Precision search for the new physics with hadronic final state at the LHC

A thesis submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

by

Anupam Ghosh

(Roll No. 18330003)

Under the guidance of

Prof. Partha Konar (Supervisor)

and

Dr. Satyajit Seth (Co-Supervisor)

Theoretical Physics Division
Physical Research Laboratory, Ahmedabad, India.

DEPARTMENT OF PHYSICS

INDIAN INSTITUTE OF TECHNOLOGY GANDHINAGAR

2023

Dedicated to
my family & teachers

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Anupam Ghosh
(Roll No: 18330003)

Date: 14/07/2023

CERTIFICATE

It is certified that the work contained in the thesis titled **"Precision search for the new physics with hadronic final state at the LHC"** by **Anupam Ghosh** (Roll No: 18330003), has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Prof. Partha Konar
(Thesis Supervisor)
Theoretical Physics Division,
Physical Research Laboratory,
Ahmedabad, India.

Date: 14/07/2023

Thesis Approval

The thesis titled

"Precision search for the new physics with hadronic final state at the LHC"

by

Anupam Ghosh
(Roll No. 18330003)

is approved for the degree of

Doctor of Philosophy

Examiner	Examiner
Co-Supervisor	Supervisor
Chairman	

Date: ____________________

Place: ____________________

List of Publications

Publications included in the thesis

1. *Precise probing of the inert Higgs-doublet model at the LHC,*
   Anupam Ghosh, Partha Konar, Satyajit Seth
   PHYSICAL REVIEW D 105, 115038 (2022) [arXiv:2111.15236]

2. *Top-philic dark matter in a hybrid KSVZ axion framework,*
   Anupam Ghosh, Partha Konar, Rishav Roshan
   JHEP 12 (2022) 167 [arXiv:2207.00487]

3. *Precise probing and discrimination of third-generation scalar leptoquarks,*
   Anupam Ghosh, Partha Konar, Debashis Saha, Satyajit Seth
   PHYSICAL REVIEW D 108, 035030 (2023) [arXiv:2304.02890]

4. *Precision prediction at the LHC of a democratic up-family philic KSVZ axion model*, **Anupam Ghosh**, Partha Konar
   [arXiv:2305.08662] (Manuscript under journal review)

Acknowledgements

I am very happy to thank my Ph.D. supervisor, Prof. Partha Konar, for his invaluable guidance throughout this journey. He always motivated me to learn something new and think out of the box, which immensely helped me in my research career. His positive attitude about everything also inspired me. His never-ending support and inspiration helped me a lot in this journey.

The work completed in this thesis would not have been possible without my co-supervisor Dr. Satyajit Seth. I sincerely thank him for guiding me and helping me whenever I faced difficulties. He not only taught me the QCD and the loop corrections, but his continuous support and fruitful discussions helped me to complete this journey.

My thesis would not have been achievable without my collaborators. I express my gratitude to my collaborators, namely Dr. Debashis Saha and Dr. Rishav Roshan. We jointly did some exciting projects. Our insightful discussions have significantly influenced my perspective on high-energy physics research. I would like to acknowledge their significant role in shaping my research journey.

I am immensely grateful for the valuable time I spent with Dr. J. Selvaganapathy during the early stage of my Ph.D. and the numerous insightful physics discussions we shared. The knowledge and insights gained from those interactions have been crucial to my research journey.

I want to seize this opportunity to express my sincere gratitude to my Doctoral Studies Committee, consisting of Prof. Srubabati Goswami and Dr. Narendra Ojha, for consistently evaluating my work throughout my doctoral studies. Their invaluable comments and feedback have played a pivotal role in enhancing the scientific outcomes of my research. I am also grateful to them for reviewing this thesis.

I extend my sincere gratitude to Dr. Ketan Patel, Prof. Namit Mahajan, Prof. Srubabati Goswami, Prof. Dilip Angom, Prof. Navinder Singh, and Prof. Jitesh R. Bhatt for their invaluable contributions in teaching various courses on high-energy physics, statistical mechanics, condensed matter physics, and cosmology. Their expertise and dedication have greatly enriched my understanding in these areas during and after my coursework.

I thank my senior, Dr. Akanksha Bhardwaj, for engaging in numerous fruitful discussions. I would also like to extend my thanks to my office mates, Dr. Debashis Saha and Dr. Sourav Pal. I would also like to thank my seniors, Dr. Vishal Singh Ngairangbam, Dr. Sudipta Show, and junior Deepanshu Srivastava. Their presence and camaraderie have made the workplace a joyful and supportive environment.

I am grateful to the Indian Institute of Technology, Gandhinagar (IITGN), for providing me with the opportunity to register for my Ph.D. program.

The last five years would not have been exciting without the joyful companionship of my batch mates. I am very thankful to Dayanand Mishra, Deepak Gaur, Supriya Pan, Devaprasad, and Yogesh, who helped me during my tough time.

I extend my heartfelt gratitude to the entire PRL (Physical Research Laboratory) family for their invaluable support in granting me access to research facilities and providing a much-needed Quartar facility during my time there.

I am fortunate to have Karabi Majhi by my side, as her unwavering support has been invaluable in helping me navigate through difficult situations. I express my heartfelt gratitude to her for being a constant source of inspiration and encouragement, urging me to strive harder and achieve my life goals.

I thank my mother, Aparna Ghosh, my father, Mr. Panchanan Ghosh, and my sister Puja Bali (Ghosh), for their constant support, encouragement, love, and care throughout my life. Completing this doctoral journey brings me immense joy, and I am certain that my mother will be even happier upon hearing the news. She always teaches me to help others and treat everyone with respect and inspires me to learn daily. Her unwavering support and belief in my abilities give me a lot of confidence during this journey.

(Anupam)

Abstract

The Standard Model (SM) has successfully described the fundamental particles and their interactions but fails to explain various natural phenomena such as dark matter, neutrino masses, the strong charge-parity (CP) problem, and matter-antimatter asymmetry. This thesis investigates Beyond Standard Model (BSM) theories, including the inert doublet model (IDM), complex scalar extended Kim-Shifman-Vainshtein-Zakharov (KSVZ) model, and exotic scalar leptoquark model. These new physics models provide potential solutions to some of the above mentioned problems, and we will explore them in the context of the Large Hadron Collider (LHC).

The LHC is a crucial experimental frontier for studying the Higgs boson and searching for signatures of BSM physics. However, the absence of prominent new physics signals at the LHC suggests that the effects of BSM phenomena may require precision studies. This thesis addresses this challenge by incorporating next-to-leading order (NLO) QCD corrections for BSM particle production processes at the 14 TeV LHC. Considering NLO-QCD corrections and including parton showering, the accuracy of cross-section estimates and reliability of differential distributions across the entire phase space are significantly improved. The presence of additional radiation at NLO-QCD alters the differential distribution of observables compared to leading-order (LO) predictions. Furthermore, the inclusion of higher-order corrections reduces theoretical uncertainties associated with renormalization and factorization scales, leading to more precise predictions of BSM signatures.

Many BSM theories predict heavy resonances that decay into particles like the W boson, Z boson, Higgs boson, or top quark. Investigating hadronic final states is crucial due to their large branching fraction. The discovery potential of TeV scale BSM particles is enhanced by incorporating boosted fatjets into the analysis. Challenges arise from QCD jets mimicking fatjets and the overwhelming SM background. To overcome these challenges, recent techniques, like jet substructure variables, are also employed in this thesis to analyze the internal structure and properties of jets, enabling the identification and characterization of underlying physics processes.

In contrast to traditional cut-based analyses, a sophisticated multivariate analysis (MVA) approach is adopted in this thesis. By combining multiple observables and constructing non-linear decision boundaries, the MVA approach enhances the efficiency of extracting the signal from the background, providing a more powerful tool for studying physics at the LHC.

The thesis presents detailed investigations into specific BSM models. Firstly, the Inert Doublet Model, which offers a viable Higgs-portal dark matter candidate, is explored. The challenging hierarchical mass spectrum of the IDM, featuring a light dark matter particle coexisting with heavier scalar states, is studied. NLO-QCD corrections are considered for pair and associated production of BSM particles, leading to substantial corrections in total cross sections and differential distributions compared to LO estimates. The heavier scalar particles decay into dark matter, accompanied by a boosted W or Z boson due to their substantial mass difference. A multivariate analysis of the di-fatjets plus missing transverse momentum (MET) signal is performed, unveiling the discovery potential of various parameter spaces within the hierarchical mass spectrum at the High-Luminosity 14 TeV LHC (HL-LHC).

The KSVZ model, which solves the strong CP problem, is extended in subsequent studies by including a complex scalar singlet (S). This two-component dark matter model provides the correct relic density without fine-tuning model parameters. The colored vector-like quark (VLQ) in the KSVZ model significantly impacts the relic density calculation by opening up new annihilation and co-annihilation channels. Additionally, it introduces new direct detection diagrams. The Yukawa interaction between the BSM scalar, VLQ, and up-type SM quarks is $f_i S\overline{\Psi}_L u_{iR} + \ h.c\ (i = u,\ c,\ t)$, which plays a significant role in dark matter and collider phenomenology. The initial study of the extended KSVZ model emphasizes the large coupling f_t while f_u and f_c are relatively small. Following pair production, each VLQ decays into the top quark associated with the scalar. The analysis involves examining two boosted top-like fatjets with large MET using multivariate analysis techniques.

In the next study of the extended KSVZ model, equal coupling strengths (democratic) are considered, but flavor constraints require extremely small values for one or both of the couplings, f_u and f_c. This leads to unique parameter spaces that satisfy relic density and other constraints. The study includes an analysis of NLO-QCD corrections for VLQ pair production. It is observed that the total NLO cross section increases by approximately 30% compared to LO estimates, and the differential distributions exhibit significant changes. Considering the $\mathcal{O}(\alpha_S)$ correction for VLQ pair production, a multivariate analysis of two top-like fatjets plus MET signal is conducted. The discovery potential for significant parameter spaces at the 14 TeV LHC, with an integrated luminosity of 139 fb^{-1}, is determined.

We also study the third-generation scalar leptoquark, an intriguing BSM particle predicted by numerous theories. Leptoquarks interact with quarks and leptons and can explain observed anomalies like W-mass, muon g-2, and $R_{D^{(*)}}$. Recent

ATLAS analysis excludes a third-generation scalar leptoquark with a mass of up to 1240 GeV. Consequently, the decay of the leptoquark produces a highly boosted top quark accompanied by a neutrino. The investigation focuses on probing the leptoquark at the 14 TeV LHC using top-like fatjets and MET signatures. The impact of NLO-QCD corrections on leptoquark pair production is significant and taken into account. The multivariate analysis is employed to assess the discovery potential of the leptoquark signal. Furthermore, polarization variables sensitive to top quark polarization are utilized to differentiate between different leptoquark models.

The analysis presented in this thesis offers a generic framework that can be applied to various BSM models and within the context of the SM. It enhances our understanding of BSM physics and offers valuable insights into the discovery potential of new physics phenomena at the LHC. The combination of theoretical investigations, precision studies, and sophisticated analysis techniques strengthens our quest to unravel the mysteries beyond the Standard Model.

Contents

List of Abbreviations

ATLAS	A Toroidal LHC Apparatus
BDT	Boosted Decision Tree
BP	Benchmark Point
BR	Branching Ratio
BSM	Beyond the Standard Model
CA	Cambridge-Achen
CM	Centre-of-Mass
CMS	Compact Muon Solenoid
CP	Charge conjugation Parity
DD	Direct detection
DGLAP	Dokshitzer–Gribov–Lipatov–Altarelli–Parisi equations
DM	Dark matter
GeV	Giga electron Volt
HEFT	Higgs Effective Field Theory
ID	Indirect detection
IDM	Inert Higgs doublet model
IR	InfraRed
KLN	Kinoshita–Lee–Nauenberg
KSVZ	Kim-Shifman-Vainshtein-Zakharov model
LEP	Large Electron-Positron
LHC	Large Hadron Collider
LO	Leading Order
LQ	Leptoquark
MET or $\not{E}_T$	Missing transverse Energy
MLM	Michelangelo L. Mangano
MVA	Multivariate Analysis
NLO	Next-to Leading Order
NNLO	Next-to-Next-to-Leading order
PDF	Parton Distribution Function
QCD	Quantum Chromodynamics
SI	Spin-independent
SM	Standard Model
TeV	Tera electron Volt
UFO	Universal FeynRules Output
UV	Ultraviolet
vev	Vacuum Expectation Value
WIMP	Weakly Interacting Massive Particle

Chapter 1

Introduction

Modern-day high-energy research community puts massive collaborative efforts both from theoretical and experimental sides to discover the mysteries of the Universe. The Large Hadron Collider (LHC) is one such endeavor that tries to detect the fundamental particles of nature and their interactions. It investigates the physics in sub-nuclear length scales, and at such length scales, nature follows the quantum field theoretical description. The Standard Model (SM) [1–10] is widely regarded as the most successful theory in particle physics, providing explanations for various phenomena. A brief overview of the SM is given in Section 1.1.

Despite the remarkable success of SM, many aspects of nature are still entirely missing or unsatisfactory in SM. For instance, experiments on neutrino oscillation have revealed that at least two active neutrinos possess mass, whereas neutrinos are massless in the SM. Other unresolved issues include the matter-antimatter asymmetry, the hierarchy problem, and the inability to account for dark matter and dark energy. These flaws are discussed in detail in Section 1.2, and they imply the existence of a more comprehensive theory. To address these shortcomings, one can broaden the SM by introducing additional particles and/or new interactions and enlarging its gauge group to mitigate some of these limitations.

The analyses of the LHC events are somewhat complicated because of the presence of colored quarks and gluons (collectively called partons). The gluons have self-interactions because of the non-abelian nature of the $SU(3)_C$ group and interact with quarks since the quarks have color charges. As a result, partons produced at high energy will emit more partons that share the energy of the mother parton. The colored partons can not be directly observed at the detectors because of the color confinement of the QCD. As a result, partons after fragmentation form colorless hadrons recorded in the various parts of the detectors. Different reconstruction techniques are used to map those thousands of hadrons in terms of a few well-defined reconstructed jets along with other objects like photons and leptons.

Those reconstructed objects are then used to correctly identify the partonic kinematics of the events that link the experimental observation and the theoretical prediction. We detail the LHC and the event reconstruction in Section 1.3.

