLAS	A Toroidal LHC Apparatus	31
AUC	Area Under a ROC Curve	102
BDT	Boosted Decision Tree	13
BR	Branching Ratio	52
BSM	Beyond the Standard Model	31
C	Charge Conjugation	30
CKM	Cabibbo-Kobayashi-Maskawa	30
CP	Charge Conjugation Parity	30
EW	Electroweak	34
EWSB	Electroweak Symmetry Breaking	32
FCNCs	Flavour-changing Neutral Currents	38
FNSM	Froggatt-Nielsen Singlet Model	73
FN	Froggatt-Nielsen	31
HEP	High Energy Physics	99
HL-LHC	High-Luminosity LHC	31
KS	Kolmogorov-Smirnov	102
LHC	Large Hadron Collider	31

LHCO	LHC Olympics	32
LFV	Lepton-Flavour-Violating	31
LFVp	Lepton Flavour Violating Processes	78
MET	Missing Transverse Energy	87
PMNS	Ponte-corvo-Maki-Nakagawa-Sakata	30
QCD	Quantum Chromodynamics	33
ROC	Receiver Operating Characteristic	102
SM	Standard Model	29
SSB	Spontaneous Symmetry Breaking	29
VEV	Vacuum Expectation Value	29
XGBoost	Extreme Gradient Boosting	100

Declaration

The work in this thesis is based on research carried out in the Department of Physics at Durham University. No part of this thesis has been submitted elsewhere for any degree or qualification. Unless otherwise stated, all original work is my own in collaboration with my co-supervisor Marco Antonio Arroyo-Ureña.

Parts of this thesis are based on joint work, presented in the publications:

- M. A. Arroyo-Ureña, E. A. Herrera-Chacón, S. Rosado-Navarro and H. Salazar, “Hunting for a charged Higgs boson pair in proton–proton collisions,” *Phys. Rev. D* **111** (2025) 015023, doi:10.1103/PhysRevD.111.015023. Available at: APS Journal.
- M. A. Arroyo-Ureña, J. Lorenzo Díaz-Cruz, E. A. Herrera-Chacón, T. A. Valencia-Pérez and J. Mejia Guisao, “Flavoring the production of Higgs boson pairs,” *Phys. Rev. D* **111** (2025) 055028, doi:10.1103/PhysRevD.111.055028. Available at: APS Journal.

Copyright © 2025 Edwin Ali Herrera-Chacon.

The copyright of this thesis rests with the author. No quotation from it should be published without the author’s prior written consent and information derived from it should be acknowledged.

Acknowledgements

This thesis is a witness to the long journey I have taken to complete this stage of my life, which essentially began in 2020 when I applied to join Durham University. Along the way, I have been blessed to be associated with many people who, in different ways, have supported me throughout these years. I would like to say thank you to most of them.

To my supervisor Frank Krauss, for giving me the freedom to work on this project.

To my co-supervisor and friend Marco Arroyo, who has supported me for years with deep discussions about physics and who always inspires me with his passion for his work. Many thanks for believing in me!

To my colleagues Tomas Valencia and Sebastian Rosado, with whom I closely collaborated on the projects presented in this thesis, thank you for your excellent teamwork and enriching discussions.

To Christoph Englert and Michael Spannowsky, for their valuable comments and corrections.

To my mum Elizabeth, who is always there for me at any moment, even from a distance, and who taught me that hard work and integrity will always help us achieve our goals. To my dad Jose, for his constant support and for showing me that it is we who shape our destiny, not our circumstances. I love you both so much!

To my *bosses* from Bhaktivedanta Manor, Parama and Acarya, who always made sure everything was going well with my studies. To the devotees at Bhaktivedanta Institute and from the Pranama course for their association. To Bhakti Vijnana

Goswami for his inspiration and instructions, and to A.C. Bhaktivedanta Swami Srila Prabhupada for all the facilities he provided to allow us to learn about all types of matter and their interactions.

To Adriana, who has supported me from the very beginning of my degree, shared her experiences inspiring me to study in the UK, and proofread this thesis at the end; to Mia, who has always been very supportive in both my personal and academic life, even sacrificing a lovely Friday going out to proofread this thesis; and to James, who probably can't imagine how useful it was for me when he asked about my work.

To Ryan, Magan, and Reza; their hospitality in their office, kitchen, and flat respectively, made my working hours productive and joyful. Thank you for sharing this time with me.

To David (Damodar Das), Nita, Ram, Samuel, Alexandre, Jorge, and Liam, who have enriched this journey with their words of support.

To Diana, Niels, Emma, and Anthony, with whom I had the privilege to work online during the pandemic at the beginning of my PhD.

To all the community at University College, especially the Cricket Club 2023-2025 and Rugby Football Club 2024-2025, from whom I learnt so much (I swear more than how to chop a drink!). To all the KCSoc members, in particular Riya, Ankith, Shatakshi, Alec, and Shamayita, who have always supported me in my efforts to share Krishna consciousness. All of you were definitely my spiritual support in this process. Thank you, Kirtida and Bhaktirasa, for your service and for supporting me and the KCSoc! Thank you also to Wendy, Ellen, Sukanya, David, Nicola, Janet, Stephanie, Jamie, and all the porters. You have all contributed to making me feel part of the Castle community.

To all my housemates, especially Yue, Joe, Lisa, Pier, Adam, Mathew, Jake, and Joseph, sharing this experience with you has been truly special.

To all my colleagues at IPPP: Sofie, Lois, Yannick, Peter, Tommy, Guillaume, Ansh, Jack, Yunji, Ery, Elliot, Thomas, Malina, Hitman, and Livia; and special thanks to

Despoina for her kindness in proofreading my thesis. To Helen, Trudy and Joanne, for their constant help and support in administrative matters.

To Across the Pond, Monica in particular, for her valuable guidance during my application process to Durham University and her support with my visa paperwork.

Finally, I gratefully acknowledge financial support from SECIHTI (previously known as CONACYT) as well as from the Student Support Fund at Durham University.

Hare Krishna!

The plurality that we perceive is only an appearance; it is not real. Vedantic philosophy, in which this is a fundamental dogma, has sought to clarify it by a number of analogies, one of the most attractive being the many-faceted crystal which, while showing hundreds of little pictures of what is in reality a single existent object, does not really multiply that object.

— from *My View of the World* by Erwin Schrödinger

Dedicated to

Sri Sri Radha Gokulananda
from
Bhaktivedanta Manor

Chapter 1

Introduction

1.1 The Standard Model

The Standard Model (SM) unifies three of the four fundamental forces of nature: electromagnetic, weak, and strong, through a renormalisable quantum field theory [1]. This model successfully explains the behaviour of all known elementary particles, from quarks and leptons to gauge bosons like gluons (g), W^\pm , Z , and the photon (γ). At the core of this framework lies the Higgs mechanism, which introduces a single scalar Higgs doublet. Through Spontaneous Symmetry Breaking (SSB), the Higgs field acquires a Vacuum Expectation Value (VEV), providing the weak bosons and charged fermions with mass.

Despite the successes in explaining the most fundamental interactions of nature, the SM has certain limitations. It leaves several questions unanswered, hinting at the existence of physics beyond its scope. These unresolved issues include:

- **Neutrino Masses and Oscillations:** The SM predicts that neutrinos are massless. However, experimental evidence from neutrino oscillation experiments shows that they possess non-zero masses. This discrepancy requires

extensions to the SM, such as the introduction of right-handed neutrinos or higher dimensional operators, [2].

- **Matter-Antimatter Asymmetry:** The observed dominance of matter over antimatter in the universe cannot be fully explained with the SM. According to the Sakharov criteria, baryon asymmetry in the early universe can only emerge if three conditions are satisfied: (i) baryon number is not an exact symmetry; (ii) Charge Conjugation (C) and Charge Conjugation Parity (CP) invariance are violated, which manifests as unequal partial probabilities for charge conjugate reactions; and (iii) the universe experiences a strong departure from thermal equilibrium during the superdense stage of its non-stationary expansion. [3]. These are conditions that the SM does not fulfil simultaneously. The amount of CP violation in the SM through the Cabibbo-Kobayashi-Maskawa (CKM) matrix is therefore insufficient to account for the observed baryon asymmetry, and new sources of CP violation or alternative mechanisms are then required. [4]
- **Flavour Structure:** The SM does not explain the hierarchical pattern of fermion masses or the specific values of the mixing angles in the CKM and Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrices. The origin of the extensive hierarchy in Yukawa couplings remains as one of the major open questions in the SM.
- **Dark Matter and Unification:** The SM lacks a candidate for dark matter. This mysterious ‘substance’ constitutes about 27% of the universe’s energy density [5]. Additionally, the gauge couplings of the SM do not unify at high energies [6].

These limitations indicate that the SM is likely an approximation of a more fundamental theory. Therefore, extending the SM with additional Higgs doublets or scalar fields like Flavons offers promising options for solving some of these issues. The presence of extra scalars can provide new sources of CP violation for the matter-

antimatter asymmetry, and a Flavon framework naturally explains the significant hierarchies in fermion masses.

1.2 Research Aims and Key Contributions

In this work, we explore the phenomenology of scalar extensions of the SM, focusing on potential signatures of hypothetical scalar particles at the Large Hadron Collider (LHC) and its upgrade, the High-Luminosity LHC (HL-LHC). We analyse two Beyond the Standard Model (BSM) frameworks: the Froggatt-Nielsen (FN) mechanism [7] which introduces a new scalar particle known as the Flavon (H_F), and the Two-Higgs Doublet Model of type III (2HDM-III), which extends the scalar sector by adding a heavy CP-even scalar (H), a CP-odd scalar (A), and a charged scalar pair (H^\pm). We investigate the production and possible detection of prospects of charged scalar pairs within the 2HDM-III framework. We identify viable regions in the model parameter space that accommodate current experimental observations, particularly the reported excess in the process $\mathcal{BR}(t \rightarrow H^\pm b) \times \mathcal{BR}(H^\pm \rightarrow cb)$ at $M_{H^\pm} = 130$ GeV, as observed by the ATLAS collaboration.

Our analysis incorporates both theoretical and experimental constraints, proposing scenarios where charged scalar bosons with masses in the 100 – 350 GeV range can be probed with significances up to 5σ at the HL-LHC. Similarly, for the Flavon framework, we study the production and decay of scalar resonances, including processes where the SM-like Higgs boson (h) decays into pairs of photons or b quarks, as well as Lepton-Flavour-Violating (LFV) decays such as ($h \rightarrow e\mu$). Using multivariate analysis techniques and Machine Learning tools, we demonstrate that, under specific parameter scenarios, these processes can also reach discovery level sensitivity at the HL-LHC, with predicted significances up to 5σ . Overall, the key contributions of this thesis include a comprehensive collider analysis of these scalar extensions, demonstrating their potential observability at current and future collider experiments. Throughout this work, we also highlight the relevance of Machine Learning

methodologies in improving the signal from background separation to enhance the discovery prospects of BSM scalar particles.

1.3 Structure of This Thesis

In Chapter 2, we review the theoretical framework of the SM, discussing its gauge structure, analysing the field content, the mechanisms behind Electroweak Symmetry Breaking (EWSB), and going deeper on the motivations to extend its scalar sector. Chapter 3 introduces the scalar extensions of the SM considered in this thesis. We give a detailed description of the 2HDM-III and the FN Flavon model, focusing on their scalar sectors, Yukawa structures, and relevant experimental constraints. Following this, in Chapter 4, we present the multivariate analysis techniques that we used. We describe the setup of the event simulation, the LHC Olympics (LHCO) file format, and the implementation of BDTs to improve signal background discrimination. In Chapter 5, we perform detailed collider analyses for both models. For the 2HDM-III, we study the production of charged Higgs boson pairs and their decays into $\mu\nu_\mu cb$ final states, identifying parameter regions compatible with current ATLAS observations. For the Flavon model, we analyse the production of heavy CP-even Flavon resonances that decay into pairs of SM-like Higgs bosons, which subsequently decay into $b\bar{b}$ or $\gamma\gamma$ final states; in addition, we study the LFV process, $h \rightarrow e\mu$. Finally, we conclude in Chapter 6, summarising the key findings of this thesis and discussing possible future directions.

Chapter 2

Theoretical Framework

This chapter provides a comprehensive review of the SM structure, from its gauge symmetries and field content to the mechanism of EWSB, setting the stage for exploring these extensions in greater detail further on.

2.1 Gauge Symmetries and the Field Content

2.1.1 Local Gauge Group

The SM is built upon the gauge group

$$SU(3)_c \times SU(2)_L \times U(1)_Y,$$

which governs the interactions of elementary particles. This product of Lie groups follows the Yang–Mills principle of local gauge invariance, originally formulated for $SU(2)$ [8]. The associated Lie algebras fix the number and quantum numbers of the gauge bosons that mediate each interaction in the SM. Each component of this group plays a role as follows:

- $SU(3)_c$ describes the strong interactions, Quantum Chromodynamics (QCD),

mediated by eight massless gluons that couple to quarks, which carry colour charge.

- $SU(2)_L$ governs the weak interactions, acting on left-handed fermion doublets through three gauge bosons: W^1 , W^2 , and W^3 .
- $U(1)_Y$ corresponds to weak hypercharge Y , mediated by the gauge boson B_μ . After EWSB, B_μ is mixed with W_μ^3 to form the γ and the Z -boson. [1]

The fermions of the SM are organised into three generations, each consisting of:

$$(u, d)_L, \quad (c, s)_L, \quad (t, b)_L \quad \text{for quarks,}$$

$$(e, \nu_e)_L, \quad (\mu, \nu_\mu)_L, \quad (\tau, \nu_\tau)_L \quad \text{for leptons.}$$

right-handed fermions, such as u_R , d_R , and e_R , are singlets under $SU(2)_L$, they do not transform under that symmetry. The gauge principle holds that the physical laws must remain invariant under certain local transformations, gauge symmetries. Alternatively, in non-Abelian gauge groups like $SU(3)_c$ and $SU(2)_L$, the gauge bosons themselves carry an associated charge. Consequently, these bosons can self interact, except in an Abelian theory such as electromagnetism, where the Z gauge boson and the γ do not carry electric charge.

2.1.2 Standard Model Lagrangian

The full Electroweak (EW) Lagrangian can be decomposed into four distinct sectors [9–11]:

$$\mathcal{L}_{SU(2)_L \times U(1)_Y} = \mathcal{L}_{\text{Gauge}} + \mathcal{L}_\Phi + \mathcal{L}_F + \mathcal{L}_{Yukawa}. \quad (2.1.1)$$

Gauge Sector $\mathcal{L}_{\text{Gauge}}$

The gauge sector describes the gauge fields of $SU(2)_L$ and $U(1)_Y$ and their interactions, determined by field strength tensors that describe how these fields evolve across space and time and, for the non-Abelian cases, incorporate their self-interactions.

This sector is described by the Lagrangian:

$$\mathcal{L}_{\text{Gauge}} = -\frac{1}{4} F_{\mu\nu}^\alpha F^{\alpha\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu},$$

where the field strength tensors are defined as:

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu,$$

$$F_{\mu\nu}^\alpha = \partial_\mu W_\nu^\alpha - \partial_\nu W_\mu^\alpha - g_2 \epsilon^{\alpha\beta\gamma} W_\mu^\beta W_\nu^\gamma.$$

Here, g_2 is the $\text{SU}(2)_L$ coupling constant and $\epsilon^{\alpha\beta\gamma}$ is the Levi-Civita symbol in the adjoint indices.

Fermion Sector \mathcal{L}_F

In the fermionic sector, the kinetic terms for quarks and leptons are written as left-handed doublets and right-handed singlets, respectively, along with their interactions through covariant derivatives that introduce the gauge fields in accordance with the gauge symmetries of the theory.

For each family m , the left-handed quarks and lepton doublets are defined as:

$$Q'_{Lm} = \begin{pmatrix} u'_m \\ d'_m \end{pmatrix}_L, \quad L'_{Lm} = \begin{pmatrix} \nu'_m \\ l'_m \end{pmatrix}_L.$$

The right-handed fields u'_R, d'_R, l'_R are singlets under $\text{SU}(2)_L$. Then the fermionic Lagrangian is given by:

$$\begin{aligned} \mathcal{L}_F = \sum_{m=1}^3 & \left(\bar{Q}'_{Lm} i\gamma^\mu D_\mu Q'_{Lm} + \bar{L}'_{Lm} i\gamma^\mu D_\mu L'_{Lm} \right. \\ & \left. + \bar{u}'_{Lm} i\gamma^\mu D_\mu u'_{Lm} + \bar{d}'_{Lm} i\gamma^\mu D_\mu d'_{Lm} + \bar{l}'_{Lm} i\gamma^\mu D_\mu l'_{Lm} \right), \end{aligned} \quad (2.1.2)$$

where D_μ is the covariant derivative, which includes the gauge fields. We define the $\text{SU}(2)_L$ generators in the fundamental representation as $\tau^a \equiv \frac{1}{2}\sigma^a$. Hence, the covariant derivative acting on the left-handed quark doublet is [1]:

$$D_\mu Q'_{Lm} = \left(\partial_\mu + i g_2 \tau^a W_\mu^a + i g_1 Y B_\mu \right) Q'_{Lm},$$

In this expression, Y denotes the weak hypercharge of the corresponding fermion multiplet. For instance, for the left-handed quark doublet, we have $Y = \frac{1}{6}$; for the right-handed up- and down-type singlets, $Y = \frac{2}{3}$ and $Y = -\frac{1}{3}$, respectively. Similarly, for the left-handed lepton doublet, $Y = -\frac{1}{2}$, and for the right-handed charged lepton singlet, $Y = -1$ [12].

Scalar Sector \mathcal{L}_Φ

The scalar sector includes the Higgs doublet Φ and its potential. This sector is the responsible for EWSB and the generation of particle masses.

This sector is described by the Lagrangian:

$$\mathcal{L}_\Phi = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi),$$

where Φ is an $SU(2)_L$ doublet with hypercharge $+\frac{1}{2}$:

$$\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix}.$$

The components of the Higgs doublet are fixed by gauge invariance and charge conservation. In the fundamental representation of $SU(2)_L$, the generator $T_3 = \frac{1}{2}\sigma_3$ has eigenvalues $+\frac{1}{2}$ and $-\frac{1}{2}$ acting on the upper and lower entries, respectively. With hypercharge $Y = +\frac{1}{2}$, the electric charge of each component is given by $Q = T_3 + Y$ [1], resulting in $Q = +1$ for the upper field and $Q = 0$ for the lower one. The Higgs potential is given by:

$$V(\Phi) = -\mu^2 (\Phi^\dagger \Phi) + \lambda (\Phi^\dagger \Phi)^2. \quad (2.1.3)$$

For a stable vacuum, the quartic coupling λ must be positive. When $\mu^2 > 0$, the Higgs field acquires a non-zero VEV,

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad v = \sqrt{\frac{\mu^2}{\lambda}}.$$

The spontaneous breaking of $SU(2)_L \times U(1)_Y$ to $U(1)_{em}$ results in the emergence of three Goldstone bosons which become the longitudinal components of the W^\pm and Z bosons. The remaining degree of freedom corresponds to the physical Higgs boson h , with a mass given by:

$$m_h^2 = 2\mu^2.$$

The masses of the gauge bosons are generated through their interactions with the Higgs field. From the covariant derivative:

$$D_\mu \Phi = \left(\partial_\mu + i g_2 \tau^a W_\mu^a + i g_1 Y B_\mu \right) \Phi,$$

the term $(D_\mu \Phi)^\dagger (D^\mu \Phi)$ produces the masses:

$$m_W = \frac{1}{2} g_2 v, \quad m_Z = \frac{1}{2} \sqrt{g_2^2 + g_1^2} v, \quad m_\gamma = 0$$

In a more explicit notation:

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (W_\mu^1 \mp i W_\mu^2), \quad (2.1.4)$$

$$Z_\mu = \frac{1}{\sqrt{g_2^2 + g_1^2}} (g_2 W_\mu^3 - g_1 B_\mu), \quad (2.1.5)$$

$$A_\mu = \frac{1}{\sqrt{g_2^2 + g_1^2}} (g_1 W_\mu^3 + g_2 B_\mu) \quad (2.1.6)$$

Thus, the W^\pm and Z bosons acquire mass, while the γ remains massless.

Yukawa Sector \mathcal{L}_{Yukawa}

This sector is introducing the interactions between the Higgs field and the fermions, giving rise to fermion masses after SB.

The Yukawa sector is described by the Lagrangian:

$$-\mathcal{L}_{Yukawa} = \sum_{m,n=1}^3 \left[Y_{mn}^u \bar{Q}'_{Lm} \tilde{\Phi} u'_{Rn} + Y_{mn}^d \bar{Q}'_{Lm} \Phi d'_{Rn} + Y_{mn}^l \bar{L}'_{Lm} \Phi l'_{Rn} \right] + h.c., \quad (2.1.7)$$

where $\tilde{\Phi} = i \sigma_2 \Phi^*$, m, n are flavour indices.

Here, $\tilde{\Phi}$ is needed to ensure electric charge conservation in the up-type Yukawa interaction. The quarks and charged leptons obtain mass matrices:

$$M^u = \frac{v}{\sqrt{2}} Y^u, \quad M^d = \frac{v}{\sqrt{2}} Y^d, \quad M^l = \frac{v}{\sqrt{2}} Y^l.$$

These matrices can be diagonalised through unitary transformations, producing the physical masses of the fermions. For example, the up-type quark mass matrix is diagonalised as:

$$\overline{M}^u = \text{diag}(M_u, M_c, M_t),$$

with similar expressions for the down-type quarks and charged leptons.

2.1.3 Flavour Diagonalisation and CKM

The diagonalisation of the mass matrices produces the CKM, which describes the mixing between quark flavours in weak interactions. This matrix arises from the difference between the transformations that diagonalise the up-type and down-type quark mass matrices. In the SM, the Higgs boson does not mediate Flavour-changing Neutral Currents (FCNCs) at tree-level, as the Yukawa couplings become flavour diagonal in the mass basis.

The mass terms for the quarks after EWSB are given by [13],

$$\mathcal{L}_{\text{mass}} = -\frac{v}{\sqrt{2}} (\bar{d}_L Y_d d_R + \bar{u}_L Y_u u_R) + \text{h.c.},$$

where Y_d and Y_u are the Yukawa matrices for the down-type and up-type quarks. These matrices can be diagonalised using the unitary matrices U_d , U_u , K_d , and K_u ,

$$Y_d = U_d M_d K_d^\dagger, \quad Y_u = U_u M_u K_u^\dagger,$$

where M_d and M_u are diagonal mass matrices. Then the CKM matrix V is,

$$V = U_u^\dagger U_d,$$

The CKM matrix can be parametrised by three mixing angles θ_{12} , θ_{23} , θ_{13} and a

complex phase δ [14]:

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix},$$

where $c_{ij} = \cos \theta_{ij}$ and $s_{ij} = \sin \theta_{ij}$.

At this point it can be seen that there is a relatively weak flavour mixing as the CKM is nearly diagonal. The Wolfenstein parametrisation [15], provides a useful approximation,

$$|V| \approx \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & \lambda^3 \\ -\lambda & 1 - \frac{\lambda^2}{2} & \lambda^2 \\ \lambda^3 & \lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4),$$

where $\lambda = \sin \theta_{12} \approx 0.22$.

The unitarity of the CKM can be tested experimentally, and any deviation from it could indicate BSM physics.

2.2 Motivations for Extending the Scalar Sector

Extending the scalar sector of the SM can address several of its limitations:

- **Extra CP Violation:** Additional Higgs doublets can introduce new CP-violating phases [16], which are crucial to explain the matter-antimatter asymmetry.
- **Stronger Phase Transition:** A more complex Higgs sector can lead to a strongly first-order EW phase transition, enabling baryogenesis.
- **Flavour Symmetries:** Introducing additional fields or mechanisms can explain the hierarchical pattern of fermion masses through SSB.
- **Seesaw Mechanisms and Dark Matter:** Additional scalars can couple to

heavy neutrinos or dark matter sectors, providing mechanisms for neutrino masses and dark matter candidates. [17]

2.3 Overview of BSM Models

The 2HDM extends the SM by adding a second scalar doublet Φ_2 , with the same hypercharge as Φ_1 . A theoretical motivation for introducing a second Higgs doublet is the possibility of spontaneous CP violation. Additionally, the model predicts a richer scalar spectrum, including charged scalars and additional neutral states, which give rise to new decay channels and distinctive collider signatures [16, 18, 19]. However, the inclusion of additional scalar fields may also affect electroweak precision observables. In particular, the scalar potential should be constructed with care to avoid significant deviations from custodial symmetry, which ensures the tree-level relation $\rho = 1$, with ρ defined as the ratio $M_W^2/(M_Z^2 \cos^2 \theta_W)$. This symmetry, often referred to as custodial isospin, protects the relation between the W and Z boson masses in the SM. Such deviations are tightly constrained by experimental data [20].

After the SSB, one obtains up to five physical Higgs-like states: h , H (CP-even), A (CP-odd), and H^\pm (charged scalars). These additional scalars can be lighter or heavier than the observed 125 GeV Higgs, and can produce exotic decays, new loop effects in flavour observables, or new CP-violating phases. To avoid large tree-level FCNC, discrete symmetries or specific Yukawa textures are often imposed (Type I, II, X, Y, or the so-called Type-III with four-zero textures) [21]. Possible solutions to SM deficiencies include the following: additional scalar fields can lead to a strong first order EW phase transition, which is an essential factor in some baryogenesis scenarios [22]; with carefully chosen textures, it is possible to incorporate realistic quark/lepton masses while controlling FCNC at tree-level, and certain anomalies in heavy flavour data might also be addressed [23]; finally, the presence of H^\pm offers direct collider signatures, e.g. $t \rightarrow H^+ b$, $H^+ \rightarrow \tau^+ \nu$, cb , tested at the LHC [24, 25].

SM fermion masses extend over five orders of magnitude. The Yukawa couplings do

not provide a reason for this vast range. A Flavon model suggests a Spontaneously Broken Flavour Symmetry, e.g. $U(1)_F$, under which, each fermion f is assigned an integer charge Q_f . The Yukawa terms appear as higher dimensional operators suppressed by $\langle \Phi_F \rangle / \Lambda$ [26]. One introduces a gauge singlet Φ_F , which is the so called Flavon, with a negative $U(1)_F$ charge. The effective operators for up type quarks look like

$$\left(\frac{\Phi_F}{\Lambda}\right)^{Q_{Q_L} + Q_{u_R}} \bar{Q}_L \tilde{\Phi} u_R,$$

where Λ is some high flavour scale. A small ratio $\epsilon = \frac{\langle \Phi_F \rangle}{\Lambda} < 1$ can then produce the observed hierarchy in m_u, m_c, m_t . The remaining physical Flavon field can be mixed with the Higgs or induce exotic decays. If $\langle \Phi_F \rangle$ is near the TeV domain, such effects might be accessible at collider experiments [27].

The phenomenological impact includes hierarchical fermion masses, as by powers of ϵ , large mass hierarchies arise naturally from integer charges. It also includes LFV, since the Flavon can couple off-diagonally to leptons, enabling processes like $\tau \rightarrow \mu\gamma$ or $h \rightarrow \tau\mu$ if there is a mixing with the SM Higgs [28]. Additionally, rare B -decays might be impacted, as quark flavour transitions can receive contributions at loop level. This possibility can test the parameter space via meson observables [29].

We have established the essential features of the SM, the gauge group $SU(3)_c \times SU(2)_L \times U(1)_Y$, the single Higgs doublet responsible for EWSB, and the Yukawa interactions that provide fermions with mass. While robustly confirmed by experiments, including the discovery of the ~ 125 GeV Higgs boson [30, 31], the SM fails to address several questions.

To confront these issues, many models extend the scalar sector. In this work, we use two distinct BSM frameworks: the 2HDM-III and a Flavon model via the FN mechanism, which we will analyse in the next chapter.

Chapter 3

Beyond Standard Models

3.1 Two-Higgs Doublet Model of Type III

In this section, we present the theoretical framework, which conducts a thorough analysis of the Yukawa Lagrangian within the context of the 2HDM-III and subsequently derives the necessary Feynman rules for our calculations. Through this in-depth exploration, we construct a robust framework in order to facilitate our ensuing analysis of the charged Higgs boson production and decay at the LHC and HL-LHC.

3.1.1 Scalar Potential

The 2HDM-III incorporates an extra Higgs doublet in addition to what is in the SM. The two scalar doublets can be written as $\Phi_a^T = (\phi_a^+, \phi_a^0)$ for $a = 1, 2$, and they have an associated hypercharge with a value of $+1$. During the SSB, the Higgs doublets acquire non-zero VEV, which can be expressed as

$$\langle \Phi_a \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_a \end{pmatrix}.$$

The introduction of additional Higgs doublets engenders new interactions between these doublets and all different types of fermions [32]. Nonetheless, such interactions may lead to the occurrence of FCNCs at the tree-level. Given that empirical observations impose rigorous restrictions on FCNCs, mechanisms must be implemented to suppress their presence within the framework of the model [10]. An effective methodology for attaining this suppression involves the implementation of a specific four-zero texture in the Yukawa sector. This is a specific arrangement of zeros within the Yukawa matrices, which is crucial to define the mass and mixing angles of the fermions. This texture acts as a simplified theoretical construct that delineates the flavour interactions between fermions and the Higgs bosons. By employing a four-zero texture, we can effectively modulate the intensity of FCNCs while ensuring alignment with the experimental data. [33–36]

The most general scalar potential that is invariant under the symmetry group $SU(2)_L \times U(1)_Y$ can be expressed as follows [37]:

$$\begin{aligned}
V(\Phi_1, \Phi_2) &= \mu_1^2(\Phi_1^\dagger\Phi_1) + \mu_2^2(\Phi_2^\dagger\Phi_2) - \mu_{12}^2(\Phi_1^\dagger\Phi_2 + h.c.) \\
&+ \frac{1}{2}\lambda_1(\Phi_1^\dagger\Phi_1)^2 + \frac{1}{2}\lambda_2(\Phi_2^\dagger\Phi_2)^2 \\
&+ \lambda_3(\Phi_1^\dagger\Phi_1)(\Phi_2^\dagger\Phi_2) + \lambda_4(\Phi_1^\dagger\Phi_2)(\Phi_2^\dagger\Phi_1) \\
&+ \frac{1}{2}\lambda_5(\Phi_1^\dagger\Phi_2)^2 + \lambda_6(\Phi_1^\dagger\Phi_1)(\Phi_1^\dagger\Phi_2) \\
&+ \lambda_7(\Phi_2^\dagger\Phi_2)(\Phi_1^\dagger\Phi_2) + h.c.
\end{aligned} \tag{3.1.1}$$

In order to prevent tree-level FCNC, a discrete symmetry is commonly imposed, under which the scalar fields transform as $\Phi_1 \rightarrow \Phi_1$ and $\Phi_2 \rightarrow -\Phi_2$ [38]. Such a symmetry forbids the λ_6 and λ_7 terms in the scalar potential, which break this discrete invariance. When this symmetry is exact, the potential is reduced by eliminating these terms. Furthermore, this discrete invariance ensures that only one Higgs doublet couples to each type of right handed fermion, avoiding flavour violation in the Yukawa sector. This symmetry also contributes to CP conservation,

as long as all parameters remain real. As a result, the scalar potential simplifies to a CP conserving and flavour safe form, consistent with phenomenological constraints.

In this study, we take all parameters in the scalar potential to be real, including the VEV's of the Higgs doublets. Under these conditions, the CP symmetry is conserved in this sector because there is no source of complex phases within the scalar potential itself. However, this does not imply global CP conservation across the model, as the CKM matrix in the fermion sector still introduces CP violation.

Furthermore, the model includes several additional independent parameters, among which the mixing angle α , emerges from diagonalising the CP-even scalar mass matrix and governs the transition from gauge eigenstates to mass eigenstates:

$$\begin{aligned} H &= \text{Re}(\phi_1^0) \cos \alpha + \text{Re}(\phi_2^0) \sin \alpha, \\ h &= \text{Re}(\phi_1^0) \sin \alpha + \text{Re}(\phi_2^0) \cos \alpha, \end{aligned} \quad (3.1.2)$$

with

$$\tan 2\alpha = \frac{2m_{12}}{m_{11} - m_{22}},$$

where $m_{11,12,22}$ correspond to the entries of the real component of the mass matrix \mathbf{M} determining how the fields acquire their masses,

$$\text{Re}(\mathbf{M}) = \begin{pmatrix} m_{11} & m_{12} \\ m_{12} & m_{22} \end{pmatrix},$$

where

$$\begin{aligned} m_{11} &= m_A^2 \sin^2 \beta + v^2(\lambda_1 \cos^2 \beta + \lambda_5 \sin^2 \beta) \\ m_{12} &= -m_A^2 \cos \beta \sin \beta + v^2(\lambda_3 + \lambda_4) \cos \beta \sin \beta \\ m_{22} &= m_A^2 \cos^2 \beta + v^2(\lambda_2 \sin^2 \beta + \lambda_5 \cos^2 \beta). \end{aligned} \quad (3.1.3)$$

Eqs. (3.1.2) define the distinct scalar states in the model, which are one physical charged scalar boson, a pseudo Goldstone boson that is absorbed by the W gauge

fields, a separate CP-odd state, and another pseudo Goldstone boson corresponding to the Z gauge boson. As detailed further in this section, h typically behaves as the SM-like Higgs boson, whereas H emerges as a heavier scalar state. In Eqs. (3.1.3), m_A represents the mass of the CP-odd scalar, A .

In addition to α , there is a second mixing angle, β , which governs how the charged components of Φ_a are combined with the imaginary parts of the neutral components being crucial to distinguish the physical charged states from the Goldstone modes absorbed by the gauge bosons determining the final spectrum of charged scalar fields.

$$\begin{aligned}
G_W^\pm &= \phi_1^\pm \cos \beta + \phi_2^\pm \sin \beta, \\
H^\pm &= -\phi_1^\pm \sin \beta + \phi_2^\pm \cos \beta, \\
G_Z &= \text{Im}(\phi_1^0) \cos \beta + \text{Im}(\phi_2^0) \sin \beta, \\
A^0 &= -\text{Im}(\phi_1^0) \sin \beta + \text{Im}(\phi_2^0) \cos \beta.
\end{aligned} \tag{3.1.4}$$

The second mixing angle β is conventionally defined through the VEV ratio, such than $\tan \beta = v_2/v_1$, playing a key role in determining how the Higgs doublets contribute to the masses of the fermions and gauge bosons.

In this theoretical framework, the primary parameters of interest include the two mixing angles α and β , as well as the physical masses M_{H^\pm} , M_h , M_H and M_A . The expressions for these masses in terms of the underlying parameters are given by

$$M_{H^\pm}^2 = \frac{\mu_{12}}{\sin \beta \cos \beta} \tag{3.1.5}$$

$$- \frac{1}{2} v^2 \left(\lambda_4 + \lambda_5 + \cot \beta \lambda_6 + \tan \beta \lambda_7 \right),$$

$$M_{h,H}^2 = \frac{1}{2} \left(m_{11} + m_{22} \mp \sqrt{(m_{11} - m_{22})^2 + 4m_{12}^2} \right), \tag{3.1.6}$$

$$M_A^2 = M_H^\pm + \frac{1}{2} v^2 (\lambda_4 - \lambda_5). \tag{3.1.7}$$

3.1.2 Yukawa Lagrangian

The Yukawa sector encapsulates the interactions between the two Higgs doublets and the fermion fields. It is described by the following Lagrangian [21]:

$$\begin{aligned} \mathcal{L}_Y = & - \left(Y_1^u \bar{Q}_L \tilde{\Phi}_1 u_R + Y_2^u \bar{Q}_L \tilde{\Phi}_2 u_R \right. \\ & + Y_1^d \bar{Q}_L \Phi_1 d_R + Y_2^d \bar{Q}_L \Phi_2 d_R \\ & \left. + Y_1^\ell \bar{L}_L \tilde{\Phi}_1 l_R + Y_2^\ell \bar{L}_L \tilde{\Phi}_2 l_R \right), \end{aligned} \quad (3.1.8)$$

here, $\tilde{\Phi}_a$ with ($a = 1, 2$) is defined by $\tilde{\Phi}_a = i\sigma_2 \Phi_a^*$. After that the EWSB takes place, the fermion mass matrices have the following form:

$$M_f = \frac{1}{\sqrt{2}} \left(v_1 Y_1^f + v_2 Y_2^f \right), \quad f = u, d, \ell. \quad (3.1.9)$$

In this step, we take both Yukawa matrices to exhibit the previously described four-zero texture structure and we also assume that they are Hermitian to reduce the parameter space and contribute to maintaining consistency with the assumption of CP conservation.