To test the SM with high precision and to recognize any deviations of the LHC data from the SM, it is essential to have theoretical predictions of the observables with the least uncertainty. Leading order (LO) predictions have large theoretical uncertainty. The inclusion of next-to-leading order (NLO) or even higher-order computation in this perturbative series expansion ensures fewer theoretical scale uncertainties. This thesis includes the NLO QCD correction of the partonic cross section at the LHC. The corrections coming from one-loop QCD have significant effects in describing the physics phenomena at the LHC. The differential NLO cross section is also necessary for accurate prediction in addition to the total NLO cross section. This is because when we perform signal-to-background analysis to isolate the tiny signal from the extensive background, we typically apply suitable selection criteria on many relevant variables. Thus the accurate prediction of differential distributions in those phase space regions is essential. At higher order, some new production channels can also open up, which are crucial for accurate prediction. We outlined, in brief, the importance of including NLO QCD corrections for the LHC phenomenology in Section 1.4. We will provide a brief outline of the thesis in Section 1.5.

1.1 Standard Model

The Standard Model successfully describes the three fundamental forces of the nature. The SM is a gauge theory, and its group structure is as follows $SU(3)_C \times SU(2)_L \times U(1)_Y$. The $SU(3)_C$ describes the strong interaction dynamics. $SU(2)_L \times U(1)_Y$ explains the electromagnetic and weak interactions at low energy. Gluons are the gauge bosons of the $SU(3)_C$ group that act as a mediator of the strong interaction. $W^\pm$ and Z are the gauge bosons of the weak interaction that behave as a mediator of charged and neutral currents, respectively. The massless photon is the mediator of the electromagnetic interaction. The heavy weak gauge boson gets its mass through the spontaneous symmetry-breaking known as Higgs mechanism [3, 4, 11, 12]. The $SU(2)_L \times U(1)_Y$ gauge group is broken down to $U(1)_{em}$ in the Higgs mechanism, whereas the $SU(3)_C$ gauge group remains unbroken. Since $SU(3)_C$ remains unbroken, gluons remain massless even after electroweak symmetry breaking. The SM is a chiral theory where left-handed and right-handed fermion carries different $SU(2)_L \times U(1)_Y$ charges, so the mass term $m\bar{\Psi}\Psi = m\bar{\Psi}_R\Psi_L + h.c$ is not gauge invariant. As a result, the fermions in the SM are massless before electroweak symmetry breaking, but after the symme-

Names	Fields	$SU(3)_C,\ SU(2)_L,\ U(1)_Y$
Leptons	$L_L^i = \begin{pmatrix} \nu_L^i \\ e_L^i \end{pmatrix}$	(1, 2, -1)
	e_R^i	(1, 1, -2)
Quarks	$Q_L^i = \begin{pmatrix} u_L^i \\ d_L^i \end{pmatrix}$	(3, 2, 1/3)
	u_R^i	(3, 1, 4/3)
	d_R^i	(3, 1, $-2/3$)

Table 1.1: The chiral fermion (spin 1/2) of the Standard Model. The electromagnetic charge is $Q = T_3 + \frac{Y}{2}$.

try breaking, the fermions get mass through the Yukawa interaction of the Higgs field. The Yukawa Lagrangian is as follows:

$$-\mathcal{L}_{\text{Yukawa}} = y_e \bar{e}_R \Phi^\dagger L_L + y_d \bar{d}_R \Phi^\dagger Q_L + y_u \bar{u}_R \tilde{\Phi}^\dagger Q_L + h.c\ , \tag{1.1}$$

where $\tilde{\Phi} = i\sigma_2\Phi^*$, $\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ is the Pauli matrix, and Φ is the Higgs doublet. $SU(2)_L$ doublet lepton and quark fields are L_L, Q_L, respectively. The fermion sector of the SM and their charges are given in Table 1.1. The fermions of the SM are classified as quarks or leptons depending on their charge under the color group $SU(3)_C$. The quarks are color triplets, but the leptons are singlet under $SU(3)_C$.

1.1.1 QCD Lagrangian

Hadrons are made up of quarks, anti-quarks, and gluons. The dynamics of the strong interaction between partons are described by Quantum ChromoDynamics (QCD), and its Lagrangian is as follows.

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4} F^a_{\mu\nu} F^{a\mu\nu} + \sum_f \bar{\Psi}_f\ i\Big(\mathbb{1}\partial\!\!\!/ - ig_s A\!\!\!/^a T^a\Big)\Psi_f\ , \tag{1.2}$$

where $A\!\!\!/^a = \gamma^\mu A^a_\mu$, and f is the quark flavor index, $f \subset \{u, d, c, s, b, t\}$. The Lagrangian is invariant under the $SU(3)_C$ group. Interestingly the mass term $-\bar{\Psi}_f m_f \mathbb{1}\Psi_f$ is also invariant under the $SU(3)_C$, but we do not write this term. This is because, as discussed earlier, this mass term is not allowed in the full SM gauge group because of the chiral nature of the SM fermions. The fermionic field Ψ_f is a three-component vector where each component has a color index.

$SU(3)_C$ is a non-abelian gauge group, and its gauge fields are denoted by A^a_μ, called gluons. There are eight gluons, $a = \{1, 2, \cdots, 8\}$. The generators T^a in the fundamental representation are 3×3 matrices and can be written in terms of Gell-Mann matrices. The generators satisfy the following commutation relation,

$$\left[T^a, T^b\right] = i f^{abc} T^c \ , \tag{1.3}$$

where f^{abc} is the structure constant. The kinetic energy part of the above Lagrangian contains a second-rank field strength tensor $F^a_{\mu\nu}$ is given as

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g_s f^{abc} A^b_\mu A^c_\nu \ . \tag{1.4}$$

The kinetic term gives the interaction between the gauge bosons among themselves, which is the characteristic of any non-abelian gauge theory. The Abelian theory (like QED, which is invariant under local $U(1)$ symmetry) has no self-interaction between its gauge bosons. The usual gauge-fixing process introduces ghost fields while quantizing a non-abelian theory. The ghost fields interact with gauge bosons and can appear only through quantum loops. As a result, beyond leading order, one has to take the contributions from the ghost loops.

The QCD Lagrangian is renormalizable; hence, the strong coupling constant is not a constant number and changes with the scale. In other words, the value of the strong coupling constant measured at two scales will be different. The renormalization group equation (RGE) governs scale dependence, which has the following form,

$$\mu_R^2 \frac{d\alpha_S}{d\mu_R^2} = \beta(\alpha_S(\mu_R)) = -\sum_{i=0}^{\infty} \alpha_S^{i+2} \beta_i \ , \tag{1.5}$$

where μ_R is the renormalized scale, $\alpha_S = g_s^2/16\pi^2$, and $\beta(\alpha_S(\mu_R))$ is the beta function. β_0 comes from the one-loop contribution, and it is as below,

$$\beta_0 = \frac{11}{3} C_A - \frac{4}{3} n_f T_f \ , \tag{1.6}$$

where $C_A = 3$, $T_f = 1/2$, and n_f is the number of quark flavors. In the SM $n_f = 6$, therefore, β_0 is a positive number. β_1, β_2, and higher terms come from the NNLO and higher loops. In the literature, the value of the beta function is known up to five loops [13]. Integrating Equation 1.5, taking only one-loop contribution, gives the following:

$$\begin{aligned} \frac{1}{\alpha_S(Q^2)} - \frac{1}{\alpha_S(\mu_0^2)} &= \beta_0 \log\left(\frac{Q^2}{\mu_0^2}\right) \\ \text{or,} \quad \alpha_S(Q^2) &= \frac{\alpha_S(\mu_0^2)}{1 + \alpha_S(\mu_0^2)\beta_0 \log\left(\frac{Q^2}{\mu_0^2}\right)} + \mathcal{O}(\alpha_S^2) \ . \end{aligned} \tag{1.7}$$

The above equation suggests that the strong coupling constant decreases when the energy scale Q increases. As a result, it is exceedingly tiny at $Q \to \infty$. This feature of the QCD is known as asymptotic freedom. This means the strong interactions at very high energy become a free theory, while at low energy, it is a strongly interacting theory.

1.2 Motivation to go beyond Standard Model

The SM of particle physics has successfully described many natural phenomena we observed in the experiment. For example, precise electron magnetic dipole moment measurements agree well with the SM predictions [14], including higher-order loop corrections. The requirement of the SM to be anomaly free [15, 16] predicts the existence of the third-generation quark after the discovery of the tau lepton [17]. Another example is the prediction of the top quark mass [18] before its discovery from the $W^{\pm}$ and Z boson mass measurements, fermi constant G_F, and the fine structure constant. The last missing piece of the SM is the Higgs boson, which was discovered at the LHC in 2012 [19, 20].

Despite the tremendous success of SM, many aspects of nature are still entirely missing or unsatisfactory. Some of these shortcomings are discussed in this section.

Various neutrino oscillation experiments established that at least two active neutrinos are massive, and different flavor eigenstates mix among themselves. The mass eigenstates and the flavor eigenstates of the neutrinos are related through the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix. Although the exact masses of the neutrinos are unknown, the observations of the cosmic microwave background have provided an upper bound on the total neutrino masses, $\sum_i m_i < 0.12\ eV$ [21]. In the SM framework, the gauge-invariant Dirac mass for neutrinos is not possible because of the absence of right-handed neutrinos.

Another important flaw of the Standard Model is the matter anti-matter asymmetry. The particle and its anti-particle differ only by the equal and opposite quantum charges; therefore, they will always pair produced and annihilate each other. As a result of this process, the present Universe is expected to have an equal amount of matter and anti-matter. However, our existence and the structures of the Universe indicate the excess matter abundance over anti-matter. This asymmetry is called baryonic asymmetry. The Standard Model of particle physics has no correct explanation for this imbalance. To explain the baryon asymmetry, Sakharov proposed [22] three necessary conditions: C and CP violation, baryon number violation, and out of the thermal equilibrium. However, CP violation in the SM quark sector alone can not provide the required asymmetry.

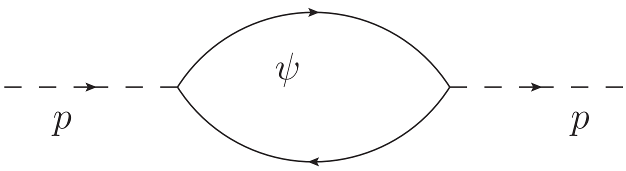


Figure 1.1: Fermion loop contributing to the Higgs boson mass

1.2.1 Hierarchy problem

Apart from the above mentioned shortcomings of the SM, several flaws are present from the theoretical side, like the hierarchy problem. Because of the quantum nature of the SM, many particles can contribute to the physical mass of a particle through virtual loop corrections. The fermion and the gauge boson masses in the SM are protected by the chiral and gauge symmetry, respectively. As a result, the corrections of their masses from the higher loops are proportional to their mass itself. However, any symmetry of the SM does not protect the mass of the Higgs boson. Therefore, if new particles exist in the UV scale, and the SM Higgs boson interacts with them, then the mass of the Higgs boson receives an enormous quantum correction. Hence, the fundamental value of the renormalized mass squared parameter of the Higgs boson should be of the order of the quantum correction for a delicate cancelation between them. Therefore, the hierarchy problem is closely related to fine-tuning, known as the naturalness problem.

Below we describe if there is a Yukawa-type interaction between the SM Higgs boson with the UV scale fermions; the Higgs boson tree-level mass receives a correction quadratic in M_f. Since the mass $M_f \sim$ UV scale, the correction is enormous; therefore, to match its mass with the experimentally measured value, very high fine-tuning is needed, which is unnatural.

Consider a fermion ψ, which has mass M_f, couples to the SM Higgs boson (h) through Yukawa interaction $yh\psi\bar{\psi}$, where y is the strength of the interaction. The one-loop $\psi - \bar{\psi}$ contribution to the Higgs self-energy (Figure 1.1) can be written as below:

$$i\Sigma_2(p^2) = \int \frac{d^4k}{(2\pi)^4}(iy)^2 \frac{\mathrm{Tr}\Big[(\not{p} + \not{k} + M_f)(\not{k} + M_f)\Big]}{((p+k)^2 - M_f^2)(k^2 - M_f^2)} \;, \tag{1.8}$$

where k is the loop momentum and $\mathrm{Tr}\Big[(\not{p}+\not{k}+M_f)(\not{k}+M_f)\Big] = 4(k^2+M_f^2+k.p)$. Using Feynman parameterization, we can write the above equation as

$$i\Sigma_2(p^2) = -4y^2 \int_0^1 dx \int \frac{d^4k}{(2\pi)^4} \frac{k^2 + \Delta}{(k^2 - \Delta)^2} \;, \tag{1.9}$$

where x is Feynman parameter and $\Delta = M_f^2 - p^2 x(1-x)$.

The first term of the above equation has quadratic UV divergences. Such quadratic divergences can be kept away by doing the renormalization, but we will see it has the finite correction of the following form $\frac{y^2}{4\pi^2}M_f^2$. This correction is still manageable for an SM fermion like the top quark. However, any heavy UV fermion that couples with the SM Higgs will make tremendous contributions.

After doing the integration [23] in dimensional regularization, we have,

$$\Sigma_2(p^2) = -\frac{3y^2}{4\pi^2}\mu^\epsilon \int_0^1 dx \Big(\frac{1}{\bar{\epsilon}} - ln\ \Delta + \frac{1}{3}\Big)\Delta\ , \tag{1.10}$$

where $\frac{1}{\bar{\epsilon}} = \frac{2}{\epsilon} - \gamma + ln\ 4\pi$. Factor μ^ϵ comes since the integration is done in the $d = 4-\epsilon$ dimensions. For $p^2 << M_f^2$, one can write the following:

$$\Delta\ ln\ \Delta = \Delta\ ln\Big[M_f^2\Big(1 - \frac{p^2 x(1-x)}{M_f^2}\Big)\Big] = \Delta\ lnM_f^2 - p^2 x(1-x) + \mathcal{O}(p^4). \tag{1.11}$$

Substituting Equation 1.11 into Equation 1.10 and doing the integration we have,

$$\Sigma_2(p^2) = -\frac{3y^2}{4\pi^2}\Big[\Big(\frac{1}{\bar{\epsilon}} + \frac{1}{3} - ln\frac{M_f^2}{\mu^2}\Big)(M_f^2 - \frac{p^2}{6}) + \frac{p^2}{6} + \mathcal{O}(p^4)\Big]\ . \tag{1.12}$$

To renormalize the UV divergences, we have to add the counter terms. Therefore, the total one-loop two-point function becomes,

$$\Sigma(p^2) = \Sigma_2(p^2) + \delta_h p^2 - (\delta_h + \delta_m)m_R^2\ , \tag{1.13}$$

where m_R is the renormalized physical mass of the Higgs boson. In the $\overline{MS}$ scheme, the counter terms absorb only the UV divergent parts and are defined to have no finite parts. Therefore, the counter terms are as follows:

$$\delta_h = -\frac{y^2}{8\pi^2\bar{\epsilon}}\ \ (\text{coeff. of } p^2)\ , \text{and}\ \ (\delta_h + \delta_m)m_R^2 = -\frac{3y^2 M_f^2}{4\pi^2\bar{\epsilon}} \tag{1.14}$$

Therefore, we get

$$\Sigma(p^2) = -\frac{3y^2}{4\pi^2}\Big[\Big(\frac{1}{3} - ln\frac{M_f^2}{\mu^2}\Big)(M_f^2 - \frac{p^2}{6}) + \frac{p^2}{6} + \mathcal{O}(p^4)\Big]\ , \tag{1.15}$$

Finally, the physical Higgs boson propagator can be written as,

$$\frac{1}{p^2 - m_h^2} = \frac{1}{p^2 - m_R^2 + \Sigma(p^2)} \tag{1.16}$$

$$m_h^2 \approx m_R^2 + \frac{y^2}{4\pi^2} M_f^2 - \frac{3y^2}{4\pi^2} M_f^2 \; ln \frac{M_f^2}{\mu^2} \; . \tag{1.17}$$

The left-hand side is the Higgs boson's mass, equated with the experimentally observed value 125 GeV, where m_R^2 is the renormalized (divergence-free) Higgs boson mass. At tree-level $m_h^2 = m_R^2$. One important point is that the mass-squared correction of the Higgs boson is independent of m_R^2; hence m_R^2 is an unnatural parameter. The above equation shows that tree-level Higgs boson mass gets a quadratic correction $\frac{y^2}{4\pi^2} M_f^2$. If $y = \mathcal{O}(1)$, and M_f is large ($\sim$ UV scale), then Higgs boson mass receives a large correction, leading to the hierarchy problem.

Various theoretical solutions, including Supersymmetry (SUSY) and extra dimensions, are put forward to address the hierarchy problem. In an unbroken SUSY scenario, all bosons have fermion counterparts and vice versa. Since the fermion loops have a negative sign, the quadratic divergences of the Higgs self-energy precisely cancel out when considering contributions from both bosonic and fermionic loops. In the broken SUSY scenario, all quadratic divergences disappear, but log terms present, which are still manageable.

1.2.2 Dark Matter (DM)

One of the most compelling drawbacks of the SM is that it has no explanation for the dark matter. In 1930 Zwicky observed that the constituents of the Coma Cluster do not follow the virial theorem [24]. In order to explain this, the dark matter was initially postulated. Vera Rubin and her collaborators made a crucial measurement of the galactic rotation curve in 1970. The orbital/rotational speed of a visible star/galaxy is plotted against the radial distance of the star/galaxy from the galactic center. In this observation, they found that the rotation speed of stars (or galaxies) remains relatively constant at large radial distances. This contradicts Kepler's third law, which predicts a decrease in orbital speed as the radial distance increases. Therefore, the explanation of the rotation curve demands the presence of more matter than what is visible. Later, the gravitation lensing of the bullet cluster also demands a significant amount of non-luminous matter in the Universe. Finally, Cosmic Microwave Background (CMB) radiation observations, most recently by the Planck collaboration [25], provide the relic density of the DM, which is as below:

$$\Omega_{\rm DM} h^2 = \; 0.120 \pm 0.001 \; . \tag{1.18}$$

All the evidence indicates that DM has gravitational interactions. Since DM does not interact with photons, it is non-luminous and electrically neutral. Therefore, the SM has no particle that fits all the properties of the DM.

Eighty percent of the total amount of matter in the Universe is made up of DM, which accounts for twenty-six percent of its total energy budget. Despite its matter contents, we do not yet know its mass, its interaction with the particles in the Standard Model, or how many DM candidates are in the Universe. It is also unknown whether the DM is a fermion, a scalar, or something else. In order to accommodate the DM in the context of particle physics, the SM of particle physics must be extended because it lacks a good candidate for the DM.

Depending on how dark matter is produced and interacts, several different dark matter paradigms exist in the literature. The most well-known is the weakly interacting massive particle (WIMP) paradigm. Another interesting paradigm is the feebly interacting massive particle (FIMP) paradigm, an alternative to the WIMP.

1.2.2.1 Boltzmann Equation:

Before taking a short introduction of different paradigms of dark matter, let us talk about the Boltzmann equation, which describes how the number density of a particle changes with time in a thermodynamic system. The number density can be written in terms of phase space distribution function $f(p^\mu, x^\mu)$, as below,

$$n(t) = \frac{g}{(2\pi)^3}\int d^3p f(p^\mu, x^\mu) = \frac{g}{(2\pi)^3}\int d^3p f(E,t)\ , \tag{1.19}$$

where g denotes the internal degree of freedom. The equilibrium distribution function of a particle of energy E is $f_{eq} = \exp(-E/T)$, which can be used to calculate the equilibrium number density, as shown below.

$$n_{\chi,eq} = \frac{g_\chi}{(2\pi)^3}\int d^3p\ \exp(-E_\chi/T) \tag{1.20}$$

After substituting energy $E_\chi = \dfrac{p^2}{2m_\chi} + m_\chi$ (natural unit) for the non-relativistic case ($T << m_\chi$) or $E_\chi = p^2 + m_\chi^2$ for the relativistic case ($T >> m_\chi$) in the above equation, we get the following:

$$\begin{aligned} n_{\chi,eq} &= g_\chi\Big(\frac{m_\chi T}{2\pi}\Big)^{3/2}\exp(-m_\chi/T) = g_\chi\Big(\frac{m_\chi^2}{2\pi\ x}\Big)^{3/2}\exp(-x)\ \text{(non-relativistic)} \\ &= g_\chi\frac{T^3}{\pi^2} = g_\chi\frac{m_\chi^3}{\pi^2}\frac{1}{x^3}\quad \text{(for relativistic case)}\ . \end{aligned} \tag{1.21}$$

Particle physicists commonly assume that DM exhibits particle-like behavior and is stable. When calculating the relic density of a stable DM particle, the calculations involve considering $2 \leftrightarrow 2$ scattering processes. Consider that a stable particle χ is in thermal equilibrium with the bath particles in the early Universe. All the processes like $\chi\bar{\chi} \leftrightarrow \psi\bar{\psi}$ (ψ is any SM particle in thermal equilibrium) will contribute to the Boltzmann equation. The evolution of the number density of the DM is as follows:

$$\frac{dY_\chi}{dx} = -\frac{\langle \sigma v \rangle s}{Hx}(Y_\chi^2 - Y_{\chi,eq}^2) \ . \tag{1.22}$$

In the preceding equation, we employ a dimensionless parameter $Y_\chi = n_\chi/s = n_{\bar{\chi}}/s$ instead of n_χ to absorb the effect of the Universe's expansion. Entropy density is denoted by s. We vary using a dimensionless variable $x = m_\chi/T$ instead of time since the temperature is a better variable to study the evolution of the Universe. $\langle \sigma v \rangle$ denotes the thermally averaged cross section.

The Hubble rate and the entropy density of the radiation-dominated early Universe are given below.

$$\begin{aligned} H &= \sqrt{\frac{\pi^2 g_*}{90}}\,\frac{T^2}{M_{pl}} = \sqrt{\frac{\pi^2 g_*}{90}}\,\frac{m_\chi^2}{M_{pl}}\frac{1}{x^2} \ , \\ s &= \frac{2\pi^2}{45} g_{*s} T^3 = \frac{2\pi^2}{45} g_{*s}\frac{m_\chi^3}{x^3} \ . \end{aligned} \tag{1.23}$$

g_{*s} and g_* are the relativistic degree of freedom that contribute to the entropy and energy density, respectively. Now, by inserting the Hubble rate and entropy density expressions into Equation 1.22, we obtain

$$\frac{dY_\chi}{dx} = -\frac{\lambda}{x^2}\langle \sigma v \rangle (Y_\chi^2 - Y_{\chi,eq}^2) \ , \tag{1.24}$$

where,

$$\lambda = \frac{sx}{H} = \frac{2\pi\sqrt{90}}{45}\frac{g_{*s}}{\sqrt{g_*}} m_\chi M_{pl} \ . \tag{1.25}$$

1.2.2.2 WIMP Dark Matter:

The WIMP paradigm assumes that the DM in the early Universe is in thermal equilibrium with the rest of the bath particles, which implies at $x = 0$, $Y_\chi = Y_{\chi,eq}$. The interaction rate of any $2 \leftrightarrow 2$ processes is $\Gamma_{\rm an} = n_\chi \langle \sigma v \rangle$. As the Universe expands, the interaction rate decreases, and at $\Gamma_{\rm an} \simeq H$, the DM decouples from the thermal plasma. This mechanism is known as freeze-out, and this point is characterized by $x = x_{fo}$. In the WIMP scenario, if the DM mass lies between GeV to TeV range, the typical value of the freeze-out point is $x_{fo} = 20 - 25$. After decoupling, DM interactions become insignificant, and its abundance freezes. As a result, we have the following scenarios:

$$\begin{aligned} Y_\chi(x) &= Y_{\chi,eq} \qquad \text{for} \quad x \leq x_{fo} \ , \\ Y_\chi(x) &= Y_\chi(x_{fo}) \qquad \text{for} \quad x > x_{fo} \ . \end{aligned} \tag{1.26}$$

The abundance of the DM in today's Universe is Y_∞, defined as $Y_\infty = Y(x = x_\infty)$, where $x_\infty = m_\chi/T_\infty$ and T_∞ is today's temperature ($= 2.725 \pm 0.001\ K$). For any mass of the DM, its relic density can be obtained by solving the following expression numerically:

$$\Omega_\chi h^2 = 2.744 \times 10^8 \frac{m_\chi}{\text{GeV}} Y_\infty \ . \tag{1.27}$$

The experimentally measured value of the DM relic density is given in Equation 1.18. Extraction of the analytical solution of Equation 1.24 is difficult because, in most cases, many annihilation processes contribute to the annihilation cross section, making integration over x challenging. However, one can obtain a numerical solution. Now we will find an approximate analytic solution of the evolution equation by employing a few assumptions as below.

Equilibrium density falls exponentially with x (Equation 1.21). Therefore, for $x >> x_{fo}$, we can safely ignore $Y_{\chi,eq}$ compared to Y_χ in Equation 1.24. However, this assumption is invalid if x is less than x_{fo} since, in that case, $Y_\chi(x) = Y_\chi(x_{fo})$.

$$\frac{dY_\chi}{dx} \approx -\frac{\lambda}{x^2}\langle\sigma v\rangle Y_\chi^2 \tag{1.28}$$

Integrating the above equation, we get,

$$\frac{1}{Y_\infty} - \frac{1}{Y_{fo}} = \int_{x_{fo}}^{\infty} \frac{dx}{x^2}\lambda\langle\sigma v\rangle \tag{1.29}$$

Here λ varies with x since g_{*s} and g_* vary with temperature. To obtain an approximate analytic equation for the relic density, we assume that λ is temperature independent, take it outside the integral, and replace it with λ_{fo}, the value of λ at the freeze-out point.

$$\frac{1}{Y_\infty} - \frac{1}{Y_{fo}} = \lambda_{fo} \int_{x_{fo}}^{\infty} \frac{dx}{x^2}\langle\sigma v\rangle \ , \tag{1.30}$$

In order to obtain the expression of λ_{fo}, we have to use the Hubble rate during the freeze-out point, which is given below.

$$H_{fo} = \sqrt{\frac{4\pi^3}{45}}\sqrt{g_*}\,\frac{m_\chi^2}{M_{pl}}\frac{1}{x_{fo}^2} \tag{1.31}$$

Therefore, the expression of the λ_{fo} becomes,

$$\lambda_{fo} = \frac{s x_{fo}}{H_{fo}} = \sqrt{\frac{\pi}{45}} \frac{g_{*s}}{\sqrt{g_*}} m_\chi M_{pl} \ . \tag{1.32}$$

These two relativistic degrees of freedom are roughly the same for the WIMP scenario at freeze-out. Therefore, we replace $g_{*s} = g_* (= 80 - 100)$ in the preceding equation and get the following.

$$\lambda_{fo} = \sqrt{\frac{\pi}{45}} \sqrt{g_*} \ m_\chi M_{pl} \ . \tag{1.33}$$

After neglecting Y_{fo} in Equation 1.30,

$$Y_\infty \approx \frac{1}{\lambda_{fo} \ J(x_{fo})} \ , \tag{1.34}$$

where,

$$J(x_{fo}) = \int_{x_{fo}}^{\infty} \frac{dx}{x^2} \langle \sigma v \rangle \ . \tag{1.35}$$

Substitution Equations 1.33 and 1.34 into Equation 1.27, the approximate relic density becomes,

$$\Omega_\chi h^2 = 1.09 \times 10^9 \frac{\mathrm{GeV}^{-1}}{\sqrt{g_*} M_{pl}} \frac{1}{J(x_{fo})} \ . \tag{1.36}$$

Although this is the approximate analytic solution of DM relic density, we write our model in the FeynRules [26] package and then implement it in `micrOMEGAs` -v5 [27] to get the actual numerical result of the relic density.

If DM annihilation happens only via the s-wave, we can obtain a trivial solution for the DM relic density. S-wave means σv does not depend on the annihilating DM particles' relative velocity, therefore no temperature dependence. In that case, $J(x_{fo})$ becomes $J(x_{fo}) = \langle \sigma v \rangle \frac{1}{x_{fo}}$. As a result, the relic density of DM becomes,

$$\Omega_\chi h^2 = 1.09 \times 10^9 \frac{\mathrm{GeV}^{-1}}{\sqrt{g_*} M_{pl}} \frac{x_{fo}}{\langle \sigma v \rangle} \ . \tag{1.37}$$

If the mass of DM lies in the range of $0.1 - 10^4$ GeV, then x_{fo} varies between 20 and 30. Therefore for this mass range of the DM, if $\langle \sigma v \rangle \sim 2 \times 10^{-26}$ cm$^3/s \sim$ 1 pb , then the DM relic density obtained from the above equation matches the observed value given by the Planck collaboration. 1 pb is the typical cross section of the electroweak interaction. It is known as the WIMP miracle because the required cross section to achieve the correct relic density surprisingly matches the electroweak cross section.

Although we study only the WIMP scenario in the thesis, we briefly mention an alternative idea of the WIMP paradigm below.

1.2.2.3 FIMP Dark Matter:

In the FIMP paradigm, the interaction between the DM and bath particles (visible sector) is assumed to be extremely small $\sim \mathcal{O}(10^{-7})$; hence, the DM never in thermal equilibrium with bath particles in the early Universe. The production of DM in the early Universe occurred through the annihilation of the Standard Model particles. At high temperatures, the SM particles annihilate and produce DM. This production process continues for a while, so the abundance of the DM increases. Since the Universe expands, the rate of these annihilation processes slows down. As a result, after a while, the Universe temperature becomes low enough such that the DM production becomes insignificant. Therefore, the abundance of dark matter does not increase further and remains constant until today. This is known as the freeze-in mechanism of the feebly interacting massive particles.