Following the diagonalisation process, we obtain

$$\bar{M}_f = V_{fL}^\dagger M_f V_{fR} = \frac{1}{\sqrt{2}} \left(v_1 \tilde{Y}_1^f + v_2 \tilde{Y}_2^f \right), \quad (3.1.10)$$

$$\tilde{Y}_a^f = V_{fL}^\dagger Y_a^f V_{fR} \quad (3.1.11)$$

from these relations, we can deduce the following expressions,

$$\left[\tilde{Y}_a^f \right]_{ij} = \frac{\sqrt{2}}{v_a} \delta_{ij} \bar{M}_{ij}^f - \frac{v_b}{v_a} \left[\tilde{Y}_b^f \right]_{ij} \quad (3.1.12)$$

where f refers to massive fermions and the parameters $\left[\tilde{\chi}_a^f \right]_{ij}$ represent unknown dimensionless quantities in the model. It is important to mention that from Eq. 3.1.12, various types of interactions may be obtained. By selecting specific structures [21], the following models can be defined as:

- 2HDM-III I-like

$$\begin{aligned}
[\tilde{Y}_2^d]_{ij} &= \frac{\sqrt{2}}{v \sin \beta} \delta_{ij} \bar{M}_{ij}^d - \cot \beta [\tilde{Y}_1^d]_{ij} \\
[\tilde{Y}_2^u]_{ij} &= \frac{\sqrt{2}}{v \sin \beta} \delta_{ij} \bar{M}_{ij}^u - \cot \beta [\tilde{Y}_1^u]_{ij} \\
[\tilde{Y}_2^\ell]_{ij} &= [\tilde{Y}_1^d]_{ij} \quad (d \rightarrow \ell).
\end{aligned} \tag{3.1.13}$$

- 2HDM-III II-like

$$\begin{aligned}
[\tilde{Y}_1^d]_{ij} &= \frac{\sqrt{2}}{v \cos \beta} \delta_{ij} \bar{M}_{ij}^d - \tan \beta [\tilde{Y}_2^d]_{ij} \\
[\tilde{Y}_2^u]_{ij} &= \frac{\sqrt{2}}{v \sin \beta} \delta_{ij} \bar{M}_{ij}^u - \cot \beta [\tilde{Y}_1^u]_{ij} \\
[\tilde{Y}_1^\ell]_{ij} &= [\tilde{Y}_1^d]_{ij} \quad (d \rightarrow \ell).
\end{aligned} \tag{3.1.14}$$

- 2HDM-III Lepton Specific-like

$$\begin{aligned}
[\tilde{Y}_2^d]_{ij} &= \frac{\sqrt{2}}{v \sin \beta} \delta_{ij} \bar{M}_{ij}^d - \cot \beta [\tilde{Y}_1^d]_{ij} \\
[\tilde{Y}_2^u]_{ij} &= \frac{\sqrt{2}}{v \sin \beta} \delta_{ij} \bar{M}_{ij}^u - \cot \beta [\tilde{Y}_1^u]_{ij} \\
[\tilde{Y}_1^\ell]_{ij} &= [\tilde{Y}_1^d]_{ij} \quad (d \rightarrow \ell).
\end{aligned} \tag{3.1.15}$$

- 2HDM-III Flipped-like

$$\begin{aligned}
[\tilde{Y}_1^d]_{ij} &= \frac{\sqrt{2}}{v \cos \beta} \delta_{ij} \bar{M}_{ij}^d - \tan \beta [\tilde{Y}_2^d]_{ij} \\
[\tilde{Y}_2^u]_{ij} &= \frac{\sqrt{2}}{v \sin \beta} \delta_{ij} \bar{M}_{ij}^u - \cot \beta [\tilde{Y}_1^u]_{ij} \\
[\tilde{Y}_2^\ell]_{ij} &= [\tilde{Y}_2^d]_{ij} \quad (d \rightarrow \ell).
\end{aligned} \tag{3.1.16}$$

Where v is defined as $v = \sqrt{v_1^2 + v_2^2}$, representing the combined VEVs of the two Higgs doublets. These models correspond to the following types of Higgs doublet couplings [11, 39]:

- **Type I:** A single Higgs doublet provides mass to both up-type and down-type quarks.
- **Type II:** The neutral component of one Higgs doublet couples to up-type

Type	Couples to up-type quarks	Couples to down-type quarks	Couples to charged leptons
I	ϕ_2	ϕ_2	ϕ_2
II	ϕ_2	ϕ_1	ϕ_1
III	ϕ_1, ϕ_2	ϕ_1, ϕ_2	ϕ_1, ϕ_2
X	ϕ_2	ϕ_2	ϕ_1
Y	ϕ_2	ϕ_1	ϕ_2

Table 3.1: The fact that the up-type quark couples to the doublet ϕ_2 is by convention. Type III exhibits a flavour violation.

quarks, while the neutral component of the other doublet couples to down-type quarks.

- **Type III:** Both Higgs doublets can generate masses for up-type and down-type quarks simultaneously.
- **Type X (Lepton-Specific):** The quark couplings are as in Type I, while the lepton couplings are as in Type II.
- **Type Y (Flipped):** The quark couplings are as in Type II, while the lepton couplings are as in Type I.

From Eqs. (3.1.8)-(3.1.12), we obtain

$$\mathcal{L}_Y^\phi = \phi \bar{f}_i (S_{ij}^\phi + iP_{ij}^\phi \gamma^5) f_j, \quad (3.1.17)$$

where $\phi = h, H, A$. The minus sign in Yukawa terms is absorbed into S_{ij}^ϕ and P_{ij}^ϕ in Eq.(3.1.17), simplifying the notation. Those CP-conserving and CP-violating factors S_{ij}^ϕ and P_{ij}^ϕ , respectively, incorporate flavour dynamics and are written as:

$$\begin{aligned} S_{ij}^\phi &= \frac{gm_f}{2M_W} c_f^\phi \delta_{ij} + d_f^\phi [\tilde{Y}_a^f]_{ij}, \\ P_{ij}^\phi &= \frac{gm_f}{2M_W} e_f^\phi \delta_{ij} + g_f^\phi [\tilde{Y}_a^f]_{ij}. \end{aligned} \quad (3.1.18)$$

The coefficients $c_f^\phi, d_f^\phi, e_f^\phi, g_f^\phi$ depend on the presence of new physics in the Higgs sector. Within the SM $c_f^{\phi=h} = 1$ and $d_f^{\phi=h} = e_f^{\phi=h} = g_f^{\phi=h} = 0$, in contrast, within the 2HDM-III these coefficients, in the CP conserving case, are as shown in Table 3.2. In Eqs. (3.1.18) we can choose a for the Yukawa matrices according to (3.1.13)-

Coefficient	c_f^h	c_f^A	c_f^H	d_f^h	d_f^A	d_f^H
d -type	$-\frac{\sin \alpha}{\cos \beta}$	$-\tan \beta$	$\frac{\cos \alpha}{\sin \beta}$	$\frac{\cos(\alpha-\beta)}{\cos \beta}$	$\csc \beta$	$\frac{\sin(\alpha-\beta)}{\cos \beta}$
u -type	$\frac{\cos \alpha}{\sin \beta}$	$-\cot \beta$	$\frac{\sin \alpha}{\sin \beta}$	$-\frac{\cos(\alpha-\beta)}{\sin \beta}$	$\sec \beta$	$\frac{\sin(\alpha-\beta)}{\sin \beta}$
leptons ℓ	$-\frac{\sin \alpha}{\cos \beta}$	$-\tan \beta$	$\frac{\cos \alpha}{\sin \beta}$	$\frac{\cos(\alpha-\beta)}{\cos \beta}$	$\csc \beta$	$\frac{\sin(\alpha-\beta)}{\cos \beta}$

Table 3.2: The coefficients for the ϕ -Fermion couplings in the 2HDM-III with a \mathcal{CP} -conserving Higgs potential.

(3.1.16). Although we adopt a 2HDM-II like scenario within the broader 2HDM-III framework, for simplicity, we will refer to the setup simply as the 2HDM-III.

The couplings of the charged scalar boson with quarks are given by

$$\begin{aligned}
\mathcal{L}_Y^{H^\pm q_i q_j} = & \frac{\sqrt{2}}{v} \left\{ \left[\bar{d}_i \left(m_u \cot \beta - \frac{v}{\sqrt{2}} g(\beta) [\tilde{Y}_2^u]_{ij} \right) u_j H^- V_{\text{CKM}}^{ij*} \right. \right. \\
& - \bar{u}_i \left(m_d \tan \beta - \frac{v}{\sqrt{2}} f(\beta) [\tilde{Y}_1^d]_{ij} \right) d_j H^+ V_{\text{CKM}}^{ij} \left. \right] P_R \\
& - \left[\bar{d}_i \left(m_d \tan \beta - \frac{v}{\sqrt{2}} f(\beta) [\tilde{Y}_1^d]_{ij} \right) u_j H^- V_{\text{CKM}}^{ij*} \right. \\
& \left. \left. + \bar{u}_i \left(m_u \cot \beta - \frac{v}{\sqrt{2}} g(\beta) [\tilde{Y}_2^u]_{ij} \right) d_j H^+ V_{\text{CKM}}^{ij} \right] P_L \right\}. \quad (3.1.19)
\end{aligned}$$

where

$$\left[\tilde{Y}_a^f \right]_{ij} = \frac{\sqrt{m_{f_i} m_{f_j}}}{v} \chi_{ij}, \quad (a = 1, 2), \quad (3.1.20)$$

$$f(\beta) = \sqrt{1 + \tan^2 \beta}, \quad (3.1.21)$$

$$g(\beta) = \sqrt{1 + \cot^2 \beta}, \quad (3.1.22)$$

while those couplings of the charged scalar boson with leptons are given by

$$\mathcal{L}_Y^{H^\pm \ell \nu_\ell} = \frac{\sqrt{2} m_{\ell_i}}{v} \bar{\nu}_L \left(\tan \beta \frac{m_{\ell_i}}{m_{\ell_j}} \delta_{ij} - \frac{f(\beta)}{\sqrt{2}} \sqrt{\frac{m_{\ell_i}}{m_{\ell_j}}} \chi_{ij}^\ell \right) \ell_R^- H^+ + H.c. \quad (3.1.23)$$

3.2 Constraints on the 2HDM-III Parameter Space

In order to obtain realistic predictions, we carry out a comprehensive analysis of multiple experimental constraints, which can be grouped into two main categories:

- Process induced for the neutral scalar (pseudo scalar) bosons h, H, A :
 - LHC Higgs boson data [40, 41].
 - Neutral meson physics $B_s^0 \rightarrow \mu^+ \mu^-$ [42], $B_d^0 \rightarrow \mu^+ \mu^-$ [42].
 - Lepton flavour violating decays $\ell_i \rightarrow \ell_j \ell_k \bar{\ell}_k$ and radiative decays $\ell_i \rightarrow \ell_j \gamma$ [43].
 - Muon anomalous magnetic moment a_μ ¹ [44].
 - Double Higgs production cross-section $\sigma(pp \rightarrow H \rightarrow hh \rightarrow b\bar{b}\gamma\gamma)$ [45].
 - Single scalar production $\sigma(gb \rightarrow \phi \rightarrow \tau\tau)$ ($\phi = H, A$) [46, 47].
- Process induced for the charged scalar boson H^\pm :
 - Limits on $\sigma(pp \rightarrow tbH^\pm) \times \text{BR}(H^\pm \rightarrow \tau^\pm \nu)$ [48].
 - Limits on $\mathcal{BR}(t \rightarrow H^\pm b) \times \mathcal{BR}(H^\pm \rightarrow cb)$ [49].
 - Radiative b quark decay $b \rightarrow s\gamma$ [50, 51].

The free model parameters that directly influence our predictions can be summarised as follows:

1. The cosine of the difference of mixing angles: $\cos(\alpha - \beta)$.
2. The ratio of the VEV's: $\tan \beta$.
3. The masses of the additional bosons: M_H, M_A, M_{H^\pm} .

¹ a_μ also receives contributions from the charged scalar boson, however, its contribution is subdominant.

4. The matrix elements $\chi_{cb}, \chi_{tb}, \chi_{\mu\mu}, \chi_{tt}, \chi_{bb}$.

There are several parameters χ_{ij} needed in the calculation of $\mathcal{BR}(H^+ \rightarrow bc)$ and $\mathcal{BR}(t \rightarrow H^+b)$. Unless otherwise specified, all of these parameters are taken to be $\chi_{ij} = 1$. Here the indices i and j denote fermions, and in general $i \neq j$.

We also consider basic theoretical constraints, such as perturbativity, vacuum stability, tree-level unitarity [52], and the oblique parameters (see Section 3.2.2 for further discussion). Although these constraints are typically derived for \mathbb{Z}_2 -symmetric models with $\lambda_6 = \lambda_7 = 0$, we adopt them in the Type-III case. Nevertheless, they do not impose significant restrictions on the explored parameter space, compared to the experimental ones.

3.2.1 Constraint on $\tan \beta$ and $\cos(\alpha - \beta)$

LHC Higgs boson data: Signal strength modifiers

For a decay $S \rightarrow X$ (where X denotes a specific final state), or a production process $\sigma(pp \rightarrow S)$, the signal strength is defined by the following parameterisation [53]

$$\mu_X = \frac{\sigma(pp \rightarrow h) \cdot \mathcal{BR}(h \rightarrow X)}{\sigma(pp \rightarrow h^{\text{SM}}) \cdot \mathcal{BR}(h^{\text{SM}} \rightarrow X)}, \quad (3.2.1)$$

here, $\sigma(pp \rightarrow S)$ denotes the production cross-section of S , where $S = h, h^{\text{SM}}$; in this notation, h represents the SM-like Higgs boson coming from an extension of the SM, whereas h^{SM} is the SM Higgs boson. Furthermore, $\mathcal{BR}(S \rightarrow X)$ is the Branching Ratio (BR), i.e., the probability of the decay $S \rightarrow X$, with $X = \bar{b}b, \tau^-\tau^+, \mu^-\mu^+, WW^*, ZZ^*, \gamma\gamma$ [54, 55].

Neutral meson physics

To complement the constraints provided by LHC Higgs boson data, we also examine LFV processes. These are mediated by the neutral scalar bosons H and h , a pseudo-

scalar A , (In principle, each of these can induce FCNC at tree-level), and a charged scalar boson H^\pm . These observables are:

- The decay $B_s^0 \rightarrow \mu^+ \mu^-$.
- The muon anomalous magnetic moment a_μ .
- Radiative processes $\ell_i \rightarrow \ell_j \gamma$ ($\ell_i = \tau, \mu$; $\ell_j = \mu, e$; $i \neq j$).
- $\ell_i \rightarrow \ell_j \ell_k \bar{\ell}_k$ ($k = \mu, e$) and the flavour changing decay of the Higgs boson, $h \rightarrow \ell_i \ell_j$.

Decays $B_s^0 \rightarrow \mu^+ \mu^-$, $B_d^0 \rightarrow \mu^+ \mu^-$ and $B \rightarrow B_s^0, B_d^0$

In the SM, neutral mesons that decay into muons are strongly suppressed. This suppression arises due to three primary reasons: (i) these processes proceed via loop level diagrams, which are inherently weaker than tree-level processes; (ii) helicity suppression further reduces the interaction probability; and (iii) certain CKM matrix elements are very small, thereby lowering the decay rate even more. As a result, the corresponding BRs are extremely small. Although other decay channels such as those involving electrons or τ leptons, can occur in principle, they are heavily suppressed in one case and challenging to reconstruct in the other, respectively. Fig. 3.1 presents a representative Feynman diagram for this process.

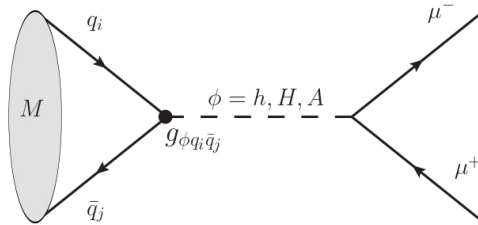


Figure 3.1: Generic Feynman diagram for the decays of a neutral meson M (such as B^0 , K^0 , or D^0) into $\mu^+ \mu^-$. The black circle indicates a Flavour-changing vertex in the quark sector.

The decay of $B_{s,d}^0$ mesons into a $\mu^+ \mu^-$ pair is both compelling and highly constraining, given its sensitivity to BSM physics. Within the theoretical framework of the SM,

the BRs are [56]:

$$\mathcal{BR}(B_s^0 \rightarrow \mu^+ \mu^-) = (3.66 \pm 0.14) \times 10^{-9}$$

and

$$\mathcal{BR}(B_d^0 \rightarrow \mu^+ \mu^-) = (1.03 \pm 0.05) \times 10^{-10}.$$

According to measurements by the CMS collaboration [57], the current experimental value is,

$$\mathcal{BR}(B_s^0 \rightarrow \mu^+ \mu^-) = \left(4.02_{-0.38}^{+0.40}(\text{stat})_{-0.23}^{+0.28}(\text{syst})_{-0.15}^{+0.18}(\text{ext})\right) \times 10^{-9},$$

where *stat* denotes the statistical uncertainty, *syst* the experimental systematic uncertainty, and *ext* the external uncertainty from the branching fractions $\mathcal{B}(B^+ \rightarrow J/\psi K^+)$, $\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)$, and from the fragmentation fraction ratio of the probabilities for a *b* quark to hadronise into a B_s^0 or B^+ meson. While

$$\mathcal{BR}(B_d^0 \rightarrow \mu^+ \mu^-) < 1.9 \times 10^{-10}$$

at 95% C.L. [42]. In this context of the 2HDM-III, these decays, $B_{s,d}^0 \rightarrow \mu^+ \mu^-$, can be mediated at tree-level by the SM-like Higgs boson, a heavy scalar H , or the pseudoscalar A . The relevant Feynman diagram for these decays is depicted in Fig. 3.1. In the case of $B_s^0 \rightarrow \mu^+ \mu^-$ (or $B_d^0 \rightarrow \mu^+ \mu^-$), the quark indices are $q_i = s$, $\bar{q}_j = \bar{b}$ (or $q_i = d$, $\bar{q}_j = \bar{b}$).

The effective Hamiltonian governing the $B_s^0 \rightarrow \mu^+ \mu^-$ transition is,

$$\begin{aligned} \mathcal{H}_{\text{eff}}^{B_s^0 \rightarrow \mu^+ \mu^-} = & -\frac{G_F^2 m_W^2}{\pi^2} \left(C_A^{bs} \mathcal{O}_A^{bs} + C_S^{bs} \mathcal{O}_S^{bs} + C_P^{bs} \mathcal{O}_P^{bs} \right. \\ & \left. + C_A'^{bs} \mathcal{O}_A'^{bs} + C_S'^{bs} \mathcal{O}_S'^{bs} + C_P'^{bs} \mathcal{O}_P'^{bs} \right) + h.c., \end{aligned} \quad (3.2.2)$$

where C_i^{bs} and $C_i'^{bs}$ (with $i = A, S, P$) are the Wilson coefficients encoding the short distance contributions. The Wilson operators \mathcal{O}_i^{bs} describe the low energy effective

interactions between the quark and lepton fields, they take the form,

$$\begin{aligned}\mathcal{O}_A^{bs} &= (\bar{b}\gamma_\mu P_L s) (\mu^+ \gamma_\mu \gamma_5 \mu^-), \\ \mathcal{O}_S^{bs} &= (\bar{b} P_L s) (\mu^+ \mu^-), \\ \mathcal{O}_P^{bs} &= (\bar{b} P_L s) (\mu^+ \gamma_5 \mu^-).\end{aligned}\tag{3.2.3}$$

The primed operators are obtained by interchanging $P_L \leftrightarrow P_R$. The BR for this decay is then given by

$$\begin{aligned}\mathcal{BR}(M \rightarrow \ell^+ \ell^-) &= \frac{G_F^4 m_W^4}{8\pi^5} \sqrt{1 - 4 \frac{m_\ell^2}{m_M^2}} m_M f_M^2 m_\ell^2 \tau_M \\ &\times \left[\left| \frac{m_M^2 (C_P^{ij} - C_P'^{ij})}{2(m_i + m_j) m_\ell} - C_A^{\text{SM}} \right|^2 + \left| \frac{m_M^2 (C_S^{ij} - C_S'^{ij})}{2(m_i + m_j) m_\ell} \right|^2 \left(1 - 4 \frac{m_\ell^2}{m_M^2}\right) \right],\end{aligned}\tag{3.2.4}$$

here, $\ell^{+(-)} = \mu^{+(-)}$ and $i(j) = s(\bar{b})$. The parameter $m_M = m_{B_s^0} = 5.36692$ GeV denotes the B_s^0 meson mass, while $f_M = f_{B_s^0} = 0.2303$ GeV is the B_s^0 meson decay constant. The lifetime of the B_s^0 meson is given by $\tau_M = \tau_{B_s^0} = 2.311 \times 10^{12}$ GeV⁻¹ [58]. The Fermi constant is represented by G_F , and the SM contribution at one loop, C_A^{SM} , is given by,

$$C_A^{\text{SM}} = -V_{tb}^* V_{ts} Y \left(\frac{m_t^2}{m_W^2} \right) - V_{cb}^* V_{cs} Y \left(\frac{m_c^2}{m_W^2} \right),\tag{3.2.5}$$

where Y is written as $Y = \eta_Y Y_0$ to include NLO QCD effects through $\eta_Y = 1.0113$ [59]. The loop Inami-Lim function [60] is given by

$$Y_0(x) = \frac{x}{8} \left[\frac{4-x}{1-x} + \frac{3x}{(1-x)^2} \ln(x) \right].\tag{3.2.6}$$

Finally, the form factors are

$$\begin{aligned}C_S^{ij} &= \frac{\pi^2}{2G_F^2 m_W^2} \sum_{\phi=h,H} \frac{2g_{\phi\ell^+\ell^-} g_{\phi ij}}{M_\phi^2}, \\ C_P^{ij} &= \frac{\pi^2}{2G_F^2 m_W^2} \frac{2g_{\phi\ell^+\ell^-} g_{\phi ij}}{M_A^2}, \\ C_S'^{ij} &= C_S^{ij} (g_{\phi ij} \leftrightarrow g_{\phi ji}),\end{aligned}\tag{3.2.7}$$

$$C_P^{ij} = C_S^{ij} (g_{\phi ij} \leftrightarrow g_{\phi ji}).$$

The symbol \leftrightarrow indicates that the couplings $g_{\phi ij}$ and $g_{\phi ji}$ must be exchanged in the expression. This substitution accounts for the contribution where the external fermions are reversed, corresponding to a change in the flavour indices of the scalar or pseudoscalar interaction vertex. To obtain the corresponding $\mathcal{BR}(B_d^0 \rightarrow \mu^+ \mu^-)$, we make the following substitutions in the BR expression Eq. (3.2.4) the replacements $m_{B_s^0} \rightarrow m_{B_d^0} = 5.27966$ GeV, $f_{B_s^0} \rightarrow f_{B_d^0} = 0.190$, $\tau_{B_s^0} \rightarrow \tau_{B_d^0} = 2.312 \times 10^{12}$ GeV; In Eqs. (3.2.7) $g_{\phi \bar{b}s} \rightarrow g_{\phi \bar{b}d}$ and in Eqs. (3.2.3) $s \rightarrow d$.

Lepton Flavour violating processes

Muon anomalous magnetic moment a_μ

Considering the persistent discrepancy between the experimental measurement of a_μ and its theoretical prediction within the SM, we use this observable to constrain the relevant free parameters of the model. The Feynman diagrams contributing to a_μ are shown in Fig. 3.2 [61]. The one-loop contributions are given by

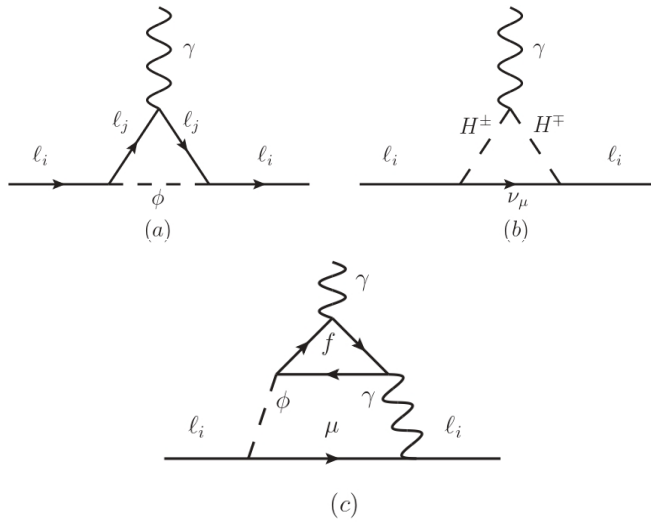


Figure 3.2: Feynman diagrams contributing to a_μ . Here ϕ represents a CP-even scalar, CP-odd scalar, or the SM-like Higgs boson. H^\pm denotes charged scalar bosons. In these diagrams $l_i = \mu$.

$$\delta a_\mu^\phi = \frac{G_F m_\mu^2}{4\pi^2 \sqrt{2}} \sum_{\ell_i} g_{\mu\ell_i}^2 R_\phi F_\phi(R_\phi), \quad (3.2.8)$$

with

$$F_{h,H} = \int_0^1 dx \frac{x^2(2-x)}{R_{h,H}x^2 - x + 1}, \quad (3.2.9)$$

$$F_A = \int_0^1 dx \frac{-x^3}{R_A x^2 - x + 1}, \quad (3.2.10)$$

$$F_{H^\pm}^\pm = \int_0^1 dx \frac{-x^2(1-x)}{R_{H^\pm}x^2 + (1 - R_{H^\pm})x}. \quad (3.2.11)$$

where $R_\phi = m_\mu^2/M_\phi^2$ ($\phi = h, H, A, H^\pm$). The dominant two-loop effect arises from a diagram in which A circulates inside the loop kind box, and it is given by

$$\delta a_\mu^{2-loops} = \frac{\alpha^2}{8\pi^2 s_W^2} \frac{m_\mu^2 g_{A\mu\mu}}{m_W^2} \sum_f N_c^f Q_f^2 R_A \bar{F}_A g_{Af_i\bar{f}_j}, \quad (3.2.12)$$

where

$$\bar{F}_A = \int_0^1 dy \frac{\log\left(\frac{R_A}{y(1-y)}\right)}{R_A - y(1 - R_A)}, \quad (3.2.13)$$

here, $g_{Af_i\bar{f}_j}$ represents the Yukawa coupling of the CP-odd scalar A to fermions f_i and \bar{f}_j . Only diagonal terms are considered, i.e., $i = j = \mu$.

$\ell_i \rightarrow \ell_j \gamma$ decays

The effective Lagrangian describing $\ell_i \rightarrow \ell_j \gamma$ processes is

$$\mathcal{L}_{\text{eff}} = C_L Q_{L\gamma} C_R Q_{R\gamma} + h.c., \quad (3.2.14)$$

the relevant dim-5 electromagnetic penguin operators are

$$Q_{L\gamma, R\gamma} = \frac{e}{8\pi^2} (\bar{\ell}_j \sigma^{\alpha\beta} P_{L,R} \ell_i) F_{\alpha\beta}, \quad (3.2.15)$$

here $F_{\alpha\beta}$ denotes the electromagnetic field strength tensor. The Feynman diagram for the process $\ell_i \rightarrow \ell_j \gamma$ is shown in Fig. 3.3.

The Wilson coefficients $C_{L,R}$ receive contributions at one-loop order and also significant contributions from the Barr-Zee diagrams at two-loops. For the specific

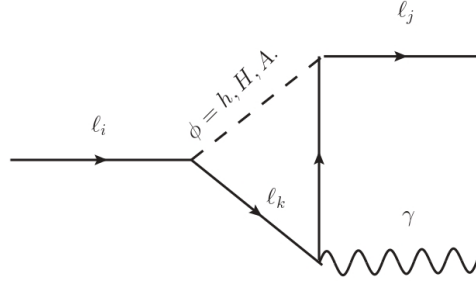


Figure 3.3: Feynman diagram for process $\ell_i \rightarrow \ell_j \gamma$ at one-loop level induced by $\phi = h, H, A$.

case in which $\ell_i = \tau$ and $\ell_j = \mu$, we adopt the approximations $g_{\phi\mu\mu} \ll g_{\phi\tau\tau}$ and $m_\mu \ll m_\tau \ll m_\phi$. Based on these assumptions, the one-loop Wilson coefficients $C_{L,R}$ simplify as follows [62, 63]

$$\begin{aligned}
 C_L^{1-loop} &\simeq \sum_{\phi} \frac{g_{\phi\tau\tau} g_{\phi\tau\mu}}{12m_\phi^2} \left(-4 + 3 \log \frac{m_\phi^2}{m_\tau^2} \right), \\
 C_R^{1-loop} &\simeq \sum_{\phi} \frac{g_{\phi\tau\tau} g_{\phi\tau\mu}}{12m_\phi^2} \left(-4 + 3 \log \frac{m_\phi^2}{m_\tau^2} \right).
 \end{aligned} \tag{3.2.16}$$

The numerical expressions for the 2-loop contributions are given by

$$\begin{aligned}
 C_L^{2-loops} &= \sum_{\phi=h,H,A} g_{\phi\tau\mu}^* (-0.082g_{\phi tt} + 0.11)/(m_\phi \text{GeV})^2, \\
 C_R^{2-loops} &= C_L^{2-loops} (g_{\phi\tau\mu}^* \rightarrow g_{\phi\tau\mu}).
 \end{aligned} \tag{3.2.17}$$

The rate for $\tau \rightarrow \mu \gamma$ is

$$\Gamma(\tau \rightarrow \mu \gamma) = \frac{\alpha m_\tau^2}{64\pi^4} (|C_L|^2 + |C_R|^2). \tag{3.2.18}$$

To obtain the corresponding width decays for the processes $\mu \rightarrow e \gamma$ and $\tau \rightarrow e \gamma$, we make the substitutions $\tau \rightarrow \mu$, $\mu \rightarrow e$ in the first case, and $\mu \rightarrow e$ in the second one, applying them consistently from Eqs. (3.2.14) to (3.2.18).

$$\ell_i \rightarrow \ell_j \ell_k \bar{\ell}_k$$

Within the 2HDM-III framework, these types of decays can proceed at tree-level via the exchange of h , H , A , as shown in Fig. 3.4. Nonetheless, the process is suppressed by the LFV Yukawa couplings $Y_{\ell_i \ell_j}$ and by the Flavour-conserving couplings $Y_{\ell_k \ell_k}$. There are also higher order contributions at one-loop and two-loops level. The partial width for the corresponding flavour violating decay is

$$\Gamma(\tau \rightarrow 3\mu) \sim \frac{\alpha m_\tau^5}{6(2\pi)^5} \left| \log \frac{m_\mu^2}{m_\tau^2} - \frac{11}{4} \right| (|C_L|^2 + |C_R|^2), \quad (3.2.19)$$

here any terms suppressed by the muon mass are neglected. The Wilson coefficients $C_{L,R}$ are given in Eqs. (3.2.16)-(3.2.17).

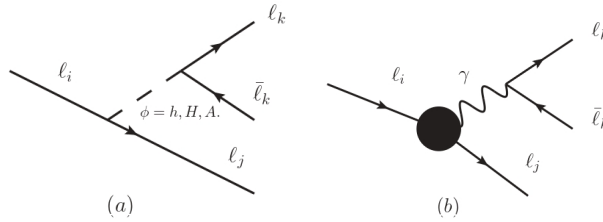


Figure 3.4: Feynman diagrams contributing to the process $\ell_i \rightarrow \ell_j \ell_k \bar{\ell}_k$. Panel (a) shows the tree-level and (b) one-loop level, where the black circle denotes a loop of the type as Feynman diagram of Fig. 3.3.

$$h \rightarrow \ell_i \ell_j$$

The LFV processes $h \rightarrow \ell_i \ell_j$ ($\ell_{i,j} = \ell_{i,j}^-, \ell_{i,j}^+$) where $\ell_i \ell_j = e\mu, e\tau, \tau\mu$ may occur at tree-level in various SM extensions, as shown in Fig. 3.5.

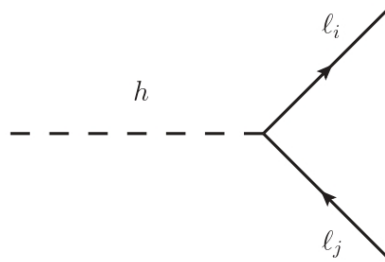


Figure 3.5: Feynman diagram for the tree-level decay $h \rightarrow \ell_i \ell_j$.

The relevant interactions can be derived from the Yukawa Lagrangian

$$\mathcal{L}_Y \supset -Y_{ij} \bar{\ell}_L^i \ell_R^j h + h.c. \quad (3.2.20)$$

The total decay width for $h \rightarrow \bar{f}_i f_j$ is given by

$$\Gamma(h \rightarrow \bar{f}_i f_j) = \frac{N_c g_{h\bar{f}_i f_j}^2 m_h}{128\pi} \left[4 - (\sqrt{\tau_{f_i}} + \sqrt{\tau_{f_j}})^2 \right]^{3/2} \sqrt{4 - (\sqrt{\tau_{f_i}} - \sqrt{\tau_{f_j}})^2}, \quad (3.2.21)$$

where $g_{h\bar{f}_i f_j}$ is the coupling associated with $h\bar{f}_i f_j$ arising from an SM extension, $N_c = 3(1)$ is the colour factor for quarks (leptons), m_h denotes the Higgs boson mass and $\tau_i = 4m_i^2/m_h^2$.

Having presented all the analytical expressions needed to calculate the BRs for the observables described, we proceed to evaluate all processes discussed in Sec. 3.2.1 with the `Mathematica` package called `SpaceMath`¹ [64], in order to find the permissible values of the free parameters in our analysis. In Fig. 3.6, we present the plane $\cos(\alpha - \beta) - \tan \beta$, where each point corresponds to a parameter combination allowed by experimental constraints. The orange points satisfy the measured (or upper limits) BRs of $B_{s,d} \rightarrow \mu^+ \mu^-$ and the LFV processes described above. Meanwhile, the blue points accommodate the LHC Higgs boson data. By considering all the observables jointly, it is possible to identify a parameter region that is consistent with every constraint, thus enabling the exploration of different scenarios to predict further observables. As an illustration, two such realistic parameter scenarios are:

- $-0.04 \lesssim \cos(\alpha - \beta) \lesssim 0.025$ for $1 \lesssim \tan \beta \lesssim 10$,
- $-0.01 \lesssim \cos(\alpha - \beta) \lesssim 0.01$ for $1 \lesssim \tan \beta \lesssim 50$.

These bounds extend to the decoupling limit $\alpha - \beta \rightarrow \pi/2$, which suppresses flavour-violating decays of the SM-like Higgs boson. Simultaneously, having $\cos(\alpha - \beta) \sim 0$ can enhance such processes in decays of the heavy Higgs boson H .

¹This software implements all the experimental constraints considered in this research.

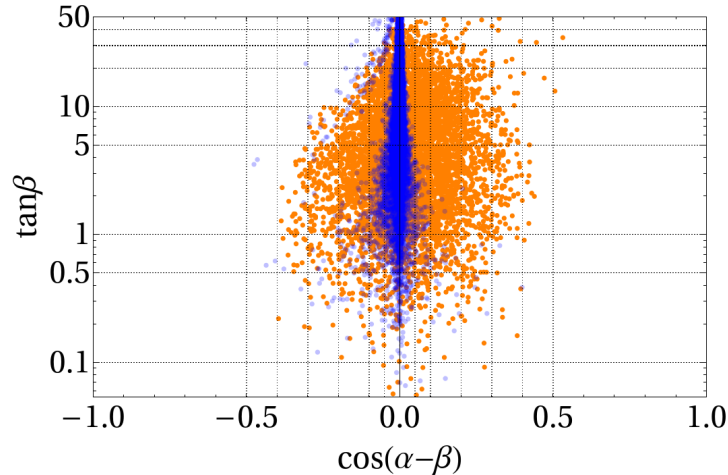


Figure 3.6: In the $\cos(\alpha-\beta) - \tan\beta$ plane, the blue points represent parameter values allowed by all signal strength modifiers μ_X , whereas the orange points indicate those consistent with the LFV processes. The dataset was generated using the SpaceMath framework.

Further details on the individual allowed parameter values for each observable are included in Appx. A.

3.2.2 Constraint on M_H , M_A , and M_{H^\pm}

Collider Constraints

The ATLAS and CMS collaborations have conducted several searches to look for additional neutral Higgs bosons in different channels. Among these is the di-tau channel $gb \rightarrow \phi \rightarrow \tau\tau$, with $\phi = A, H$. The corresponding Feynman diagram is shown in Figure 3.7. Although no evidence of extra Higgs bosons was found, upper limits were imposed on the production cross-section $\sigma(gb \rightarrow \phi)$ multiplied by $\mathcal{BR}(\phi \rightarrow \tau\tau)$.

Fig. 3.8a shows the production cross-section $\sigma(gb \rightarrow Ab) \times \mathcal{BR}(A \rightarrow \tau\tau)$ as a function of M_A . The curves are shown for representative values of $\tan\beta = 5, 10, 20$ and $\cos(\alpha-\beta) = 0.01$. Meanwhile, Fig. 3.8b presents the analogous results for $\phi = H$. In both plots, the black points correspond to the expected 95% CL upper limits, whereas the red crosses indicate the observed limits at the same confidence level.

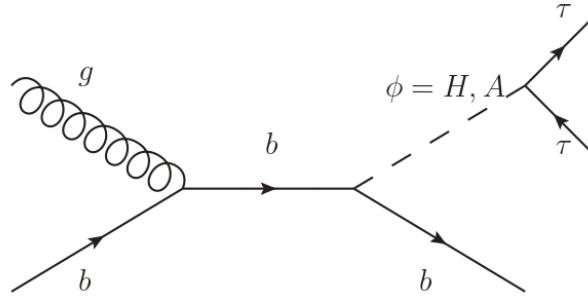
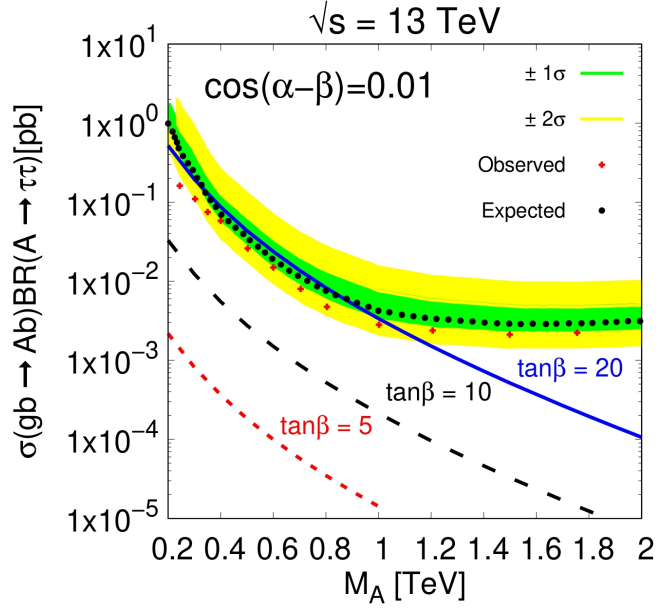


Figure 3.7: Feynman diagram for the production of ϕ in association with a bottom quark at the LHC, followed by its decay into a $\tau\tau$ pair.

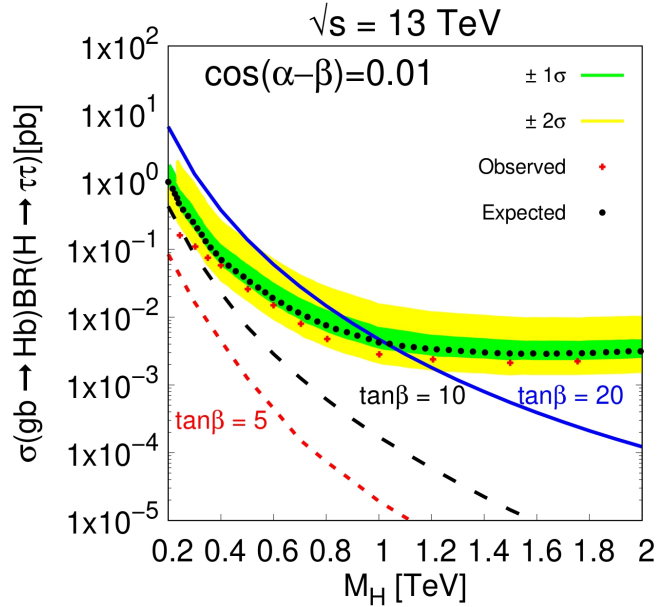
Additionally, the green (yellow) shaded bands illustrate the $\pm 1\sigma$ ($\pm 2\sigma$) intervals around the expected limits, providing a visual measure of the experimental and theoretical uncertainties.