For the freeze-in scenario, the initial amount of dark matter abundance is negligibly small or zero; therefore, the Y_χ^2 term in Equation 1.24 can be safely ignored compared to $Y_{\chi,eq}^2$.

$$\frac{dY_\chi}{dx} \simeq \frac{\lambda}{x^2}\langle\sigma v\rangle Y_{\chi,eq}^2 \tag{1.38}$$

Unlike the freeze-out mechanism, the abundance of FIMP-type DM increases with temperature until it freezes in. The above equation suggests that a larger annihilation cross section of the visible particles provides larger abundances of the FIMP DM, unlike the WIMP scenario.

There are three main categories for experimental dark matter detection: direct detection, indirect detection, and collider searches at the LHC. We will now briefly explore them in this section.

1.2.2.4 Direct Detection (DD)

The main idea of the direct detection experiment is to measure the recoil energy of the nucleus deposited in a detector due to the elastic scattering between DM and the nucleus. The underground detectors of the direct detection experiment use various targets nucleus like Xenon, argon, and many others. Prediction of the nucleus-DM interaction rate requires different astrophysical and cosmological inputs.

Event Rate: One of the essential quantities in the DD experiment is the differential event rate [28–30], calculated per count, day, kilogram, and keV, as below.

$$\frac{dR}{dE_{NR}} = \frac{\rho_0}{m_N m_\chi} \int_{v_{min}} dv\ v f(v) \frac{d\sigma}{dE_{NR}}(v, E_{NR}) \ , \tag{1.39}$$

where E_{NR} is the nuclear recoil energy, σ is the dark matter nucleon scattering cross section, and m_N and m_χ are the nucleon and dark matter masses, respectively. The typical value of the local density of dark matter is ρ_0 = 0.3 GeV/cm^3 [31], and most direct detection experiments use Gaussian velocity distribution for the dark matter halo,

$$f(v) = \frac{1}{\sigma_v \sqrt{2\pi}} \exp\left(-\frac{v^2}{2\sigma_v^2}\right) \ . \tag{1.40}$$

The velocity dispersion σ_v of the DM gas cloud can be expressed by the galaxy's circular velocity, $\sigma_v = \sqrt{3/2}\ v_c$, with $v_c = 220 \pm 20$ km/s [32]. One obtains the differential event rate by integrating Equation 1.39 from the minimum velocity v_{min} to all possible velocities up to v_{esc}. The minimum velocity is the following,

$$v_{min} = \sqrt{\frac{m_N E_{NR}}{2\mu_{\chi N}^2}} \ . \tag{1.41}$$

where $\mu_{\chi N} = \dfrac{m_\chi m_N}{m_\chi + m_N}$ represents the reduced mass of the DM-nucleon system. $v_{esc} = 544$ km/s [33] is the escape velocity; if $v > v_{esc}$ DM escapes from the DM halo.

The total event rate per day per kg can be obtained after integrating Equation 1.39 over all the possible energy ranges of the nuclear recoil as follows.

$$R = \int_{E_{NR,min}}^{E_{NR,max}} dE_{NR}\ \epsilon(E_{NR})\ \frac{dR}{dE_{NR}} \ . \tag{1.42}$$

$\epsilon(E_{NR})$ denotes the efficiency of the detector. The minimum recoil energy, $E_{NR,min}$, is the detector's threshold, and the maximum energy can be obtained from Equation 1.41 by replacing v_{min} with v_{esc},

$$E_{NR,max} = \frac{2\mu_{\chi N}^2 v_{esc}^2}{m_N} \ . \tag{1.43}$$

Dark Matter Nucleon Cross Section: To get the differential event rate from Equation 1.39, we need to know the DM nucleon cross section, which can be spin-dependent or spin-independent. Nucleon (proton or neutron) comprises light quarks (u, d, s) and gluon. Spin-dependent cross section arises when elastic scattering between DM and light quarks/gluon occurs through the axial-vector mediator. The spin-dependent cross section relies on the spin of DM and the total angular momentum of the nucleus. On the contrary, the spin-independent cross section does not depend on the nucleus angular momentum and the DM's spin [34]. The cross section becomes spin-independent when the elastic scattering between the DM and quarks/gluon occurs through a scalar, vector, or fermion mediator. All the DM-motivated models in this thesis have only spin-independent direct detection channels; therefore, we only discuss the spin-independent scattering cross section below.

The spin-independent (SI) scattering cross section expression can be expressed as,

$$\left(\frac{d\sigma}{dE_{NR}}\right)_{SI} = \frac{2m_N}{\pi v^2}\left(\{Zf^p + (A-Z)f^n\}^2 + \frac{B_N^2}{256}\right)F^2(E_{NR}) \; . \tag{1.44}$$

$(A - Z)$ and Z represent the number of neutrons and protons in the nucleus. $B_N = \alpha_u^V(A + Z) + \alpha_d^V(2A - Z)$ will be non-zero if the scattering between DM and the quark occurs through a vector mediator. α_u^V (α_d^V) is the coupling strength of the interaction between the vector mediator, dark matter, and u (d) quark. $F(E_{NR})$ is the experimental form factor [35, 36].

The structure of the quantity $f^{p,n}$ is as follows [37]

$$\frac{f^{p,n}}{m_{p,n}} = \sum_{q=u,d,s}\frac{\alpha_q^S}{m_q}f_{T,q}^{p,n} + \frac{2}{27}f_{T,g}^{p,n}\sum_{q=u,d,s}\frac{\alpha_q^S}{m_q} \; , \tag{1.45}$$

where α_q^S is the coupling strength of the scalar interaction. For example, if the DM is a fermion and the elastic scattering between the fermionic DM and Standard Model quarks occurs via a scalar mediator, then α_q^S denotes the coupling strength of that interaction.

The coefficient of the nucleon matrix element is defined below,

$$f_{T,q}^{p,n} = \frac{m_q}{m_{p,n}}\langle N|\bar{q}q|N\rangle \; . \tag{1.46}$$

These quantities are computed either experimentally or accurately using the pion-nucleon sigma term's measurements [36] or the Lattice QCD. The contribution of the gluon is denoted by $f_{T,g}^{p,n}$ and is defined below,

$$f_{T,g}^{p,n} = 1 - \sum_{q=u,d,s} f_{T,q}^{p,n} \; . \tag{1.47}$$

Over the last few years, many experimental groups have made numerous attempts to find evidence of dark matter. However, they have not successfully identified the WIMP or any other type of dark matter. Based on the absence of any observed events, experimental constraints are placed on the elastic cross section of dark matter-nucleon interactions, with respect to the dark matter mass,

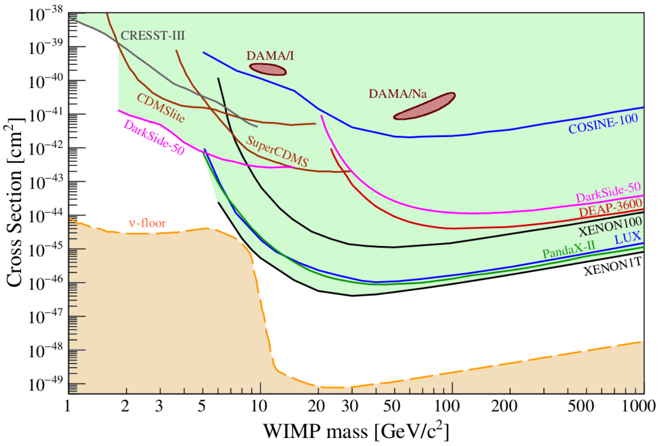


Figure 1.2: Different exclusion contours of Direct detection experiments are presented. The regions above the contours are excluded. The yellow region denotes the neutrino floor. The image is taken from [38].

for various theoretical models. Some current limits obtained from the DD experiments are displayed in Figure 1.2.

As previously stated, direct detection scattering cross sections can be either spin-dependent or spin-independent. However, the strongest constraints for most models come from the spin-independent cross section. The coherent neutrino-nucleus scattering gives the same type of signal in the detector, resulting in an irreducible background for the WIMP search. The yellow dashed line, known as the neutrino floor, bounds the parameter space from below in Figure 1.2.

1.2.2.5 Indirect Detection (ID)

The basic idea of indirect detection is to observe particles generated due to the annihilation and/or Co-annihilation processes of dark matter. The final state particles include electrons, positrons [39, 40], protons, anti-protons [41, 42], photons [43, 44], and neutrinos [45]. In most BSM scenarios, DM particles annihilate and produce the SM particles, out of which electromagnetically neutral photons and neutrinos have good chances of reaching the detector without getting deflected from the source. Various experiments, such as PAMELA [46], MAGIC [47], Fermi-LAT [48], and others, actively search for indirect detection signatures. The stable particles reach the earth, and one can limit the parameter spaces of various DM models by measuring their fluxes.

1.2.2.6 Collider Searches

At colliders like the LHC, dark matter particles can be produced at the hard interactions accompanied by QCD radiations or other entities. Alternatively, DM can also be produced from the decay of heavy BSM particles, where heavy BSM particle (or pair of heavy BSM particles) are produced in the hard interaction and subsequently decays into DM.

Dark matter is an electromagnetically neutral particle and has no strong interactions. Therefore, it does not leave any signature on the detectors. However, the presence of DM in the final state causes an imbalance of the transverse momentum between the initial and final state particles. Therefore, the usual method of searching for DM at the LHC involves mono-x search, where x can be a photon, a QCD jet, a heavy gauge boson, a Higgs boson, or any other particle, along with a large missing transverse momentum.

Precise measurement of the Z/W bosons decay width from the LEP experiment also restricts parameter spaces of many BSM models. The current measurement of the Higgs invisible decay branching ratio [49] also strongly constrains the parameter spaces of the Higgs portal models, where DM couples with the Higgs boson for DM mass $m_{DM} < m_h/2$.

In Chapter 3, we explore the inert Higgs doublet model, a simple extension of the SM that has a suitable DM candidate and provides the correct relic density, and satisfies all theoretical and experimental constraints. Chapters 4 and 5 explore a two-component dark matter scenario in an extended Kim-Shifman-Vainshtein-Zakharov (KSVZ) framework. The KSVZ model also offers a natural solution to the strong-CP problem, which is covered in the next section.

1.2.3 Strong CP

The strong charge-parity (CP) problem is another robust shortcoming of the SM. The $SU(3)_C$ symmetry of the SM allows a term like $\theta \frac{g_S^2}{32\pi^2}\tilde{G}_{\mu\nu}G^{\mu\nu}$, where $G^{\mu\nu}$ is the gluon field strength tensor, and $\tilde{G}^{\mu\nu} = 1/2\ \epsilon^{\mu\nu\rho\sigma}G_{\rho\sigma}$. This term contributes to the neutron electric dipole moment (eDM), and the experimental measurement [50] constraints the parameter $\overline{\theta} \leq 10^{-10}$. These two parameters $\overline{\theta}$ and θ are related through quark field chiral rotation. When the variable $\overline{\theta} \to 0$, it does not lead to an enhancement of the symmetry in the theory. Therefore, it is expected that $\overline{\theta}$ would take a value $\overline{\theta} \sim \mathcal{O}(1)$. This is known as a strong charge-parity (CP) problem [51–54].

In the literature, many solutions exist that can explain the strong CP problem, and some of them are listed here. If the quarks are massless, then it is easy to get rid of the strong CP problem, as follows. Under chiral transformation, the quark field transforms as $q \to e^{i\alpha\gamma_5/2}q$ and $m\bar{q}q \to m\bar{q}e^{i\alpha\gamma_5}q$; therefore, chiral symmetry will be a good symmetry if the quarks are massless. Under the chiral transformation, the change of the QCD action will be as follows:

$$\delta S_{QCD} = -i \int d^4x \; \alpha \partial_\mu J^5_\mu = -i \int d^4x \; \alpha \frac{g_S^2}{32\pi^2} \tilde{G}_{\mu\nu} G^{\mu\nu} \; . \tag{1.48}$$

The total Lagrangian becomes,

$$\mathcal{L} = \mathcal{L}_{QCD} + (\theta - \alpha) \frac{g_S^2}{32\pi^2} \tilde{G}_{\mu\nu} G^{\mu\nu} \; . \tag{1.49}$$

One can eliminate the original θ term by choosing $\alpha = \theta$. Thus strong CP problem is resolved.

The next most straightforward solution is the RG running of the $\bar{\theta}$. In SM, RG running of $\bar{\theta}$ starts at seven loops [55]. Therefore, if one sets $\bar{\theta} = 0$ at some RG scale and considers the SM is a low-energy effective theory, then $\bar{\theta}$ will be very small at low energies.

Although there are several solutions to the Strong CP problem, the QCD axions are the most popular solution. The effective field theory (EFT) of the axion is simple. EFT considers one axion field $a(x)$ and a single coupling F_a and introduces an effective Lagrangian as below.

$$\mathcal{L} \supset \mathcal{L}_{QCD} + (\frac{a}{F_a} + \theta) \frac{g_S^2}{32\pi^2} \tilde{G}_{\mu\nu} G^{\mu\nu} \; . \tag{1.50}$$

The above interaction leads to an axion potential of the following form [51],

$$V = -m_\pi^2 f_\pi^2 \sqrt{1 - \frac{4 m_u m_d}{(m_u + m_d)^2} \sin^2\Big(\frac{a}{2F_a} + \frac{\bar{\theta}}{2}\Big)} \; . \tag{1.51}$$

The potential will be minimum when axion vev $\langle a \rangle = -\bar{\theta} F_a$. The neutron electric dipole moment will be $\propto \frac{a}{F_a} + \bar{\theta} = 0$. Thus, when the axion relaxes to the minimum potential, the neutron eDM is dynamically adjusted to zero, and the Strong CP problem is resolved.

Since axion has no symmetry property, the effective axion Lagrangian in Equation 1.50 can not be written without introducing a host of other couplings. Therefore, we will discuss the simplest UV complete axion model known as Kim-Shifman-Vainshtein-Zakharov (KSVZ) model [56,57]. This model includes a complex scalar that is singlet under the SM gauge groups $\eta \sim (1, 1, 0)$ and a vector-like quark (VLQ) $\Psi = \Psi_L + \Psi_R$ that is an $SU(3)_C$ color triplet but $SU(2)_L$ singlet with hypercharge zero, with an approximate $U(1)$ symmetry. The $U(1)$ symmetry is historically known as Peccei-Quinn (PQ) symmetry, $U(1)_{PQ}$. The Lagrangian is as follows

$$\mathcal{L}_{\Psi,\eta} = i\bar{\Psi}\not{D}\Psi + f_{\Psi}(\eta\overline{\Psi}_L\Psi_R + \eta^{\dagger}\overline{\Psi}_R\Psi_L) + (\partial_{\mu}\eta^{\dagger})(\partial^{\mu}\eta) + m^2(\eta^{\dagger}\eta) - \lambda(\eta^{\dagger}\eta)^2 \ . \quad (1.52)$$

The scalar field can be written as,

$$\eta = \frac{1}{\sqrt{2}}(F_a + \sigma_0)\ e^{\frac{ia(x)}{F_a}}\ , \quad (1.53)$$

where, σ_0 and $a(x)$ are the radial mode and the axion field, respectively. F_a is known as the axion decay rate or the PQ breaking scale. The above Lagrangian contains the $\bar{\Psi}\Psi a$ and $\bar{\Psi}\Psi g$ interactions. After the spontaneous PQ symmetry breaking, VLQ gains mass proportionate to the PQ breaking scale, which is very large. Integrating out the VLQ loop in its infinite mass limit, the effective Lagrangian becomes [57],

$$\frac{a}{F_a}\frac{g_S^2}{32\pi^2}\tilde{G}_{\mu\nu}G^{\mu\nu}\ . \quad (1.54)$$

Like the preceding discussion, the neutron eDM becomes zero as the vev of the axion field cancels the initial $\bar{\theta}$ term. Thus we get rid of the strong CP problem.