Moreover, from Fig. 3.8a, it is observed that $M_A \lesssim 1$ TeV ($M_A \lesssim 1.2$ TeV) is excluded at 1σ (2σ) CL for $\tan\beta = 20$. On the other hand, for lower values of $\tan\beta \lesssim 5, 10$, the upper limit on $\sigma(gb \rightarrow \phi b) \times \mathcal{BR}(\phi \rightarrow \tau\tau)$ is readily satisfied, indicating that such scenarios are not significantly constrained by the current experimental data. Similarly, from Fig. 3.8b, it is found that $M_H \lesssim 1.1$ TeV ($M_H \lesssim 1.3$ GeV) is excluded at 1σ (2σ) for $\tan\beta = 20$. For $\tan\beta \lesssim 15$, a wide range of masses is allowed, particularly for $M_A > 300$. Additionally, another process used to constrain the mass of the heavy scalar M_H involves its decay into a pair of Higgs bosons, which subsequently decay into $b\bar{b}$ and $\gamma\gamma$, i.e., $pp \rightarrow H \rightarrow hh, h \rightarrow b\bar{b}, h \rightarrow \gamma\gamma$, as shown in Fig. 3.9. We focus on this final state because both ATLAS and CMS have performed dedicated searches in this channel [45]. The $\gamma\gamma$ final state provides excellent mass resolution, allowing the comparison of our results with the existing experimental constraints in a well studied channel. This decay, $h \rightarrow \gamma\gamma$, receives loop-level contributions from the W^\pm bosons and the top quark, as in the Standard Model, and an additional contribution from the charged scalar boson H^\pm present in the 2HDM-III. The partial width can be computed using the effective vertex approximation [18],

$$\Gamma(h \rightarrow \gamma\gamma) = \frac{\alpha^2 m_h^3}{1024\pi^3 m_W^2} |A^{h\gamma\gamma}|^2, \quad (3.2.22)$$



(a)



(b)

Figure 3.8: Observed and expected 95% CL upper limits on the production cross-section multiplied by the di-tau BR for a scalar boson produced in association with a bottom quark. The results are shown as a function of (a) M_A and (b) M_H , where we consider $\tan\beta = 5, 10, 20$ and $\cos_{\alpha\beta} = 0.01$.

where the total amplitude includes contributions from all particles running in the loop,

$$A^{h\gamma\gamma} = \sum_s A_s^{h\gamma\gamma}(\tau_s), \quad (3.2.23)$$

with s denoting the spin of the loop particle. For the scalar contribution, $s = 0$,

and $\tau_0 = \tau_{H^\pm} = 4m_{H^\pm}^2/m_h^2$. The amplitude $A_0^{h\gamma\gamma}$ is then given by

$$A_0^{h\gamma\gamma}(\tau_{H^\pm}) = \frac{m_W g_{hH^+H^-}}{m_{H^\pm}^2} F_0(\tau_{H^\pm}), \quad (16)$$

where $g_{hH^+H^-}$ is the trilinear coupling between the Higgs boson h and the charged Higgs bosons. The loop functions $F_0(x)$ and $f(x)$ are defined as

$$F_0(x) = x [1 - xf(x)], \quad (3.2.24)$$

$$f(x) = \begin{cases} \left[\arcsin\left(\frac{1}{\sqrt{x}}\right) \right]^2, & x \geq 1, \\ -\frac{1}{4} \left[\log\left(\frac{1+\sqrt{1-x}}{1-\sqrt{1-x}}\right) - i\pi \right]^2, & x < 1. \end{cases} \quad (3.2.25)$$

The vertex used in the simulation for the interaction between the H and the top quarks follows the effective Lagrangian

$$\mathcal{L}_Y^\phi = \phi \bar{f}_i (S_{ij}^\phi + iP_{ij}^\phi \gamma^5) f_j,$$

where S_{ij}^ϕ and P_{ij}^ϕ denote the scalar and pseudoscalar couplings, respectively. In the CP-conserving case, the pseudoscalar term vanishes for H , so $P_{tt}^H = 0$.

For the scalar component, the relevant term is

$$S_{tt}^H = \frac{g m_t}{2M_W} c_u^H + d_u^H [\tilde{Y}_a^u]_{tt}, \quad \text{with } \frac{g m_t}{2M_W} = \frac{m_t}{v}.$$

Using the expressions $c_u^H = \sin \alpha / \sin \beta$ and $d_u^H = \sin(\alpha - \beta) / \sin \beta$, along with the Yukawa texture $[\tilde{Y}_a^u]_{tt} = (m_t/v) \chi_{tt}$, the coupling becomes

$$S_{tt}^H = \frac{m_t}{v} \left[\frac{\sin \alpha}{\sin \beta} + \frac{\sin(\alpha - \beta)}{\sin \beta} \chi_{tt} \right], \quad (3.2.26)$$

defining a correction factor

$$\kappa_{Htt} = \frac{\sin \alpha}{\sin \beta} + \frac{\sin(\alpha - \beta)}{\sin \beta} \chi_{tt}.$$

Thus, the SM-like coupling $(m_t/v) \bar{t}t h$ is replaced by $(m_t/v) \kappa_{Htt} \bar{t}t H$.

The cross-section $\sigma(pp \rightarrow H \rightarrow hh, h \rightarrow b\bar{b}, h \rightarrow \gamma\gamma)$ as a function of M_H is shown in Fig. 3.10, where the upper limits at 1σ (green band) and 2σ (yellow band) are displayed alongside the expected limit (dashed black line) and the observed

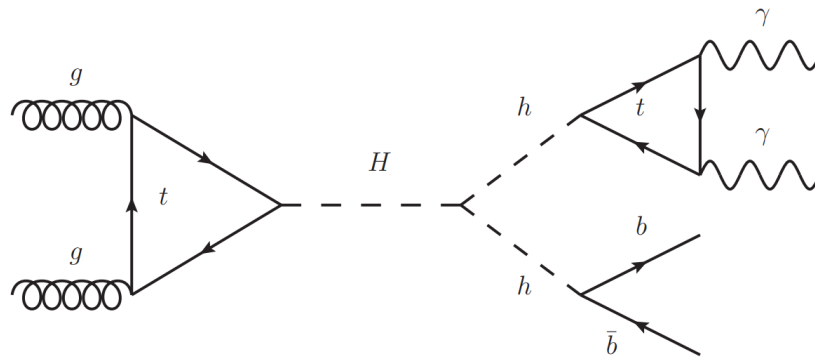


Figure 3.9: Feynman diagram illustrating the production process of H and its subsequent decay into a pair of Higgs bosons hh .

limit (blue line). These results, reported by the ATLAS collaboration [45], are compared with the theoretical predictions of the 2HDM-III for $\tan\beta = 5, 10, 20$ and $\cos(\alpha - \beta) = 0.01$. We note that the aforementioned limits on M_H are readily satisfied within the 2HDM-III framework for values of $\tan\beta \lesssim 18$. The $H \rightarrow VV$, ($V = WZ$) channel is also examined [65, 66]. However, the constraints derived from this channel are not particularly stringent in the current model due to the proportionality of the coupling g_{HVV} to $\cos(\alpha - \beta)$. Additionally, the $t\bar{t}$ channel is analysed and its predictions are compared with experimental results from a model independent analysis [67]. The cross-section predicted by the 2HDM-III is found to be heavily suppressed, being 4 to 8 orders of magnitude smaller, as it scales inversely with $\tan^4\beta$.

$b \rightarrow s\gamma$ decay

The detection of the charged scalar boson H^\pm would provide a clear signature of new physics. Constraints on its mass M_{H^\pm} have been obtained from collider searches focusing on the production of H^\pm and its subsequent decay into a $\tau^\pm\nu_\tau$ pair [48]. However, we find that such processes are not particularly effective in constraining the charged scalar mass M_{H^\pm} within the framework of the 2HDM-III. In contrast, the decay $b \rightarrow s\gamma$, imposes stringent lower limits on M_{H^\pm} due to the contributions from the charged scalar boson [50, 51]. In Fig. 3.12(a)-(b), a scatter plot in the

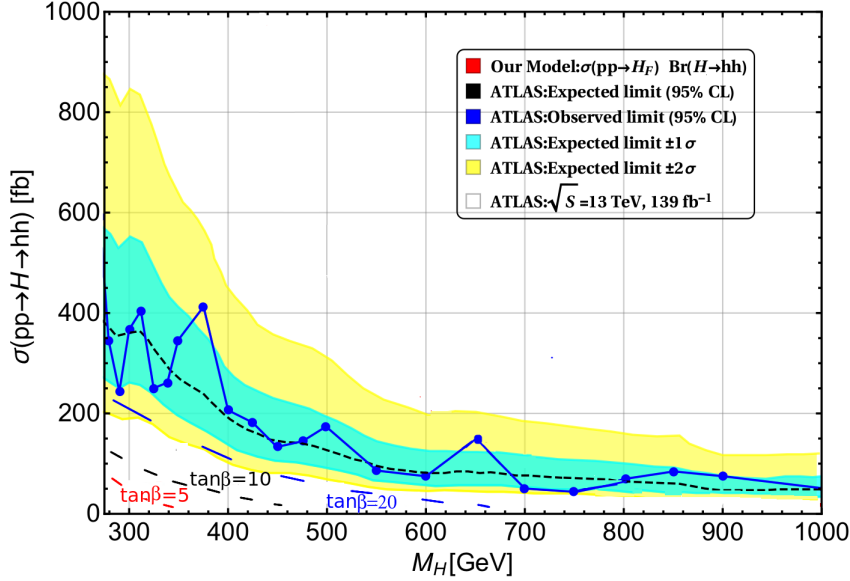


Figure 3.10: The 95% CL upper limits on the production cross-section multiplied by the di-Higgs BR are presented for a scalar boson produced in pp collisions as a function of M_H . We observed and expected limits are shown, considering $\tan\beta = 5, 10, 20$ and $\cos(\alpha - \beta) = 0.01$.

M_{H^\pm} - $\tan\beta$ plane is presented, where red and blue points represent regions allowed by the ratio $2.77 \times 10^{-3} < R_{\text{quark}} < 3.67 \times 10^{-3}$ [51], defined as,

$$R_{\text{quark}} = \frac{\Gamma(b \rightarrow X_s \gamma)}{\Gamma(b \rightarrow X_c e \nu_e)}, \quad (3.2.27)$$

here, X_s and X_c denote inclusive hadronic final states containing a strange and charm quark, respectively.

A preference for low $\tan\beta$ is observed in both mass ranges, $114 \leq M_{H^\pm} \leq 140$ GeV and $m_t + m_b \leq M_{H^\pm} \leq 1000$ GeV. This preference arises from the structure of the contributions in the 2HDM-III, where the couplings $g_{\bar{u}dH^-} \sim 1/\tan\beta(1 - \chi_{tt}/\sqrt{2})$ and $g_{\bar{d}uH^+} \sim \tan\beta(1 - \chi_{bb}/\sqrt{2})$ indicate the large values of $\tan\beta$ could lead to significant contributions to $b \rightarrow s\gamma$ potentially conflicting with experimental constraints. The partial width is proportional to the squared amplitude, $\Gamma(b \rightarrow s\gamma) \propto |C_7^{(R)}|^2 + |C_7^{(L)}|^2$, where $C_7^{(R)}$ and $C_7^{(L)}$ are the Wilson coefficients of the magnetic dipole operators contributing to the effective Hamiltonian. The contribution to $C_7^{(L)}$ [68], takes the

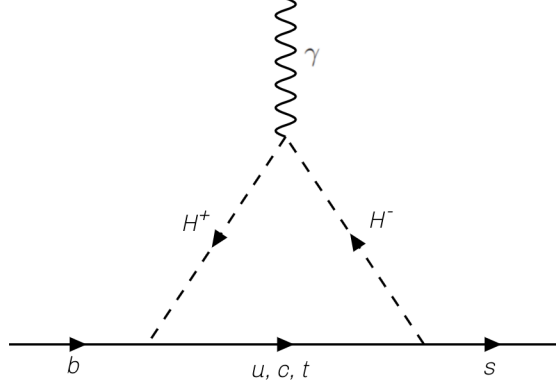


Figure 3.11: Radiative decay $b \rightarrow s\gamma$ mediated by charged Higgs bosons. The loop involves up-type quarks (u, c, t), with the dominant contribution coming from the top quark. The photon is emitted from the charged scalar line.

form

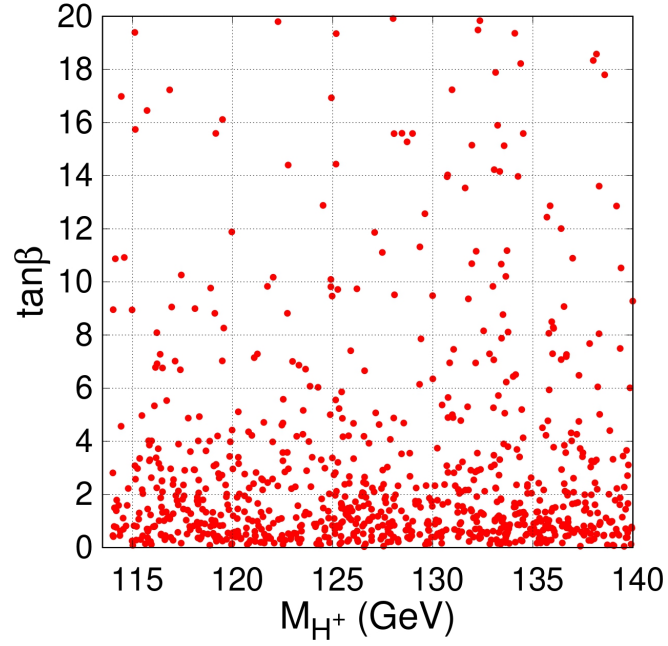
$$C_7^{(L)} = \frac{v^2}{\lambda_t m_b} \sum_j \frac{\Gamma_{u_j d_2}^{RLH^+} \Gamma_{u_j d_3}^{LRH^+}}{m_{u_j}} C_{7,XY}^0(y_j) + \frac{v^2}{\lambda_t} \sum_j \frac{\Gamma_{u_j d_2}^{RLH^+} \Gamma_{u_j d_3}^{LRH^+}}{m_{u_j}^2} C_{7,YY}^0(y_j),$$

where the charged Higgs couplings are given by

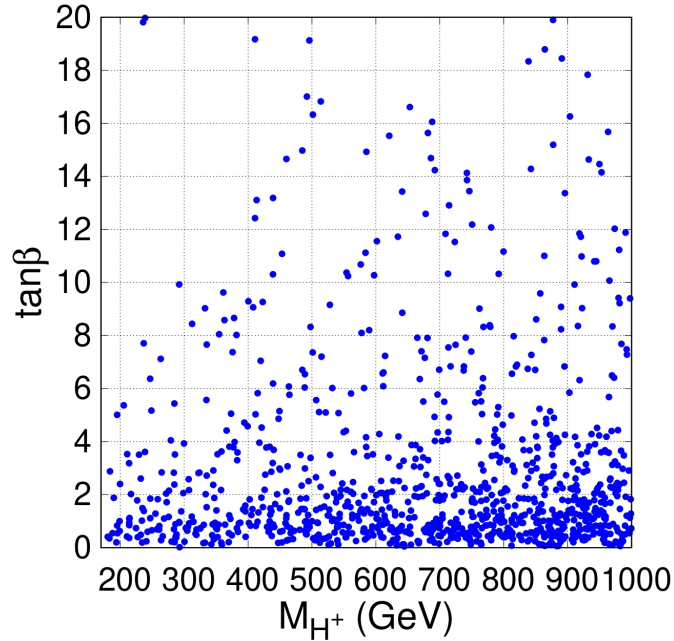
$$\Gamma_{u_f d_i}^{LRH^+} = \sum_j \sin \beta V_{fj} \left(\frac{m_{d_j}}{v_d} \delta_{ji} - \epsilon_{ji}^d \tan \beta \right),$$

$$\Gamma_{d_f u_i}^{LRH^+} = \sum_j \cos \beta V_{jf}^* \left(\frac{m_{u_j}}{v_u} \delta_{ji} - \epsilon_{ji}^u \tan \beta \right).$$

Here, $\lambda_t = V_{tb} V_{ts}^*$. The functions $C_{7,XY}^0(y)$ and $C_{7,YY}^0(y)$ are defined in [69], with $y_j = m_{u_j}^2/m_{H^\pm}^2$. The corresponding contribution to $C_7^{(R)}$ is suppressed by light quark masses and flavour-changing couplings. The lower mass range, $114 \leq M_{H^\pm} \leq 140$, is particularly motivated by the recent excess of events reported by the ATLAS collaboration [49]. It is worth noting that the parameters χ_{tt} and χ_{bb} play a crucial role in attenuating the magnitude of the $g_{\bar{d}uH^+}$ and $g_{\bar{u}dH^-}$ couplings. This attenuation allows for lighter charged scalar masses in the 2HDM-III compared to the 2HDM-I and -II, where the couplings are fixed as $g_{\bar{u}dH^-}^I = g_{\bar{d}uH^+}^I = 1/\tan \beta$, $g_{\bar{u}dH^-}^{II} = 1/\tan \beta$ and $g_{\bar{d}uH^+}^{II} = \tan \beta$. In Fig. 3.13, the $\chi_{bb} - \chi_{tt}$ plane is displayed, with red points rep-



(a)



(b)

Figure 3.12: Scatter plot in the $M_{H^\pm} - \tan\beta$ plane. (a) Red points stand for these regions allowed by R_{quark} for the mass interval $114 \leq M_{H^\pm} \leq 140$ GeV. (b) The same as in (a) but for the mass range $m_t + m_b \leq M_{H^\pm} \leq 1000$. The R_{quark} is defined in Eq. (3.2.27).

representing parameter combinations satisfying the experimental constraints on R_{quark} as defined in Eq. (3.2.27). Meanwhile, the scanned parameters are summarised in Tab. 3.3.

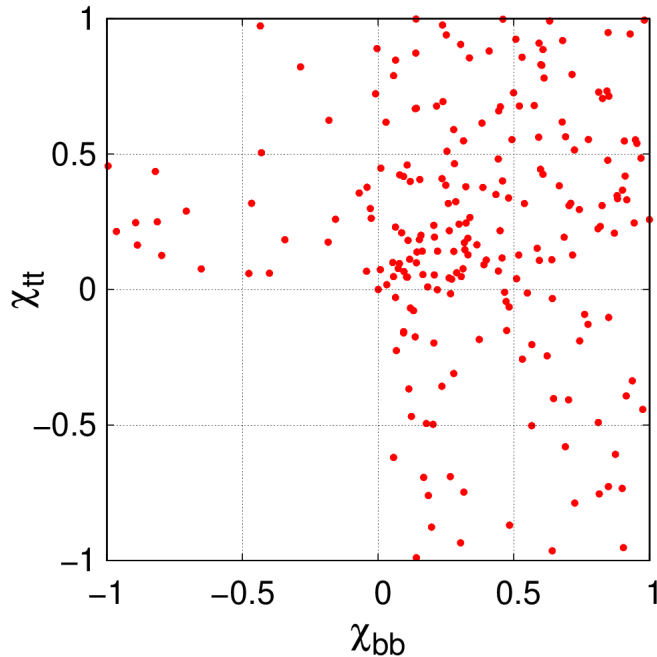


Figure 3.13: Scatter plot in the $\chi_{bb} - \chi_{tt}$ plane. Red points represent parameter combinations allowed by R_{quark} as defined in Eq. (3.2.27).

Parameter	Scanned range
$\tan \beta$	[0.1, 20]
χ_{tt}	[-1, 1]
χ_{bb}	[-1, 1]
M_{H^\pm}	[114, 140] GeV

Table 3.3: Scanned range of the parameters. The values of χ_{ij} are set to 1 for fermions i and j not included in the Table (in general $i \neq j$).

Given that this analysis focuses on light charged scalar masses ($M_{H^\pm} < m_t + m_b$), it is important to highlight that the model under consideration could potentially explain the recent slight excess in data, with a significance of around 3σ , reported by the ATLAS collaboration for $M_{H^\pm} = 130$ [49]. This analysis used a data set of collisions pp collected at a centre of mass energy of $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of 139 fb^{-1} . The search targeted a data sample enriched in top quark pair production, where one top quark decays into a leptonically decaying W boson and a bottom quark, while the other top quark may decay into a H^\pm boson and a bottom quark. The model independent exclusion at the 95% CL on the product of BRs, $\mathcal{BR} = \mathcal{BR}(t \rightarrow H^\pm b) \times \mathcal{BR}(H^\pm \rightarrow cb)$ was reported as a function of

Parameter	Scanned range
$\tan \beta$	[0.1, 20]
χ_{tb}	[-10, 10]
χ_{cb}	[-10, 10]
$\chi_{\mu\mu}$	[-10, 10]
M_{H^\pm}	[114, 140] GeV

Table 3.4: Scanned range of the parameters. The values of χ_{ij} are set to 1 for fermions i and j not included in the Table (in general $i \neq j$).

M_{H^\pm} [49].. In Fig. 3.14 we present the $\chi_{cb} - \chi_{tb}$ plane for $0.1 < \tan \beta < 20$, $\chi_{ij} = 1$, $-10 < \chi_{\mu\mu} < 10$. Blue, red, and green points correspond to parameter values that can accommodate the current excess for $M_{H^\pm} = 120$ (130, 140) GeV, respectively.

We highlight two key characteristics:

1. The density of red points is significantly higher compared to blue and green points. This is because the red region corresponds to parameter values associated with the excess for $M_{H^\pm} = 130$ GeV, which represents the largest excess reported by the ATLAS collaboration,. In contrast, the excess for $M_{H^\pm} = 140$ GeV is less pronounced, as illustrated in Fig. 8 of Ref. [49].
2. Based on our scan over the model parameter space, as shown in Table 3.4, the parameters χ_{tb} and χ_{cb} play a crucial role in accommodating the current excess due to their high sensitivity. This is a distinctive feature of the 2HDM-III, as the 2HDM-I, II, Lepton Specific and Flipped, which lack these parameters. It is worth noting that several parameters χ_{ij} , associated with fermions i and j , influence the total width decay of the charged scalar boson H^\pm . In our analysis, we have set $\chi_{ij} = 1$ for simplicity.

Oblique parameters

We require that the Peskin-Takeuchi parameters S , T and U lie within the experimental results [70],

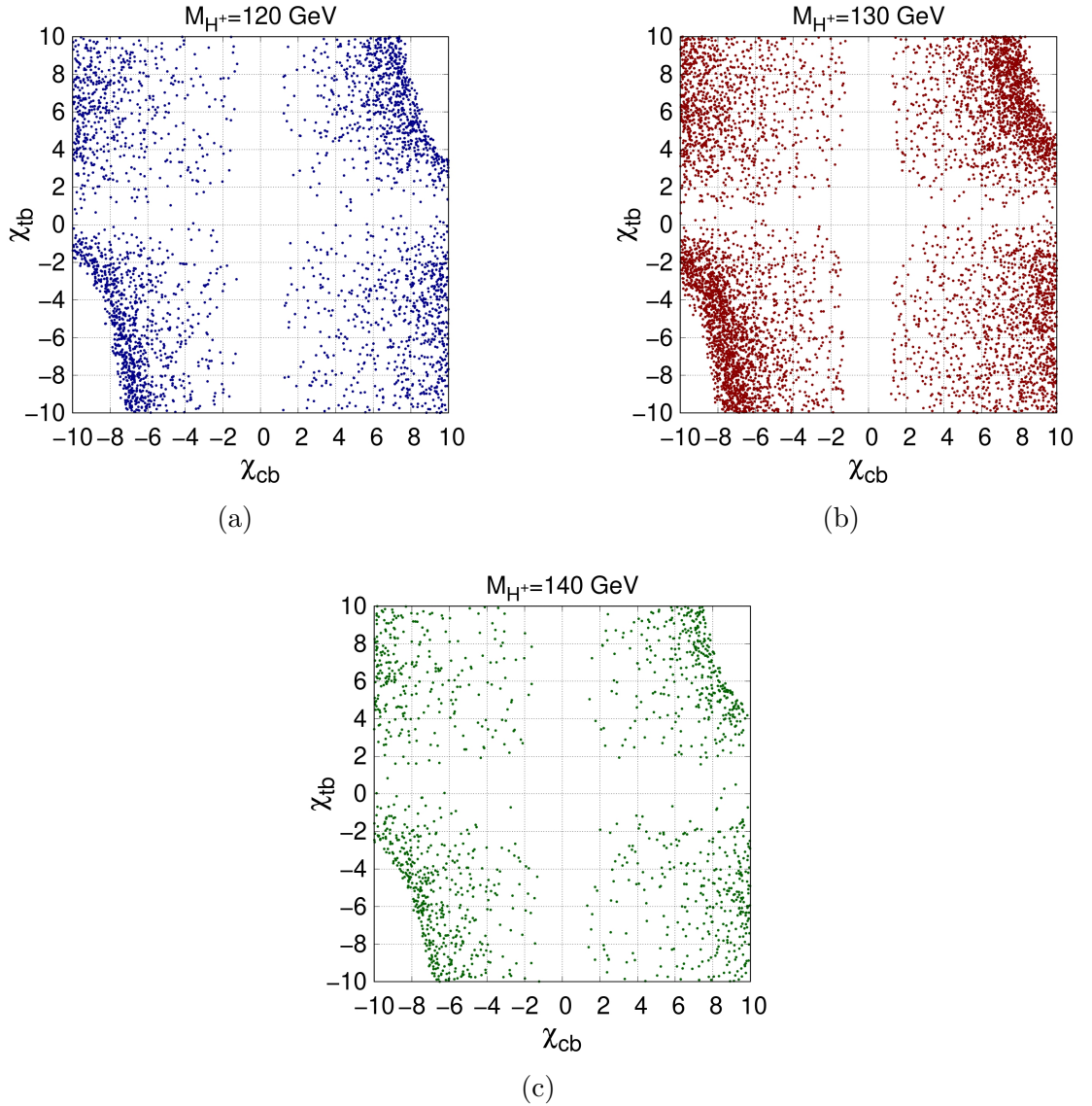


Figure 3.14: χ_{cb} - χ_{tb} plane. The coloured points indicate the values of χ parameters that accommodate the current excess of events reported by the ATLAS collaboration. (a) Blue points represent parameter values that explain the excess for $M_{H^\pm} = 120$ GeV. (b) and (c) show the same as in (a) but for $M_{H^\pm} = 130$ GeV and $M_{H^\pm} = 140$ GeV, respectively.

- $S = -0.04 \pm 0.1$,
- $T = 0.01 \pm 0.12$,
- $U = -0.01 \pm 0.09$.

These oblique parameters quantify the effects of new physics on electroweak precision observables by modifying the propagation of the gauge bosons. The parameter S reflects new contributions to neutral current processes via the Z boson self energy; T measures the breaking of custodial symmetry through differences in the W and Z propagators at zero momentum; and U accounts for momentum dependent corrections to charged current processes involving the W boson. These parameters are defined with the SM contributions subtracted and offer a model independent framework for constraining extended scalar sectors [71]. The expressions for S , T , and U are provided in Appx. B. For a scenario with $\cos(\alpha - \beta) = 0.01$ and $0.1 \leq \tan \beta \leq 15$, the oblique parameters impose stringent constraints on the mass differences among the scalars predicted by the model.

We present in Fig. 3.15 a scattering plot where the points correspond to values of M_H , M_A and M_{H^\pm} that simultaneously satisfy the constraints imposed by the oblique parameters U , S , T . We observe that the dominant contribution to the production cross-section $\sigma(pp \rightarrow H^- H^+)$ arises from the on-shell decay $H \rightarrow H^+ H^-$. This could lead to a suppression of the signal cross-section if the masses M_{H^\pm} , M_H , M_A are nearly degenerate. To avoid this, we analyse scenarios where $M_{H^\pm} - M_H \leq -250$ GeV. Specifically, we consider three mass values for the neutral scalar boson $M_H = 500, 800, 1000$ GeV. For the case $M_H = 500$ GeV, the charged scalar mass is constrained to $M_{H^\pm} \leq 250$ GeV corresponding to $M_H - M_A \approx 250$ GeV and $M_{H^\pm} \approx M_A$. Although the oblique parameters U , S , and T impose stringent constraints on the mass differences, the scenarios proposed below allow us to evade their constraints, as well as restrictions from collider searches and LFV processes.

In summary, based on our systematic exploration of the 2HDM-III parameter space, we propose three phenomenologically viable scenarios for collider studies, which are

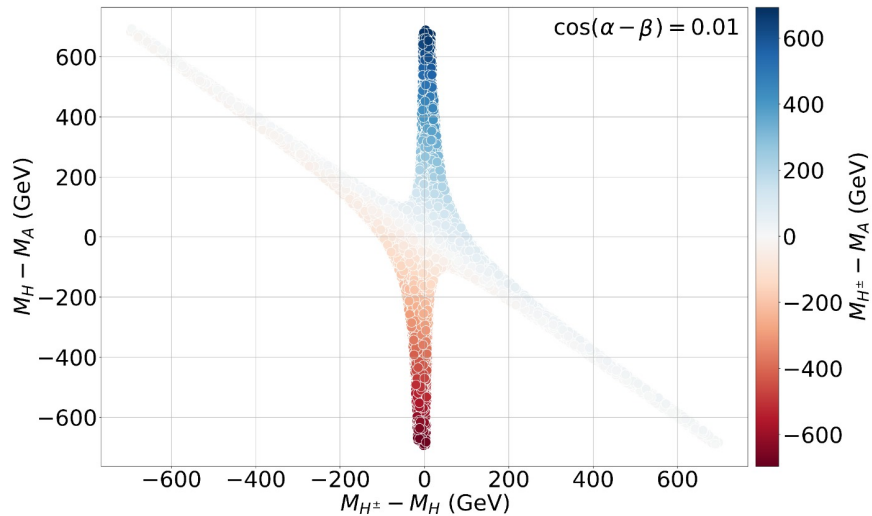


Figure 3.15: Allowed points (simultaneously) for the difference of scalar masses by oblique parameters.

outlined in Table 3.5. These scenarios will serve as benchmarks for the simulations discussed in the subsequent section.

Scenario	$\tan \beta$	χ_{tt}	χ_{bc}	$\chi_{\mu\mu}$	$\cos(\alpha - \beta)$	$M_H(\text{GeV})$
$S1$	1	0.1	10	10	0.01	500, 800, 1000
$S2$	5	0.1	5	5	0.01	500, 800, 1000
$S3$	10	0.1	1	1	0.01	500, 800, 1000

Table 3.5: Benchmark scenarios to be used in the subsequent calculations.

Through these simulations, we seek to explore the phenomenological consequences of the model and identify potential signatures that could be detected in forthcoming collider experiments.

3.3 Flavon Model

In addition to the 2HDM-III, we explore the Flavon Singlet Model or Froggatt-Nielsen Singlet Model (FNSM), a different extension of the SM that introduces a complex singlet S_F , referred to here as the Flavon. In this FNSM, the VEV of the complex singlet $\langle S_F \rangle = v_s/\sqrt{2}$ spontaneously breaks the global $U(1)_F$ Flavour

symmetry. As a result, the physical spectrum includes a CP-even Flavon, H_F , and a CP-odd Flavon, A_F , alongside the SM-like Higgs boson, h . In this section we present an analysis of the scalar and Yukawa sectors within the FNSM framework. We derive constraints from perturbativity, unitarity, and vacuum stability in a similar manner to the 2HDM-III treatment. We then incorporate experimental bounds on Higgs observables, including signal strengths, LFV processes, and rare meson decays, to highlight the viable parameter space of the FNSM.

3.3.1 Scalar Potential

The scalar sector of the FNSM extends the SM by introducing the singlet complex FN scalar S_F , which can be expressed as,

$$S_F = \frac{(v_s + S_R + iS_I)}{\sqrt{2}}, \quad (3.3.1)$$

where v_s denotes the VEV of the FN singlet. The scalar potential remains invariant under the $U(1)_F$ Flavour symmetry associated with the FN mechanism. Under this Flavour symmetry, the S_F transforms as $S_F \rightarrow e^{i\alpha} S_F$, while the SM Higgs doublet $\Phi = \left(\frac{v+\phi^0}{\sqrt{2}}, \phi^+\right)^T$ remains unchanged, i.e., $\Phi \rightarrow \Phi$, it is important to note the convention used here differs from the definition above. The scalar potential allows a complex VEV of the form $(S_F)_0 = \frac{v_s}{\sqrt{2}} e^{i\alpha}$. However, in this work, we restrict our analysis to the CP-conserving scenario by setting $\alpha = 0$. In this case, the CP-conserving Higgs potential takes the following form:

$$\begin{aligned} V_0 = & -\frac{1}{2} m_1^2 \Phi^\dagger \Phi - \frac{1}{2} m_2^2 S_F^* S_F + \frac{1}{2} \lambda_1 (\Phi^\dagger \Phi)^2 \\ & + \lambda_2 (S_F^* S_F)^2 + \lambda_3 (\Phi^\dagger \Phi) (S_F^* S_F). \end{aligned} \quad (3.3.2)$$

After that the VEVs of the spin-0 fields (Φ, S_F) spontaneously break the $U(1)_F$ Flavour symmetry, a massless Goldstone boson appears in the physical spectrum. We introduce a soft $U(1)_F$ breaking term into the scalar potential to give it a mass,

$$V_{\text{soft}} = -\frac{m_3^2}{2} (S_F^2 + S_F^{*2}). \quad (3.3.3)$$

Thus, the full scalar potential is expressed as,

$$V = V_0 + V_{\text{soft}}. \quad (3.3.4)$$

The parameter λ_3 in Eq. (3.3.2) enables the mixing between the Flavon and Higgs fields once the $U(1)_F$ Flavour and EW symmetries are spontaneously broken, thereby determining the masses of these two states. In contrast, the soft $U(1)_F$ breaking term V_{soft} is responsible for generating the pseudoscalar Flavon (S_I) mass. Upon minimising V , one obtains the following relations among its parameters

$$m_1^2 = v^2\lambda_1 + v_s^2\lambda_3, \quad (3.3.5)$$

$$m_2^2 = -2m_3^2 + 2v_s^2\lambda_2 + v^2\lambda_3. \quad (3.3.6)$$

Since all the parameters of the scalar potential are taken to be real, the imaginary and real components remain separate. Consequently, the CP-even mass matrix can be expressed in the (ϕ_0, S_R) basis as

$$M_S^2 = \begin{pmatrix} \lambda_1 v^2 & \lambda_3 v v_s \\ \lambda_3 v v_s & 2\lambda_2 v_s^2 \end{pmatrix}, \quad (3.3.7)$$

whose mass eigenstates are derived via the 2×2 rotation,

$$\phi^0 = \cos \alpha h + \sin \alpha H_F, \quad (3.3.8)$$

$$S_R = -\sin \alpha h + \cos \alpha H_F, \quad (3.3.9)$$

where α denotes the mixing angle. We identify h with the SM-like Higgs boson with a mass $M_h = 125.5$ GeV. While the other CP-even eigenstate H_F is associated with the Flavon. The CP-odd Flavon is denoted as $A_F \equiv S_I$ and acquires its mass through the soft breaking term V_{soft} such that $M_{A_F}^2 = 2m_3^2$. It is assumed that both the CP-even H_F and the CP-odd A_F are heavier than h . In this analysis, we treat the physical masses $M_S (S = h, H_F, A_F)$ and the mixing angle α as free parameters. Their relations with the quartic couplings of the scalar potential in Eq. (3.3.2) can

be expressed as,

$$\begin{aligned}
\lambda_1 &= \frac{\cos \alpha^2 M_h^2 + \sin \alpha^2 M_{H_F}^2}{v^2}, \\
\lambda_2 &= \frac{M_{A_F}^2 + \cos \alpha^2 M_{H_F}^2 + \sin \alpha^2 M_h^2}{2v_s^2}, \\
\lambda_3 &= \frac{\cos \alpha \sin \alpha}{vv_s} (M_{H_F}^2 - M_h^2).
\end{aligned} \tag{3.3.10}$$

3.3.2 Yukawa Lagrangian

The effective $U(1)_F$ invariant Yukawa Lagrangian, which yields the Yukawa couplings once the $U(1)_F$ Flavour symmetry is spontaneously broken, is given by [7]

$$\begin{aligned}
\mathcal{L}_Y &= \rho_{ij}^d \left(\frac{S_F}{\Lambda} \right)^{q_{ij}^d} \bar{Q}_i d_j \tilde{\Phi} + \rho_{ij}^u \left(\frac{S_F}{\Lambda} \right)^{q_{ij}^u} \bar{Q}_i u_j \Phi \\
&\quad + \rho_{ij}^\ell \left(\frac{S_F}{\Lambda} \right)^{q_{ij}^\ell} \bar{L}_i \ell_j \Phi + \text{h.c.},
\end{aligned} \tag{3.3.11}$$

In this expression, ρ_{ij}^f ($f = u, d, \ell$) are dimensionless parameters, typically of order $\mathcal{O}(1)$. The quantities q_{ij}^f denote the Abelian charges that are assigned in order to reproduce the observed fermion masses, and Λ represent the ultraviolet mass scale. In order to derive the Yukawa coupling from the Lagrangian in (3.3.11), it is necessary to spontaneously break both the $U(1)_F$ Flavour symmetry and the EW symmetry. After this symmetry breaking, the resulting $S f_i \bar{f}_i$ interactions emerge as it is shown in Table 3.6. It is important to note that to avoid significant deviations from the SM couplings, the conditions $v_s \approx \Lambda$ and $\cos \alpha \approx -1$ must be satisfied.

Meanwhile, to induce non-diagonal $S f_i \bar{f}_j$ interactions, we proceed as follows. In the unitary gauge, one can perform a first-order expansion of the neutral component of the neutral component of the heavy Flavour field S_F around its v_s .

$$\begin{aligned}
\left(\frac{S_F}{\Lambda_F} \right)^{q_{ij}} &= \left(\frac{v_s + S_R + iS_I}{\sqrt{2}\Lambda_F} \right)^{q_{ij}} \\
&\simeq \left(\frac{v_s}{\sqrt{2}\Lambda_F} \right)^{q_{ij}} \left[1 + q_{ij} \left(\frac{S_R + iS_I}{v_s} \right) \right],
\end{aligned} \tag{3.3.12}$$

This expansion leads to the following interaction Lagrangian upon replacing the

Vertex SXX	Coupling
$hf_i\bar{f}_i$	$\frac{v_s m_f}{v\Lambda^2} (v \sin \alpha - v_s \cos \alpha)$
hZZ	$g_{c_W}^{m_Z} \cos \alpha$
hWW	$gm_W \cos \alpha$
$H_F f_i \bar{f}_i$	$-\frac{v_s m_f}{v\Lambda^2} (v_s \sin \alpha + v \cos \alpha)$
$H_F ZZ$	$g_{c_W}^{m_Z} \sin \alpha$
$H_F WW$	$gm_W \sin \alpha$
$A_F f_i \bar{f}_i$	$-\frac{v_s m_f}{\Lambda^2}$
$A_F ZZ$	0
$A_F WW$	0

Table 3.6: Diagonal SXX interactions, ($S = h, H_F, A_F$).

mass eigenstates,

$$\begin{aligned}
\mathcal{L}_Y &= \frac{1}{v} [\bar{U} M^u U + \bar{D} M^d D + \bar{L} M^\ell L] (c_\alpha h + s_\alpha H_F) \\
&+ \frac{v}{\sqrt{2}v_s} [\bar{U}_i \tilde{Z}_{ij}^u U_j + \bar{D}_i \tilde{Z}_{ij}^d D_j + \bar{L}_i \tilde{Z}_{ij}^\ell L_j] \\
&\times (-\sin \alpha h + \cos \alpha H_F + iA_F) + \text{h.c.}, \tag{3.3.13}
\end{aligned}$$

Here, M^f denotes the diagonal fermion mass matrix. We encapsulate the Higgs-Flavon couplings in the matrices $\tilde{Z}_{ij}^f = U_L^f Z_{ij}^f U_L^{f\dagger}$. In the flavour basis, the Z_{ij}^f matrix elements are given by

$$Z_{ij}^f = \rho_{ij}^f \left(\frac{v_s}{\sqrt{2}\Lambda_F} \right)^{q_{ij}^f} q_{ij}^f, \tag{3.3.14}$$

which, in general, remains non-diagonal even after the mass matrices are diagonalised, thereby giving rise to flavour-violating interactions.

Finally, in addition to the standard Yukawa couplings, we also require the $H_F h f_i \bar{f}_j$ couplings for our calculations. In the FNSM, these interactions are described by

$$H_F h f \bar{f} = \frac{m_f v_s}{\sqrt{2}\Lambda^2} (1 - 2 \cos^2 \alpha). \tag{3.3.15}$$

As a particular case, we explore the scenario where $f = b$. This choice is motivated by experimental reports on Higgs pair searches [72–76].

3.4 Constraints on the FNSM Parameter Space

To compute a realistic numerical analysis of the signals proposed in this project, namely

$$pp \rightarrow H_F \rightarrow h b \bar{b} \quad (\text{with } h \rightarrow b \bar{b}, \gamma\gamma),$$

it is necessary to constrain the free FNSM parameters involved in the forthcoming calculations. These free parameters are:

- The mixing angle α of the real components of the doublet Φ and the FN singlet S .
- FN singlet VEV v_s .
- The ultraviolet mass scale Λ .
- CP-even scalar mass M_{H_F} .

These parameters can be constrained by several theoretical requirements, such as absolute vacuum stability, triviality, perturbativity, and unitarity of scattering matrices and by various experimental data, mainly the upper limits on the production cross-sections of additional Higgs states from LHC measurements. We also consider bounds from Lepton Flavour Violating Processes (LFVp), such as $L_i \rightarrow \ell_j \ell_k \bar{\ell}_k$ and $\ell_i \rightarrow \ell_j \gamma$, as well as measurements of $\mathcal{BR}(B_{s,d}^0 \rightarrow \mu^+ \mu^-)$. Finally, the anomalous magnetic moments of the muon and electron Δa_μ and Δa_e , respectively, are also taken into account.

3.4.1 Theoretical Constraints

Stability of the Scalar Potential

It is essential to ensure that the scalar potential in Eq. (3.3.2) is bounded from below, meaning that it does not approach negative infinity in any direction of the field space

(h, H_F, A_F) at large field values. At high field strengths, the quartic terms dominate over the quadratic ones. Therefore, the conditions for absolute stability read [77],

$$\left(\lambda_1, \lambda_2, \lambda_3 + \sqrt{2\lambda_1\lambda_2}\right) > 0.$$

The quartic couplings are evaluated at the scale Λ using the Renormalisation Group Evolution (RGE) equations. If the scalar potential in Eq. (3.3.2) exhibits a meta-stable EW vacuum, these conditions must be appropriately modified [77]. In order to constrain the scalar field masses M_ϕ , the VEV of the complex singlet v_s , and the mixing angle α , one can use Eqs. (3.3.10) to translate the stability limits into constraints on the model parameters.

Perturbativity and Unitarity Constraints

The upper limits provided in Eq. (3.4.1) are essential to ensure that the radiatively corrected scalar potential of the FNSM remains perturbative at all energy scales [78],

$$|\lambda_1, \lambda_2, \lambda_3| \leq 4\pi. \quad (3.4.1)$$

The quartic couplings emerging from the scalar potential are subject to stringent constraints arising from the unitarity of the S -matrix. In practice, even at large field values, one can derive the S -matrix by analysing various $2 \rightarrow 2$ scattering processes involving interactions between pseudo scalar bosons, scalars, and gauge bosons—denoted here as $(P)S$, V , and $(P)S - V$ interactions. Unitarity requires that all eigenvalues of the S -matrix remain below 8π [77, 79]. In the context of the FNSM, and by employing the equivalence theorem, the resulting unitary bounds can be expressed as,

$$\lambda_1 \leq 16\pi \quad \text{and} \quad \left| \lambda_1 + \lambda_2 \pm \sqrt{(\lambda_1 - \lambda_2)^2 + (2/3\lambda_3)^2} \right| \equiv |\lambda_U^\pm| \leq 16/3\pi. \quad (3.4.2)$$

By using the Eqs. (3.3.10), (3.4.1) and (3.4.2), we can derive constraints on the scalar singlet VEV v_s , the masses of the heavy Higgs bosons M_{H_F} and M_{A_F} , as well as on the mixing angle α .

Parameter	Interval
$\cos \alpha$	$[-1, 1]$
M_{H_F}	800-1500 GeV
M_{A_F}	800-1500 GeV
v_s	v -5000 GeV

Table 3.7: Scanned parameters for theoretical viability in the FNSM.

Figure 3.16 displays the $\cos \alpha$ - v_s plane, where each point represents a parameter combination that satisfies all the theoretical constraints, including perturbativity and the unitarity conditions imposed on the S -matrix. For this analysis, we generated a set of random points that comply with the relations given in Eqs. (3.3.10), (3.4.1), and (3.4.2). The ranges of the scanned parameters to achieve that proposal are summarised in Table 3.7.

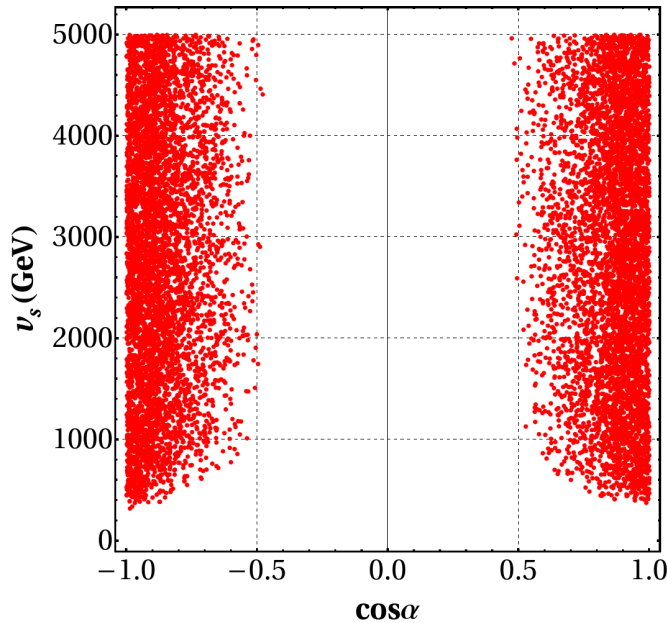


Figure 3.16: The VEV of the FN singlet v_s as a function of the cosine of the mixing angle α . The red points denote the region of parameter space that is consistent with all the theoretical constraints discussed in the text.

We find that the upper bound $|\lambda_U^+| \leq \frac{16\pi}{3}$ represents the most stringent constraint on the scalar quartic couplings. This bound translates into a lower limit on the scalar singlet VEV, namely $v_s \geq (276, 345, 415, 519)$ GeV for specific choices of the heavy

Higgs masses $M_{H_F} = M_{A_F} = (800, 1000, 1200, 1500)$ GeV and for $\cos \alpha = -0.995$. It should be noted that our analysis is performed only at the EW scale, and a detailed Renormalisation Group Evolution (RGE) analysis is beyond the scope of this thesis. Moreover, we observe a clustering of allowed points near $\cos \alpha \approx 1$ and $\cos \alpha \approx -1$, this is expected, since α must remain close to zero in order to avoid large deviations from the SM $h f \bar{f}$ coupling.

It is also important to emphasise that the range of masses considered in Table 3.7 was chosen based on a previous study [78], in which it was shown that $M_{H_F} > 800$ GeV is necessary to alleviate the constraints imposed by the cross-section limits on the process $pp \rightarrow S \rightarrow hh$ reported by the ATLAS Collaboration [45], where S represents a resonant pseudo-scalar particle.

3.4.2 Experimental Constraints

LHC Higgs boson data

We complement the theoretical constraints by incorporating experimental measurements from Higgs boson physics [70]. In particular we consider the *signal strengths* defined as

$$\mathcal{R}_X = \frac{\sigma(pp \rightarrow h) \cdot \mathcal{BR}(h \rightarrow X)}{\sigma(pp \rightarrow h^{\text{SM}}) \cdot \mathcal{BR}(h^{\text{SM}} \rightarrow X)}, \quad (3.4.3)$$

where $\sigma(pp \rightarrow H_i)$ denotes the production cross-section of H_i (with $H_i = h, h^{\text{SM}}$); here, h represents the SM-like Higgs boson arising in the extended model, while h^{SM} is the SM Higgs boson. The BR $\mathcal{BR}(H_i \rightarrow X)$ corresponds to the decay of H_i into final states X , with X taking values such as $b\bar{b}$, $c\bar{c}$, $\tau^-\tau^+$, $\mu^-\mu^+$, WW^* , ZZ^* , or $\gamma\gamma$.

Lepton Flavour Violating Processes

In addition, we examine several LFVp to impose further constraints on the parameters pertinent to our subsequent calculations. Specifically, we utilise (i) the upper limits on $\mathcal{BR}(L_i \rightarrow \ell_j \ell_k \bar{\ell}_k)$ and $\mathcal{BR}(\ell_i \rightarrow \ell_j \gamma)$, and (ii) the measurements

Parameter	Interval
$\cos \alpha$	$[-1, -0.7]$
Λ	v -5000 GeV
v_s	v -5000 GeV

Table 3.8: Scanned parameters for \mathcal{R}_X evaluation in the FNSM.

of $\mathcal{BR}(B_{s,d}^0 \rightarrow \mu^+ \mu^-)$ together with the anomalous magnetic moment of the muon, Δa_μ . To facilitate the exploration of the FNSM parameter space, the model is implemented also in the `Mathematica` package `SpaceMath`.

In Fig. 3.17, the ultraviolet mass scale Λ is presented as a function of the VEV of the complex singlet v_s . Different symbols in this plot denote the individual signal strengths \mathcal{R}_X : blue circles represent \mathcal{R}_W , green triangles correspond to \mathcal{R}_Z , yellow diamonds indicate \mathcal{R}_γ , green squares denote \mathcal{R}_τ , orange triangles correspond to \mathcal{R}_b , and green rectangles represent additional channels. The overlapping region common to all these individual measurements is highlighted by the area enclosed within solid black lines. Additionally, the green rectangle marks the region permitted by the LFVp constraints, while the cyan area reflects the intersection of all the theoretical and experimental constraints discussed.

Motivated by the analysis presented in Sec. 3.4.1, we once again perform a scan over the model parameters involved in the evaluation of \mathcal{R}_X , as summarised in Table 3.8. It is noteworthy that $v_s \approx \Lambda$ is expected since the $hf\bar{f}$ coupling scales as v_s^2/Λ^2 when $\cos \alpha \approx -1$, as detailed in Table 3.6. We have also explored the allowed ranges for \mathcal{R}_μ and \mathcal{R}_c (although these results are not depicted in Fig. 3.17). However, these observables do not impose particularly stringent constraints on the FNSM.

We explored the theoretical frameworks of 2HDM-III as well as FNSM and reviewed the experimental and theoretical constraints. In the following chapter, we introduce our data analysis methodology by describing how simulated collider events are processed and subsequently analysed using BDTs. This Machine Learning based approach is aimed at optimising the signal versus background discrimination and in that way, refining our predictions for the potential discovery of new scalar particles.

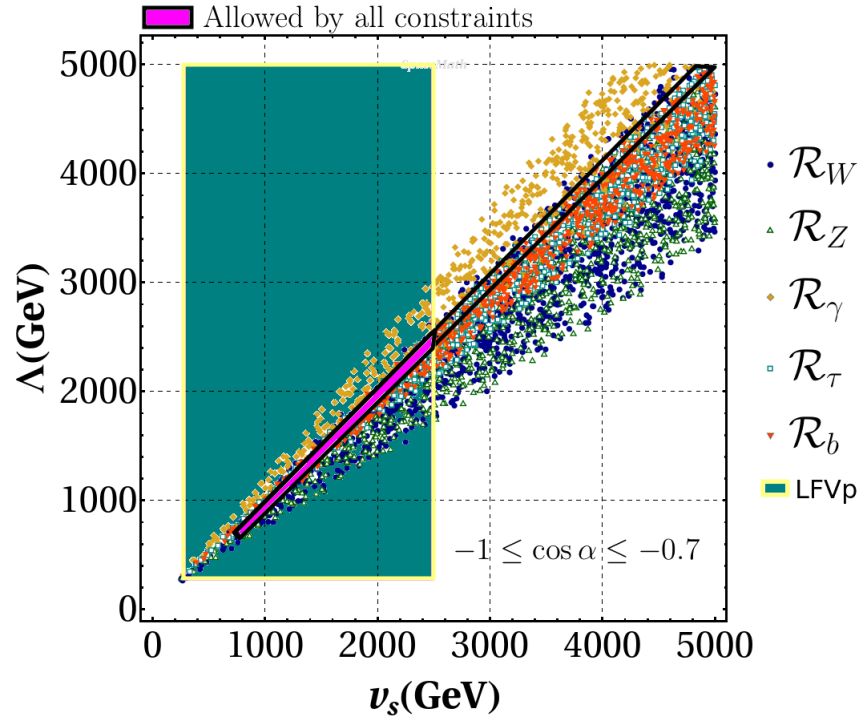


Figure 3.17: This plot shows the ultraviolet mass scale Λ as a function of the VEV of the complex singlet v_s . The area enclosed by the solid black lines corresponds to the intersection of all constraints, while the different points represent the individual \mathcal{R}_X measurements. The green rectangle indicates the parameter space allowed by the LFVp constraints.

Chapter 4

Methodology BDT Data Analysis

In the study of BSMs, simulations are an essential tool to explore phenomenology and test theoretical models. This chapter describes the methodology used to process the simulated data, extract observables information and train **BDTs!** (**BDTs!**) to separate signal from background in the phenomenological analysis of BSM physics. BDT is a Machine Learning algorithm that applies a series of conditions on the input variables to progressively divide a dataset into smaller groups. In this approach, the decision trees are designed to correct the misclassifications made in previous steps [80, 81].

Monte Carlo event generators, such as MadGraph5 [82] produce output files in the “.root” format, containing simulated collider events with detailed kinematic information. We have to distinguish the “.root” file extension from the CERN developed software framework called ROOT. The ROOT framework is widely used in high energy physics to read, process and analyse “.root” files. However, these data files can be very large and unwieldy. While ROOT or other analysis tools can be employed to extract kinematic observables, the internal structure of .root files is not directly readable.

To address this challenge and get direct information from .root files, it is a common practice to convert .root files into the LHC Olympics (LHCO) format [83]. The LHCO

format was originally introduced in the ‘LHC Olympics’, a series of community challenges around the start of LHC operations where researchers were trying to reconstruct signals from ‘black box’ detector level events. This format is well known for being lightweight and text based, while preserving essential information for phenomenological studies such as the particle types, pseudorapidity, azimuthal angle, transverse momentum, and invariant mass. As described in the manual [84], each event in that LHCO file is structured as a block of rows.

For conveniently reading and processing these LHCO files, there is a Python module called *LHCO_reader* [85], which analyses detector level events stored in LHCO format. Nevertheless, *LHCO_reader* does not perform the final analysis or extract physical observables, rather, it provides an interface to read the LHCO file data quickly and easily into Python structures.

The Python [86] based framework presented here, called PrakritiMLPrep [87], is designed to process the large volumes of simulation data by extracting primitive and derived physical observables from the events. These computed observables are then used to classify the events, enabling the creation of datasets that can be used to train Machine Learning models and thereby improving signal versus background discrimination. This approach typically achieves a higher statistical significance than obtaining it by empirical cuts in the observables.

By converting .root files to LHCO format, we can significantly reduce data size while retaining relevant kinematic details needed for Machine Learning training and analysis.

Its main functionalities include user-friendly interaction with LHCO files, complete with basic .py examples to guide users in reading LHCO files and defining custom observables. This makes it easy to extend the analysis to specific decay modes or processes. Additionally, it supports data preparation for Machine Learning, focusing on the creation of datasets that blend simulated signal and background events for training models such as BDTs. It also features a flexible architecture with a modular design, allowing users to add more observables.

4.0.1 The LHCO File Format

The LHCO format is a simple, text based format used to show collider event data. It consists of a series of block events where each block contains rows. Every row corresponds to a physics object, such as a lepton, photon, jet, or Missing Transverse Energy (MET).

4.0.2 Column Definitions and Structure

Each row in an LHCO file contains several columns. The Table 4.1 summarises the standard column definitions.

4.0.3 Example for Reading a LHCO

In the following example (Table 4.2), we show how to read an LHCO file. Despite the apparent simplicity of the LHCO format, each event block contains kinematic variables and object classifications necessary for a robust analysis. The code example below walks through these steps one at a time so we can clearly see how it works.

In this event blocks, each row corresponds to a distinct object recorded during a simulated proton–proton collision. The structure follows a standardised layout:

Event Block 0

- **Row 0 (Event Header):** The first row in the block (with counter 0). It contains the event number and trigger information. In this case, all entries are 0, indicating that no additional details are provided. This row is primarily used for identification purposes and can generally be ignored during analysis. However, it is useful to identify as a block in our Python framework.
- **Row 1 (Lepton Entry):** The second row is labelled with counter 1 and has a type value of 2, which indicates that the object is a lepton (specifically, a muon). The kinematic properties are as follows:

Column #	Name/Type	Description
1	Counter	Sequential index labelling the object in the event.
2	Type	Object type indicator: 0 = photon, 1 = electron, 2 = muon, 3 = hadronically decaying tau, 4 = jet, 6 = MET.
3	η (Pseudorapidity)	The pseudorapidity, defined as $\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]$, where θ is the polar angle relative to the beam axis. For massless objects, η is the rapidity y .
4	ϕ (Azimuthal Angle)	The azimuthal angle around the beam axis in radians.
5	p_T (Transverse Momentum)	The momentum component perpendicular to the beam direction, measured in GeV.
6	Invariant Mass ($jmass$)	For jets, the invariant mass calculated from the energy deposits within the jet.
7	$ntrk$ (Number of Tracks)	Number of tracks associated with the object. For leptons, this number is multiplied by the charge (e.g., -1 for a muon, +1 for a positron).
8	$btag$	B-tagging information: values 1 or 2 indicate different levels of heavy-flavour tagging for jets; for other objects (e.g. muons), this column can carry special meaning such as the index of the closest jet.
9	Had/EM Ratio	The ratio of hadronic to electromagnetic energy deposition in the associated calorimeter cells. Typically, jets yield values greater than 1, electrons or photons yield values much less than 1. For muons this field is formatted as <code>xxx.yy</code> , where <code>xxx</code> is the isolation p_T^{iso} (the sum of transverse momenta in a cone of radius $R = 0.4$, excluding the muon) and <code>yy</code> is the energy ratio in a surrounding grid.
10 and 11	Dummy	Reserved for future use, currently set to zero.

Table 4.1: LHCO File Column Definitions.

1	2	3	4	5	6	7	8	9	10 & 11
Counter	Type	η	ϕ	p_T	Jmass	Ntrk	Btag	Had/em	Dummies
Event Block 0									
0		0	0						
1	2	-1.925	-1.406	138.29	0.11	1.0	6.0	138.00	0.0
2	4	-0.773	2.467	58.03	11.73	7.0	0.0	0.27	0.0
3	4	2.136	0.930	46.42	8.49	7.0	0.0	2.86	0.0
4	4	-2.052	0.987	45.30	2.94	3.0	0.0	0.00	0.0
5	4	0.275	2.311	30.45	6.27	6.0	0.0	0.00	0.0
6	4	-0.142	-1.995	29.20	7.34	5.0	0.0	0.42	0.0
7	4	1.537	0.003	23.67	11.55	5.0	0.0	4.10	0.0
8	6	0.000	2.020	28.44	0.00	0.0	0.0	0.00	0.0
Event Block 1									
0		1	0						
1	1	1.037	2.787	47.10	5.21	7.0	0.0	3.62	0.0
2	1	0.049	2.905	32.17	5.99	-7.0	0.0	0.00	0.0
3	4	-0.647	-2.051	24.05	3.49	3.0	0.0	999.90	0.0
4	6	0.000	-0.081	74.31	0.00	0.0	0.0	0.00	0.0
Event Block 2									
0		2	0						
1	2	-1.649	0.979	232.50	0.11	1.0	5.0	237.03	0.0
2	4	-2.088	-2.437	172.57	17.59	15.0	0.0	0.07	0.0
3	4	-1.500	-2.354	49.42	7.43	10.0	0.0	1.92	0.0
4	4	0.018	-2.715	44.89	4.55	4.0	0.0	0.00	0.0
5	4	-1.119	0.429	30.85	6.95	7.0	0.0	0.25	0.0
6	4	2.757	-0.564	23.64	5.09	0.0	0.0	999.90	0.0
7	6	0.000	-1.166	45.86	0.00	0.0	0.0	0.00	0.0

Table 4.2: Detailed example of LHCO event data formatted for multiple event blocks.

- **Pseudorapidity (η):** -1.925, which describes the particle’s angle relative to the beam axis.
- **Azimuthal Angle (ϕ):** -1.406 radians, specifying the orientation around the beam.
- **Transverse Momentum (p_T):** 138.29 GeV, representing the momentum perpendicular to the beam axis.
- **Invariant Mass (**Jmass**):** 0.11 GeV. For leptons, the invariant mass is typically very small.

The seventh column (**Ntrk**), is given as 1.0. According to the LHCO conventions, for leptons this number is multiplied by the charge. A positive value here is indicating that the lepton is positively charged (i.e., a μ^+). The eighth column, (**Btag**) shows a value of 6.0, which, although primarily used for jets, is repurposed in some analyses to encode additional identification or isolation information for leptons. Finally, the ninth column (**Had/em**) is 138.00; for muons, this field is formatted as **xxx.yy**, where the integer part represents the isolation transverse momentum (p_T^{iso} , measured in GeV within a cone of $R = 0.4$ around the muon, excluding the muon itself) and the decimal part gives a percentage ratio of the surrounding transverse energy. Here, the large value emphasises the isolation characteristic of this object.

- **Rows 2 to 7 (Jet Entries):** Rows 2 through 7 correspond to jets (objects with type 4). For example, considering Row 2:
 - **Pseudorapidity (η):** -0.773
 - **Azimuthal Angle (ϕ):** 2.467 radians
 - **Transverse Momentum (p_T):** 58.03 GeV
 - **Invariant Mass (**Jmass**):** 11.73 GeV, which is the mass computed from the energy and momentum constituents of the jet.

- **Number of Tracks (Ntrk):** 7.0, indicating the jet is composed of several charged tracks.
- **Btag:** 0.0, meaning the jet was not tagged as containing a b-quark.
- **Had/em Ratio:** 0.27, suggesting that the energy deposited in the hadronic calorimeter is relatively low compared to that in the electromagnetic calorimeter.

The remaining jet rows (Rows 3 through 7) are structured similarly, with variations in the kinematic quantities reflecting the different characteristics of each jet.

- **Row 8 (MET):** The final row in the block is marked with a type value of 6, which denotes the MET. In this row:
 - **p_T (MET):** 28.44 GeV, representing the imbalance of transverse energy in the event.
 - All other quantities (invariant mass, number of tracks, b-tag, etc.) are set to 0, as MET does not correspond to a reconstructed particle.

Event Block 1

- **Row 0 (Event Header):**

This row is marked by a counter value of 0 and serves as the event header. In this block, the header shows a value of 1 in the third column, which can be interpreted as the event label.

- **Row 1 (First Lepton):**

The row with counter 1 contains an object of *Type 1*, which corresponds to an electron. Its kinematic and other properties are:

- **Pseudorapidity (η):** 1.037, which quantifies the angle relative to the beam axis.

- **Azimuthal Angle (ϕ):** 2.787 radians, indicating the particle's orientation in the transverse plane.
- **Transverse Momentum (p_T):** 47.10 GeV.
- **Invariant Mass (Jmass):** 5.21 GeV.

Ntrk is recorded as 7.0. This value is multiplied by the particle's charge; therefore, a positive value indicates a positively charged electron (i.e., a positron). The b-tag column is 0.0 since b-tagging is not applicable to electrons. The Had/em ratio is given as 3.62, which in the context of leptons, is interpreted using a special format where the integer part corresponds to the isolation transverse momentum (p_T^{iso}) and the decimal part to the relative energy in a surrounding grid.

- **Row 2 (Second Lepton):**

This row also represents an electron (Type 1). Its kinematic values are:

- **Pseudorapidity (η):** 0.049.
- **Azimuthal Angle (ϕ):** 2.905 radians.
- **Transverse Momentum (p_T):** 32.17 GeV.
- **Invariant Mass (Jmass):** 5.99 GeV.

Here, Ntrk is listed as **-7.0**. The negative sign indicates that this electron is negatively charged. As in Row 1, the b-tag value is 0.0 and the Had/em ratio is 0.00.

- **Row 3 (Jet):**

Row 3 contains an object with *Type 4*, corresponding to a jet. Its properties include:

- **Pseudorapidity (η):** -0.647.
- **Azimuthal Angle (ϕ):** -2.051 radians.
- **Transverse Momentum (p_T):** 24.05 GeV.

- **Invariant Mass (Jmass):** 3.49 GeV.

It is associated with 3 tracks, and the b-tag value is 0.0, indicating no heavy-flavour tagging. The Had/em ratio is 999.90, an unusually high value that likely serves as a flag to indicate a special jet characteristic or a reconstruction anomaly.

- **Row 4 (Missing Transverse Energy – MET):**

The final row in this block is marked by type 6, denoting MET. For this entry:

- **Azimuthal Angle (ϕ):** -0.081 radians.
- **Transverse Momentum (p_T):** 74.31 GeV.

Event Block 2

In this block, the structure is as follows:

- **Row 0 (Event Header):** 3 provides the event number.
- **Row 1 (Lepton Entry):** Type 2 value, indicating that the object is a lepton.

The kinematic and identification information is:

- **Pseudorapidity (η):** -1.649.
- **Azimuthal Angle (ϕ):** 0.979 radians.
- **Transverse Momentum (p_T):** 232.50 GeV.
- **Invariant Mass (Jmass):** 0.11 GeV.

Here Ntrk is 1.0, indicating that the lepton is positively charged (i.e., a positron). Row 6 and Row 7 have a similar interpretation as the last block analysed previously.

4.0.4 Setup and Usage of Our Python Framework

The framework requires a fully functional Python 3 environment and depends on the `pandas` library for data manipulation. These dependencies must be installed to ensure proper functionality.

Obtaining the Project

The PrakritiMLPrep framework can be cloned from this repository:

```
git clone https://gitlab.com/ed_in/prakritimlprep.git
```

Running the Tool

To run the tool, it is necessary to run it as:

```
python PrakritiMLPrep.py $Signal_File.lhco $Background_File.lhco
```

Here, `$Signal_File.lhco` is the LHCO file containing the signal events and the LHCO file with the background events, is `$Background_File.lhco`. The tool will then begin processing these files and generate the output dataset for further Machine Learning analysis.

Output Details

After executing the command, the directory `$Signal_File_$Background_File` is created, containing the dataset file `training-$Signal_File.csv`. This `.csv` file includes both, the signal and background information, randomly mixed and formatted to be ready for training Machine Learning models. The structure of the `.csv` file is as follows:

- The first column, `New_ID`, is used to uniquely identify each row (or event).

- The subsequent columns contain the calculated observables contained in the `Modules` directory. By convention, the raw quantities measured by the detector (the *Primitive Variables*) are labeled with the prefix `PRI_`, while the computed quantities (the *Derived Variables*) are labeled with the prefix `DER_`.
- The final column indicates whether the row corresponds to a signal or a background event.

4.0.5 Adding New Observables

The code is designed using a modular approach which allows, new physical observables be easily added by creating new `.py` modules within the `Modules/` directory.

Each observable is computed in a separate `.py` file, which allows to add or modify observables without affecting the core functionality.

To include a new observable in the framework, one begins by creating a dedicated `.py` file (for example, `DER_NewObservable.py`) within the `Modules/` directory. Next, a function is defined in this file to accept the raw event data, calculate the required observable, and record the resulting values as an additional column in a `.csv` output.

As a basic example, the Code Block 1 shows the Python code used to process LHCO data for lepton entries. This code extracts *primitive variables* and computes *derived variables* for pairs of leptons.

```
1 import csv
2 import sys
3 from math import cosh, cos, sqrt
4
5 def process_data(input_file, output_file):
6     # List to store results for each lepton pair.
7     results = []
8
```

```
9      # Read all lines from the input LHCO file.
10     with open(input_file, 'r') as file:
11         lines = file.readlines()
12         i = 0
13
14         # Process each event.
15         while i < len(lines):
16             line = lines[i].strip()
17
18             # Identify the beginning of a new event with a header row
19             ↪ (starting with '0').
20             if line.startswith('0'):
21                 leptons = [] # List to store lepton data for this
22                 ↪ event.
23                 i += 1
24
25                 # Read subsequent lines until the next event header is
26                 ↪ encountered.
27                 while i < len(lines):
28                     line = lines[i].strip()
29                     words = line.split()
30
31                     # If a new event header is reached and lepton data
32                     ↪ exists, break.
33                     if line.startswith('0') and leptons:
34                         break
35
36                     # Process rows representing electrons (type '1') or
37                     ↪ muons (type '2').
38                     if len(words) >= 8 and words[1] in ['1', '2']:
```

```
34         # Append a tuple: (object type, pT, eta, phi)
35         leptons.append((words[1], float(words[4]),
36             ↪ float(words[2]), float(words[3])))
37
38     # If at least two leptons are found, compute kinematic
39     ↪ variables for every pair.
40     if len(leptons) >= 2:
41         for j in range(len(leptons)):
42             for k in range(j+1, len(leptons)):
43                 type1, pt1, eta1, phi1 = leptons[j]
44                 type2, pt2, eta2, phi2 = leptons[k]
45
46                 # Calculate differences in pseudorapidity
47                 ↪ and azimuthal angle.
48                 delta_eta = abs(eta1 - eta2)
49                 delta_phi = abs(phi1 - phi2)
50
51                 # Compute the angular separation (Delta R)
52                 ↪ in (eta, phi) space.
53                 delta_r = sqrt(delta_eta**2 + delta_phi**2)
54
55                 # Compute the invariant mass using the
56                 ↪ formula for massless leptons:
57                 #  $m = \sqrt{2 * pt1 * pt2 * (\cosh(\delta_{\eta}) - \cos(\delta_{\phi}))}$ 
58                 ↪
59                 invariant_mass = sqrt(2 * pt1 * pt2 *
60                 ↪ (cosh(delta_eta) - cos(delta_phi)))
61
62                 # Save the computed variables for this
63                 ↪ lepton pair.
```

```
57         results.append({
58             'pt1': pt1,
59             'eta1': eta1,
60             'phi1': phi1,
61             'pt2': pt2,
62             'eta2': eta2,
63             'phi2': phi2,
64             'delta_eta': delta_eta,
65             'delta_r': delta_r,
66             'invariant_mass': invariant_mass
67         })
68     else:
69         i += 1
70
71     # Write the results to a .csv file with extended columns.
72     with open(output_file, 'w', newline='') as csvfile:
73         fieldnames = ['pt1', 'eta1', 'phi1', 'pt2', 'eta2', 'phi2',
74             ↪ 'delta_eta', 'delta_r', 'invariant_mass']
75         writer = csv.DictWriter(csvfile, fieldnames=fieldnames)
76         writer.writeheader()
77         for res in results:
78             writer.writerow(res)
```

Code Block 1: Example of Python code for processing lepton data.

Explanation of the Main Code Steps

1. File Reading and Event Loop:

The function `process_data` opens the input LHCO file and reads all lines. It then iterates over these lines to detect the beginning of each event, which is

indicated by a row starting with '0'.

2. Extracting Lepton Data:

For each event, the code initializes a list `leptons` and processes subsequent rows. If the row represents an electron (Type '1') or a muon (Type '2'), the code extracts the transverse momentum (`pT`), pseudorapidity (`eta`), and azimuthal angle (`phi`), storing these values as a tuple.

3. Computing Derived Variables:

Once at least two leptons have been collected for an event, the code computes kinematic variables for each unique pair:

- $\Delta\eta$: The absolute difference in pseudorapidity.
- $\Delta\phi$: The absolute difference in azimuthal angle.
- ΔR : The angular separation computed as

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}.$$

- **Invariant Mass** [88]:

$$m = \sqrt{2 p_{T1} p_{T2} [\cosh(\Delta\eta) - \cos(\Delta\phi)]},$$

which is standard for massless leptons.

These derived quantities are stored in a list of dictionaries.

4. Output Generation:

Finally, the results are written to a .csv file. The .csv file includes a header and rows for each lepton pair, with the following fields: `pt1`, `eta1`, `phi1`, `pt2`, `eta2`, `phi2`, `delta_eta`, `delta_r`, and `invariant_mass`. In this example, the main column of interest is the invariant mass, column number 9.

This codes can be tested individually by running:

```
python \Observable.py \LHCOFile.lhco \CSVOutputfile.csv
```

Where `$Observable.py` refers to the Python script itself. `$LHCOFile.lhco` is the LHCO event file (whether signal or background), while `$CSVOutputfile.csv` is the user specified output filename where the computed observable will be stored.

4.0.6 BDT Training

Introduction to BDT Training

A decision tree is a flowchart structure that begins with a root node (representing the entire dataset) and subsequently splits into internal nodes, like branches, based on threshold values of various features. In our case these features are the kinematic observables. At each branching, events that satisfy a split criterion (e.g. “Observable X is above some cutoff”) move along one branch, while the others follow a different and separated path. This process continues down multiple levels ending in terminal nodes, as “leafs”, where events are assigned a predicted class. Because of each threshold split attempts to make the classes as pure as possible, in our case, the signal vs. the background, the entire tree maps out a series of observable based decisions that aim to classify new events accurately.

BDTs are a form of ensemble Machine Learning algorithm that combine many decision trees to improve classification performance [89]. By utilising a boosting technique, which is a method where several simple models such as decision trees, are trained one after the other, the trees are built sequentially. Each tree is focusing on the mistakes of the previous one, thereby reducing the overall error. This ensemble approach yields a powerful discriminator that is capable of separating signal which are our desired events from background, with higher accuracy than any single tree alone.

In High Energy Physics (HEP) analyses, BDTs have become an important tool. Instead of applying a series of one dimensional cuts, a BDT learns an optimal combination of many observables at once, improving sensitivity to new phenomena.

Modern implementations, such as Extreme Gradient Boosting (XGBoost) [90], further enhance performance and efficiency. XGBoost uses gradient boosting, which is a boosting technique using the gradients of a loss function to guide the training of each new learner, ensuring that each step is aiming to minimise the general error. Also XGBoost works with built in regularisation, internal methods to control the model complexity, such as, penalising overly deep or complex trees. This helps to prevent overfitting by ensuring that the ensemble does not become too finely tuned to the training data. Actually, XGBoost gained significant attention in HEP following its successful application in the Higgs Boson Machine Learning Challenge (2014) [91].

Dataset Preparation

Ensuring a balanced number of signal and background events is crucial to avoid systematic error in the training process causing that the model could be performing well on the training data but poorly on new, unseen data, leading to inaccurate predictions or unfair decisions. If one class dominates, the BDTs might learn to simply label all events as belonging to that class.

Before training, we also screened the input observables for redundancy. For this purpose, a Pearson correlation matrix was computed on the training sample, and whenever a pair of variables exhibited a very strong linear correlation ($|\rho| > 0.95$), the second member of the pair in the correlation matrix was discarded. By eliminating such highly correlated features, we mitigate multicollinearity and improve control of overfitting.

Feature Importance and Cuts

One advantage of using BDTs is the ability to assess the feature importance of each observable. The BDT calculates the contribution of each one to the classification task, thereby providing a hierarchy of observables that are discriminate signal from background.

For example, key observables identified can include:

- $p_T(b)$: Transverse momentum of the b -jet.
- MET: Missing transverse energy.
- $p_T(c)$: Transverse momentum of the c -jet.
- $M_T(\mu)$: Transverse mass of the muon.
- $M_{\text{inv}}(bc)$: Invariant mass of the b and c -jet pair.

These observables could then be examined to see if simple cut selections could isolate the signal. However, rather than applying multiple cuts individually, the BDT can create a combined discriminant (referred to as the `xgb` output when we are using XGBoost, as aforementioned) that encapsulates the discriminating feature of all observables simultaneously. This composite variable enables a more precise selection by applying a single cut on the BDT output score.

Hyperparameter Tuning and Optimisation

Training a BDT involves setting several hyperparameters, which are parameters that control aspects such as model complexity and learning efficiency, and are set before the learning process begins. The ones that we primarily use are:

- **Number of trees:** The total number of decision trees in the ensemble.
- **Maximum depth per tree:** The maximum number of splits in each tree, controlling the complexity of the model. Deeper trees can capture more detail but may also lead to overfitting.
- **Learning rate:** A scaling factor that determines how much each new tree contributes to correcting the errors of previous trees. A smaller learning rate generally requires more trees but can lead to have a better performance.
- **Subsampling ratio:** The fraction of the training data used to build each tree.

Performance Evaluation

The performance of the BDT is primarily evaluated using the Receiver Operating Characteristic (ROC) curve, which plots the true positive rate (signal efficiency) against the false positive rate (background contamination) for different thresholds on the BDT output. The Area Under a ROC Curve (AUC) provides a single measure of classifier performance, where an AUC of 1 corresponds to perfect classification and 0.5 corresponds to random guessing. [92]

We also perform checks for overtraining by comparing the BDT output distributions between the training and testing sets. The Kolmogorov-Smirnov (KS) test [93], is used to quantify the similarity between these distributions. A KS statistic close to zero indicates that the cumulative distributions of the two samples being compared are very similar. In the context of model generalisation, this indicates that the model predictions on the training data are very similar to its predictions on new, unseen data suggesting that the model has not overfitted to the training set and is performing well on new data [94].

XGB Output Variable and Final Significance Estimation

A cut on the `xgb` variable is then applied to separate signal from background events. The optimal cut is determined by maximising the *Signal Significance*, defined as [95]

$$Z = \frac{S}{\sqrt{S+B}}, \quad (4.0.1)$$

where S and B are the number of signal and background events passing the cut, respectively, and they are defined as,

$$S = L \times \sigma_S, \quad B = L \times \sigma_B.$$

Incorporating the integrated luminosity, cross-sections, and BRs allows the estimation of the discovery potential under realistic experimental conditions. Thus, the

BDT training process involves preparing and balancing a representative dataset, selecting key observables and understanding their relative importance, adjusting hyperparameters to optimise performance while avoiding overfitting, evaluating the model using ROC curves, KS tests, and significance calculations, and using the combined `xgb` variable to perform the final event selection.