Equation 1.52 suggests that the axion is massless at tree level. However, the interaction in Equation 1.54 can provide a small non-zero axion mass ($\mathcal{O}(KeV)$) through loop induced process [57], and its mass is inversely proportional to the breaking scale F_a. The same interactions also help the axion to decay into gluons, where the decay rate is also inversely proportional to F_a. As a result, if the braking scale is tuned correctly, the axion lifetime can be larger than the age of the Universe and behaves as a dark matter. The bound on F_a is as follows: 10^{10} GeV $\leq F_a \leq 10^{12}$ GeV, where supernova cooling data [58] provides the lower bound, whereas the upper bound results from the overproduction of the axion.

So, the KSVZ model simultaneously handles two problems of the SM: the first is the solution to the strong CP problem, and the second is that the axion behaves as dark matter. Chapters 4 and 5 will further explore this model from two perspectives: dark matter phenomenology and probing at the LHC.

Experiments have verified that electromagnetic and weak interactions merge at high energies into a single unified electroweak interaction. The SM has gauge group $SU(3)_C \times SU(2)_L \times U(1)_Y$ and many free parameters. Therefore many theorists construct a theory where all those parameters come naturally. The underline theory has a bigger lie group, and the SM gauge groups become its subgroup. It has few coupling constants corresponding to that simple lie group. Therefore instead of three different couplings in the SM corresponding to the three different gauge groups, the unified theory has only one coupling constant at very high energy. The coupling constant splits at low energy due to spontaneous symmetry breaking [59]. Therefore in this unified theory, leptons, quarks, and their anti-particles coexist in the same multiplets.

Many interesting scenarios beyond the Standard Model, for example, the Grand unified theory [59, 60], Pati-Salam model [61, 62], Composite model [63], etc., predict the existence of new heavy particles, known as leptoquarks (LQ). Searching for leptoquarks is an active research area of high-energy physics. The leptoquark is a hypothetical particle that couples quark and lepton together. It carries both the baryon and lepton numbers and provides a means to unify quarks and leptons. In Chapter 6, we will probe different scalar leptoquark models at the LHC and show that they can be discovered at the LHC. Once discovered, the next goal will be distinguishing two scalar leptoquark models. Using polarization variables, we will try to distinguish two leptoquarks of the same electromagnetic charge at the LHC.

The following section will cover the LHC experiment and some features of its detectors that will help to describe phenomenological research in the subsequent chapters.

1.3 The Large Hadron Collider

The Large Hadron Collider (LHC) is a particle accelerator where protons of multi-TeV energies collide, located at CERN [64], Geneva. Two Proton beams are accelerated in the circular tunnels in opposite directions where the tunnels have a circumference of 27 km and about 100 meters underneath the Franch-Switzerland border. Two beams collide head-on at four interaction points - covered by sophisticated particle detectors. 7 and 8 TeV at run 1, 13 TeV at run 2, and presently running nearly at 14 TeV at run 3. There are differences between instantaneous luminosity and integrated luminosity. 3000 fb^{-1} is the total expected integrated luminosity. The primary purpose of the LHC is to know the fundamental building blocks of the Universe and to understand the different fundamental forces that our Universe has. The LHC contains six different experiments in search of different purposes. ATLAS [65] and CMS [66] are two general-purpose detectors designed independently to study the Higgs boson's and top quark's properties and search for new physics beyond the SM. The experiment ALICE [67] is designed to study the heavy-ion collision to understand the structure of quark-gluon plasma, a very dense and hot mixture of quarks and gluons. The temperature in the early Universe, just after Big Bang, would be very high such that it could overcome the strong binding energy resulting from the interaction between quarks and gluons and make a new phase of matter known as quark-gluon plasma. The experiment, LHCb [68], studies heavy flavor physics with the system that contains the b and

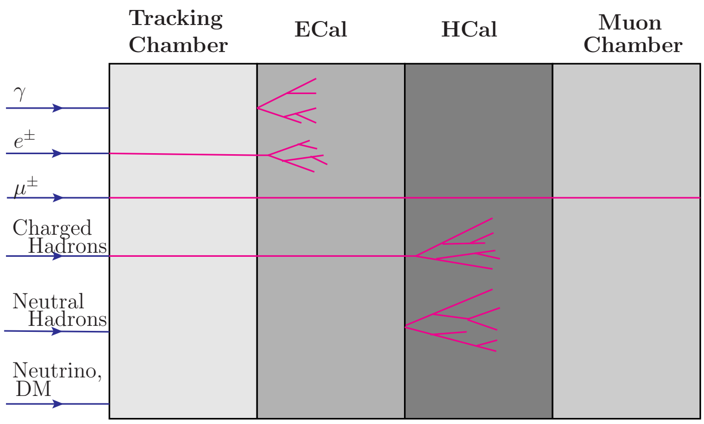


Figure 1.3: Different particles, including dark matter (DM), and their signatures in the various parts of the detector are shown. Neutrino and DM do not leave any signature on any detector parts.

c quarks.

During the collision of the proton beams at LHC, only a tiny fraction of these interactions are hard enough (energy transfer between partons is high) so that part of the energy gets converted into producing new particles according to their nature of interactions. Those new particles open up physics which was unexplored before and can give new interactions. Newly produced heavy fundamental particles are unstable and decay into lighter particles unless some of them are stable following some symmetry of the model and constitute a dark matter candidate. In chapters 3, 4 and 5, we probed different DM-motivated models at the LHC.

The produced particles can have many more possibilities like it can be exotic. One example of an exotic particle is Leptoquark which couples with quark and lepton simultaneously, and we probed different Leptoquark models in Chapter 6. After the cascades of decay of the heavy fundamental particles, the final outcome at the detector is expected to be in the form of some light SM particles like the photon, light leptons, or stable hadrons. The detectors record the produced particles for further study. Below we will briefly describe different parts of the detectors and the reconstruction of objects detected at the hadron collider.

1.3.1 Components of a detectors

The detectors contain the following main parts for measuring the four-momentum of the particles produced at the LHC.

Tracking Chamber is the innermost part of the detector, made of silicon pixels and strips. All the particles produced after the collision go through the Tracking Chamber, which determines the charged particles' trajectories and their electromagnetic energy losses. The curvature of the trajectory is proportional to $\frac{QB}{p}$, where p and Q are the particles' momentum and charge, respectively. The momentum measurement of a highly energetic particle using curvature requires a powerful magnetic field. B is around 3.8 (2) Tesla for the CMS (ATLAS) detectors. The pseudo-rapidity range of the tracking chamber is $|\eta| < 2.4$.

Electromagnetic Calorimeter (ECAL): Electrons and photons deposit their total energy in ECAL. The high-energetic electron and photons show electromagnetic showering through bremsstrahlung and pair production. Bremsstrahlung is a process where a charged particle, when deflected by another charged particle (here, it is the atomic nucleus of the calorimeter), the moving charged particle loses its kinetic energy through the radiation of photons. ECAL measures the energy of the electron and photon with high accuracy and the pseudo-rapidity (η) and azimuthal angle (ϕ) where they deposit their energy with a resolution around 0.025×0.025 in the central region of ECAL. The rapidity coverage of the ECAL is $|\eta| < 3.0$. The photon and electron can be further isolated using the tracker information. The electron exhibits some tracks of the inner tracking chamber while the photons are not.

Hadronic Calorimeter (HCAL): The hadronic calorimeter measures the energy of the neutral and charged hadrons. After hadronization, all colorless hadrons pass through the ECAL and HCAL. The showering of the hadrons involves two distinct parts: (I) an electromagnetic shower and (II) a nuclear shower. Neutral pions π^0s, eta-mesons η^0s, and other mesons decay into photons, which develop an electromagnetic shower. Since π^0 decays almost entirely into two photons, so deposits its energy in ECAL. Few charged light mesons like charged-pions and kaons can lose their few fractions of energy by the bremsstrahlung mechanism and deposit it in ECAL, but most of their energy is deposited at HCAL.

The neutral and charged hadrons that reach the hadronic calorimeter will go into nuclear showering due to the interaction between the incoming hadrons and the HCAL material's atomic nucleus. Elastic and non-elastic nuclear reactions occur between the incoming hadron and the nucleus. If the incoming hadron is highly energetic, then a non-elastic nuclear reaction can happen where new hadrons are produced, and the internal structure of the nucleus is changed. The showering lengths are much longer, and the shape has significant event-by-event fluctuation. Each hadron deposits its energy in one HCAL cell, and since the showering length is large, so to allow complete showering, HCAL cell has to be larger in size. So the resolution of the HCAL is lower than the ECAL. The rapidity coverage of the HCAL is $|\eta| < 5.0$.

Muon Chambers: Muon loses very few fractions of its energy due to the bremsstrahlung mechanism in ECAL and HCAL. Because muon is heavy, it loses less energy and penetrates a larger path into the matter because of its low deceleration rate. The energy loss due to the bremsstrahlung of a charged particle moving in a circular path of radius R is $\Delta E \propto \frac{Q^2}{R}\frac{E^4}{m^4}$, where Q, E, m are the particle's charge, energy, and mass respectively. So, for a muon with the same energy as an electron, the ratio of energy loss of the muon to that of an electron is $\left(\frac{m_e}{m_\mu}\right)^4 = \left(\frac{0.5}{105.66}\right)^4 \approx 5.0 \times 10^{-10}$. The muon chamber is the detector's outermost layer, enclosing the hadronic calorimeter. Muon and its anti-particles will leave tracks in the tracking system, deposit an almost negligible amount of energy in both ECAL and HCAL, and show tracks in the muon chamber. The muon's energy and the momentum direction (η and ϕ of the deposited energy) are obtained from the muon chamber. Sometimes momentum of less-energetic muon can also be obtained from the innermost tracking system.

Figure 1.3 describes the signatures of the objects like electrons, muons, photons, and charged and neutral hadrons in the detector. Dark matter candidates and neutrinos do not show any footprint on the detector. They still keep an imprint from the measurement of the imbalance of momentum components between the initial and final states of such interactions. In a hadron collider, missing transverse momenta (MET or $\not{E}_T$) accounts for such missing particles. It is not possible to reconstruct the longitudinal component of this missing component here.

1.3.2 Reconstructed objects

In the previous subsection, we talked about muons, photons, and $e^{\pm}$ and their different signatures in the detector. In this section, we will briefly talk about other interesting objects seen at the detector. For the identification of an object, isolation of that object is required. We are doing phenomenology using pseudo-data obtained from the package MadGraph5_aMC@NLO [69], and the detector simulation is carried out by the software package Delphes3 [70]. A straightforward definition of the isolation criteria used by Delphes3 is as follows:

$$I(P) = \frac{\sum\limits_{i \neq P}^{\Delta R_{ip} < R,\ P_T(i) > P_T^{\min}} P_T(i)}{P_T(P)} , \tag{1.55}$$

where $P = (e, \mu, \gamma)$ is the particle of interest, and its transverse momentum is the denominator. $P_T(i)$ is the transverse momentum of the i-th particle above $P_T^{\min}$, excluding particle P. ΔR_{ip} represents the angular separation in the $\eta - \phi$ plane between the i-th particle and P. When $I(P) < I_0$, the particle P is considered isolated, and a smaller value for $I(P)$ indicates better isolation. $P_T^{\min}$, R, and I_0 are the three input parameters that users can specify. The default values in DELPHES3 are $P_T^{\min} = 0.5$ GeV, $R = 0.5$, and $I_0 = 0.12$ (for e, γ) or $I_0 = 0.25$ (for μ).

Jets: The fundamental degree of freedom of the SM Lagrangian are leptons, quarks, Higgs, and gauge bosons. However, most can not be seen directly at a hadron collider since heavy particles like W, Z, and top quarks will decay instantly to the quarks and leptons, quarks and gluons fragments and hadronized to form colorless hadrons because of the QCD confinement. The time scale of the QCD confinement is roughly $1/\Lambda_{\text{QCD}} = \frac{1}{(200 \text{ MeV})} \approx 3.3 \times 10^{-24}$ *sec.* So instead of a single fundamental quark or gluon, we see a collimated spray of neutral and charged hadrons. Neutral pion deposits almost its total energy at ECAL, as mentioned earlier. While a few light charged-hadrons lose a small fraction of their energy in ECAL. However, most hadrons deposit their energy in HCAL after a subsequent nuclear shower. Each hadron enters one ECAL and one HCAL cell, as ECAL and HCAL cells are perfectly overlaid. The corresponding HCAL and ECAL cells are grouped to form a calorimeter tower. Each particle corresponds to a single tower. Different towers are used as input for jet formation. The calorimetric towers are clustered in a single object called a jet, using some jet algorithm and recombination scheme, which will be discussed in detail in the next chapter. Those jets are utilized to determine the characteristics of the original gluons and quarks.

Fatjet: The hard interactions with large momentum transfer at the hadronic collider contain many events where massive particles like Z/W/Higgs bosons and top quarks are present in the final state with a large Lorentz boost. Because of the large boost, the hadronic decay products of these massive particles will be highly collimated, and all the decay products can be combined within a single large radius jet called fatjet. Since the QCD jets mimic the fatjets and, at the LHC, the QCD multi-jet production cross section is very large, studying that fatjet is challenging. Jet substructure variables, which examine the internal structure of the jets, are very helpful for the extraction of tiny signals. Further discussion about those variables will be covered in the next chapter.

b-tagging: B-tagging is an essential and effective tool for studying processes with one or more b-jets in the final state. The b quark in the final state can appear in the hard collisions associated with other particles, from the decay of the top quark and Higgs boson and the decay of many BSM particles. For identifying a top-fatjet, b-tagging within the top-fatjet is very effective in enhancing the signal efficiency while decreasing the background efficiency. The b quark forms a B-hadron, which further decays to lighter hadrons. Before decaying into charged particles, a B-hadron traverses a distinguishable distance from the primary collision point, called the secondary or displaced vertex. By identifying those charged tracks and the displaced vertex, one can identify a b-jet. The b-tagging efficiency of the detector depends on the transverse momentum of the b-jet and can reach up to 0.73 [71]. The efficiency lies between 0.7 and 0.73 for b-jet momentum in the 80 GeV to 260 GeV range. Sometimes other particles, like charm quarks, can be misidentified as b-jets. The misidentification rate can reach up to 0.21.