This approach shows that the BDT is not only capable of effectively distinguishing signals from background, but also provides a robust tool to estimate the potential discovery of new physics processes.

In the following chapter, we are applying our previous methodology to perform a detailed collider analyses within the frameworks of the 2HDM-III and the FNSM focusing specifically on the identification of relevant collider signatures.

Chapter 5

Collider Analysis

In this chapter, we run a detailed collider analysis within the two different BSM frameworks that we are working on. We examine the decay processes and production mechanisms pertinent to these models, emphasising the identification of key signals and their differentiation from the SM background. Furthermore, we utilise multivariate analysis techniques to enhance the discrimination power and improve the sensitivity for a potential experimental discovery of the processes. The chapter thus sets the stage for a systematic exploration of the collider signatures associated with our proposed scenarios.

5.1 Collider Analysis for the 2HDM-III

We present the analytical expressions required to compute the decay widths for the processes $H^\pm \rightarrow u_i d_j$ and $H^\pm \rightarrow \ell^\pm \nu_\ell$. These formulas are essential for describing the underlying dynamics of these decays and are given by,

$$\Gamma(H^\pm \rightarrow u_i d_j) = \frac{3G_F M_{H^\pm}}{\sqrt{2}4\pi} \left(m_{u_i}^2 |Y_{ij}|^2 + m_{d_j}^2 |X_{ij}|^2 \right),$$

$$\Gamma(H^\pm \rightarrow \ell^\pm \nu_\ell) = \frac{G_F M_{H^\pm}}{\sqrt{2} 4\pi} \left(m_\ell^2 |Z_{ij}^\ell|^2 \right),$$

where

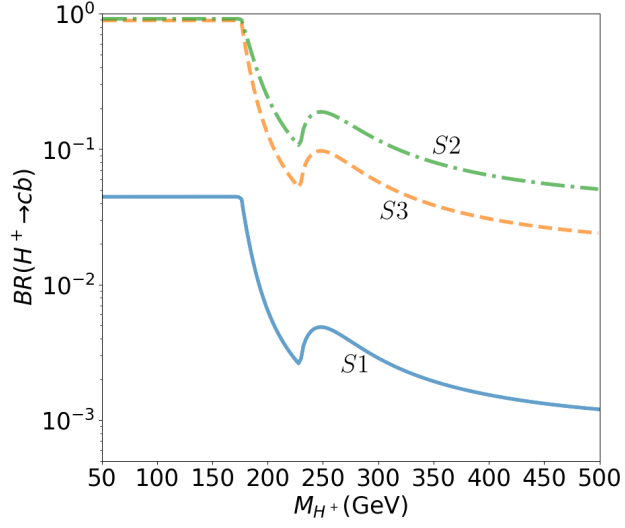
$$\begin{aligned} X_{ij} &= \sum_{l=1}^3 (V_{\text{CKM}})_{il} \left[t_\beta \frac{m_{d_l}}{m_{d_j}} \delta_{lj} - \frac{\sqrt{1+t_\beta^2}}{\sqrt{2}} \sqrt{\frac{m_{d_l}}{m_{d_j}}} \chi_{lj}^d \right], \\ Y_{ij} &= \sum_{l=1}^3 \left[\frac{1}{t_\beta} \delta_{il} - \frac{\sqrt{1+\frac{1}{t_\beta^2}}}{\sqrt{2}} \sqrt{\frac{m_{u_l}}{m_{u_j}}} \chi_{lj}^u \right] (V_{\text{CKM}})_{lj}, \\ Z_{ij}^\ell &= \left[t_\beta \frac{m_{\ell_i}}{m_{\ell_j}} \delta_{ij} - \frac{\sqrt{1+t_\beta^2}}{\sqrt{2}} \sqrt{\frac{m_{\ell_i}}{m_{\ell_j}}} \chi_{ij}^\ell \right], \end{aligned}$$

where $t_\beta \equiv \tan \beta$. In Fig. 5.1, we present the BRs for the corresponding decay channels. $H^+ \rightarrow cb$ and $H^+ \rightarrow \mu \nu_\mu$, plotted as functions of the charged scalar mass M_{H^+} , with the analysis carried out for the scenarios **S1**, **S2**, and **S3**, which are defined in Table 5.1.

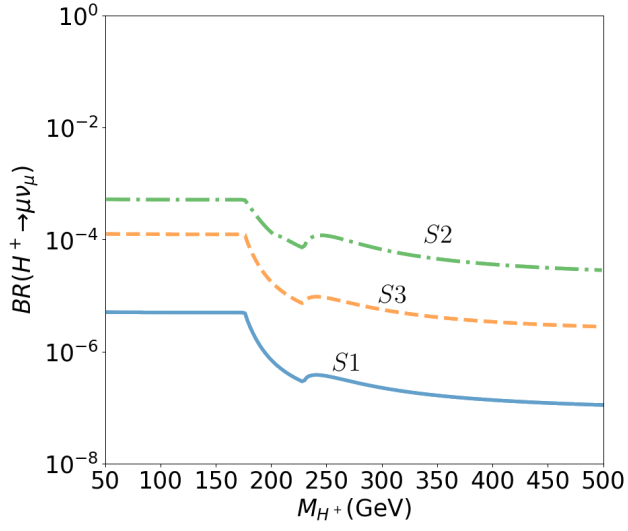
5.1.1 Signal and Background

We examine the signal and background processes arising from pp collisions at the LHC, carefully assessing their contributions to the observed event rates.

- **SIGNAL:** We focus on the search for a specific final state, $cb\mu\nu_\mu$, which arises from the pair production of charged Higgs bosons in proton-proton collisions. This process can be expressed as $pp \rightarrow H^- H^+ \rightarrow \mu^+ \nu_\mu \bar{c} b + \mu^- \bar{\nu}_\mu c \bar{b}$, and we denote this final state as $\mu \nu_\mu c b$. For our analysis, we assume a b -tagging efficiency of $\epsilon_b = 80\%$, a c -jet mis-tagging rate of $\epsilon_c = 10\%$, and a misidentification rate for other jets of $\epsilon_j = 1\%$. The dominant contributions to this final state within the framework of the 2HDM-III are illustrated in Fig. 5.2.



(a)



(b)

Figure 5.1: Branching Ratios (a) $BR(H^+ \rightarrow cb)$, (b) $BR(H^+ \rightarrow \mu\nu)$, as a function of the charged scalar mass M_{H^+} .

- **BACKGROUND:** The dominant SM background comes from the final state $bj\ell\nu_\ell$, which is produced by

$$- Wjj + Wb\bar{b},$$

$$- tb + tj,$$

$$- t\bar{t}.$$

In the case of the $t\bar{t}$ background process, either one of the two leptons is missed in semi-leptonic top quark decays, or two of the four jets are not reconstructed when

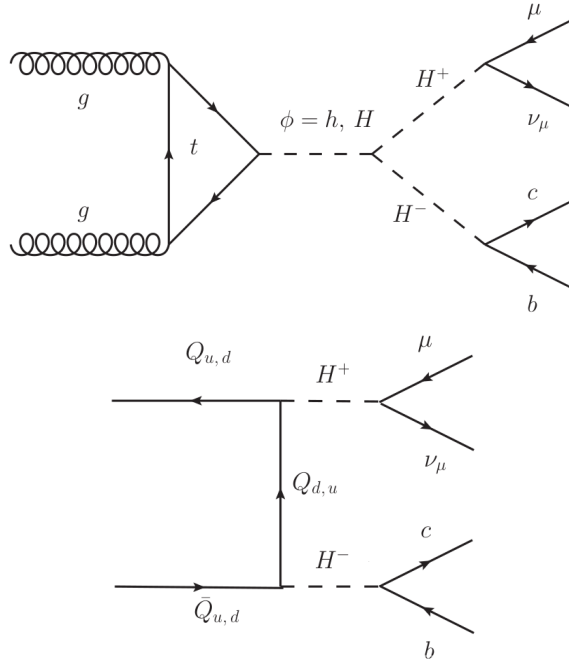


Figure 5.2: Feynman diagrams of the production cross-section of the signal $pp \rightarrow H^+ H^- \rightarrow \mu \nu_\mu c b$.

one of the top quarks undergoes a semi-leptonic decay. The numerical values of the cross-sections and BRs for the signal, corresponding to $M_H = 500, 800, 1000$ GeV in scenario **S2**, are summarised in Table 5.1. Meanwhile, the cross-sections of the dominant SM background processes are presented in Table 5.2.

M_H (GeV)	M_{H^\pm} (GeV)	$\sigma(gg \rightarrow H)$	BRs			$\sigma(pp \rightarrow \mu^+ \nu_\mu b \bar{c})$	Events (3000 fb $^{-1}$)
			$H \rightarrow H^- H^+$	$H^- \rightarrow b \bar{c}$	$H^+ \rightarrow \mu^+ \nu_\mu$		
500	100	1200 fb	0.32	0.944	5.27×10^{-4}	0.189 fb	567
	150	1200 fb	0.54	0.943	5.26×10^{-4}	0.324 fb	972
800	100	151 fb	0.27	0.944	5.27×10^{-4}	0.02 fb	59
	150	151 fb	0.40	0.943	5.26×10^{-4}	0.03 fb	100
1000	100	42 fb	1.67×10^{-3}	0.944	5.27×10^{-4}	3.5×10^{-5} fb	0.1
	150	42 fb	2.83×10^{-3}	0.943	5.26×10^{-4}	5.9×10^{-5} fb	0.2

Table 5.1: Leading contributions to the cross-section and BRs for scenario **S2**.

SM backgrounds	Cross-section [fb]	Events (3000 fb $^{-1}$)
$pp \rightarrow Wjj + Wbb (W \rightarrow l\nu_\ell)$	3745960	$\mathcal{O}(10^{10})$
$pp \rightarrow tb + tj (t \rightarrow l\nu_\ell b)$	1734	$\mathcal{O}(10^6)$
$pp \rightarrow t\bar{t} (t \rightarrow l\nu_\ell b, t \rightarrow q_i q_j b)$	431001	$\mathcal{O}(10^9)$

Table 5.2: Cross-section of the dominant SM background processes.

Figure 5.3 provides an overview of the production cross-section as a function of M_{H^\pm} for the scenarios **S1**, **S2**, **S3**, with $M_H = 500$ GeV. In contrast, Fig. 5.4 illustrates

the corresponding plane for $M_H = 800, 1000$ GeV, focusing exclusively on scenario **S2**.

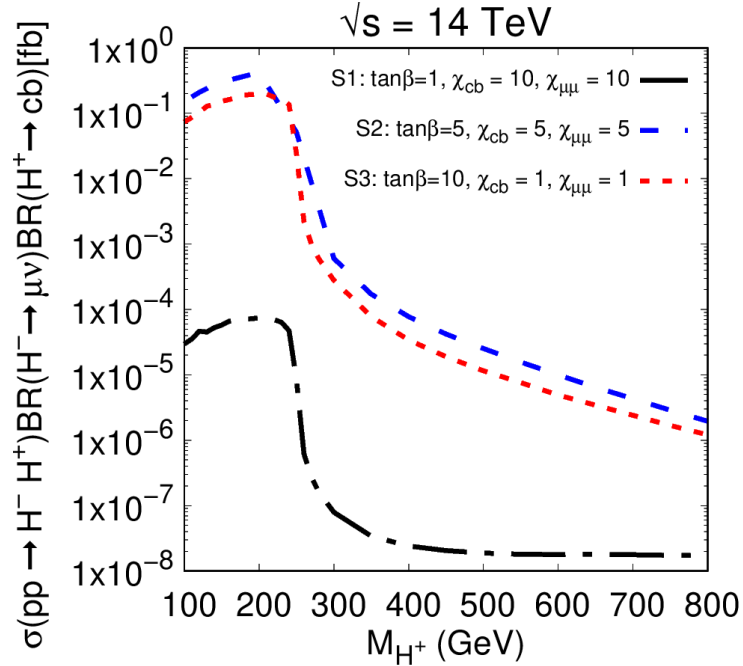


Figure 5.3: Production cross-section of a charged Higgs pair with their subsequent decays into $\mu\nu_\mu cb$ for $M_H = 500$ GeV.

For the three scenarios, we can observe that the cross-sections are significantly higher when the charged Higgs boson masses lie in the range of $100 \leq M_{H^\pm} \leq 250$ GeV. This is due to the dominant contribution from the neutral heavy Higgs boson ($M_H = 500$ GeV), which is capable of producing two real charged Higgs bosons in this mass range. As M_{H^\pm} increases beyond 250 GeV, the mediation of the heavy Higgs boson becomes virtual, leading to a suppression of the cross-section with increasing M_{H^\pm} . For $M_{H^\pm} \gtrsim 400$ GeV, the cross-section tends to stabilise, particularly in scenario **S1**, where a plateau emerges. This behaviour may be attributed to the increasing relevance of alternative decay channels, which dominate the total rate once the neutral Higgs bosons are produced off-shell. In contrast, scenarios **S2** and **S3** do not exhibit a similar plateau, likely due to their specific parameter choices, which prevent the cross-section from levelling off at higher masses.

From Tables 5.1 and 5.2, we can note that for $M_{H^\pm} = 100$ GeV and $M_H = 500$ GeV, the signal cross-section exceeds the corresponding background production cross-

section by up to seven orders of magnitude. Assuming an integrated luminosity of 3000 fb^{-1} , this corresponds to a difference of approximately 10^8 events. However, for $M_H = 1000 \text{ GeV}$ no events are expected to be produced even at the maximum integrated luminosity achievable at HL-LHC. This effectively rules out the possibility of observing a massive scalar in this mass range for the given luminosity.

In contrast, for $M_H = 800 \text{ GeV}$, although the number of expected signal events is small, our analysis suggests that the HL-LHC may still have the potential to see the proposed signal. This possibility will be explored further in this work.

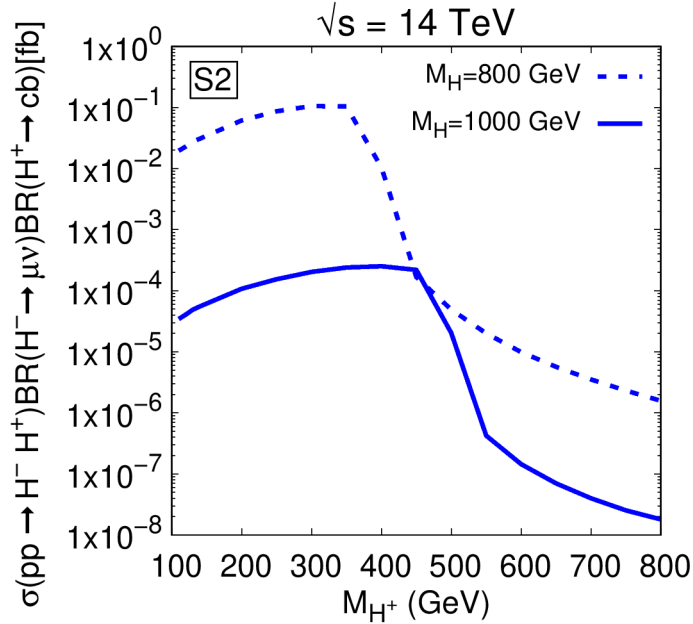


Figure 5.4: Production cross-section for a pair of charged Higgs bosons followed by their subsequent decays into $\mu\nu_\mu cb$, for $M_H = 800, 1000 \text{ GeV}$ in scenario **S2**.

In our computational framework we begin by implementing the full model using `FeynRules` [96] within `MadGraph5` [82], employing the NNPDF2.3LO [97] Parton Distribution Functions (PDFs). Subsequently, the model is interfaced with `Pythia8` [98] for parton showering and hadronisation and with `Delphes 3` [99] for detector simulations. For jet reconstruction, we use the `FastJet` package [100] in conjunction with the `anti- k_t` algorithm [101].

5.1.2 Signal Significance

Traditional strategies based on kinematic cuts applied to observables often reject a significant number of background events. However, these cuts also discard a substantial fraction of signal events reducing the overall sensitivity. To address this limitation, we employ a Multivariate Analysis (MVA) approach, which provides a more efficient way to discriminate between signal and background events. Specifically, as mentioned in Chapter 4, we compute a BDT [102] training it on the set of variables listed in Table B.1 of App. B.

The relevant hyperparameters for the BDT are chosen as follows: number of trees (`NTree`) is set to 50, the maximum depth of each decision tree (`MaxDepth`) is set to 5, and the maximum number of leaves per tree (`MaxLeaves`) is set to 8. All other parameters are left at their default values.

In Fig. 5.5, we present the discriminant distributions for both the signal and background processes. The goodness of fit is evaluated using the KS test. We find that the KS value lies within the permissible range of $[0,1]$, with values of 0.47 for the signal and 0.59 for the background, indicating a reasonable agreement between the observed and expected distributions.

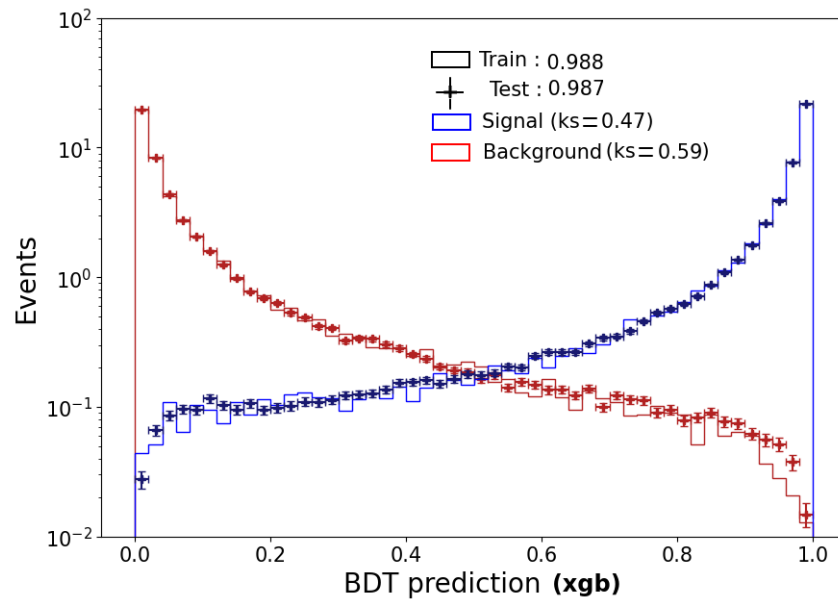


Figure 5.5: Plot of the discriminant for signal and background data.

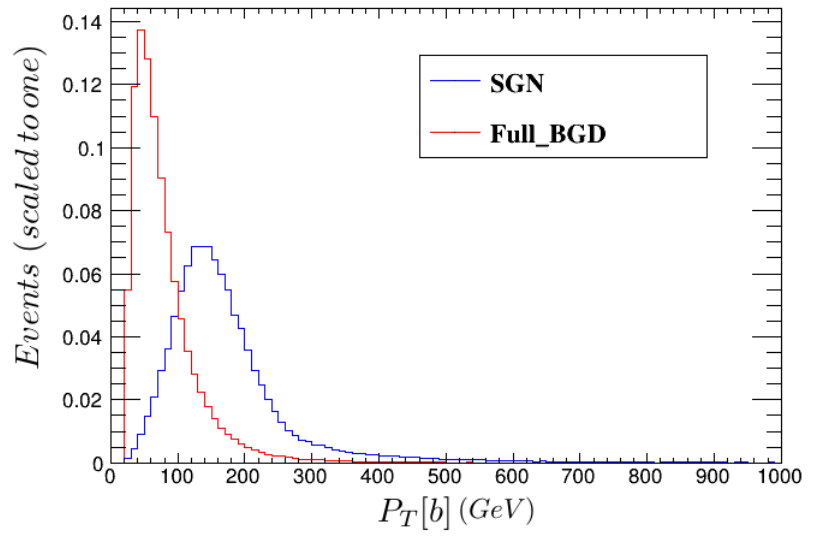
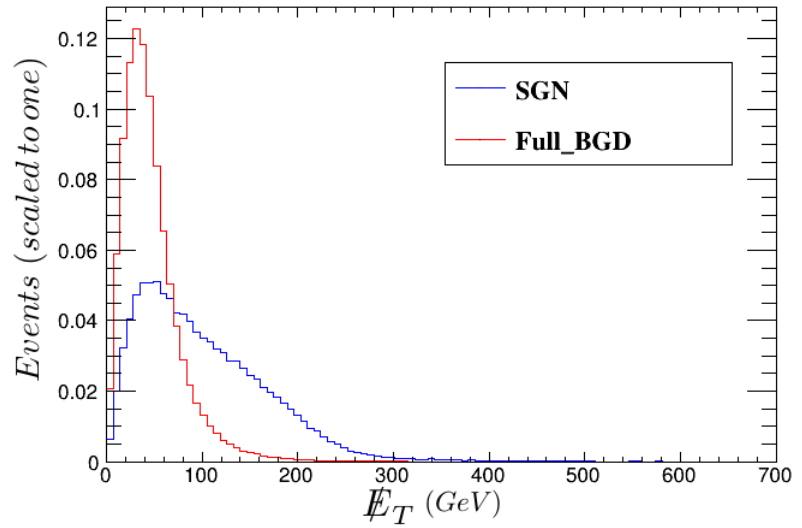
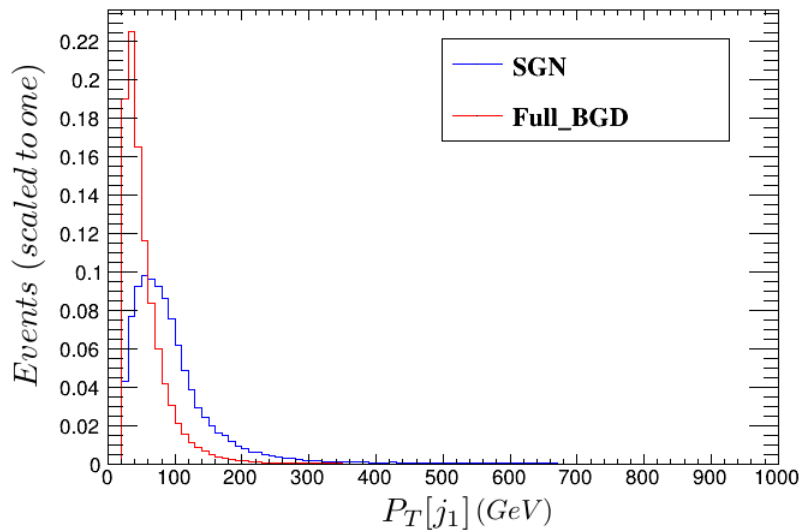
Once the classifier was trained, it outputs the single variable denoted as \mathbf{xgb} , which effectively separates signal events from background events, as illustrated in Fig. 5.5. Subsequently, the signal-to-background ratio is optimised based on this discriminant variable.

Our analysis reveals that the most discriminative observables are as follows: (i) the transverse momentum of the b -jet, (ii) the missing transverse energy \cancel{E}_T , (iii) the transverse momentum of the c -jet, (iv) the transverse mass of the muon $M_T[\mu]$, and (v) the invariant mass of one of the charged Higgs bosons decaying into a bc -jet pair, $M_{\text{inv}}[bc]$. For the scenario **S2** with $M_{H^\pm} = 110$ GeV and $M_H = 500$ GeV, the distributions of these observables are presented in Fig. 5.7.

A notable feature emerges from the distributions of $M_T[\mu]$ and $M_{\text{inv}}[bc]$, a clear distinction between the behaviour of the signal and background processes. Specifically, both observables exhibit a resonant peak around $M_{H^\pm} = 110$ GeV, which corresponds to the presence of the charged scalar bosons. One of these bosons decays into a bc -jet pair, while the other decays into $\mu + \cancel{E}_T$. These features represent the most distinctive signatures of our signal. Based on these observations, one could directly impose cuts on these variables as follows,

1. $M_{H^\pm} - 20 < M_T[\mu] < M_{H^\pm} + 20$ GeV,
2. $M_{H^\pm} - 30 < M_{\text{inv}}[bc] < M_{H^\pm} + 20$ GeV,
3. $P_T[b] > 50$ GeV,
4. MET > 20 GeV,
5. $P_T[j_1] > 30$ GeV.

With the aforementioned kinematic cuts (which we rigorously verify using the `MadAnalysis5` package [103]), a good *Signal Significance* can be achieved. Nevertheless, we instead choose to impose cuts on the variable \mathbf{xgb} as detailed below.

(a) Transverse moment of the b -jet(b) Missing energy transverse \cancel{E}_T (c) Transverse moment of the c -jet.

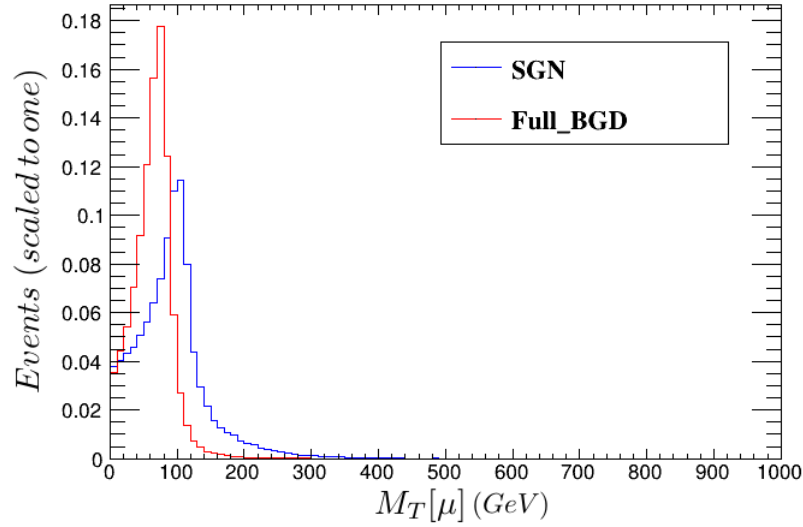
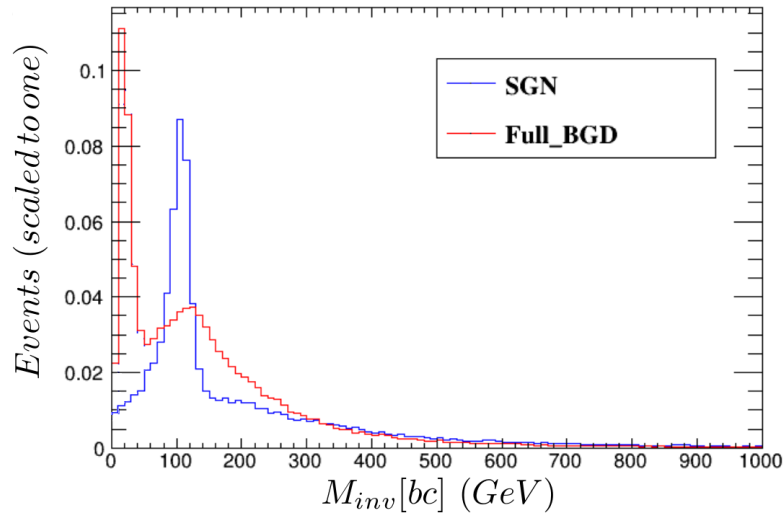
(d) Transverse mass of the muon $M_T[\mu]$.(e) Invariant mass of one of the charged Higgs bosons decaying into a bc -jet pair.

Figure 5.7: Kinematic distributions for the signal and SM background processes are shown for the following observables: (a) transverse momentum of the b -jet, (b) missing transverse energy \cancel{E}_T , (c) transverse momentum of the c -jet, (d) transverse mass of the muon $M_T[\mu]$ and (e) invariant mass of one of the charged Higgs bosons decaying into a bc -jet pair, $M_{inv}[bc]$. In all plots the signal is labeled as SGN, while the total background is labeled as Full_BGD.

Our procedure involves maximising the *Signal Significance*, defined as the ratio given in Eq. (4.0.1). This maximum significance is achieved by applying a cut on the BDT output, requiring $\mathbf{xgb} > 0.95$, as clearly illustrated in Fig. 5.5.

Figure 5.8 presents contour plots of the *Signal Significance* as a function of the charged Higgs boson mass M_{H^\pm} and the integrated luminosity for scenarios **S2**¹ and **S3**. We find that scenario **S1** poses significant experimental challenges as can be seen in Fig. 5.1, which is why it has been excluded from our analysis.

We find that scenario **S2** is the most promising, predicting a statistical significance of 5σ in the mass range $100 \leq M_{H^\pm} \leq 160$ GeV for an integrated luminosity of 250–300 fb^{-1} . This scenario could potentially be explored at the LHC and subsequently tested with greater precision at the HL-LHC.

Regarding scenario **S3**, it also holds potential for experimental investigation, with the possibility of exploring charged Higgs boson masses up to $M_{H^\pm} \sim 250$ GeV. We anticipate achieving a statistical significance of 5σ for $100 \lesssim M_{H^\pm} \lesssim 250$ GeV once an integrated luminosity in the range 220–1000 fb^{-1} is accumulated.

5.2 Collider Analysis for the FNSM

Having completed the collider analysis for the 2HDM, we proceed to perform an analogous analysis within the FNSM framework. We first present the decay width $\Gamma(H_F \rightarrow h\bar{f}f)$ for the three-body process. The study of such processes is of particular interest, as they can exhibit a sizeable BR. The Feynman diagrams contributing to these reactions are shown in Fig. 5.9 [104].

¹We explore masses up to $M_{H^\pm} = 200$ GeV. However, the *Signal Significance* remains negligible in this range.

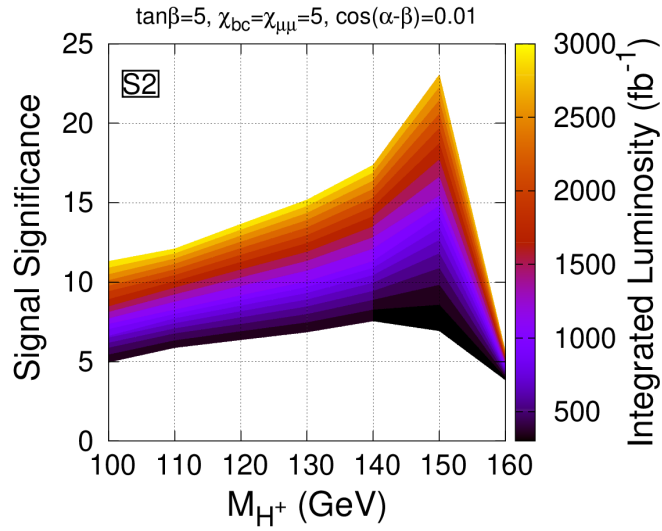
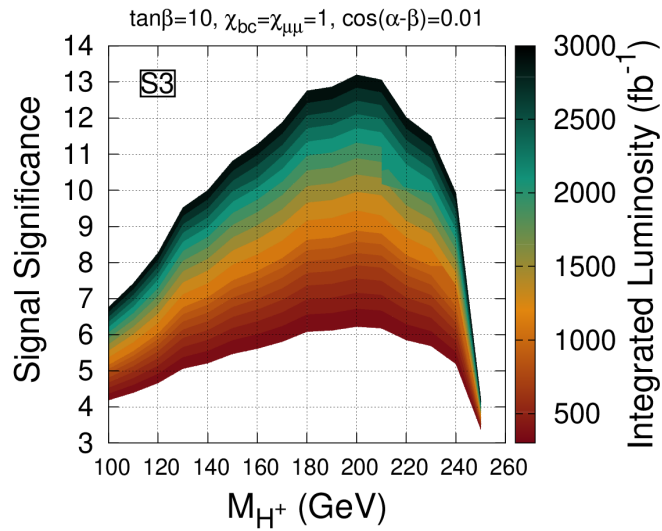
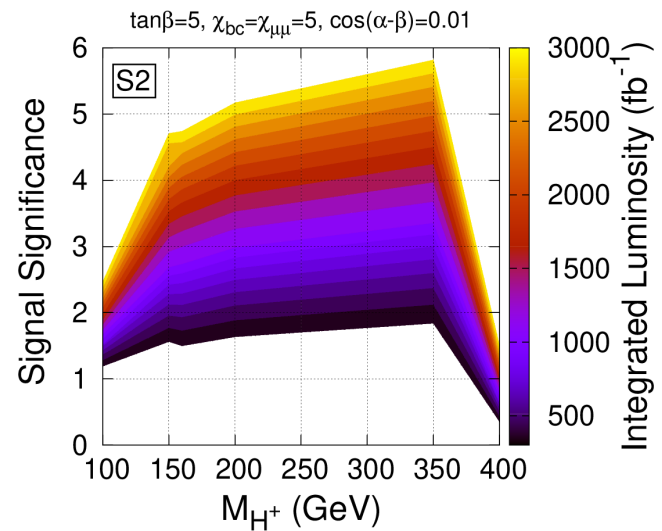
(a) **S2**, $M_H = 500$ GeV(b) **S3**, $M_H = 500$ GeV(c) **S2**, $M_H = 800$ GeV

Figure 5.8: *Signal Significance* as a function of the charged Higgs boson mass M_{H^\pm} and the integrated luminosity for the following scenarios: (a) **S2**, (b) **S3** with $M_H = 500$ GeV, and (c) **S2** with $M_H = 800$ GeV.

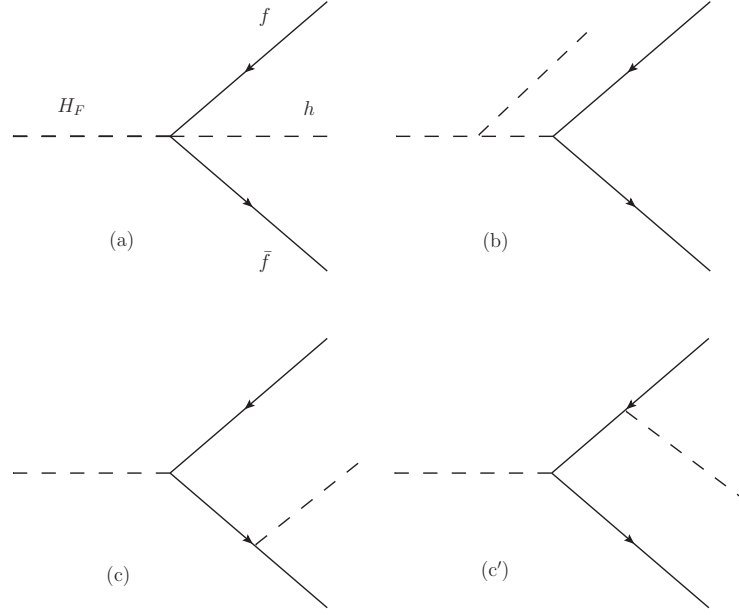


Figure 5.9: Feynman diagrams inducing the $H_F \rightarrow \bar{f}fh$ decay in the FNSM. Diagram (a) shows an effective vertex corresponding to a dimension-5 operator suppressed by $1/\Lambda$. Such interactions are non renormalisable and are treated within the effective field theory framework, where higher dimensional operators are systematically incorporated and suppressed by powers of the new physics scale Λ , as shown in Eq. (3.3.15). This ensures the consistency of the model at energies below Λ .

The decay width can be written as follows

$$\Gamma(H_F \rightarrow h\bar{f}f) = \frac{M_{H_F}}{256\pi^3} \int dx_a \int dx_b |\overline{\mathcal{M}}|^2, \quad (5.2.1)$$

where the average square amplitude is given by

$$\begin{aligned} |\overline{\mathcal{M}}|^2 = & \frac{1}{2(x_a + x_b + x_h - 2)^2} (x_a + x_b + x_h - 4x_t - 1) ((x_a + x_b + x_h - 2)C_a + C_b)^2 \\ & + \frac{2}{(x_a - 1)^2(x_b - 1)^2} \left((x_a - 1)(x_b - 1)(x_a - x_b)^2 \right. \\ & - 16(x_a + x_b - 2)^2 x_t^2 \\ & + 4(x_a + x_b - 2)(2 - 3x_b + x_a(4x_b - 3))x_t \\ & \left. + x_h \left(4(x_a + x_b - 2)^2 x_t - (x_a - x_b)^2 \right) \right) C_c^2 \end{aligned}$$

$$\begin{aligned}
& - \frac{4\sqrt{x_t}}{(x_a - 1)(x_b - 1)} \left(x_a^2 + 2(3x_b + x_h - 4x_t - 3)x_a + x_b^2 - 4x_h \right. \\
& \quad \left. + 2x_b(x_h - 4x_t - 3) + 16x_t + 4 \right) C_a C_c \\
& - \frac{4\sqrt{x_t}}{(x_a - 1)(x_b - 1)(x_a + x_b + x_h - 2)} \\
& \quad \times \left(x_a^2 + 2(3x_b + x_h - 4x_t - 3)x_a + x_b^2 - 4x_h \right. \\
& \quad \left. + 2x_b(x_h - 4x_t - 3) + 16x_t + 4 \right) C_b C_c. \tag{5.2.2}
\end{aligned}$$

with $x_a = (m_a/M_{H_F})^2$. The factors $C_a = g_{H_F h f f}$, $C_b = g_{H_F h h} g_{h f f}/m_h^2$, and $C_c = g_{H_F f f} g_{h f f}/m_h$ represent the coupling constants involved in the Feynman diagrams of Fig. 5.9.

Finally, the integration domain is given by

$$2\sqrt{x_t} \leq x_a \leq 1 - x_h - 2\sqrt{x_t x_h}, \tag{5.2.3}$$

$$x_b \begin{matrix} \geq \\ \leq \end{matrix} \frac{2(1 - x_h + 2x_t) + x_a(x_a + x_h - 2x_t - 3) \mp \sqrt{x_a^2 - 4x_t} \sqrt{(x_a + x_h - 1)^2 - 4x_h x_t}}{2(1 - x_a + x_t)}. \tag{5.2.4}$$

Meanwhile, the production cross-section of the heavy CP-even Flavon H_F (or the pseudo scalar A_F , for that matter) depends primarily on the $g_{H_F t \bar{t}} = \frac{c_\alpha v + s_\alpha v_s}{v_s} \frac{y_t}{\sqrt{2}}$ ($g_{A_F t \bar{t}} = \frac{v}{v_s} \frac{y_t}{\sqrt{2}}$) coupling. The corresponding term in the effective Lagrangian reads [105]:

$$\mathcal{L}_{\text{eff}} = \frac{1}{v} g_{hgg} h G_{\mu\nu} G^{\mu\nu}, \tag{5.2.5}$$

$$g_{Sgg} = -i \frac{\alpha_S}{8\pi} \tau (1 + (1 - \tau) f(\tau)) \quad \text{with} \quad \tau = \frac{4M_t^2}{M_h^2}, \tag{5.2.6}$$

$$f(\tau) = \begin{cases} (\sin^{-1} \sqrt{\frac{1}{\tau}})^2, & \tau \geq 1, \\ -\frac{1}{4} [\ln \frac{1 + \sqrt{1 - \tau}}{1 - \sqrt{1 - \tau}} - i\pi]^2 & \tau < 1. \end{cases} \tag{5.2.7}$$

In FNSM, the ggh , ggH_F and ggA_F couplings are given, respectively, by:

$$g_{hgg} = \left(\frac{c_\alpha v_s - s_\alpha v}{v_s} \right) g_{Sgg},$$

$$\begin{aligned}
g_{H_F gg} &= \left(\frac{c_\alpha v + s_\alpha v_s}{v_s} \right) g_{S gg}, \\
g_{A_F gg} &= \frac{v}{v_s} (-i \alpha_S / \pi) \tau f(\tau),
\end{aligned}
\tag{5.2.8}$$

The scalar A_F is CP-odd, while h and H_F are CP-even scalars. When the couplings involving left and right-handed fields are expressed in terms of Dirac fields, the Hermitian part of the coupling in Eq. (3.3.11) introduces an imaginary factor $i = \sqrt{-1}$ for h and H_F and a γ_5 coupling for A_F . Consequently, the result of the top quark loop integral differs for h , H_F , and A_F . [106, 107]. It is to be noted that for $M_{H_F, A_F} > 2 M_t$, $f(\tau) = -\frac{1}{4} [\ln \frac{1+\sqrt{1-\tau}}{1-\sqrt{1-\tau}} - i\pi]^2$.

Following a similar procedure as described previously for the 2HDM-III framework, we first use `FeynRules` [96] to derive the FNSM model in our computational framework and generate the corresponding UFO files for `MadGraph5` [108]. We compute the production cross-sections for the aforementioned production and decay processes using the resulting particle spectrum.

Background events within the SM were generated using the `MadGraph_aMC@NLO` framework [108]. Simulations of parton showering and hadronisation were subsequently performed using `Pythia-8` [109]. For this analysis, the default ATLAS card provided in the `Delphes-3.4.2` package was employed as well as to emulate the detector response [110]. We consider the Leading Order cross-sections calculated by `MadGraph_aMC@NLO` for both signal and background processes.

Having outlined the methodology for evaluating collider observables, we now define in Table 5.3 three different scenarios to be analysed in the following analysis. For practicality, we label these scenarios in the same way as for the 2HDM-III collider analysis: **S1**, **S2**, and **S3**.

Parameter	S1	S2	S3
$\cos \alpha$	0.995	0.995	0.995
Λ	1 TeV	1.5 TeV	2.5 TeV
v_s	1 TeV	1.5 TeV	2.5 TeV

Table 5.3: Scenarios (**S1**, **S2**, **S3**) used in the calculations.

In Table 5.4 we present the numerical cross-sections of the proposed signal, along with the corresponding number of events generated for the scenarios **S1**, **S2** and **S3**.

Scenario	M_{H_F} (GeV)	$\sigma(pp \rightarrow H_F \rightarrow hb\bar{b})$ (fb)	Events ($\mathcal{L}_{\text{int}} = 300\text{fb}^{-1}$)
S1	(800, 900, 1000)	(6.8, 3.3, 1.7)	(2040, 990, 510)
S2	(800, 900, 1000)	(4.3, 2.3, 1.3)	(1290, 690, 390)
S3	(800, 900, 1000)	(3.8, 2.1, 1.1)	(1140, 630, 330)

Table 5.4: Cross-section of the signal for scenarios **S1**, **S2**, **S3**.

Meanwhile, Fig. 5.10 provides an overview of the production cross-section for the signal process $pp \rightarrow H_F \rightarrow hb\bar{b}$, considering the three scenarios **S1**, **S2**, and **S3**.

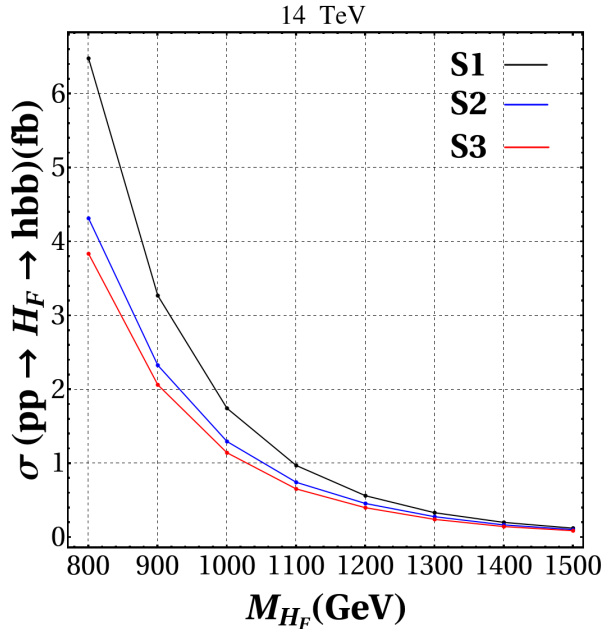


Figure 5.10: Production cross-section of the signal $pp \rightarrow H_F \rightarrow hb\bar{b}$. The centre-of-mass energy was set to 14 TeV.

On the other hand, we analyse two specific decay channels of the Higgs boson: (i) $pp \rightarrow H_F \rightarrow hb\bar{b}(h \rightarrow b\bar{b})$ and (ii) $pp \rightarrow H_F \rightarrow hb\bar{b}(h \rightarrow \gamma\gamma)$.

The identification of b -jets produced from the fragmentation and hadronisation of bottom quarks plays a crucial role in distinguishing the signal from background processes, which typically involve gluons, light-flavour jets (u , d , s) and charm-quark fragmentation. To address this challenge we employ the **FastJet** package [100] (via

SM background process	Cross-section (fb)
$pp \rightarrow t\bar{t}$	26910
$pp \rightarrow Wh$	0.5463
$pp \rightarrow ZZ$	231.5
$pp \rightarrow Zh$	79.75
$pp \rightarrow bbjj$	5.441×10^8

Table 5.5: Cross-section of the SM background processes for the $h \rightarrow b\bar{b}$ channel.

MadAnalysis [103]) and utilise the anti- k_T algorithm [101] for jet reconstruction. Additionally, we incorporate a b -tagging efficiency of $\epsilon_b = 90\%$. The probabilities of misidentifying a c -jet or any other light jet (j) as a b -jet are set to $\epsilon_c = 5\%$ and $\epsilon_j = 1\%$, respectively [111].

$pp \rightarrow H_F \rightarrow hb\bar{b}(h \rightarrow b\bar{b})$

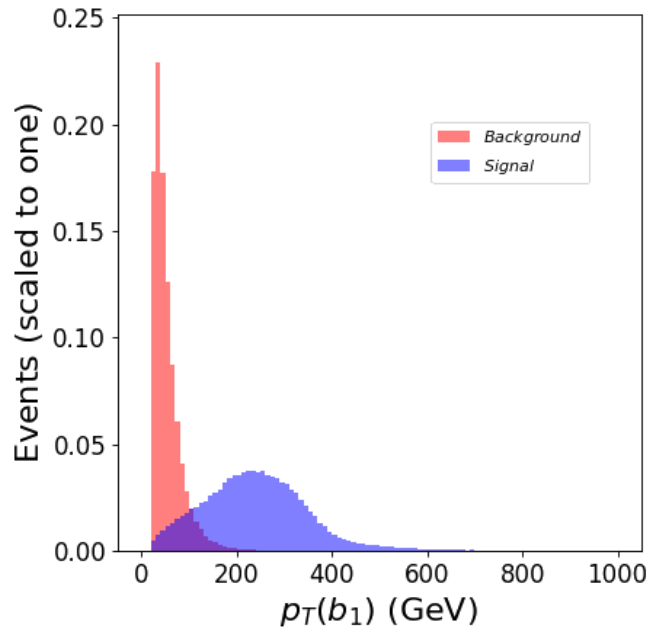
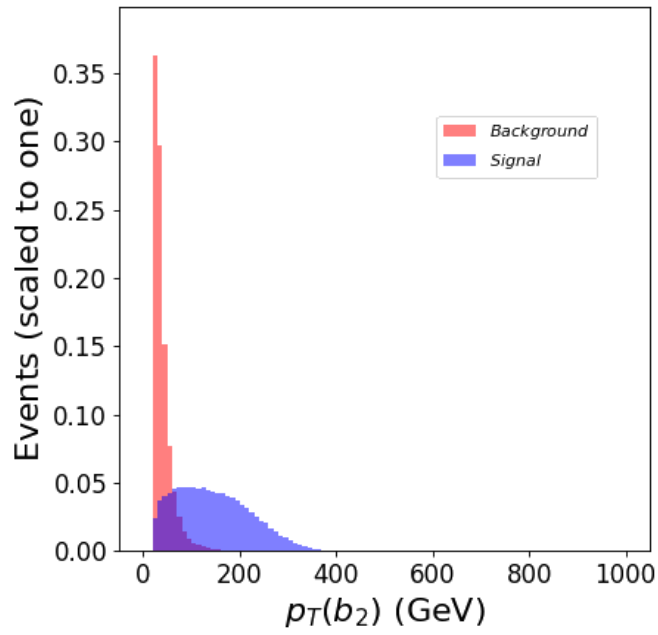
The SM background processes for the $h \rightarrow b\bar{b}$ channel are given by

- $pp \rightarrow t\bar{t}$, ($t \rightarrow W^+b$, $W^+ \rightarrow c\bar{b}$, $\bar{t} \rightarrow W^-b$, $W^+ \rightarrow b\bar{c}$),
- $pp \rightarrow Wh$, ($W \rightarrow cb$, $h \rightarrow b\bar{b}$),
- $pp \rightarrow ZZ$, ($Z \rightarrow b\bar{b}$, $Z \rightarrow b\bar{b}$),
- $pp \rightarrow Zh$, ($Z \rightarrow b\bar{b}$, $h \rightarrow b\bar{b}$),
- $pp \rightarrow bbjj$, where j denotes non-bottom-quark jets.

The numerical cross-section of the SM background processes is presented in Table 5.5.

For this channel, the b -jets originating from the primary vertex are expected to exhibit high transverse momentum, denoted as $p_T(b_1, b_2)$. In contrast, the b -jets produced via the decay $h \rightarrow b\bar{b}$ possess lower transverse momentum, $p_T(b_3, b_4)$, compared to the primary b -jets. A key feature of our signal, and its most distinctive signature, is the resonant effect arising from the decay chain $H_F \rightarrow hb\bar{b}(h \rightarrow b\bar{b}) \rightarrow b\bar{b}b\bar{b}$.

In Fig. 5.13 we present the invariant mass $M_{\text{inv}}(b_1b_2b_3b_4)$ for the scenario **S1** with $M_{H_F} = 800$ GeV.

(a) $p_T(b_1)$ (b) $p_T(b_2)$

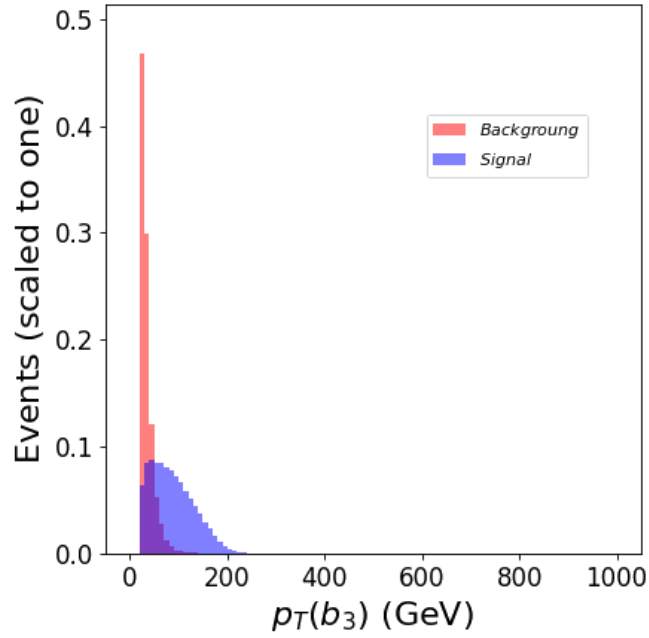
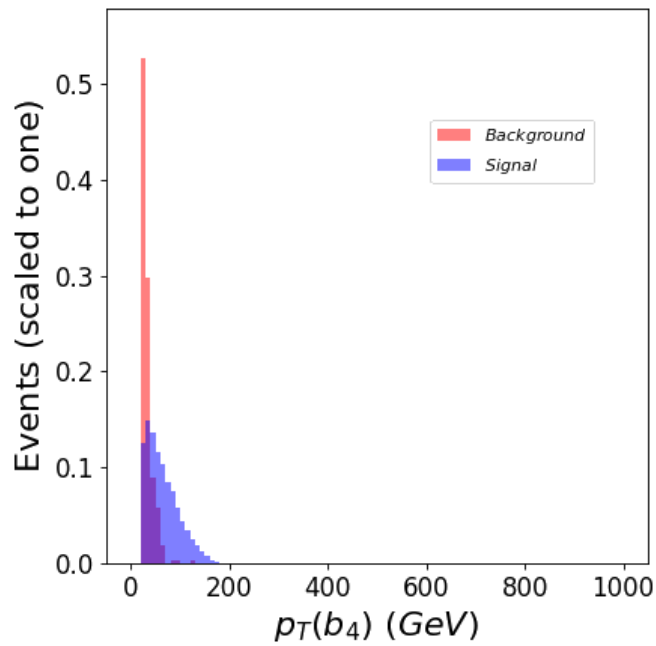
(c) $p_T(b_3)$ (d) $p_T(b_4)$

Figure 5.12: b -jet transverse momentum normalised distributions for the signal and total background following the application of the acceptance cuts.

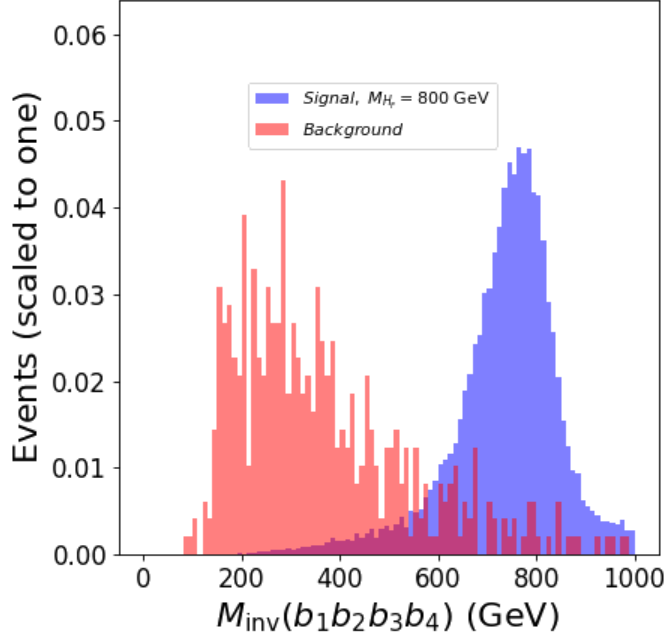


Figure 5.13: Normalised distribution of the reconstructed invariant mass $M_{\text{inv}}(b_1b_2b_3b_4)$ for the signal and background processes.

Meanwhile, Fig. 5.12 displays the $p_T(b_i)$ distributions ($i = 1, 2, 3, 4$) for both the signal and the SM background processes. The subscript 1 (4) corresponds to the b -jet with the highest (lowest) transverse momentum, while the subscripts 2 and 3 represent the second and third most energetic b -jets, respectively.

From Fig. 5.12, we observe that $p_T(b_1, b_2)$ is higher than $p_T(b_3, b_4)$, as the former originates from the primary vertex, while the latter arises from the decay of the Higgs boson. Both $p_T(b_i)$ and $M_{\text{inv}}(b_1b_2b_3b_4)$ are the most critical variables for isolating the signal from the background.

$$pp \rightarrow H_F \rightarrow hb\bar{b}(h \rightarrow \gamma\gamma)$$

Regarding the di-photon channel, the SM background processes are as follows,

- $pp \rightarrow ht\bar{t}$, ($h \rightarrow \gamma\gamma$, $t \rightarrow W^+\bar{b}$, $W^+ \rightarrow \ell^+\nu_\ell$, $\bar{t} \rightarrow W^-b$, $W^+ \rightarrow \ell^-\bar{\nu}_\ell$),
- $pp \rightarrow t\bar{t}\gamma\gamma$, ($t \rightarrow W^+\bar{b}$, $W^+ \rightarrow \ell^+\nu_\ell$, $\bar{t} \rightarrow W^-b$, $W^+ \rightarrow \ell^-\bar{\nu}_\ell$),
- $pp \rightarrow Wh$, ($W \rightarrow cb$, $h \rightarrow \gamma\gamma$),

SM background process	Cross-section (fb)
$pp \rightarrow htt\bar{t}$	3.2×10^{-2}
$pp \rightarrow tt\gamma\gamma$	0.57
$pp \rightarrow Wh$	9.7×10^{-4}
$pp \rightarrow Zh$	0.14
$pp \rightarrow hjj$	13.62
$pp \rightarrow \gamma\gamma jj$	1.1×10^5
$pp \rightarrow bb\gamma\gamma$	5113

Table 5.6: Cross-section of the SM background processes for the $h \rightarrow \gamma\gamma$ channel.

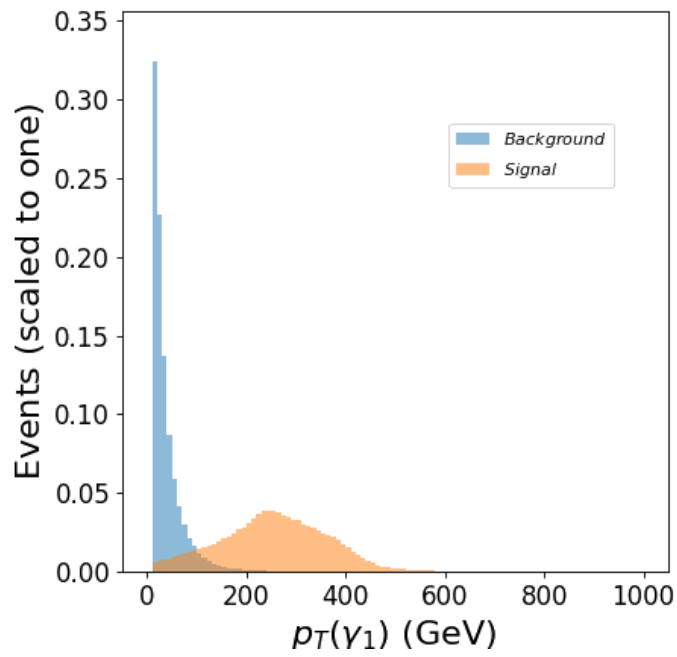
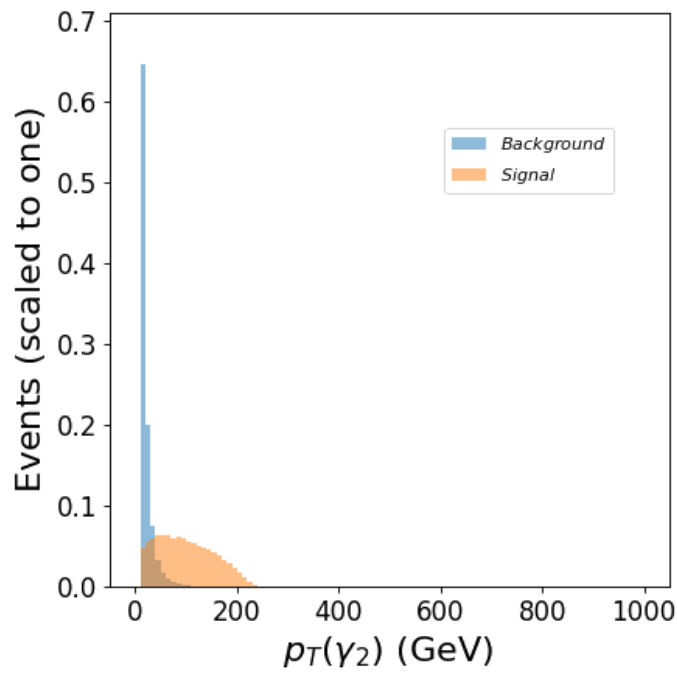
- $pp \rightarrow Zh, (Z \rightarrow b\bar{b}, h \rightarrow \gamma\gamma),$
- $pp \rightarrow hjj, (h \rightarrow \gamma\gamma),$
- $pp \rightarrow \gamma\gamma jj$
- $pp \rightarrow \gamma\gamma b\bar{b}.$

The numerical cross-section of the SM background processes is presented in Table 5.6.

In this case we observe a scenario similar to the previous channel with respect to the resonant effect arising from the decay $H_F \rightarrow b\bar{b}\gamma\gamma$. Figure 5.16 displays the invariant mass distribution $M_{\text{inv}}(bb\gamma\gamma)$ for $M_{H_F} = 900$ GeV. As in the previous channel, we also present in Fig. 5.15 the the $p_T(b_i), p_T(\gamma_i)$ ($i = 1, 2$).

5.2.1 Multivariate Analysis

Following the kinematic analysis, we found that most of the observables used to distinguish the signal from the background exhibit relatively weak discriminating power. Consequently, we perform the final candidate selection using MVA discriminators, combining these observables into the single, more powerful classifier. As for the 2HDM-III, we employ a BDT algorithm implemented via the XGBoost library.

(a) $p_T(\gamma_1)$ (b) $p_T(\gamma_2)$

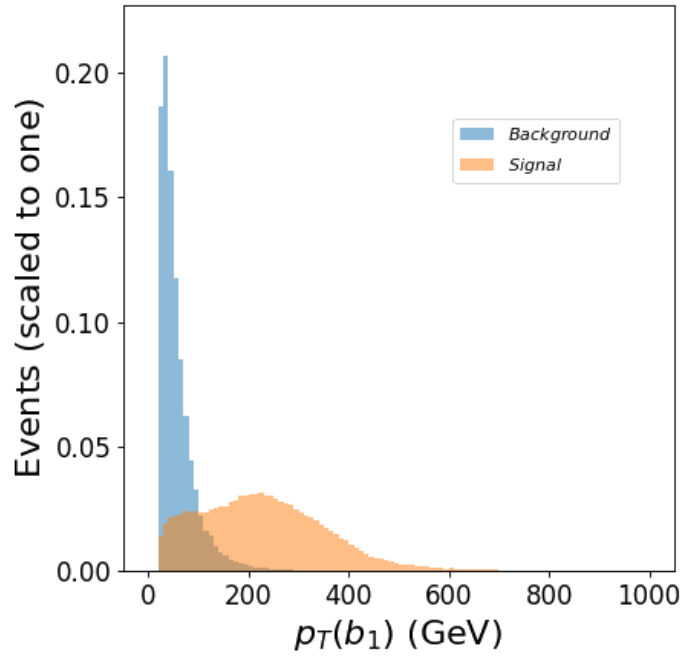
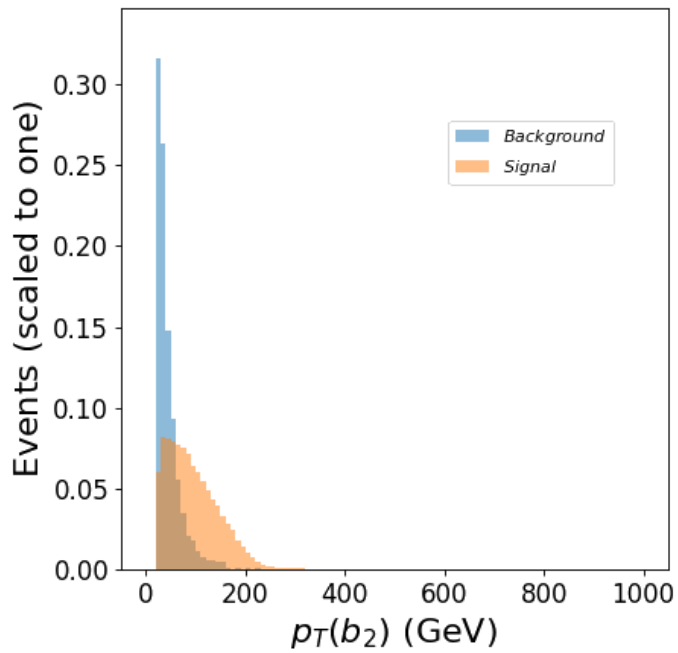
(c) $p_T(b_1)$ (d) $p_T(b_2)$

Figure 5.15: Normalised p_T distributions for photons (γ) and b -jets, comparing signal and total background after acceptance cuts.

For this framework, we train the BDT classifiers using variables associated with the kinematics of both final-state and intermediate-state particles, including the following:

Case 1: Four b -jets (b_1, b_2, b_3, b_4):

- Transverse momentum (p_T): $p_T(b_1), p_T(b_2), p_T(b_3), p_T(b_4)$
- Rapidity (η): $\eta(b_1), \eta(b_2), \eta(b_3), \eta(b_4)$
- Combined transverse momentum: $p_T(b_1 b_2 b_3 b_4)$

Case 2: Two b -jets (b_1, b_2) and two photons (γ_1, γ_2):

- Transverse momentum (p_T): $p_T(b_1), p_T(b_2), p_T(\gamma_1), p_T(\gamma_2)$
- Rapidity (η): $\eta(b_1), \eta(b_2), \eta(\gamma_1), \eta(\gamma_2)$
- Invariant mass of the photons: $M(\gamma_1 \gamma_2)$
- Combined transverse momentum: $p_T(b_1 b_2 \gamma_1 \gamma_2)$

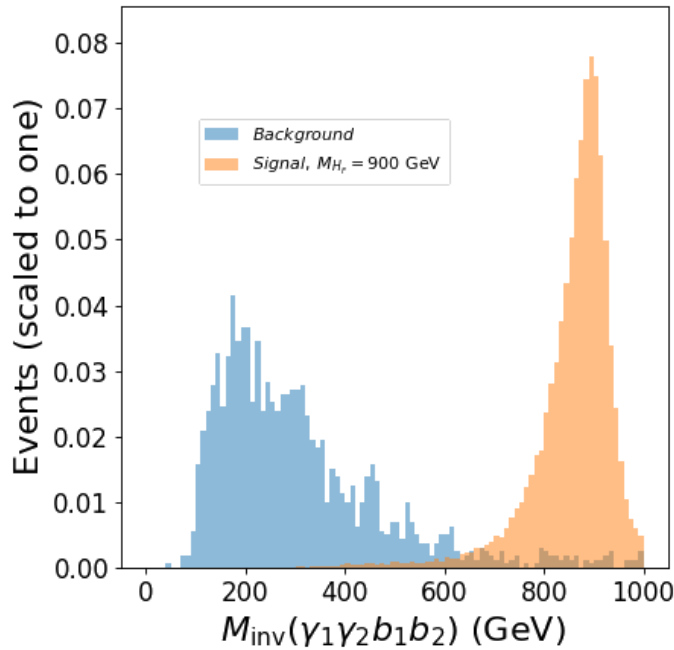


Figure 5.16: Normalised distribution of the reconstructed invariant mass $M_{\text{inv}}(\gamma_1 \gamma_2 b_1 b_2)$ for the signal and background processes.

The training of the BDT classifier is carried out using Monte Carlo simulated samples. The dataset comprises 200,000 signal events and an equivalent number of background

events, distributed across the various background processes described earlier, each contributing 200,000 events. The data is partitioned such that 70% of the total events are allocated to the training dataset, while the remaining 30% are reserved for testing. To optimise the performance of the BDT model, this time we employ *HyperOpt* [112, 113] for hyperparameter tuning. *HyperOpt* library employs a Bayesian optimisation strategy to maximise classification performance. Rather than exhaustively testing every possible combination, this library treats the performance metric, such as accuracy, as a function of the hyperparameters. It builds a probabilistic model that estimates how changes in hyperparameters affect the performance.

This iterative process continues until it finds a set of hyperparameters that maximise the classifier's performance on validation data. This optimisation process helps prevent overfitting by finding the optimal balance between the model complexity and generalisation. Regularisation parameters in XGBoost also reduce overfitting by penalising overly complex trees. As an example of the resulting performance, Fig. 5.18 presents the ROC curve for the $h \rightarrow b\bar{b}$ channel. Moreover, in Fig. 5.19 we show the relationship between the $1/\text{Background Acceptance}$ and the Signal Acceptance for the same decay mode.

The expected number of candidates is determined based on the integrated luminosity and the corresponding cross-sections, and the signal and background samples are scaled accordingly. To maximise the *Signal Significance*, Eq. (4.0.1), the BDT selection is optimised independently for each channel. Figure 5.17 illustrates the *Signal Significance* (for the three scenarios) as a function of luminosity and the Flavon mass M_{H_F} for the process $pp \rightarrow H_F \rightarrow hb\bar{b}(h \rightarrow b\bar{b})$.

For scenario **S1**, we observe strong sensitivity to various choices of M_{H_F} at a 14 TeV pp collider with an integrated luminosity of 3000 fb^{-1} . Specifically, in the range $800 < M_{H_F} < 950 \text{ GeV}$, a *Signal Significance* between $3\text{-}5.6\sigma$ is achieved. Consequently, if nature favours scenario **S1**, the Flavon particle could potentially be discovered at the HL-LHC. In contrast, scenarios **S2** and **S3** exhibit a depletion in significance around $M_{H_F} = 1.2 \text{ TeV}$, followed by an increase near $M_{H_F} = 1.3 \text{ TeV}$.

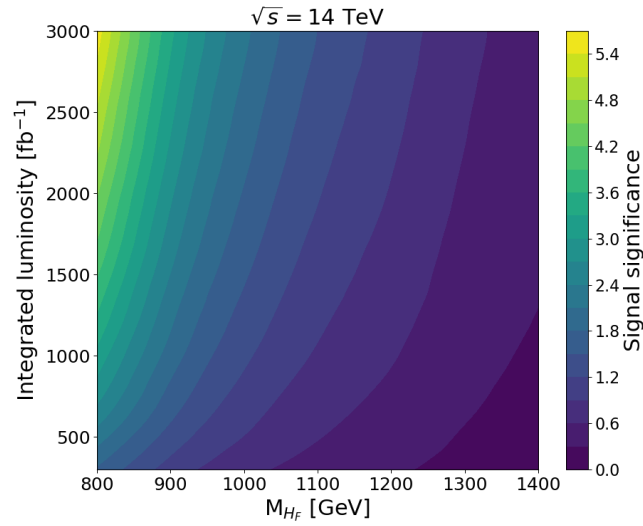
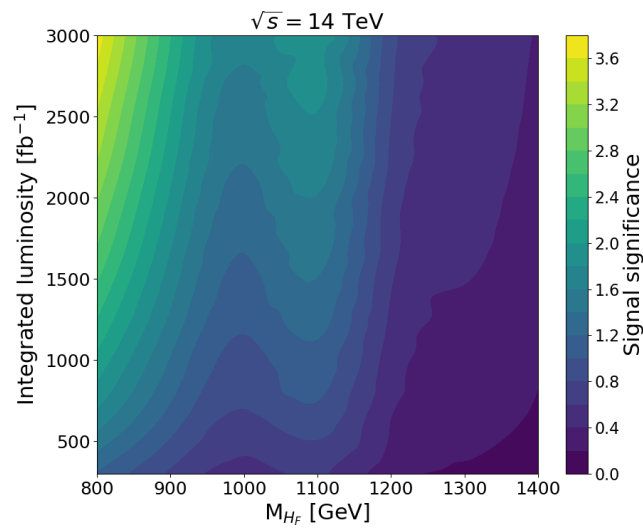
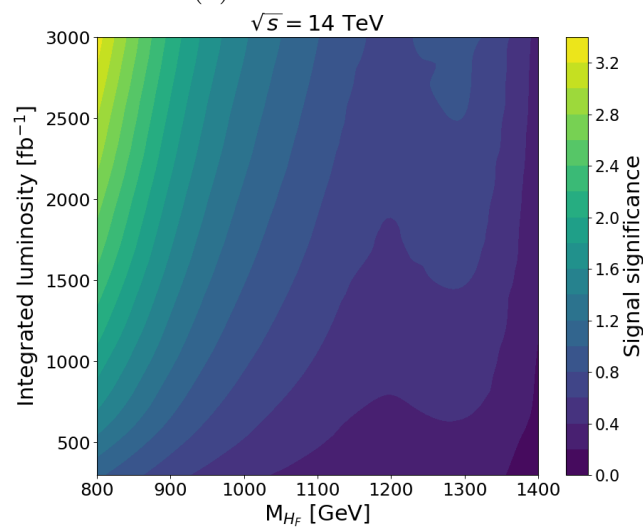
(a) **S1**(b) **S2**(c) **S3**

Figure 5.17: Density plot showing the *Signal Significance* for the $h \rightarrow b\bar{b}$ channel as a function of the integrated luminosity and the Flavon mass M_{H_F} , the subplots correspond to the following benchmark scenarios: (a) **S1**, (b) **S2**, and (c) **S3**.

M_{H_F} (GeV)	<i>Signal Significance</i>
800	1.75 σ
900	0.87 σ
1000	0.4872 σ

Table 5.7: The *Signal Significance* for the decay channel $h \rightarrow b\bar{b}$ evaluated at $\mathcal{L}_{\text{int}}=3000 \text{ fb}^{-1}$.

This behaviour can be attributed to the slightly enhanced performance of the BDT classifier for $M_{H_F} = 1.3 \text{ TeV}$ compared to 1.2 TeV.

Fig. 5.20 illustrates the *Signal Significance* as a function of luminosity and the Flavon mass M_{H_F} for the process $pp \rightarrow H_F \rightarrow hb\bar{b}(h \rightarrow \gamma\gamma)$. A key distinction in this channel is its reduced sensitivity compared to the earlier case, which is primarily attributed to the Branching Ratio $\mathcal{BR}(h \rightarrow \gamma\gamma) \sim 10^{-3}$. This suppression factor of 10^{-3} significantly diminishes the observable signal.

Despite this limitation, the prospect of a future 100 TeV collider [114] offers renewed optimism. Our calculations suggest that, with an integrated luminosity of 30 ab^{-1} , the sensitivity could achieve up to 5σ , making this channel a promising avenue for exploration in next generation experiments.

In Table 5.7 we show the *Signal Significance* values for the $h \rightarrow b\bar{b}$ channel, where we have taken into account the basic cut, included in the HL-LHC card [115], along with the following kinematic cuts applied for $M_{H_F} = 800 \text{ GeV}$:

1. $p_T(b_1b_2b_3b_4) > 80, 60, 50, 25 \text{ GeV}$,
2. $650 < M_{\text{inv}}(b_1b_2b_3b_4) < 900 \text{ GeV}$,
3. Combined transverse momentum $p_T(b_1b_2b_3b_4) > 550 \text{ GeV}$.

We can notice that the use of BDT substantially improves the isolation and identification of the proposed signals, resulting in a more effective and accurate analysis.

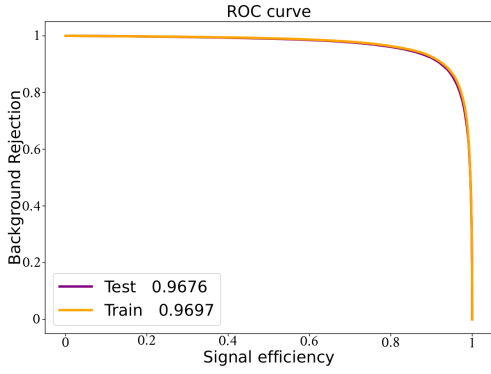
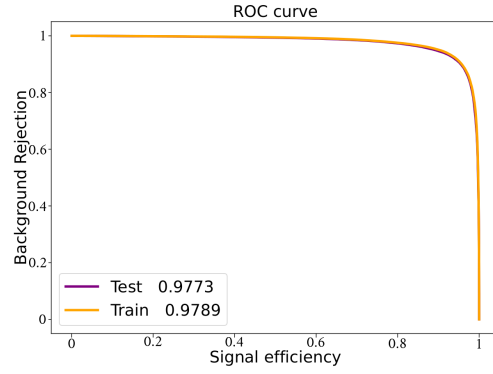
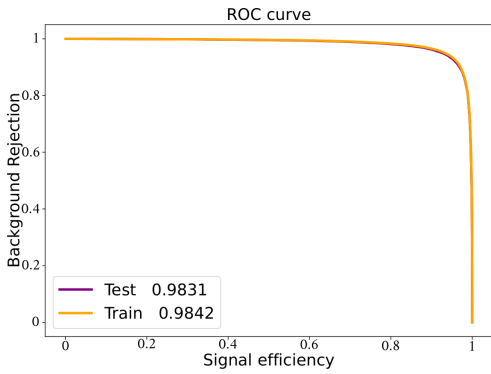
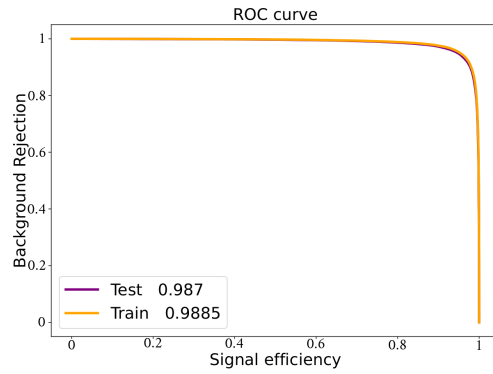
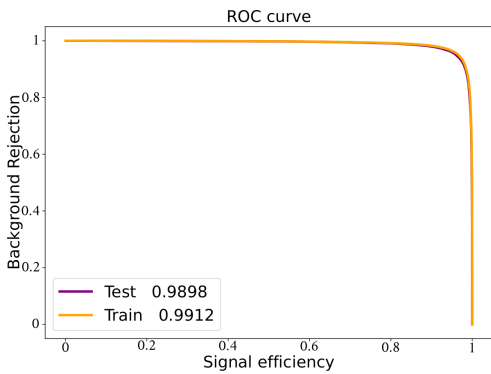
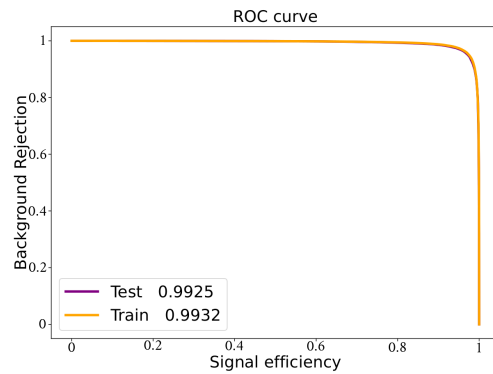
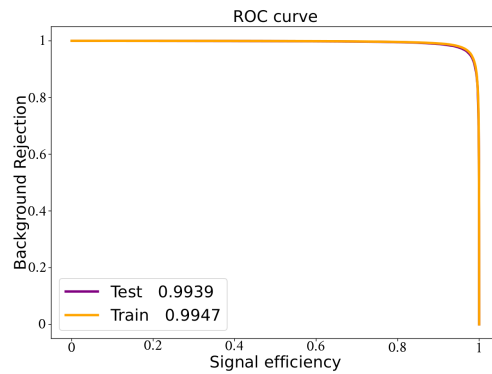
(a) $M_{H_F} = 800$ GeV(b) $M_{H_F} = 900$ GeV(c) $M_{H_F} = 1000$ GeV(d) $M_{H_F} = 1100$ GeV(e) $M_{H_F} = 1200$ GeV(f) $M_{H_F} = 1300$ GeV(g) $M_{H_F} = 1400$ GeV

Figure 5.18: ROC curve for $h \rightarrow b\bar{b}$ channel showing the model's ability to predict training data (used to fit the model) and independent test data (evaluating performance on new, unseen data). Each plot corresponds to a different Flavon mass as indicated.