Other heavy states, like tau leptons, decay into lighter leptons ($e^{\pm}$ or $\mu^{\pm}$) or mesons and exhibit displaced vertex signatures. So observing the displacement from the primary collision point of the decay particles, one can identify a τ-jet.

Missing (transverse) energy: As we have seen, the particles with strong interaction will fragment and, after hadronization, produce a collimated spray of hadrons that deposit their energy in HCAL. The electrically charged particles will emit photons that will capture in ECAL. So, the particles with no strong or electromagnetic interactions will leave the detector without leaving any signatures on the detector and will be considered missing particles. The neutrinos of the SM and the dark matter candidates are examples of the missing particles. To study the processes that contain missing particles in the final state, missing transverse energy is a good variable, defined as the negative sum of the transverse momentum of all the visible reconstructed objects at the detector.

$$\not{E}_T = -\sum_i \vec{P_T}(i) \ , \tag{1.56}$$

where i runs over all the visible objects. Throughout the thesis, we will use $\not{E}_T$ or MET as Missing (transverse) energy.

1.3.3 Coordinates of hadron collider

Protons are composite particles. At the LHC, the center of mass frame (CM) of two colliding protons (lab frame) is not the same as that of two partons participating in the hard interaction, as shown below. Consider $p_1 = x_1 P_A$ and $p_2 = x_2 P_B$ are the 4-momentum of two partons, where $P_A = (E_{\rm CM}/2, 0, 0, P_A)$ and $P_B = (E_{\rm CM}/2, 0, 0, -P_A)$ are the 4-momentum of the incoming protons. In the lab frame, the parton system moves with a four-momentum

$$P_{\text{parton}} = \left[(x_1 + x_2)\frac{E_{\text{CM}}}{2}, 0, 0, (x_1 - x_2)P_A\right] . \tag{1.57}$$

Therefore the partonic system has rapidity in the lab frame

$$y_{\text{boost}} = \frac{1}{2} ln \frac{x_1}{x_2} \quad (P_A \approx \frac{E_{\text{CM}}}{2}) . \tag{1.58}$$

The rapidity y of the particle is defined as,

$$y = \frac{1}{2} ln \frac{E + p_z}{E - p_z} . \tag{1.59}$$

Since the initial state partons have no transverse momentum (or negligible), all the quantities in the transverse plane of the parton system are the same as the lab frame. The transverse plane ($X - Y$ or $y - \phi$) is perpendicular to the beam axis, which is considered along the Z direction. ϕ is the azimuthal angle about the Z axis. The particle's transverse momentum is $p_T = |\vec{p}| \sin\theta$, where θ is the polar angle. The particle's components of momentum can be expressed as follows:

$$p_x = p_T \cos\phi,\ p_y = p_T \sin\phi,\ p_z = |\vec{p}| \cos\theta = E_T \sinh y , \tag{1.60}$$

where $E_T = \sqrt{m^2 + p_T^2}$, and y is the rapidity of the particle. The rapidity difference between two particles is invariant under any boost along the longitudinal component of particle momentum, $\Delta y = y_2 - y_1 = y_2^{'} - y_1^{'}$. For massless particles or in the massless limit ($|\vec{p}| \approx E$), the pseudo-rapidity and rapidity of the particle are the same,

$$y = \frac{1}{2} ln \frac{E + |\vec{p}| \cos\theta}{E - |\vec{p}| \cos\theta} = \frac{1}{2} ln \frac{1 + \cos\theta}{1 - \cos\theta} = -ln\left[\tan\frac{\theta}{2}\right] \equiv \eta . \tag{1.61}$$

In the $\eta - \phi$ plane, the 4-momentum of a particle can be written as,

$$p^\mu = (E, \vec{p}) = (E_T \cosh y, p_T \cos\phi, p_T \sin\phi, E_T \sinh y) . \tag{1.62}$$

Another important variable is the angular separation between two particles in the $\eta - \phi$ plane,

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} . \tag{1.63}$$

In this section, we discussed the hadron collider and the minimal aspects of event reconstruction while ignoring many important experimental aspects.

The first important fact is that we discussed the OFFLINE event reconstruction, recorded in the storage. However, around 40 collisions at the LHC between two proton brunches happen every microsecond. A small momentum transfer happens between the partons in most of the collisions, producing the known SM particles and hence uninteresting. Nevertheless, few collision occurs through large momentum transfer, and hence interesting. We generally do not store the uninteresting collisions because of the storage capabilities, but we store the interesting events. So, experimentalists need an online trigger [72–74] for an unbiased selection of important events.

The other important fact is that several collisions per bunch-crossing occur in high luminosity at the LHC. Among those collisions, mostly one is the hard interaction, and other collisions between two protons occur through a soft exchange of gluon called pileup events. So the jets of the hard interaction is significantly affected by the soft particles originating from the pileup events. The charged particles that originate from the pileup events can be removed from the particle lists of the hard interaction using the tracking information and vertex reconstruction if the secondary collision points are resolved from the primary vertex. Collision points at the LHC are identified from where tracks originate; among those, the primary vertex is identified that is associated with the highest energy constituents. Neutral particles do not show tracks, so extracting the neutral particles of the pileup events is very challenging.

Moreover, extracting charged particles is also challenging if the secondary vertex is close to the primary vertex. Pileup subtraction [75–80] is vital for accurately predicting the QCD jets originating from hard interaction. Pileup subtraction has a great experimental and phenomenological interest.

1.4 Importance of doing QCD corrections

The colliding particles at the LHC are the protons composed of quarks and gluons, collectively called partons. The partons interact through strong interactions. The dynamics governing the strong interactions among gluons, quarks, and antiquarks are described by QCD, as discussed in Subsection 1.1.1. The strong coupling constant becomes small at large energy, whereas it is very strong at low energy. At the LHC, the energy scale of the hard interaction is large, so the coupling constant is a small parameter. Hence, we can safely apply the perturbation method in the series expansion in terms of α_s for the commutation of QCD observables.

The high energetic collision between two protons at the LHC usually breaks the protons into their constituent partons. Therefore, the high momentum transfer hadronic cross section can be factorized into parton level scattering cross section convoluted with the parton distribution function. At the hadron collider, the inclusive cross section for the final state F (for example, a Higgs boson or a pair of top quarks) can be written as follows.

$$\sigma(h_1 h_2 \to F{+}X) = \sum_{a,b\in\{q,\bar{q},g\}} \int_0^1 dx_1 \int_0^1 dx_2\, f_a^{h_1}(x_1,\mu_F^2) f_b^{h_2}(x_2,\mu_F^2)\hat{\sigma}(a,b \to F{+}x) \tag{1.64}$$

h_1 and h_2 are two colliding protons at the LHC, and X is any extra QCD radiation that appears with the final state F. In collinear factorization, the function $f_a^h(x,\mu_F^2)$ is called the parton distribution function (PDF), which is the resolving probability of a parton of species a carrying a momentum fraction x of the parent hadron. The scale at which the collinear factorization occurs is called factorization scale μ_F. Since PDFs contain a low energy part of the cross section, their computation can not be done within the perturbative technique as the strength of α_s is large. They are typically fitted from the experiment for various x and μ_F values. $\hat{\sigma}$ is the partonic scattering cross section calculable in the perturbation method and can be written as follows.

$$\begin{aligned}\hat{\sigma}(a,b \to F+x) = \hat{\sigma}^{(0)}&(a,b \to F)\\ &+ \hat{\sigma}^{(1)}(a,b \to F + \text{ up to 1 parton, } \mu_F, \alpha_s(\mu_R))\\ &+ \hat{\sigma}^{(2)}(a,b \to F + \text{ up to 2 parton, } \mu_F, \alpha_s(\mu_R))\\ &+ \cdots\end{aligned} \tag{1.65}$$

The first non-zero term is called the tree-level partonic cross section, the first term of the above equation. The second term produces the final state F with one extra parton, called the next-to-leading (NLO) cross section. NLO-QCD cross cross section has coupling order $\mathcal{O}(\alpha_s^{n+1})$ given LO cross section has order $\mathcal{O}(\alpha_s^n)$; NLO is the first correction term to the LO cross section. The third and rest of the terms are NNLO and higher order correction terms in the perturbation series.

The hadronic cross section is always inclusive, even for the LO. In the case of LO calculation, any extra QCD radiation associated with the final state F, up to transverse momentum μ_F, will be accounted for by PDFs of the initial state partons. However, any hard radiation (resolvable partons) is neglected as those are of higher order in α_s.

NLO or higher order computations will be more accurate than the LO cross section, so the computation of the partonic cross section beyond LO is required to compare the theoretical prediction with the experimental data. NLO correction often modifies the LO cross section by a factor of two. The K factor, the NLO to LO cross section ratio, can reach up to two for some SM processes, like the production of the Higgs boson [81–83] or $Wb\bar{b}$ [84,85]. Additionally, at higher orders, some new production channels can open up. Therefore, the NLO enhancement is vital for accurately predicting the SM processes and discoveries beyond SM.

The LO cross section has no divergence, but starting from NLO, the amplitude-representing Feynman diagrams have different divergences. The first type is the ultraviolet (UV) divergences because of the large loop momentum of the Feynman diagram. UV divergences can be renormalized by assuming all the quantities (couplings, masses, and fields) present in the Lagrangian are the bare parameters that absorb the UV divergences to give renormalized (finite) couplings and masses. As a result, starting from NLO, the strong coupling constant becomes the function of the renormalized scale μ_R. However, the LO Feynmann diagrams have no UV divergences, so all the bare quantities equal their renormalized version. Therefore leading-order partonic cross section is not a function of μ_R.

The second types of divergences are the infrared (IR) and collinear divergences. The collinear splitting of partons provides collinear divergences. Any incoming or outgoing quarks and gluons can undergo many collinear splittings, and each splitting has a large logarithm $\log(Q/\mu_F)$, where Q is the hard interaction scale. For example, if a parton (quark or gluon) splits and produces a collinear gluon, the contribution would be as follows.

$$\hat{\sigma}(1 \text{ splitting}) = \hat{\sigma}^{(0)} \frac{\alpha_s(Q)}{2\pi} \log \frac{Q^2}{\mu_F^2} \int_0^1 dz \; P_{g \leftarrow a}(z) \;, \tag{1.66}$$

where $P_{g\leftarrow a}(z)$ is the Altarelli-Parisi splitting function, given in Appendix A. The strong running coupling constant $\alpha_s(Q)$ is evaluated at Q. The resolution scale μ_F will always be greater than the Λ_{QCD} as at scale $\Lambda_{QCD} = 0.2\ GeV$, hadronization starts, and the perturbative method does not work below that scale. For multiple collinear splittings, the above equation modifies to

$$\hat{\sigma}(n \text{ splittings}) = \hat{\sigma}^{(0)} \frac{1}{n!} \Big(\frac{\alpha_s(Q)}{2\pi}\Big)^n \; \log^n \frac{Q^2}{\mu_F^2} \Big[\int_0^1 dz \; P_{g\leftarrow a}(z)\Big]^n \;. \tag{1.67}$$

The Parton distribution functions account for all of those large logarithms. Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) [86–89] equations provide the evolutions of the PDFs from one factorization scale to another factorization scale.

$$\mu_F^2 \frac{df_a(x,\mu_F^2)}{d\mu_F^2} = \frac{df_a(x,\mu_F^2)}{d\log(\mu_F^2)} = \sum_{b\in\{q,\bar{q},g\}} \int_x^1 \frac{dz}{z} \frac{\alpha_s(\mu_F)}{2\pi} \; P_{a\leftarrow b} \; f_b(x/z, \mu_F^2) \tag{1.68}$$

Therefore, re-summation of all the large logarithms of collinear splitting (both virtual and real splittings) of the colliding partons is taken care of by the DGLAP equations and captured in the PDFs, making the partonic scattering cross section into collinear finite.

Collinear finite NLO or higher-order partonic cross sections have other soft divergences because of the soft virtual and real emissions. The next chapter will cover an explicit example of soft and/or collinear divergences that appear at NLO. The KLN theorem states that when we add the virtual and real corrections, the divergences cancel order by order in α_s, leaving a finite correction at NLO.

1.4.1 Parton Shower

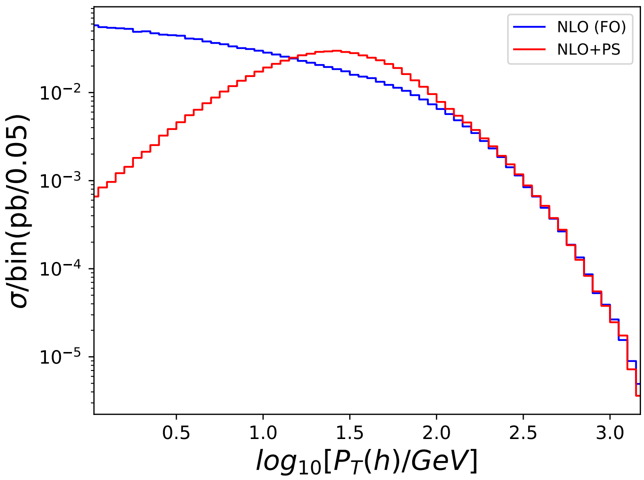


Figure 1.4: The transverse momentum of the Higgs boson (h) for the processes $pp \to h$ at fixed order NLO (blue) and at NLO+PS (red) are shown. $P_T(h)$ is equal to the transverse momentum of the NLO radiation. The NLO (FO) result has divergence at a small P_T.