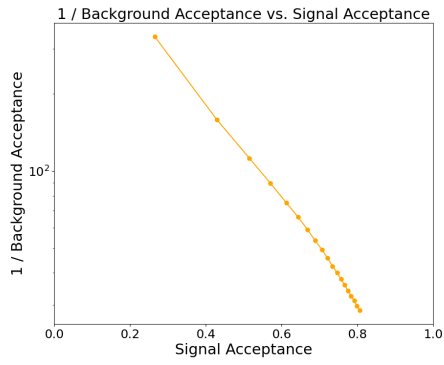
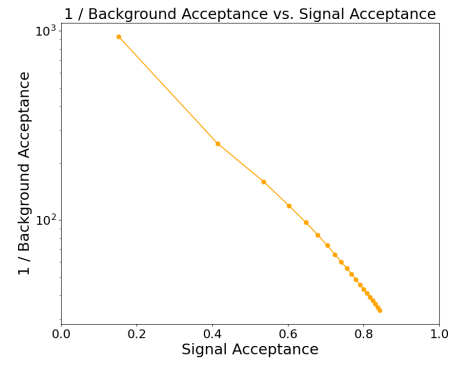
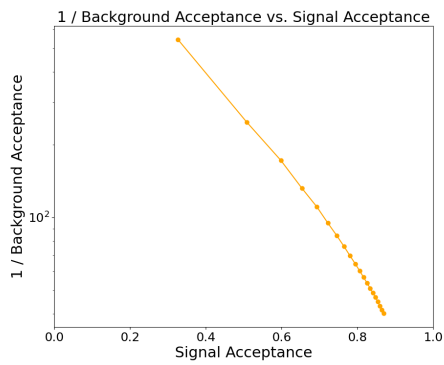
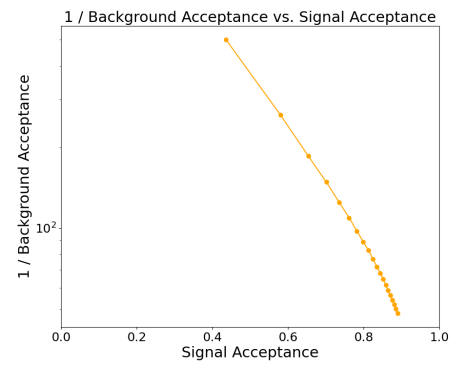
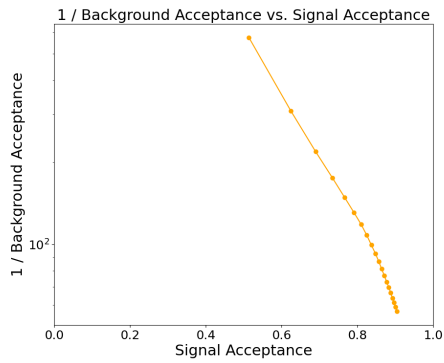
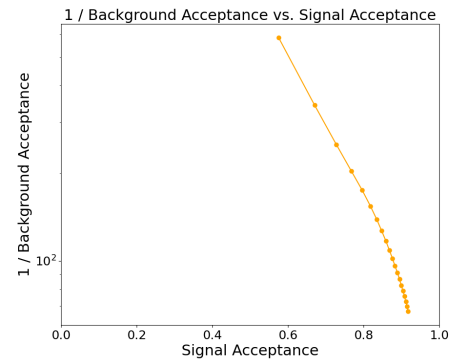
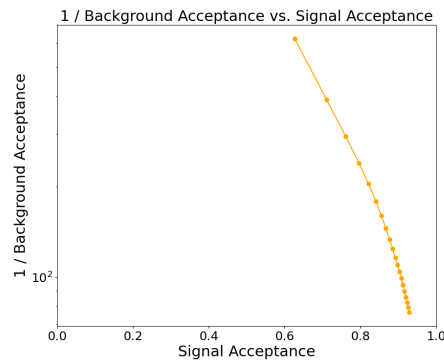
(a) $M_{H_F} = 800$ GeV(b) $M_{H_F} = 900$ GeV(c) $M_{H_F} = 1000$ GeV(d) $M_{H_F} = 1100$ GeV(e) $M_{H_F} = 1200$ GeV(f) $M_{H_F} = 1300$ GeV(g) $M_{H_F} = 1400$ GeV

Figure 5.19: $1/\text{Background Acceptance}$ against the Signal Acceptance for the $h \rightarrow b\bar{b}$ channel. Each plot corresponds to a different Flavon mass as indicated.

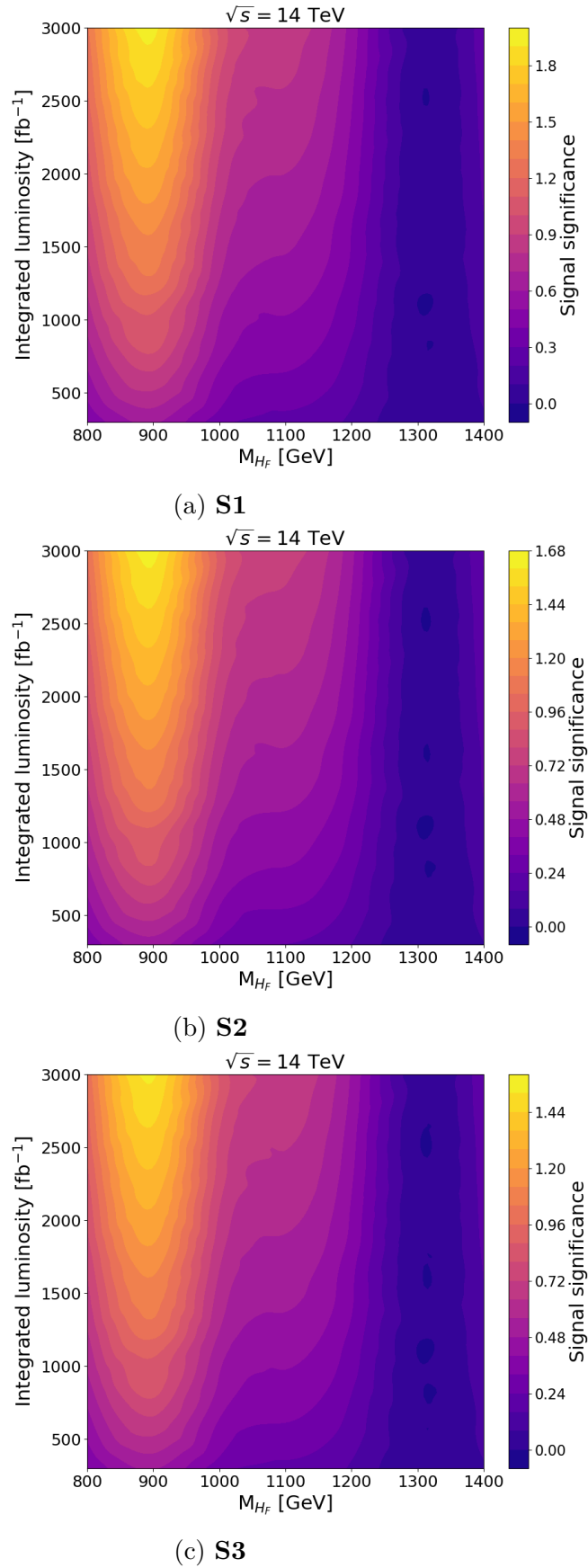


Figure 5.20: Density plot displaying the *Signal Significance* as a function of the integrated luminosity and the Flavon mass M_{H_F} for the $h \rightarrow \gamma\gamma$ decay channel, the subplots correspond to the following benchmark scenarios: (a) **S1**, (b) **S2**, and (c) **S3**.

5.2.2 Lepton-Flavour-Violating Decays $h \rightarrow e\mu$

Finally, we turn our attention to a distinctive signature of LFV decay of the SM-like Higgs boson, focusing on the channel $h \rightarrow e\mu$. In the scenarios we analyse, such a process is enabled by the extended scalar sector and associated Yukawa couplings, which offer a direct probe of flavour dynamics within the FNSM framework.

The signal process is

$$pp \rightarrow h \rightarrow e\mu,$$

As illustrated in Fig. 5.21, where h is the Higgs-like scalar but now assumed to have LFV couplings. As in the previous channels, our study targets the LHC with a centre-of-mass energy of 14 TeV (and potentially the HL-LHC).

The principal SM background processes arise primarily from:

- **Drell-Yan Production:**

- $Z/\gamma^* \rightarrow \tau^+\tau^- \rightarrow e\nu_e\bar{\nu}_\tau + \mu\nu_\mu\bar{\nu}_\tau$

- **Top Quark Pair Production ($t\bar{t}$):**

- $pp \rightarrow t\bar{t}$, with each top quark decaying as $t \rightarrow bW^+$ and $\bar{t} \rightarrow \bar{b}W^-$, followed by $W^+ \rightarrow e^+\nu_e$ and $W^- \rightarrow \mu^-\bar{\nu}_\mu$

- **Diboson Production:**

- $pp \rightarrow W^+W^-$, with $W^+ \rightarrow e^+\nu_e$ and $W^- \rightarrow \mu^-\bar{\nu}_\mu$

- $pp \rightarrow ZZ$, with each Z decaying into $Z \rightarrow \mu^+\mu^-$

- $pp \rightarrow W^+Z$, with $W^+ \rightarrow e^+\nu_e$ and $Z \rightarrow \mu^+\mu^-$

Although these SM processes can lead to final states with different flavour leptons, they lack an explicit LFV vertex within a single decay. Then in the SM with massive neutrinos, those processes could appear. However, the BR is extremely suppressed, making them practically undetectable. Consequently, their kinematic distributions, especially around a resonance in the $e\mu$ invariant mass, differ from a genuine two-body decay $h \rightarrow e\mu$.

Multivariate Analysis

Also for this decay, we generate both signal and background samples at parton level using `MadGraph5_aMC@NLO`, interfaced with `FeynRules` for the model implementation. Parton shower and hadronisation are simulated with `Pythia8`, and a fast detector simulation is performed via `Delphes3`, using an ATLAS-like parametrisation (consistent with the previous section). For the present analysis, we select a balanced dataset of 200,000 signal events and an equivalent number of background events.

To improve the separation of signal from background, we employ a Multivariate Analysis based on a BDT, similarly to the methodology applied in the previous sections. Here, we implemented `XGBoost`, optimising the hyperparameters through `Optuna` [116]. Here, we choose `Optuna` due to its more efficient exploration of large hyperparameter spaces and increased robustness against local minima compared to earlier Bayesian optimisation tools. This iterative approach enabled us to determine an optimal set of parameters, such as the number of trees (`NTree`) is set to 82, the maximum depth of each decision tree (`MaxDepth`) is set to 3, and the maximum number of leaves per tree (`MaxLeaves`) is set to 32. All other parameters are left at their default values.

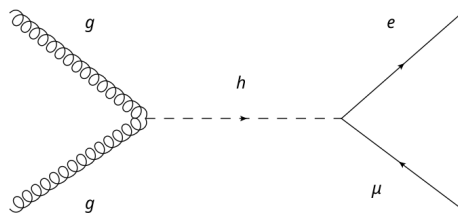


Figure 5.21: Feynman diagram of the LFV process $h \rightarrow e\mu$.

After training, the BDT outputs a single discriminant variable (`xgb`), which we can cut on in order to isolate a high purity signal region. Figure 5.22 shows the feature importances for the chosen input variables, illustrating that the invariant mass `DER_invariant_mass` (closely related to $M_{e\mu}$) emerges as the strongest discriminator, followed by other kinematic observables such as `PRI_met` and total transverse momentum `DER_pt_tot`.

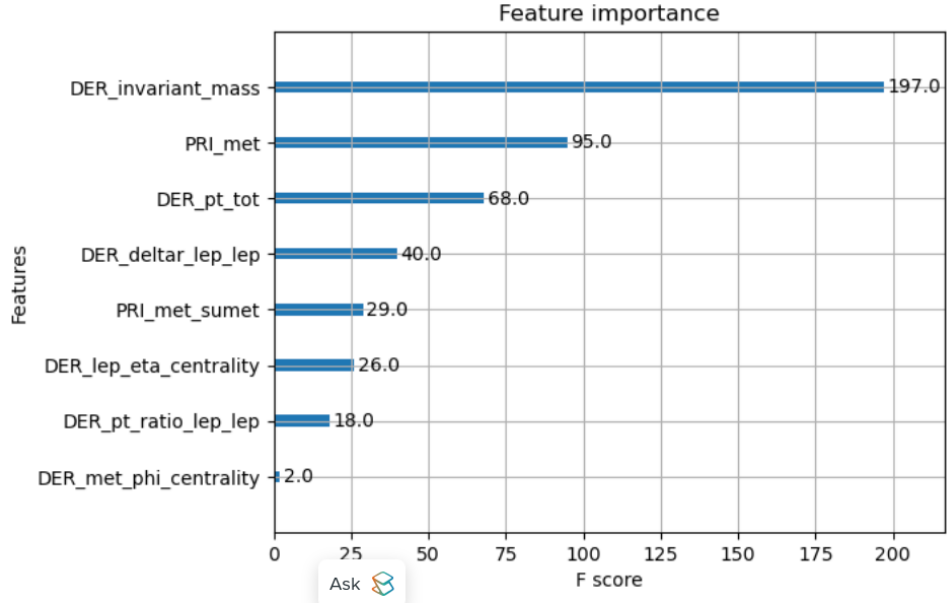


Figure 5.22: Feature importances from the XGBoost classifier, showing which variables contribute most strongly to separate signal from background for $h \rightarrow e\mu$. (Correlated features were removed prior to training, as discussed in the dataset preparation section of the previous chapter.)

We define the *Signal Significance* according to the usual approximation from Eq. (4.0.1). By scanning over different values of the model parameters (such as $\cos \alpha$, v_s , and LFV couplings like $Z_{e\mu}$), as well as varying the assumed integrated luminosity, we can identify regions in which the LFV decay $h \rightarrow e\mu$ that could be observed with evidence-level (3σ) or discovery-level (5σ) significance.

We consider the three benchmark scenarios (labelled **S4**, **S5**, and **S6**), corresponding to different choices of $(\cos \alpha, v_s, Z_{e\mu})$. We summarise below their cross-sections and the maximal significance obtained at an integrated luminosity of $\mathcal{L} = 3000 \text{ fb}^{-1}$:

Scenario	$\cos \alpha$	v_s (GeV)	$Z_{e\mu}$	σ (pb)	Luminosity (fb^{-1})
S4	0.991	350	0.0025	0.002395	3000
S5	-0.8949	2198.396	0.005601	0.003759	1800
S6	-0.9548	2461.078	0.009781	0.004404	1300

These parameters correspond to a significance of $\sim 5\sigma$.

In Figure 5.23 we show an example of distribution in the $(v_s, Z_{e\mu})$ plane [117], from which our three benchmark scenarios are selected.

Scenario	$\cos \alpha$	v_s (GeV)	$Z_{e\mu}$	σ (pb)	Luminosity (fb^{-1})
S4	0.991	350	0.0025	0.002395	1500
S5	-0.8949	2198.396	0.005601	0.003759	700
S6	-0.9548	2461.078	0.009781	0.004404	500

These parameters correspond to a significance of $\sim 3\sigma$.

Table 5.8: Parameters for scenarios **S4**, **S5** and **S6**

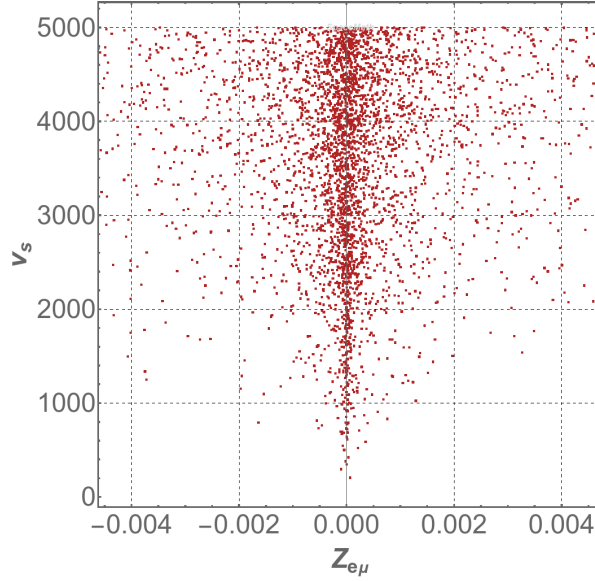


Figure 5.23: Parameter space in the plane $(v_s, Z_{e\mu})$ for $h \rightarrow e\mu$. Since this decay has not been observed, no signal strength is measured for this channel. Instead, only upper limits on the BR are available. In this analysis, we impose the experimental constraint $\mathcal{B}(h \rightarrow e\mu) < 6.2 \times 10^{-5}$.

In each case, the maximum significance is attained near a BDT discriminator threshold of $\mathbf{xgb} \approx 0.95$. Higher LFV couplings ($Z_{e\mu}$) and suitable mixing angles lead to larger cross sections and thus higher significance.

Figures 5.24 and 5.25 illustrate two key performance plots for **S6**. In Figure 5.24, the BDT prediction distributions for signal and background events in both training and test samples, along with their KS p-values. Furthermore, and in Figure 5.25, the corresponding ROC curves for the training and test datasets, indicating the AUC and confirming that the classifier exhibits robust separation power without significant overfitting.

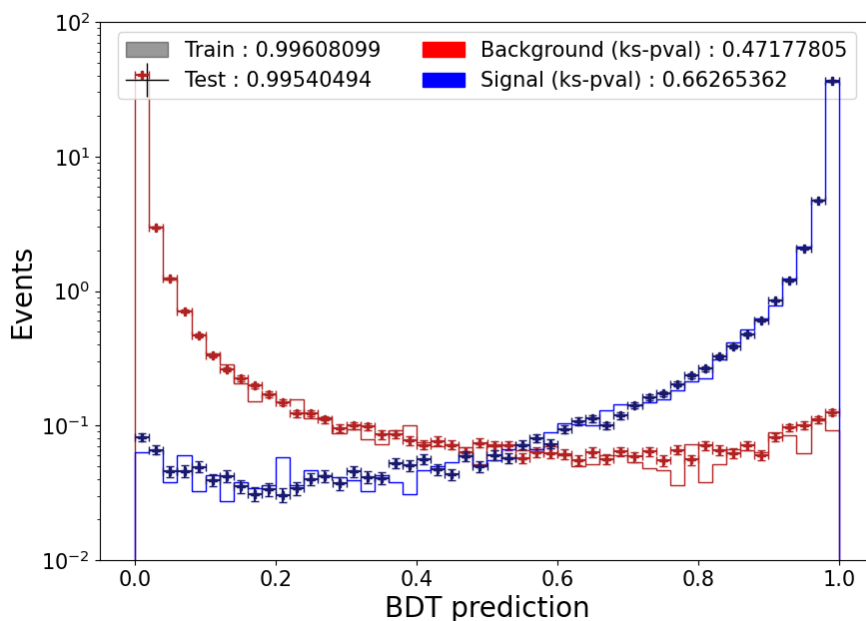


Figure 5.24: BDT classifier distribution for **S6**, comparing signal (blue) and background (red) events in both training and test samples. The legend shows the KS p-values for each distribution.

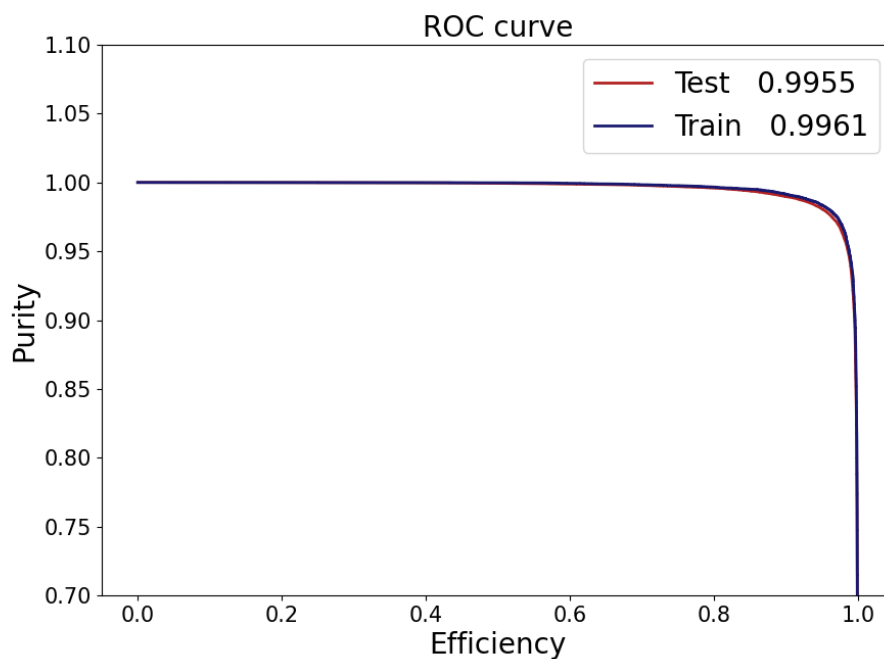
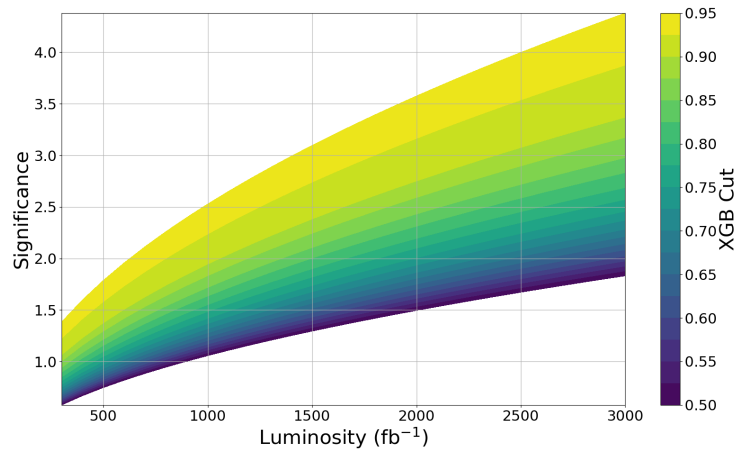


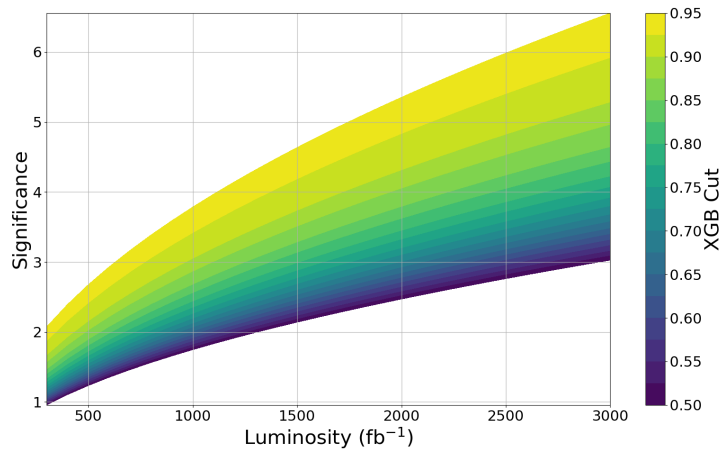
Figure 5.25: ROC curve for **S6**, illustrates purity vs. efficiency for both, the training (blue) and test (red) datasets. The AUC values are ~ 0.996 (train) and ~ 0.996 (test), reflecting strong discrimination.

Additionally, Fig. 5.26 shows how the *Signal Significance* varies with the integrated luminosity and the \mathbf{xgb} cut for **S4**, **S5** and **S6**. Finally, from the results obtained, only **S5** and **S6** would be feasible at the HL-LHC. Specifically, **S6** which can reach a 5σ discovery at around 1300 fb^{-1} , while **S5** achieves a 3σ evidence level at 700 fb^{-1} and a 5σ at 1800 fb^{-1} . Consequently, both scenarios could probe LFV in extended scalar sectors, complementing the flavour conserving searches presented in earlier sections.

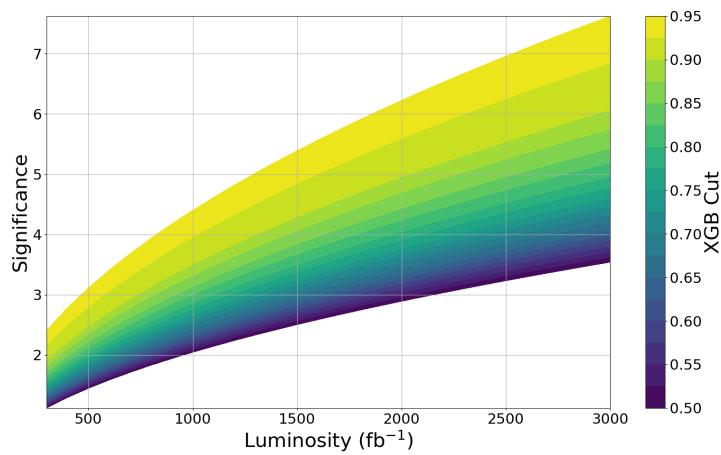
Our collider analyses within both the 2HDM-III and FNSM frameworks demonstrate the potential to uncover new physics signatures BSM. By integrating analytical decay width calculations with advanced multivariate techniques, we have identified promising regions of parameter space where the predicted signals could be distinguished from substantial SM backgrounds.



(a) S4



(b) S5



(c) S6

Figure 5.26: *Signal Significance* as a function of the integrated luminosity and the XGB cut for the decay $h \rightarrow e\mu$ in the scenarios: (a) S4, (b) S5, and (c) S6.

Chapter 6

Conclusions and Future Work

In this thesis, we worked on BSM scenarios through comprehensive phenomenological studies, using proton-proton collisions at the LHC and its future upgrade, the HL-LHC, operating at a centre of mass energy of 14 TeV. Our research addressed three principal directions: the production and potential detection of a charged scalar Higgs pair H^-H^+ decaying into the final state $\mu\nu_\mu cb$ within the 2HDM-III framework; the Flavon scalar (H_F) decaying into a $b\bar{b}h$, where the SM-like Higgs boson (h) subsequently decays into either a pair of photons ($h \rightarrow \gamma\gamma$) or a pair of b -quarks ($h \rightarrow b\bar{b}$); signals of LFV SM-like Higgs decays, specifically $h \rightarrow e\mu$ also within the framework of the Flavon Model.

Initially, we analysed the production and subsequent detection potential of charged scalar boson pairs predicted by the 2HDM-III, specifically through the process $pp \rightarrow H^-H^+ \rightarrow \mu\nu_\mu cb$. After an exhaustive scan of the parameter space, we identified two highly promising scenarios for experimental verification at the HL-LHC, the first scenario (with $\tan\beta = \chi_{\mu\mu} = \chi_{cb} = 5$, $\cos(\alpha - \beta) = 0.01$) and the second scenario (with $\tan\beta = 10$, $\chi_{\mu\mu} = \chi_{cb} = 1$, $\cos(\alpha - \beta) = 0.01$). It became evident from our analysis that the dominant contribution to the charged scalar pair production emerges from the decay of an on-shell heavy neutral scalar boson H . By employing Monte Carlo simulations and Machine Learning tools such as the BDT algorithm,

we efficiently suppressed the main background processes. This approach markedly improved the sensitivity compared to traditional kinematic analyses, reaching signal significances above 5σ for charged scalar masses in the range of $110 \lesssim M_{H^\pm} \lesssim 250$ GeV when the HL-LHC accumulates integrated luminosities of at least 300 fb^{-1} (assuming a neutral scalar mass of $M_H = 500$ GeV). Conversely, for a heavier neutral scalar boson mass ($M_H = 800$ GeV), higher luminosities ($\mathcal{L}_{\text{int}} \geq 2400 \text{ fb}^{-1}$) are required to reach comparable significance for charged Higgs masses between 180 and 360 GeV.

Furthermore, our research examined an extended scalar model based on the FN mechanism, predicting the existence of a Flavon neutral scalar particle (H_F), decaying dominantly into a SM-like Higgs boson plus a pair of b -quarks. Two decay channels of the resulting Higgs boson, $h \rightarrow b\bar{b}$ and $h \rightarrow \gamma\gamma$, were analysed under realistic scenarios obtained from theoretical and experimental constraints. At the HL-LHC (targeting an integrated luminosity of 3000 fb^{-1}), we identified specific regions within the model parameter space, yielding signal significances up to 5σ for the $h \rightarrow b\bar{b}$ decay channel and about 2σ for the $h \rightarrow \gamma\gamma$ channel, thus reinforcing the potential for experimental verification of this theoretical scenario.

Finally, we directed our focus towards the LFV SM-like Higgs decay $h \rightarrow e\mu$, enabled by the presence of extended scalar sectors and associated LFV Yukawa couplings. The same analysis method allowed us to isolate scenarios which are accessible at the HL-LHC. In particular, one scenario (with parameters $\cos \alpha = -0.9548$, $v_s = 2461.078$ GeV, $Z_{e\mu} = 0.009781$ and a production cross section of 0.004404 pb) stood out as the most promising, achieving a clear 5σ discovery potential at a luminosity of 1300 fb^{-1} . Additionally, other scenario (with parameters $\cos \alpha = -0.8949$, $v_s = 2198.396$ GeV, $Z_{e\mu} = 0.005601$ and cross section of 0.003759 pb) also proved viable, reaching a 3σ evidence level at 700 fb^{-1} and a 5σ significance threshold at 1800 fb^{-1} . Hence, these scenarios significantly enrich the phenomenological landscape of flavour dynamics at the HL-LHC.

Throughout this research, we developed and utilised a robust Python based frame-

work capable of processing LHCO files, extracting relevant physical observables, and preparing datasets suitable for advanced Machine Learning techniques. By leveraging the powerful capabilities of the XGB algorithm, our methodology effectively distinguished signal from background events, providing enhanced signal sensitivity beyond standard kinematic methods.

In conclusion, the outcomes presented here not only highlight the phenomenological prospects offered by both BSM scenarios at the LHC and HL-LHC, but also showcase the powerful combination of Monte Carlo simulation and Machine Learning techniques. Future developments of our analysis package will include integration with other widely adopted particle physics tools, as well as direct incorporation of Machine Learning training functionalities, aiming to further expand its utility.

Appendix A

Constraints

By different constraints for each observable described in Sec. 3.2, we show individual allowed regions, namely,

- Signal strengths

We present in Fig. A.1 individual planes associated to each μ_X are $\cos(\alpha - \beta) - \tan \beta$. Those allowed by μ_X correspond to the coloured points. In the evaluation of the signal strengths μ_X , the range of scanned parameters is shown in Tab. A.1.

- Decays $B_{s,d}^0 \rightarrow \mu^+ \mu^-$

The allowed points by experimental measurements on $\mathcal{BR}(B_{s,d}^0 \rightarrow \mu^+ \mu^-)$ are shown in Fig. A.2. In Table A.2, the displayed range of scanned parameters is

Parameter	Range
$\tan \beta$	[0, 50]
$\cos(\alpha - \beta)$	[-1, 1]
χ_{bb}	[-1, 1]
χ_{tt}	[-100, 100]
$\chi_{\tau\tau}$	[-1, 1]
M_{H^\pm} (GeV)	[114, 140]
M_{A^\pm} (GeV)	$[m_t + m_b, 1000]$

Table A.1: Range of scanned parameters for the signal strengths μ_X .

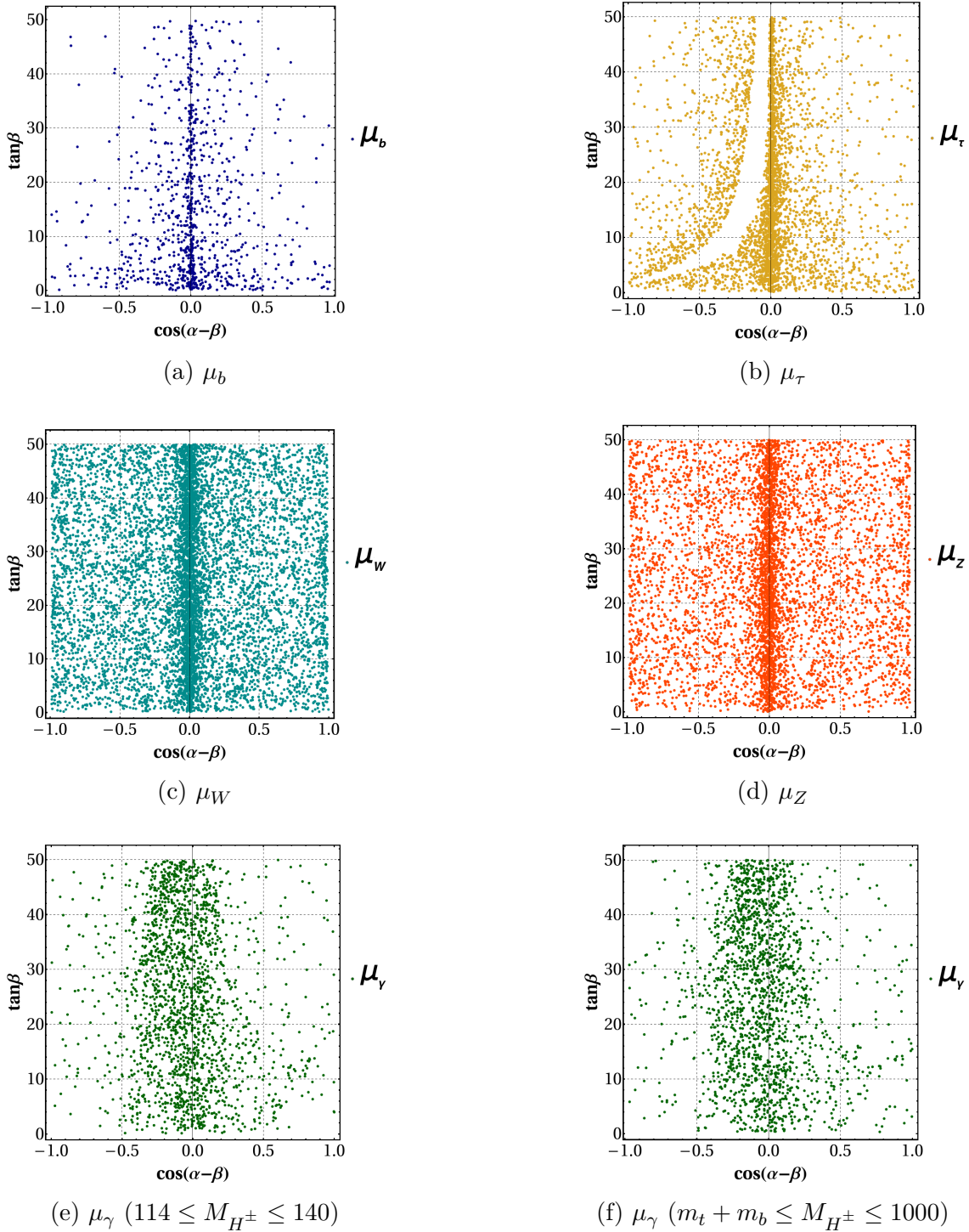


Figure A.1: Signal strengths: (a) μ_b , (b) μ_τ , (c) μ_W , (d) μ_Z , (e) μ_γ ($114 \leq M_{H^\pm} \leq 140$), and (f) μ_γ ($m_t + m_b \leq M_{H^\pm} \leq 1000$). In all cases, 50K random points are generated.

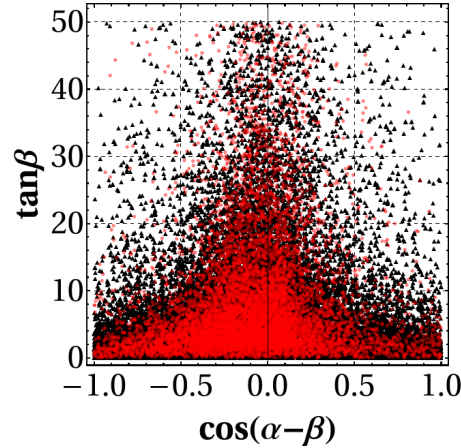


Figure A.2: Allowed values by the measurement on $\mathcal{BR}(B_d^0 \rightarrow \mu^+ \mu^-)$ (black triangles) and by the upper limit on $\mathcal{BR}(B_s^0 \rightarrow \mu^+ \mu^-)$ (light red circles). For both cases we generate 50K random points.

Parameter	Range	Parameter	Range
$\tan \beta$	$[0, 50]$	$\tan \beta$	$[0, 50]$
$\cos(\alpha - \beta)$	$[-1, 1]$	$\cos(\alpha - \beta)$	$[-1, 1]$
$\chi_{\mu\mu}$	$[-1, 1]$	$\chi_{\mu\mu}$	$[-1, 1]$
χ_{bd}	$[-1, 1]$	χ_{bs}	$[-1, 1]$
χ_{db}	$[-1, 1]$	χ_{sb}	$[-1, 1]$
M_A (GeV)	$[100, 1000]$	M_A (GeV)	$[100, 1000]$
M_H (GeV)	$[300, 1000]$	M_H (GeV)	$[300, 1000]$

Table A.2: Range of scanned parameters. On the left: $B_d^0 \rightarrow \mu^+ \mu^-$, on the right: $B_s^0 \rightarrow \mu^+ \mu^-$

- $\ell_i \rightarrow \ell_j \gamma$

The corresponding allowed points that meet upper limits on $\mathcal{BR}(\ell_i \rightarrow \ell_j \gamma)$ are presented in Fig. A.3. Displayed in Table A.3 is the range of scanned parameters.

- Decays $\ell_i \rightarrow \ell_j \ell_k \bar{\ell}_k$

We present in Fig. A.4 the points that comply with the constraints on $\mathcal{BR}(\ell_i \rightarrow \ell_j \ell_k \bar{\ell}_k)$, as far as the decays $\ell_i \rightarrow \ell_j \ell_k \bar{\ell}_k$ are concerned. The intervals of the scan on the parameters are presented in Table A.4 .