The partonic cross section at NLO contains the final state F with one extra resolvable parton. The extra NLO radiation can be collinear with the incoming partons and other quarks or gluons if the final state F has. Therefore, we must add collinear counter terms to the real emission cross section to cancel those leftover collinear divergences. For example, if the LO partonic process is $gg \to h$, then at NLO, one of the real emission diagrams will be $gg \to hg$. The collinear counter term for gluon radiation is as follows:

$$\begin{aligned}\hat{\sigma}_{ct}(gg) &= 2\hat{\sigma}^{(0)}\frac{\alpha_s(Q)}{2\pi}\Big[\Big(\frac{Q^2}{\mu_F^2}\Big)^{\epsilon}\frac{C_\Gamma}{\epsilon}\int_0^1 dz\ P_{g\leftarrow g}(z)\Big] \\ &= 2\hat{\sigma}^{(0)}\frac{\alpha_s(Q)}{2\pi}\Big[\frac{C_\Gamma}{\epsilon}\int_0^1 dz\ P_{g\leftarrow g}(z) + \log\frac{Q^2}{\mu_F^2}\int_0^1 dz\ P_{g\leftarrow g}(z)\Big]\end{aligned} \tag{1.69}$$

where $\hat{\sigma}^{(0)}$ is the LO partonic cross section, and the expression of C_Γ is given in

Equation B.8. Because the gluon can be collinear with both incoming gluons, factor 2 is present. The first part of the above equation is the divergence part that will cancel the collinear divergences, but the second term is the finite term which contains a log term. In general, if a parton in the process is collinear with the NLO radiation (j), the real cross section will have the following term for every such parton.

$$\hat{\sigma}^{(0)} \int_{\mu_F^2} \frac{dt}{t} \frac{\alpha_s(t)}{2\pi} \int_0^1 dz\ P_{j\leftarrow a}(z) = \hat{\sigma}^{(0)} \log \frac{P_T^2(j)}{\mu_F^2}\ \frac{\alpha_s(P_T(j))}{2\pi} \int_0^1 dz\ P_{j\leftarrow a}(z) \tag{1.70}$$

For the full phase space coverage, one needs to replace $P_T(j)$ with its maximum possible value, which is the hard interaction scale, Q, making the total fixed-order cross section finite since the ratio Q/μ_F is close to 1. We will set μ_F at the order of Q, as there is no other scale than the hard interaction. Here, t is the evolution variable, equal to the square of the transverse momentum of the radiation, and z is its momentum fraction taken from parton-a.

Fixed-order distributions provide correct predictions in the phase space when the ratio $P_T(j)/\mu_F$ is close to one, as log terms will have a tiny value in those regions. On the other hand, in the phase space regions where the ratio is very small or very large, which means $P_T(j)$ and μ_F are very far apart, log terms will have a large value. As a result, the perturbativity breaks down in those regions, making fixed-order results unreliable. Resumed predictions in those regions can rescue us. The leading part of soft emissions is universal. Re-summation is the process of identifying all the large logs in all orders of a perturbation series and exponentiating them. It can be carried out analytically (fully inclusive) or numerically. Parton shower (PS) is an exclusive numerical re-summation, and their merging/matching is done with the fixed-order prediction. As a result, NLO+PS results are free from all of those large logarithms, and the differential distributions of any observables are reliable over the entire phase space.

We will set μ_F to the order of Q throughout our analysis; thus, in the parts of the phase space where $P_T(j)$ is very small (threshold production), fixed order differential results contain large logs, and we require resumed predictions in those regions. However, at large $P_T(j)$ regions, the fixed-order predictions are correct.

In Figure 1.4, we plot the Higgs transverse momentum at NLO for the production of the Higgs boson at the LHC. At NLO, the Higgs transverse momentum equals the transverse momentum of the NLO radiation, and we find that the fixed-order distribution, blue line, has a divergent at low $P_T(j)$ regions, while the NLO+PS result is smooth.

The heart of the parton shower is the Sudakov form factor, which is the probability of no splitting of a parton between two scales given below.

$$\Delta_a(t_f, t_0) = \exp\Big[-\sum_{b\in\{q,\bar{q},g\}} \int_{t_0}^{t_f} \frac{dt}{t} \int dz \frac{d\phi}{2\pi} \frac{\alpha_s(t)}{2\pi} P_{b\leftarrow a}(z)\Big] \tag{1.71}$$

Here, parton-a splits into two collinear partons, b and c ($a \to bc$), and t is the evolution variable which equals P_T^2 of the parton-b or virtuality of the parton-a ($t = m_a^2 = (p_b + p_c)^2 \approx 2E_bE_c(1 - \cos\theta)$, θ is the angle between b and c). In the collinear limit, t goes to zero. In the above expression, z is the momentum fraction of the parton-b taken from parton-a. The splitting kernels control the soft behavior; parton-b is soft when $z = 0$, and parton-c is soft when $z = 1$.

Although there are formal ways of deriving the Sudakov form factor, here we are giving a simple-minded derivation. In statistics, a random variable has a Poisson distribution if a finite number of independent events occur at a constant rate in an interval. The Poisson distribution function is given below for n number of events that occur in the interval with a probability λ.

$$f(n;\lambda) = \frac{\lambda^n \exp(-\lambda)}{n!} \tag{1.72}$$

The probability of branching a parton of species a at evolution scale t_i within an interval between t_i and $t_i + \delta t_i$ is as follows:

$$\mathcal{P}_{\text{branching}}(t_i) = \sum_{b\in\{q,\bar{q},g\}} \frac{\delta t_i}{t_i} \int dz \frac{d\phi}{2\pi} \frac{\alpha_s(t_i)}{2\pi} P_{b\leftarrow a}(z)\ . \tag{1.73}$$

The probability of no branching at scale t_i can be obtained from Equation 1.72 after setting $n = 0$ and the branching rate equal to $\mathcal{P}_{\text{branching}}(t_i)$, which gives the following result.

$$f(n;\lambda) = \exp(-\lambda) = \exp[-\mathcal{P}_{\text{branching}}(t_i)] \tag{1.74}$$

We must integrate the above equation to obtain the probability of no splitting between two evolution scales, which gives the Sudakov form factor.

$$\exp[-\int_{t_0}^{t_f} \mathcal{P}_{\text{branching}}(t_i)] \equiv \Delta_a(t_f, t_0)\ . \tag{1.75}$$

Following the previous discussions, the amplitude squared of real emission diagrams (R) can be represented as $R = R^s + R^f$, where R^f represents the finite component, and R^s contains the large logarithmic terms known as the threshold-singular part. The phase space of the singular part can be split into the born phase space times the phase space of the NLO radiation $d\Phi_j = dz\frac{dt}{t}\frac{d\phi}{2\pi}$, where the evolution variable t can be written as the transverse momentum square of the radiation $P_T^2(j)$. However, this type of decomposition is not possible for the finite part. R^s can always be written as the Born times, the splitting kernel, as follows:

$$R^s = B(\Phi_B) \times \sum_{b\in\{q,\bar{q},g\}} \frac{\alpha_s(t)}{2\pi} P_{b\leftarrow a}(z) \ . \tag{1.76}$$

The differential fixed-order NLO cross section can be written as below, which has divergent at low $P_T(j)$ regions.

$$d\sigma^{\text{NLO}} = d\Phi_B\Big[B(\Phi_B) + V(\Phi_B) + d\Phi_j R^s\Big] + d\Phi_R R^f(\Phi_R) \ . \tag{1.77}$$

By integrating the above expression over the entire phase space, one gets the total NLO fixed-order cross section which is finite. The differential NLO plus parton shower cross section can be written as follows:

$$d\sigma^{\text{NLO+PS}} = d\Phi_B \bar{B}^s\Big[\Delta_a + d\Phi_j \frac{R^s}{B(\Phi_B)}\Delta_a\Big] + d\Phi_R R^f(\Phi_R) \ , \tag{1.78}$$

where the function $\bar{B}^s = B(\Phi_B) + V(\Phi_B) + \int d\Phi_j R^s$. The function $\bar{B}^s$ is finite at all the phase space regions, even at low $P_T(j)$, since the singular part R^s is integrated over all the phase space of the radiation. Using Equation 1.76, we can write,

$$d\sigma^{\text{NLO+PS}} = d\Phi_B \bar{B}^s\Big[\Delta_a + \sum_{b\in\{q,\bar{q},g\}} dz\frac{dt}{t}\frac{d\phi}{2\pi}\frac{\alpha_s(t)}{2\pi} P_{b\leftarrow a}(z)\Delta_a\Big] + d\Phi_R R^f(\Phi_R) \ . \tag{1.79}$$

Expanding the square bracket up to the first order in α_s gives the following:

$$\begin{aligned} d\sigma^{\text{NLO+PS}} = d\Phi_B \bar{B}^s\Big[1 - &\sum_{b\in\{q,\bar{q},g\}} \int \frac{dt}{t}\int dz \frac{d\phi}{2\pi}\frac{\alpha_s(t)}{2\pi} P_{b\leftarrow a}(z) \\ &+ \sum_{b\in\{q,\bar{q},g\}} \frac{\alpha_s(t)}{2\pi} dz \frac{dt}{t}\frac{d\phi}{2\pi} P_{b\leftarrow a}(z) + \mathcal{O}(\alpha_s^2)\Big] + d\Phi_R R^f(\Phi_R) \ . \end{aligned} \tag{1.80}$$

The hadronic cross section is inclusive, so the effect of the parton shower should be unitary, which means the addition of any radiation should not change the cross section. Integrating the above equation, we find that the second and third terms cancel each other, and the square bracket equals 1. Therefore, the total NLO+PS cross section is equal to the NLO cross section. Another thing is the cancellation of large logs at the differential level. To get the NLO+PS cross section at a particular phase space region, say between t and $t+dt$, and for a particular z, the second and third terms are exactly equal and opposite, so they cancel. Therefore, large logs at each phase space point cancel, making the differential distributions reliable at every phase space point. One can see that the NLO+PS distribution, the red line in Figure 1.4, has no divergence even at low $P_T(j)$, and as expected, at large $P_T(j)$, the NLO+PS distribution matches the NLO distribution.

1.4.2 Scale uncertainties

μ_F and μ_R are two spurious scales that appear because of the collinear factorization and renormalization. Hadronic cross section should not depend on the scales μ_F and μ_R as the hadronic cross section is a renormalization group invariant with these scales. However, truncating the perturbative series to a finite order is not renormalization group invariant which introduces the scale uncertainties in the cross section at any finite/fixed order computation. This dependency typically becomes significant at low orders. Therefore to get a reliable prediction, we have to calculate higher orders.

LO partonic cross is finite, so independent of any factorization and renormalized scales. Therefore, the factorization scale dependence of the LO hadronic cross section comes only from the PDFs, which is significant. However, starting from NLO or higher order, the partonic cross section depends logarithmically on μ_F and μ_R, where the dependency on μ_R comes through the renormalized strong coupling constant $\alpha_s(\mu_R)$. As a result, NLO or higher order hadronic cross section has smaller scale uncertainties than LO.

The hadronic scale, Q, is only the scale of the hard interaction; therefore, μ_F is usually chosen in order of Q. We vary these scales in the interval $Q/2 \leq \mu_F, \mu_R \leq 2Q$, which produces nine data sets for different $\{\mu_F, \mu_R\}$ combinations. The envelope of these nine data sets is given as scale uncertainty. In the upcoming chapters of this thesis, we will find that NLO scale uncertainties are always much smaller than LO uncertainty.

Therefore, the NLO correction has a series of advantages as follows.

- NLO, which computes one order more in α_s in the perturbation series than LO, gives a more precise cross section. Furthermore, at NLO, some new production channels can open up.

- Scale uncertainties of various observables at NLO are much smaller than LO, giving precise predictions of those observables at the LHC.

- To distinguish the signal from the background, we utilize multiple observables. The differential K-factor can exhibit significant variations in different kinematic regions. Moreover, the differential K-factor can differ for different observables. It is important to note that the leading-order event multiplied with the total NLO K-factor alone can not provide accurate predictions because of variations in differential distributions. Therefore, in this thesis, we consider the full NLO events to achieve more precise and reliable predictions.

- Finally, since we produce events with the parton shower, the distributions over the entire phase space are meaningful.

1.5 Brief outline of the Thesis

During this challenging time for particle physicists, tremendous efforts are pouring in to find any hints for the physics beyond the Standard Model that has remained elusive so far. On the other hand, we are at the beginning of the high luminosity era at the Large Hadron Collider when we acquire an order of magnitude more high energy data for the next decade or so. To combat this formidable adversary, we require to focus on a few clear theoretical and phenomenological directions, which are

1. precise computation and estimate of new physics processes

2. new search strategies potent enough to find the anomalous events probably hidden within the vast pool of QCD background, and finally

3. more sophisticated signal-background analysis consistent with present developments in boosted decision trees and machine learning techniques.

In this thesis, we consider different new physics models motivated by critical theoretical and experimental requirements, such as dark matter, Strong CP problems etc., and investigate their phenomenological signatures in the light of above mentioned three outlooks.

- New physics processes are examined at next-to-leading order in QCD, matched to parton shower for partonic hard processes at the LHC. Parton shower (PS) resum all the leading large logarithms terms, and the differential distributions of any observables become reliable over the entire phase space. We consider leptoquark models and some dark matter motivated BSM models like the inert Higgs doublet model and complex scalar extended KSVZ model. We find that order α_S corrections of those BSM models are very significant. Interestingly, the LO event normalized by the total NLO K-factor will not give an accurate result. This is because we use different high-level variables in signal and background analysis and apply cuts on those variables for optimization; therefore, the differential K-factor in those phase spaces is crucial in extracting the correct efficiencies. Furthermore, NLO-QCD events have much fewer factorization and renormalization scale uncertainties than the LO events, making our result more accurate.

- Numerous BSM scenarios are under investigation at the LHC and other experimental endeavors. As a result, the parameter space related to these scenarios has been subjected to substantial constraints, with particular emphasis on the masses of heavier modes and their corresponding couplings. While probing at LHC, such heavy BSM particles would naturally produce boosted W/Z boson or boosted top quark if such decay modes are present. Thus identifying such top quarks as three-prong top-like fatjet or those W/Z bosons as two-prong fatjets gives a better detection technique. Here one finds two-fold advantages in terms of hadronic branching ratios and reconstructing such a hadronic jet carrying the inherent signature of its mother particle, unlike leptonic modes with some of the missing neutrinos. Of course, that is at the expense of having a more complex object like boosted jets that can be easily mimicked by the abundance of QCD radiation, especially while working at a hadron collider. However, tremendous developments in this direction for the past decade or so can be extremely handy looking at the substructures of these jets. Therefore we consider fatjets+x, where x can be missing transverse momentum or anything else depending on the signal topology, as a potential signal in our final state. Since large radius jets (fatjets) are susceptible to soft radiation from underline or pile events, we use jet grooming techniques to remove the unassociated soft and wide-angle radiations. We also use jet-shape observable like the N-subjettiness ratio to identify the radiation pattern of a jet.

- We generated our signal and all possible background events with particle-level simulation to get a reliable estimate, including the detector-level inputs in our analysis. This is done using the publicly available package DELPHES3 [70] with the default Delphes CMS card. We also employed a sophisticated multivariate analysis of high-level variables. Jet- substructure variables are also adopted in our study. The multivariate gradient boosting technique outperforms the traditional cut-based analysis since it employs many non-linear cuts on the variables to extract the signal from the background.

The thesis will be structured as follows:

Chapter 1 presents a brief overview of the SM and its shortcomings. We listed many reasons to look for new physics beyond the Standard Model of particle physics. The importance of including higher-order corrections to predict any new physics signature precisely is discussed. We also briefly overview the parton shower and scale uncertainties.