- Decays $h \rightarrow \ell_i \ell_j$

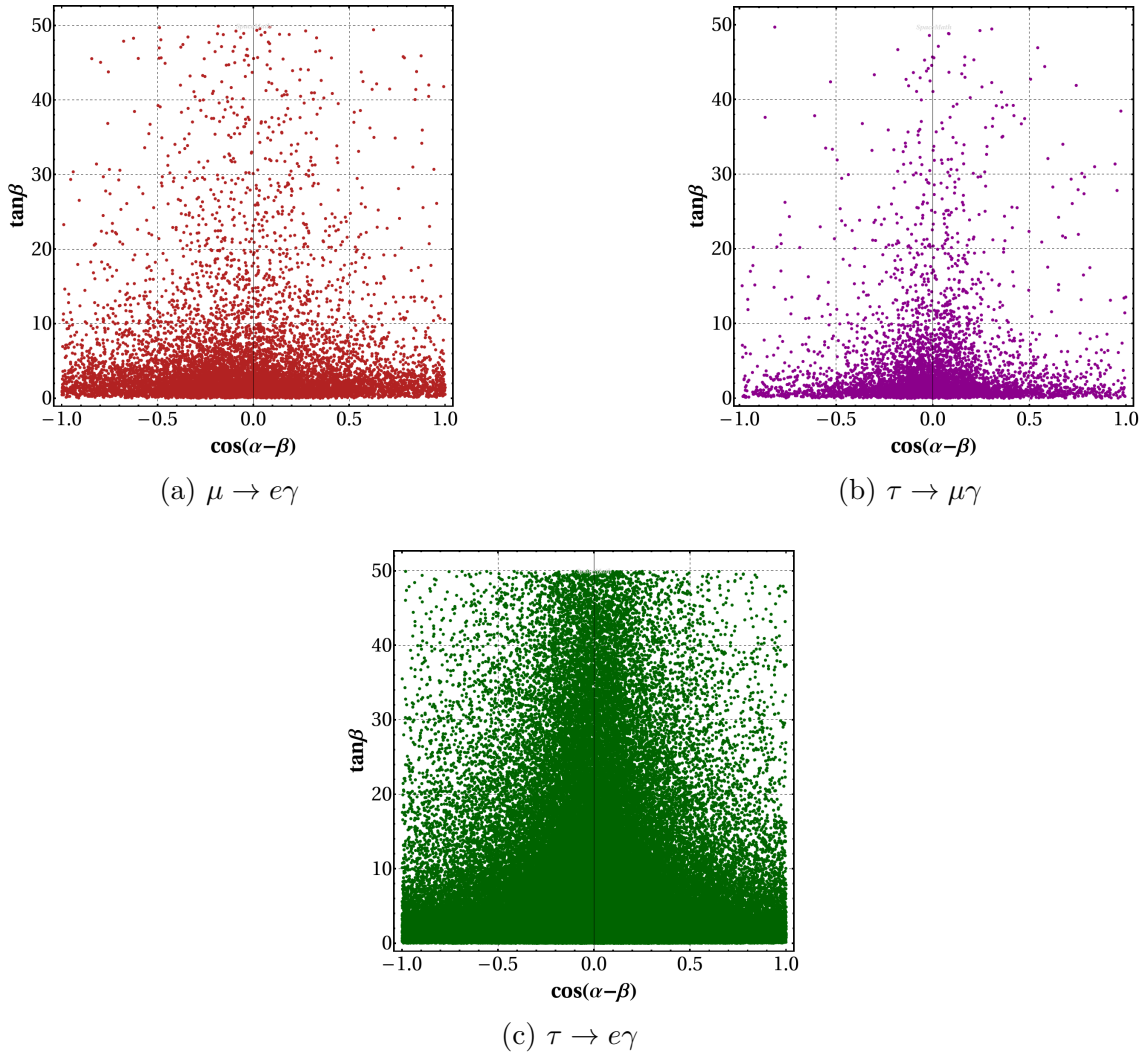


Figure A.3: Radiative decays $\ell_i \rightarrow \ell_j\gamma$: (a) $\mu \rightarrow e\gamma$, (b) $\tau \rightarrow \mu\gamma$, (c) $\tau \rightarrow e\gamma$. For all cases, we generate 100K random points.

Parameter	Range	Parameter	Range
$\tan \beta$	[0, 50]	$\tan \beta$	[0, 50]
$\cos(\alpha - \beta)$	[-1, 1]	$\cos(\alpha - \beta)$	[-1, 1]
$\chi_{\mu\mu}$	[-1, 1]	$\chi_{\tau\mu}$	[-100, 100]
$\chi_{\mu e}$	[-1, 1]	$\chi_{\tau\tau}$	[-1, 1]
χ_{tt}	[-100, 100]	χ_{tt}	[-100, 100]
M_A (GeV)	[100, 1000]	M_A (GeV)	[100, 1000]
M_H (GeV)	[300, 1000]	M_H (GeV)	[300, 1000]
M_{H^\pm} (GeV)	[110, 1000]	M_{H^\pm} (GeV)	[110, 1000]

Parameter	Range
$\tan \beta$	[0, 50]
$\cos(\alpha - \beta)$	[-1, 1]
$\chi_{\tau e}$	[-1, 1]
$\chi_{\tau\tau}$	[-1, 1]
χ_{tt}	[-100, 100]
M_A (GeV)	[100, 1000]
M_H (GeV)	[300, 1000]
M_{H^\pm} (GeV)	[110, 1000]

Table A.3: Range of scanned parameters. On the left: $\mu \rightarrow e\gamma$, on the right: $\tau \rightarrow \mu\gamma$, and below: $\tau \rightarrow e\gamma$

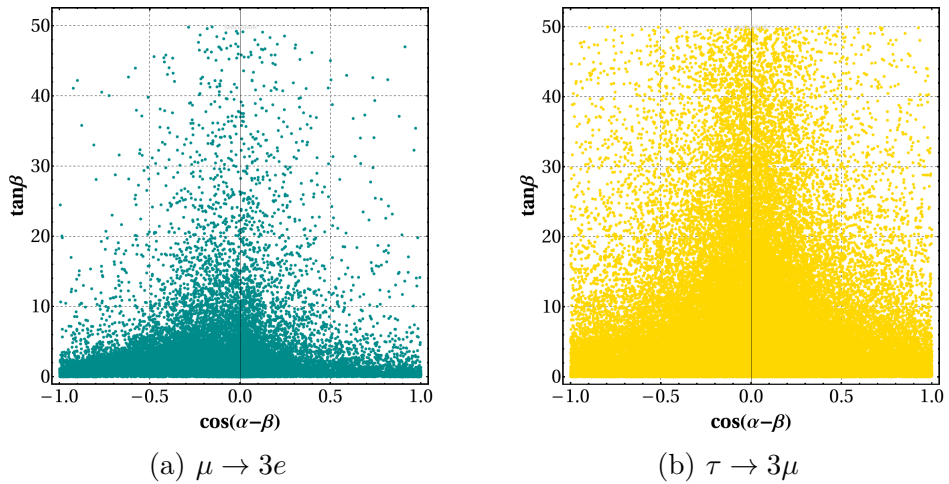


Figure A.4: Decays $\ell_i \rightarrow \ell_j \ell_k \bar{\ell}_k$: (a) $\mu \rightarrow 3e$, (b) $\tau \rightarrow 3\mu$. In all cases we generate 100K random points.

Only the allowed region coming from $h \rightarrow \tau\mu$ is presented because very weak bounds to $\tan \beta$ and $\cos(\alpha - \beta)$ are imposed by the process $h \rightarrow e\mu$. The scatter plot in the $\cos(\alpha - \beta) - \tan \beta$ plane is shown in Fig. A.5 shows the points allowed by the upper limit on $\mathcal{BR}(h \rightarrow \tau\mu)$.

Parameter	Range	Parameter	Range
$\tan \beta$	[0, 50]	$\tan \beta$	[0, 50]
$\cos(\alpha - \beta)$	[-1, 1]	$\cos(\alpha - \beta)$	[-1, 1]
χ_{ee}	[-1, 1]	$\chi_{\mu\mu}$	[-1, 1]
$\chi_{\mu e}$	[-1, 1]	$\chi_{\tau\mu}$	[-100, 100]
M_A (GeV)	[100, 1000]	M_A (GeV)	[100, 1000]
M_H (GeV)	[300, 1000]	M_H (GeV)	[300, 1000]

Table A.4: Range of scanned parameters. On the left: $\mu \rightarrow 3e$, on the right: $\tau \rightarrow 3\mu$.

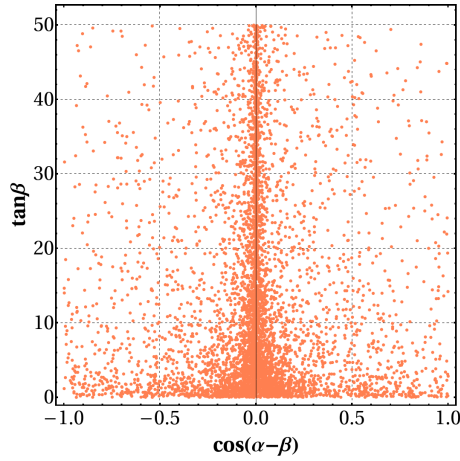


Figure A.5: Allowed values by the upper limit on $\mathcal{BR}(h \rightarrow \tau\mu)$. We generate 50K random points.

The range of scanned parameters is displayed in Table A.5.

Parameter	Range
$\tan \beta$	[0, 50]
$\cos(\alpha - \beta)$	[-1, 1]
$\chi_{\tau\mu}$	[-10, 10]

Table A.5: Range of scanned parameters associated to the decay $h \rightarrow \tau\mu$.

Appendix B

2HDM-III Variables and Parameters

In the Table , the variables used for training and testing the signal and background processes of the 2HDM-III framework are presented.

Rank	Variable	Description
1	$P_T[b]$	b -jet transverse momentum
2	\cancel{E}_T	Missing Energy Transverse
3	$P_T[j_1]$	Jet with the largest transverse momentum (c -jet)
4	$M_T[\mu]$	Transverse mass between the \cancel{E}_T and the muon.
5	$M_{inv}[b, c]$	Invariant mass of the b-jet and c-jet .
6	$\sum P_T$	Sum of the moduli of the transverse momentum of the b -jet, leading jet and the muon
7	η_b	Pseudorapidity of the b-jet
8	$P_T^{all}[jet]$	Scalar sum of the transverse momentum of all jets.
9	$\Delta\eta[j_1, j_2]$	Absolute value of the pseudorapidity separation between two jets
10	$\eta[j_1]$	Pseudorapidity of the leading jet
11	\cancel{E}_{tot}	Total transverse energy in the detector
12	$\eta[j_1] \times \eta[j_2]$	Product of the pseudorapidities of the two jets

Table B.1: Variables used to train and test the signal and background events list. The transverse mass is defined as

$$M_T[\ell] = \sqrt{2P_T^\ell E_T(1 - \cos \phi_{\ell \cancel{E}_T})}.$$

Oblique parameters

The S , T , U oblique parameters in the 2HDMs framework are given by

$$\begin{aligned}
S &= \frac{1}{\pi m_Z^2} \left\{ \sin^2(\alpha - \beta) \left[\mathcal{B}_{22}(m_Z^2; m_Z^2, m_h^2) - M_Z^2 \mathcal{B}_0(m_Z^2; m_Z^2, m_h^2) + \mathcal{B}_{22}(m_Z^2; M_H^2, M_A^2) \right] \right. \\
&+ \cos^2(\alpha - \beta) \left[\mathcal{B}_{22}(m_Z^2; m_Z^2, M_H^2) - m_Z^2 \mathcal{B}_0(m_Z^2; m_Z^2, M_H^2) + \mathcal{B}_{22}(m_Z^2; m_h^2, M_A^2) \right] \\
&\left. - \mathcal{B}_{22}(m_Z^2; M_{H^\pm}^2, M_{H^\pm}^2) - \mathcal{B}_{22}(m_Z^2; m_Z^2, m_h^2) + m_Z^2 \mathcal{B}_0(m_Z^2; m_Z^2, m_h^2) \right\},
\end{aligned} \tag{B.0.1}$$

$$\begin{aligned}
T &= \frac{1}{16\pi m_W^2 s_W^2} \left\{ \sin^2(\alpha - \beta) \left[\mathcal{F}(M_{H^\pm}^2, M_H^2) - \mathcal{F}(M_H^2, M_A^2) + 3\mathcal{F}(m_Z^2, m_h^2) - 3\mathcal{F}(m_W^2, m_h^2) \right] \right. \\
&+ \cos^2(\alpha - \beta) \left[\mathcal{F}(M_{H^\pm}^2, m_h^2) - \mathcal{F}(m_h^2, M_A^2) + 3\mathcal{F}(m_Z^2, M_H^2) - 3\mathcal{F}(m_W^2, M_H^2) \right] \\
&\left. + \mathcal{F}(M_{H^\pm}^2, M_A^2) - 3\mathcal{F}(m_Z^2, m_h^2) + 3\mathcal{F}(m_W^2, m_h^2) \right\},
\end{aligned} \tag{B.0.2}$$

$$\begin{aligned}
U &= \mathcal{H}(m_W^2) - \mathcal{H}(m_Z^2) \\
&+ \frac{1}{\pi m_W^2} \left\{ \cos^2(\alpha - \beta) \mathcal{B}_{22}(m_W^2; M_{H^\pm}^2, m_h^2) + \sin^2(\alpha - \beta) \mathcal{B}_{22}(m_W^2; M_{H^\pm}^2, M_H^2) \right. \\
&+ \mathcal{B}_{22}(m_W^2; M_{H^\pm}^2, M_A^2) - 2\mathcal{B}_{22}(m_W^2; M_{H^\pm}^2, M_{H^\pm}^2) \left. \right\} \\
&- \frac{1}{\pi m_Z^2} \left\{ \cos^2(\alpha - \beta) \mathcal{B}_{22}(m_Z^2; m_h^2, M_A^2) + \sin^2(\alpha - \beta) \mathcal{B}_{22}(m_Z^2; M_H^2, M_A^2) \right. \\
&\left. - \mathcal{B}_{22}(m_Z^2; M_{H^\pm}^2, M_{H^\pm}^2) \right\},
\end{aligned} \tag{B.0.3}$$

where

$$\begin{aligned}
\mathcal{H}(m_V^2) &\equiv \frac{1}{\pi m_V^2} \left\{ \sin^2(\alpha - \beta) \left[\mathcal{B}_{22}(m_V^2; m_V^2, m_h^2) - m_V^2 \mathcal{B}_0(m_V^2; m_V^2, m_h^2) \right] \right. \\
&+ \cos^2(\alpha - \beta) \left[\mathcal{B}_{22}(m_V^2; m_V^2, M_H^2) - m_V^2 \mathcal{B}_0(m_V^2; m_V^2, M_H^2) \right] \\
&\left. - \mathcal{B}_{22}(m_V^2; m_V^2, m_h^2) + m_V^2 \mathcal{B}_0(m_V^2; m_V^2, m_h^2) \right\},
\end{aligned} \tag{B.0.4}$$

$$\mathcal{F}(m_1^2, m_2^2) = \frac{1}{2} (m_1^2 + m_2^2) - \frac{m_1^2 m_2^2}{m_1^2 - m_2^2} \log \left(\frac{m_1^2}{m_2^2} \right), \tag{B.0.5}$$

$$\mathcal{B}_{22}(q^2; m_1^2, m_2^2) \equiv B_{22}(q^2; m_1^2, m_2^2) - B_{22}(0; m_1^2, m_2^2), \quad (\text{B.0.6})$$

$$\mathcal{B}_0(q^2; m_1^2, m_2^2) \equiv B_0(q^2; m_1^2, m_2^2) - B_0(0; m_1^2, m_2^2), \quad (\text{B.0.7})$$

$$\begin{aligned} B_{22}(q^2; m_1^2, m_2^2) &= \frac{1}{6} \left[A_0(m_2^2) + 2m_1^2 B_0(q^2; m_1^2, m_2^2) + (m_1^2 - m_2^2 + q^2) B_1(q^2; m_1^2, m_2^2) \right. \\ &\quad \left. - \frac{q^2}{3} + m_1^2 + m_2^2 \right], \end{aligned} \quad (\text{B.0.8})$$

in which A_0 , B_0 , B_1 are scalar Passarino-Veltman functions [118].

Bibliography

- [1] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*. CRC Press, FL, EE. UU., student ed., 2015.
- [2] M. Mitra, S. Saha, M. Spannowsky and M. Takeuchi, *Probing right-handed neutrinos via trilepton signals at the hl-lhc*, *Phys. Rev. D* **111** (Jan, 2025) 015005.
- [3] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, *Pisma Zh. Eksp. Teor. Fiz.* **5** (1967) 32–35.
- [4] G. White, *The sakharov conditions*, in *Electroweak Baryogenesis (Second Edition)*, 2053-2563, pp. 2–1 to 2–2. IOP Publishing, 2022. DOI.
- [5] M. Finkenberger, *Search for a dark leptophilic scalar produced in association with taupair in electron-positron annihilation at center-of-mass energies near 10.58 gev.*, .
- [6] J. Schwichtenberg, *Gauge coupling unification without supersymmetry*, *European Physical Journal C* **79** (Apr., 2019) 351.
- [7] C. D. Froggatt and H. B. Nielsen, *Hierarchy of Quark Masses, Cabibbo Angles and CP Violation*, *Nucl. Phys. B* **147** (1979) 277–298.
- [8] C. N. Yang and R. L. Mills, *Conservation of isotopic spin and isotopic gauge invariance*, *Phys. Rev.* **96** (Oct, 1954) 191–195.

-
- [9] J. H. Sánchez, *Procesos y Decaimientos de Bosones de Higgs a 1-Loop Para Futuros Colisionadores*, Ph.D. thesis, Universidad Autónoma de Puebla, Instituto de Física Luis Rivera Terrazas, 2001.
- [10] R. N. Papaqui, *Supersimetría y el Problema de la Masa*, Ph.D. thesis, Universidad Autónoma de Puebla, Instituto de Física Luis Rivera Terrazas, 2006.
- [11] J. F. Gunion, H. E. Haber, G. Kane and S. Dawson, *The Higgs Hunter's Guide*. Westview Press, 2000.
- [12] A. Pich, *The standard model of electroweak interactions*, 2012.
- [13] M. D. Schwartz, *Quantum Field Theory and the Standard Model*. Cambridge University Press, 3, 2014.
- [14] L.-L. Chau and W.-Y. Keung, *Comments on the parametrization of the kobayashi-maskawa matrix*, *Phys. Rev. Lett.* **53** (Nov, 1984) 1802–1805.
- [15] L. Wolfenstein, *Parametrization of the kobayashi-maskawa matrix*, *Phys. Rev. Lett.* **51** (Nov, 1983) 1945–1947.
- [16] T. D. Lee, *A theory of spontaneous t violation*, *Phys. Rev. D* **8** (Aug, 1973) 1226–1239.
- [17] E. Ma, *Verifiable radiative seesaw mechanism of neutrino mass and dark matter*, *Phys. Rev. D* **73** (Apr, 2006) 077301.
- [18] J. F. Gunion, H. E. Haber, G. L. Kane and S. Dawson, *The Higgs Hunter's Guide*, *Front. Phys.* **80** (2000) 1–404.
- [19] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, M. Sher and J. P. Silva, *Theory and phenomenology of two-Higgs-doublet models*, *Phys. Rept.* **516** (2012) 1–102, [1106.0034].

- [20] G. Branco, L. Lavoura and J. Silva, *CP Violation*. Oxford University Press, 07, 1999, 10.1093/oso/9780198503996.001.0001.
- [21] J. Hernández-Sánchez, S. Moretti, R. Noriega-Papaqui and A. Rosado, *Off-diagonal terms in yukawa textures of the type-iii 2-higgs doublet model and light charged higgs boson phenomenology*, *JHEP* **1307** (2013) 044, [1212.6818].
- [22] D. K. Ghosh, K. Mukherjee and S. Mukherjee, *Electroweak phase transition in two scalar singlet model with pngb dark matter*, *Journal of High Energy Physics* **2025** (Jan., 2025) .
- [23] A. Carrillo-Monteverde, R. Ramírez, S. Gómez-Ávila and L. López-Lozano, *Effective textures from a $[su(3)]^3$ flavored scalar sector*, .
- [24] A. Collaboration, *Search for a light charged higgs boson in $t \rightarrow h^\pm b$ decays, with $h^\pm \rightarrow cb$, in the lepton+jets final state in proton-proton collisions at $\sqrt{s} = 13$ tev with the atlas detector*, .
- [25] J. Kim, S. Lee, P. Sanyal, J. Song and D. Wang, *Unveiling the unexplored decay mode of a light charged higgs boson to an off-shell top quark and a bottom quark*, .
- [26] K. S. Babu, S. Chakdar, N. Das, D. K. Ghosh and P. Ghosh, *Fimp dark matter from flavon portals*, *Journal of High Energy Physics* **2023** (2023) 143.
- [27] S. Das, *Flavonstrahlung in the $b^3 - l^2 z'$ model at current and future colliders*, *Journal of High Energy Physics* **2023** (Mar., 2023) .
- [28] T.-Q. Li and C.-X. Yue, *Flavons and lfv decays and productions of pseudoscalar mesons*, *Modern Physics Letters A* **34** (Nov., 2019) 1950288.
- [29] A. Datta, *Theoretical perspective on flavour physics*, .

- [30] ATLAS collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett.* **B716** (2012) 1–29, [1207.7214].
- [31] G. Tonelli, CERN/INFN, U. of Pisa and C. Collaboration, *Update on the SM Higgs Search with CMS*, in *CMS*. 2011.
- [32] S. Weinberg, *The Quantum Theory of Fields*, vol. 2. Cambridge University Press, Cambridge, 1996.
- [33] J. Hernández-Sánchez, C. G. Honorato, S. Moretti and S. Rosado-Navarro, *Charged higgs boson production via cb -fusion at the large hadron collider*, *Phys. Rev. D* **102** (2020) 055008, [2003.06263].
- [34] M. A. Arroyo-Ureña, T. A. Valencia-Pérez, R. Gaitán, J. H. Montes De Oca and A. Fernández-Télez, *Flavor-changing decay $h \rightarrow \tau\mu$ at super hadron colliders*, *JHEP* **08** (2020) 170, [2002.04120].
- [35] M. A. Arroyo-Ureña, R. Gaitán, E. A. Herrera-Chacón, J. H. Montes de Oca Y. and T. A. Valencia-Pérez, *Search for the $t \rightarrow ch$ decay at hadron colliders*, *JHEP* **07** (2019) 041, [1903.02718].
- [36] M. A. Arroyo-Ureña, J. L. Diaz-Cruz, E. Díaz and J. A. Orduz-Ducuará, *Flavor violating higgs signals in the texturized two-higgs doublet model ($thdm-tx$)*, *Chin. Phys. C* **40** (2016) 123103, [1306.2343].
- [37] J. F. Gunion and H. E. Haber, *The CP conserving two Higgs doublet model: The Approach to the decoupling limit*, *Phys. Rev.* **D67** (2003) 075019, [hep-ph/0207010].
- [38] S. Weinberg, *Gauge theory of CP nonconservation*, *Phys. Rev. Lett.* **37** (Sep, 1976) 657–661.

- [39] S. P. Das, J. Hernández-Sánchez, S. Moretti, A. Rosado and R. Xoxocotzi, *Flavor violating signatures of lighter and heavier higgs bosons within the two higgs doublet model type iii at the lhec*, *Phys. Rev. D* **94** (Sep, 2016) 055003.
- [40] CMS collaboration, A. M. Sirunyan et al., *Search for lepton flavour violating decays of the Higgs boson to $\mu\tau$ and $e\tau$ in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **06** (2018) 001, [1712.07173].
- [41] ATLAS collaboration, G. Aad et al., *Searches for lepton-flavour-violating decays of the higgs boson in $\sqrt{s} = 13$ tev pp collisions with the atlas detector*, *Phys. Lett. B* **800** (2020) 135069, [1907.06131].
- [42] CMS collaboration, A. Tumasyan et al., *Measurement of the $b_S^0 \rightarrow \mu^+ \mu^-$ decay properties and search for the $b^0 \rightarrow \mu^+ \mu^-$ decay in proton-proton collisions at $\sqrt{s} = 13$ tev*, *Phys. Lett. B* **842** (2023) 137955, [2212.10311].
- [43] PARTICLE DATA GROUP collaboration, R. L. Workman et al., *Review of Particle Physics*, *PTEP* **2022** (2022) 083C01.
- [44] MUON G-2 collaboration, B. Abi et al., *Measurement of the positive muon anomalous magnetic moment to 0.46 ppm*, *Phys. Rev. Lett.* **126** (2021) 141801, [2104.03281].
- [45] ATLAS collaboration, G. Aad et al., *Search for Higgs boson pair production in the two bottom quarks plus two photons final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev. D* **106** (2022) 052001, [2112.11876].
- [46] A. Collaboration, *Cern rpt. atlas-conf-2017-055*, tech. rep., CERN, 2017.
- [47] CMS collaboration, A. M. Sirunyan et al., *Search for additional neutral mssm higgs bosons in the $\tau\tau$ final state in proton-proton collisions at $\sqrt{s} = 13$ tev*, *JHEP* **1809** (2018) 007, [1803.06553].

- [48] ATLAS collaboration, M. Aaboud et al., *Search for charged higgs bosons decaying via $h^\pm \rightarrow \tau^\pm \nu_\tau$ in the τ +jets and τ +lepton final states with 36 fb^{-1} of pp collision data recorded at $\sqrt{s} = 13 \text{ tev}$ with the atlas experiment*, *JHEP* **09** (2018) 139, [1807.07915].
- [49] ATLAS collaboration, G. Aad et al., *Search for a light charged higgs boson in $t \rightarrow h^\pm b$ decays, with $h^\pm \rightarrow cb$, in the lepton+jets final state in proton-proton collisions at $\sqrt{s} = 13 \text{ tev}$ with the atlas detector*, *JHEP* **09** (2023) 004, [2302.11739].
- [50] M. Ciuchini, G. Degrassi, P. Gambino and G. F. Giudice, *Next-to-leading qcd corrections to $b \rightarrow x_s \gamma$: Standard model and two higgs doublet model*, *Nucl. Phys. B* **527** (1998) 21, [hep-ph/9710335].
- [51] M. Misiak and M. Steinhauser, *Weak radiative decays of the b meson and bounds on m_{H^\pm} in the two-higgs-doublet model*, *Eur. Phys. J. C* **77** (2017) 201, [1702.04571].
- [52] M. D. Goodsell and F. Staub, *Improved unitarity constraints in two-higgs-doublet-models*, *Physics Letters B* **788** (2019) 206–212.
- [53] T. A. collaboration, *Measurements of the higgs boson production and decay rates and constraints on its couplings from a combined atlas and cms analysis of the lhc pp collision data at $\sqrt{s} = 7$ and 8 tev* , *Journal of High Energy Physics* **2016** (2016) 45.
- [54] ATLAS Collaboration, *Combined measurements of Higgs boson production and decay using up to 139 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13 \text{ TeV}$ collected with the ATLAS experiment*, tech. rep., CERN, Geneva, 2021.
- [55] CMS Collaboration, *Combined Higgs boson production and decay measurements with up to 137 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13 \text{ TeV}$* , tech. rep., CERN, Geneva, 2020.

- [56] M. Beneke, C. Bobeth and R. Szafron, *Power-enhanced leading-logarithmic qcd corrections to $b_q \rightarrow \mu^+ \mu^-$* , *JHEP* **10** (2019) 232, [1908.07011].
- [57] A. Tumasyan, W. Adam, J. Andrejkovic and T. Bergauer, *Measurement of the $bs0 \rightarrow +$ decay properties and search for the $b0 \rightarrow +$ decay in proton-proton collisions at $s=13\text{tev}$* , *Physics Letters B* **842** (2023) 137955.
- [58] M. Bauer, T. Schell and T. Plehn, *Hunting the flavon*, *Physical Review D* **94** (Sept., 2016) .
- [59] A. J. Buras, J. Girrbach, D. Guadagnoli and G. Isidori, *On the standard model prediction for $br(bs,d \text{ to } \mu^+ \mu^-)$* , *Eur. Phys. J. C* **72** (2012) 2172, [1208.0934].
- [60] T. Inami and C. S. Lim, *Effects of superheavy quarks and leptons in low-energy weak processes $k(l) \rightarrow \mu \text{ anti-}\mu$, $k^+ \rightarrow \pi^+ \text{ neutrino anti-neutrino}$ and $k0 \leftrightarrow \text{anti-}k0$* , *Prog. Theor. Phys.* **65** (1981) 297.
- [61] V. Ilisie, *New barr-zee contributions to $(g2)_\mu$ in two-higgs-doublet models*, *Journal of High Energy Physics* **2015** (2015) 77.
- [62] R. Harnik, J. Kopp and J. Zupan, *Flavor Violating Higgs Decays*, *JHEP* **03** (2013) 026, [1209.1397].
- [63] G. Blankenburg, J. Ellis and G. Isidori, *Flavour-changing decays of a 125 gev higgs-like particle*, *Phys. Lett. B* **712** (2012) 386–390, [1202.5704].
- [64] M. A. Arroyo-Ureña, R. Gaitán and T. A. Valencia-Pérez, *SpaceMath version 1.0 A Mathematica package for beyond the standard model parameter space searches*, *Rev. Mex. Fis. E* **19** (2022) 020206, [2008.00564].
- [65] ATLAS collaboration, G. Aad et al., *Search for heavy resonances decaying into a pair of Z bosons in the $\ell^+ \ell^- \ell'^+ \ell'^-$ and $\ell^+ \ell^- \nu \bar{\nu}$ final states using 139 fb⁻¹ of proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector*, *Eur. Phys. J. C* **81** (2021) 332, [2009.14791].

- [66] CMS collaboration, A. M. Sirunyan et al., *Search for a heavy Higgs boson decaying to a pair of W bosons in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **03** (2020) 034, [1912.01594].
- [67] ATLAS collaboration, G. Aad et al., *Search for $t\bar{t}H/A \rightarrow t\bar{t}t$ production in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, 2408.17164.
- [68] A. Crivellin, C. Greub and A. Kokulu, *Flavor-phenomenology of two-higgs-doublet models with generic yukawa structure*, *Phys. Rev. D* **87** (May, 2013) 094031.
- [69] F. M. Borzumati and C. Greub, *Two higgs doublet model predictions for $\bar{B} \rightarrow X_s \gamma$ in nlo qcd*, *Phys. Rev. D* **58** (Sep, 1998) 074004.
- [70] PARTICLE DATA GROUP collaboration, S. Navas et al., *Review of particle physics*, *Phys. Rev. D* **110** (2024) 030001.
- [71] W. Grimus, L. Lavoura, O. Ogreid and P. Osland, *The oblique parameters in multi-higgs-doublet models*, *Nuclear Physics B* **801** (2008) 81–96.
- [72] ATLAS collaboration, G. Aad et al., *Search for nonresonant pair production of Higgs bosons in the $bb^{-}bb^{-}$ final state in pp collisions at $s=13$ TeV with the ATLAS detector*, *Phys. Rev. D* **108** (2023) 052003, [2301.03212].
- [73] ATLAS collaboration, G. Aad et al., *Search for pair production of boosted Higgs bosons via vector-boson fusion in the $b\bar{b}b\bar{b}$ final state using pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, 2404.17193.
- [74] ATLAS collaboration, G. Aad et al., *Studies of new Higgs boson interactions through nonresonant HH production in the $b\bar{b}\gamma\gamma$ final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **01** (2024) 066, [2310.12301].
- [75] ATLAS collaboration, G. Aad et al., *Search for the non-resonant production of Higgs boson pairs via gluon fusion and vector-boson fusion in the $b\bar{b}\tau^{+}\tau^{-}$*

- final state in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, 2404.12660.
- [76] ATLAS collaboration, G. Aad et al., *Search for non-resonant Higgs boson pair production in the $2b + 2\ell + E_T^{miss}$ final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **02** (2024) 037, [2310.11286].
- [77] N. Khan and S. Rakshit, *Study of electroweak vacuum metastability with a singlet scalar dark matter*, *Phys. Rev. D* **90** (2014) 113008, [1407.6015].
- [78] Arroyo-Ureña, Marco A. and Chakraborty, Amit and Díaz-Cruz, J. Lorenzo and Ghosh, Dilip Kumar and Khan, Najimuddin and Moretti, Stefano, *Flavon signatures at the HL-LHC*, *Phys. Rev. D* **108** (2023) 095026, [2205.12641].
- [79] G. Cynolter, E. Lendvai and G. Pocsik, *Note on unitarity constraints in a model for a singlet scalar dark matter candidate*, *Acta Phys. Polon. B* **36** (2005) 827–832, [hep-ph/0410102].
- [80] T. Hastie, R. Tibshirani and J. Friedman, *The Elements of Statistical Learning*. Springer, 2009, 10.1007/978-0-387-84858-7.
- [81] Y. Coadou, *Boosted Decision Trees*, ch. Chapter 2, pp. 9–58. 2022. 2206.09645. 10.1142/9789811234033_002.
- [82] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer and T. Stelzer, *Madgraph 5: going beyond*, *JHEP* **06** (2011) 128.
- [83] G. Kasieczka, B. Nachman, D. Shih, O. Amram, A. Andreassen, K. Benkendorfer et al., *The lhc olympics 2020 a community challenge for anomaly detection in high energy physics*, *Reports on Progress in Physics* **84** (dec, 2021) 124201.
- [84] “Lhco manual.” <http://madgraph.phys.ucl.ac.be/Manual/lhco.html>, n.d.
- [85] A. Fowlie, *Lhco_reader: A new code for reading and analyzing detector-level events stored in lhco format*, *arXiv: High Energy Physics - Phenomenology* (2015) .

- [86] P. S. Foundation, “Python Language Reference, version 3.”
<https://www.python.org/>, 2025.
- [87] E. A. Herrera-Chacón and M. A. Arroyo-Ureña, “Prakritimlprep.”
https://gitlab.com/ed_in/prakritimlprep, 2025.
- [88] O. D. Education, *Calculating invariant mass*, 2025.
- [89] Leo Breiman, Jerome Friedman, R. A. Olshen and Charles J. Stone, *Classification and Regression Trees*. Chapman and Hall/CRC, New York, 1st ed., 1984,
10.1201/9781315139470.
- [90] T. Chen and C. Guestrin, *Xgboost: A scalable tree boosting system*, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, (New York, NY, USA), p. 785–794, Association for Computing Machinery, 2016, DOI.
- [91] *Higgs boson machine learning challenge*, 2014.
- [92] T. Fawcett, *An introduction to roc analysis*, *Pattern Recognition Letters* **27** (2006) 861–874.
- [93] S. M. Ross, *Introduction to Probability and Statistics for Engineers and Scientists*. Academic Press, 5 ed., 2014.
- [94] Y. Coadou, *Boosted decision trees*, in *Artificial Intelligence for High Energy Physics* (P. Calafiura, D. Rousseau and K. Terao, eds.). World Scientific Publishing, 2022. 2206.09645. DOI.
- [95] I. Roma, “Logic of an epp experiment.” Lecture slides, Experimental Elementary Particle Physics, 2017.
- [96] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr and B. Fuks, *FeynRules 2.0 - A complete toolbox for tree-level phenomenology*, *Comput. Phys. Commun.* **185** (2014) 2250–2300, [1310.1921].

- [97] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. Del Debbio, S. Forte et al., *Parton distributions with lhc data*, *Nuclear Physics B* **867** (2013) 244–289.
- [98] T. Sjostrand, *Pythia 8 status report*, tech. rep., DESY, 2009.
10.3204/DESY-PROC-2009-02/41.
- [99] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaitre, A. Mertens et al., *Delphes 3, a modular framework for fast simulation of a generic collider experiment*, *JHEP* **02** (2014) 057.
- [100] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, *Eur. Phys. J. C* **72** (2012) 1896, [1111.6097].
- [101] M. Cacciari, G. P. Salam and G. Soyez, *The anti- k_t jet clustering algorithm*, *JHEP* **04** (2008) 063, [0802.1189].
- [102] K. Woodruff, *Bdt tutorial*, tech. rep., GitHub repository, 2018.
- [103] E. Conte, B. Fuks and G. Serret, *MadAnalysis 5, A User-Friendly Framework for Collider Phenomenology*, *Comput. Phys. Commun.* **184** (2013) 222–256,
[1206.1599].
- [104] C. P. Burgess, *Introduction to Effective Field Theory: Thinking Effectively about Hierarchies of Scale*. Cambridge University Press, 2020.
- [105] T. Plehn, *Lectures on LHC Physics*, *Lect. Notes Phys.* **844** (2012) 1–193,
[0910.4182].
- [106] A. Pak, M. Rogal and M. Steinhauser, *Production of scalar and pseudo-scalar Higgs bosons to next-to-next-to-leading order at hadron colliders*, *JHEP* **09** (2011) 088,
[1107.3391].
- [107] LHC HIGGS CROSS SECTION WORKING GROUP collaboration, D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, 1610.07922.

- [108] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079, [1405.0301].
- [109] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159–177, [1410.3012].
- [110] DELPHES 3 collaboration, J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens et al., *DELPHES 3, A modular framework for fast simulation of a generic collider experiment*, *JHEP* **02** (2014) 057, [1307.6346].
- [111] ATLAS collaboration, G. Aad et al., *Fast b-tagging at the high-level trigger of the ATLAS experiment in LHC Run 3*, *JINST* **18** (2023) P11006, [2306.09738].
- [112] J. Bergstra, D. Yamins and D. D. Cox, *Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures*, in *Proceedings of the 30th International Conference on Machine Learning (ICML 2013)*, pp. I–115–I–23, June, 2013, <http://proceedings.mlr.press/v28/bergstra13.pdf>.
- [113] J. Bergstra, D. Yamins and D. Cox, *Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures*, in *Proceedings of the 30th International Conference on Machine Learning* (S. Dasgupta and D. McAllester, eds.), vol. 28 of *Proceedings of Machine Learning Research*, (Atlanta, Georgia, USA), pp. 115–123, PMLR, 17–19 Jun, 2013, <https://proceedings.mlr.press/v28/bergstra13.html>.
- [114] N. Arkani-Hamed, T. Han, M. Mangano and L.-T. Wang, *Physics opportunities of a 100 tev proton–proton collider*, *Physics Reports* **652** (2016) 1–49.
- [115] M. Selvaggi, *Beta card for HL-LHC and HE-LHC studies*, 2017.

-
- [116] T. Akiba, S. Sano, T. Yanase, T. Ohta and M. Koyama, *Optuna: A next-generation hyperparameter optimization framework*, in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '19, (New York, NY, USA), p. 2623–2631, Association for Computing Machinery, 2019, DOI.
- [117] G. Aad, B. Abbott, D. Abbott, A. Abed Abud and e. a. Abeling, *Search for the higgs boson decays $h \rightarrow ee$ and $h \rightarrow e\mu$ in pp collisions at $\sqrt{s} = 13$ tev with the atlas detector*, *Physics Letters B* **801** (Feb., 2020) 135148.
- [118] G. Passarino and M. Veltman, *One-loop corrections for $e+e$ annihilation into $+$ in the weinberg model*, *Nuclear Physics B* **160** (1979) 151–207.