The higher-order calculations involve IR and UV divergences. In any renormalized theory, UV divergences are absorbed by the counter terms, and soft and collinear divergences cancel when we combine the virtual and real corrections, leaving only the finite parts. Chapter 2 starts with an explicit example of soft and collinear divergences in a one-loop virtual diagrams calculation. This chapter also includes the methodology for dealing with various collider searches. A brief discussion of jet algorithms and jet substructure variables is covered. We also discuss the decision tree and the boosted decision tree algorithms that we employ in our analysis.

Chapter 3 explores the inert Higgs-doublet model (IDM), a simple extension of the Standard Model. IDM provides a viable Higgs portal, scalar dark matter candidate, and we probe its hierarchical mass spectrum. The effects of next-to-leading order (NLO) QCD corrections are considered. We find such corrections significantly impact various kinematic distributions and reduce scale uncertainties substantially. Fixed order NLO results are matched to the Pythia8 parton shower (PS), and the di-fatjet signal associated with the missing transverse momentum is analyzed, as this channel has the ability to explore its entire parameter space during the next phase of the LHC run.

In Chapters 4 and 5, we study the extended KSVZ model that, in addition to providing a natural solution to the strong-CP problem by including a global Peccei-Quinn symmetry, also furnishes two components of dark matter that satisfy observer relic density without fine-tuning of model parameters. This hybrid setup incorporates an extra $SU(2)_L$ complex singlet scalar whose lightest component plays the role of one of the dark matter, while the QCD axion of the KSVZ model acts as a second dark matter candidate. In those chapters, we focus on accentuating the role of vector-like quark that naturally emerges in the KSVZ model on dark matter and collider phenomenology. The presence of this colored particle can significantly affect the allowed parameter space of the scalar dark matter by opening up additional co-annihilation and direct detection channels.

This model has Yukawa interaction between VLQ, scalar DM, and the up-type quarks of the SM, $f_i(S\overline{\Psi}_L u_{i,R} + h.c)$, where $i = u,\ c,\ t$. Chapter 4 focuses on the top-philic Yukawa interaction, $f_{u,c} << f_t \simeq 1$. In contrast, in Chapter 5, we look for all Yukawa couplings to have equal strength (democratic). However,

$D^0 - \bar{D}^0$ oscillation forced us to take one of the f_u, f_c is zero (or tiny) while the other two are democratic. Two different options, top-philic and democratic, lead to very different parameter spaces of the scalar DM that gives the correct relic density and are allowed by direct and indirect detection experiments.

The collider search of the top-philic case is almost model-independent, as pair production of VLQ at the LHC has a negligible dependence on those BSM couplings since those couplings ($f_{u,c}$) are very small, therefore, are solely driven by the strong coupling. After production, each VLQ decay to the top quark associated with the scalar DM with a branching ratio of one. However, for the democratic case, the pair production depends on strong coupling and those Yukawa couplings. Furthermore, each VLQ has two decay channels: decay into a top or up quark with the scalar DM. Hence, the branching ratio of VLQ in the top quark is less than 0.5.

The Yukawa interactions provide a unique topology, two boosted-top fatjets with considerable missing transverse momentum, as a promising signature to probe the parameter spaces of this model at the LHC. Moreover, in Chapter 5, we consider the next-to-leading order NLO-QCD correction for VLQ pair production for more precise predictions.

In Chapter 6, we explore the pair production of third-generation scalar leptoquark at the LHC to next-to-leading order accuracy in QCD, matched to parton shower for a precise probing of the stemming model. We propose to tag two boosted top-like fatjets produced from the decay of heavy leptoquarks in association with significant missing transverse momentum and consider them the potential signal. Such a signal demonstrates the capability of a robust discovery prospect in the multivariate analysis. Various scalar leptoquark models predict different chirality of the top quark appearing from the decay of the leptoquark carrying the same electromagnetic charge. We use the polarization variables sensitive to the top quark polarization to identify the underlying theory.

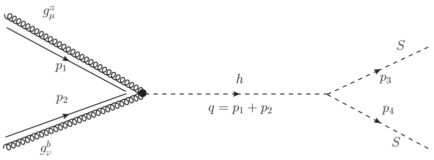
In Chapter 7, a summary of the thesis and future prospects will be presented.

Chapter 2

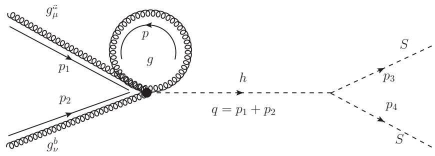
Methodology

In the previous chapter, we saw the importance of higher-order corrections to facilitate the discovery of any new physics beyond the Standard Model and any precise prediction within the SM. We discussed that the higher-order corrections contain the divergences, so in this chapter, we start by giving an example of how soft and collinear divergences appear at higher-order virtual corrections. After that, we discuss the theoretical basis of the hadronic final states. As mentioned earlier, high-energetic quarks and gluons at the LHC decrease their energy through soft and/or collinear radiations, and low-energy partons hadronize to form colorless hadrons, resulting in a multitude of particles in the final state. Therefore, to analyze the many-particle final state at the LHC, we need to define jets, which need a set of rules to group a certain number of particles within a jet and assign a four-momentum to the resulting jet. Therefore, we give infrared and collinear (IRC) safe jet algorithms in Section 2.2.

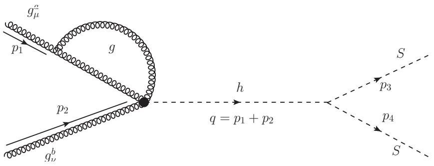
Searching for new physics beyond the SM using ordinary jets is extremely difficult because of the large cross section of the QCD multi-jet production and other SM backgrounds. Hence, in search for these BSM scenarios, we need some excellent variables that can separate those tiny signals from the overwhelming SM background. Such suitable variables can be jet substructure variables. The heavy BSM particles subsequently decay into top quark/Higgs bosons or weak bosons, and in TeV-scale BSM theories, those top, Higgs, and weak bosons have enough boost as they come from the decay of the heavy particles. Consequently, the hadronic decay products of the top quarks, Higgs, and weak bosons are collimated and form a single large-radius jet called fatjet. Jet substructure variables look inside those large-radius jets and distinguish them from the ordinary QCD jet background. The different jet substructure variables and their definitions are covered in this chapter. Moreover, we do the multivariate analysis described in Section 2.4 to optimize the collider search.



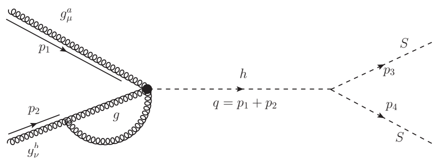
(a) LO, $\mathcal{M}_0$



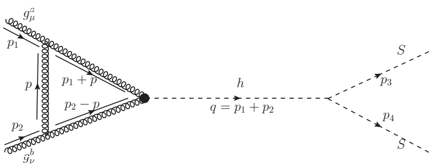
(b) $\mathcal{M}_1$



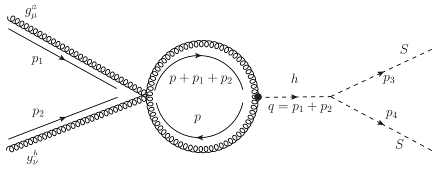
(c) $\mathcal{M}_2$



(d) $\mathcal{M}_3$



(e) $\mathcal{M}_4$



(f) $\mathcal{M}_5$

Figure 2.1: Pair Production diagrams of two scalars through gluon fusion at the LHC are shown – (a) LO diagram, and the rest are the virtual diagrams. LO diagram with an extra gluon radiation from the initial state gluons are the real-emission diagrams, which are not shown.

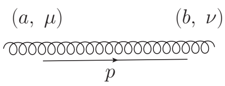
2.1 An explicit example of soft and collinear divergences

Higher-order calculations for Higgs production through gluon fusion are available in the literature. In this section, we will present a case study of pair production of scalars through an s-channel Higgs boson at the LHC at the next-to-leading order and demonstrate the shape of soft/infrared (IR) and collinear divergences of the virtual diagrams (the results of those calculations are also used in Chapter 3). Higgs boson (h) can interact with a pair of BSM scalars (S) (an example of such a model is outlined in Chapter 3). The hSS coupling can be written as $-i\Lambda v$, where Λ and v are constants. Since Higgs boson is a color singlet, it can not interact directly with gluons. The leading-order (LO) diagram of ggh occurs through the quark loops. The top-quark loop contributes the most because of its large Yukawa coupling. NLO computation of this process involves technically complicated two-loop integrals. An NNLO computation is carried out in the large top mass limit since computing the two-loop integral is challenging. We can integrate out the top quark degrees-of-freedom from the loop in the large top mass limit, which results in an effective Lagrangian with Higgs boson-gluon interactions at the tree level. Hence, the NNLO computation eventually reduces to a one-loop computation.

The effective Lagrangian is given in Equation C.13. New tree-level ggh, $gggh$, and $ggggh$ vertices appear from the effective Lagrangian, and the Feynman rules are outlined in Equations C.15 -C.17. The LO and virtual diagrams of $gg \to SS$ are shown in Figure 2.1.

We will calculate the square amplitudes in $d = 4 - 2\epsilon$ dimensions, where ϵ is infinitesimal. To keep the action dimensionless in d dimension, one must replace $g_s \to (\mu)^\epsilon\, g_s$, where μ is an arbitrary scale with mass dimension one. We will do the calculations in the axial gauge. However, one can choose any other gauge since the amplitude square should be independent of the choice of gauge. In axial gauge, $SU(n)$- non-abelian theory contains no ghost fields. The summation of gluon polarization vectors and the gluon propagator in the axial gauge are given in Equations 2.1 and 2.2.

$$\sum_{\text{polarization}} \epsilon^{a\,\mu}(p)\,\epsilon^{*\,b\,\nu}(p) \quad = \quad \delta^{ab}\,(-g^{\mu\nu} + \frac{p^\mu n^\nu + p^\nu n^\mu}{p.n}) \tag{2.1}$$



$$= \delta^{ab}\,\frac{i}{p^2 + i\epsilon}\,(-g^{\mu\nu} + \frac{p^\mu n^\nu + p^\nu n^\mu}{p.n}), \tag{2.2}$$

where n_μ is an arbitrary light-like four-vector ($n^2 = 0$), and gauge invariance demand amplitude square should be independent of n_μ.

Leading-order in d dimensions The LO amplitude (Figure 2.1a) in $d = 4-2\epsilon$ dimension can be written as below.

$$i\mathcal{M}_0 = -iC_{eff}\delta^{ab}\,(p_1^\nu p_2^\mu - g^{\mu\nu}\frac{\hat{s}_{12}}{2})(-i\Lambda v)\frac{1}{(q^2 - m_h^2) + i\Gamma_h}\epsilon_\mu^a(p_1)\epsilon_\nu^b(p_2), \tag{2.3}$$

where m_h is the mass of the Higgs boson, and Γ_h is the mass times width of the Higgs boson.

$$\hat{s}_{12} = (p_1 + p_2)^2 = q^2 = 2p_1.p_2 \quad (p_1^2 = p_2^2 = 0) \tag{2.4}$$

$$|\mathcal{M}_0|^2 = \frac{1}{2}\frac{\mu^{2\epsilon}\,C_{eff}^2\Lambda^2 v^2}{(\hat{s}_{12} - m_h^2)^2 + \Gamma_h^2}(p_1^\nu p_2^\mu - g^{\mu\nu}\frac{\hat{s}_{12}}{2})(p_1^\alpha p_2^\beta - g^{\alpha\beta}\frac{\hat{s}_{12}}{2})\epsilon_\mu^a(p_1)\epsilon_\nu^a(p_2)\epsilon_\beta^{*b}(p_1)\epsilon_\alpha^{*b}(p_2) \tag{2.5}$$

Two is the symmetry factor. The numerator includes the $\mu^{2\epsilon}$ factor because the cross section has mass dimension $[\sigma] = [M]^{2-d}$ in d dimensions, and phase space does not include any additional ϵ term. Gluon is the gauge boson of the $SU(3)$ gauge group and has eight different color states, and in the d dimension, gluon has $d - 2$ polarization states. We have the following after taking the color and polarization average of the initial state gluons.

$$\begin{aligned}\overline{|\mathcal{M}_0|^2} =& \frac{1}{8}\frac{1}{8}\frac{1}{2}\frac{\mu^{2\epsilon}\ C_{eff}^2\Lambda^2 v^2}{(\hat{s}_{12}-m_h^2)^2+\Gamma_h^2}(p_1^\nu p_2^\mu - g^{\mu\nu}\frac{\hat{s}_{12}}{2})(p_1^\alpha p_2^\beta - g^{\alpha\beta}\frac{\hat{s}_{12}}{2}) \\ & [\frac{1}{d-2}\sum_{\text{pol}}\epsilon_\mu^a(p_1)\epsilon_\beta^{*b}(p_1)]\ [\frac{1}{d-2}\sum_{\text{pol}}\epsilon_\nu^a(p_2)\epsilon_\alpha^{*b}(p_2)]\end{aligned} \tag{2.6}$$

We obtain the following using Equation 2.1 and $\delta_\mu^\mu = d$, where $\delta_\nu^\mu = g^{\mu\rho}g_{\rho\nu}$.

$$\begin{aligned}\overline{|\mathcal{M}_0|^2} &= \frac{1}{16(d-2)^2}\frac{\mu^{2\epsilon}\ C_{eff}^2\Lambda^2 v^2}{(\hat{s}_{12}-m_h^2)^2+\Gamma_h^2}\ (d-2)\frac{\hat{s}_{12}^2}{4} \\ &= \frac{1}{128(1-\epsilon)}\frac{\mu^{2\epsilon}\ C_{eff}^2\Lambda^2 v^2\ \hat{s}_{12}^2}{(\hat{s}_{12}-m_h^2)^2+\Gamma_h^2} \\ &= \frac{1}{128}\frac{\mu^{2\epsilon}\ C_{eff}^2\Lambda^2 v^2\ \hat{s}_{12}^2}{(\hat{s}_{12}-m_h^2)^2+\Gamma_h^2}\ (1-\epsilon)^{-1}\end{aligned} \tag{2.7}$$

To do the calculation in the Feynman gauge, replace $\sum_{\text{pol}}\epsilon_\mu^a\,(p)\epsilon_\nu^{*\,b}\,(p) \to -\delta^{ab}g_{\mu\nu}$, yielding the same expression as Equation 2.7. This is evident due to the independence of gauge choices. hypercube

Virtual Corrections The possible virtual diagrams at one loop are shown in Figures 2.1b - 2.1f. Diagrams 2.1b, 2.1c, and 2.1d do not contribute since they have a scaleless integral, which is zero in dimensional regularization. The non-zero contributions come only from diagrams 2.1e and 2.1f.

Vertex correction (Figure 2.1e): Using the triple gluon vertex from the SM Lagrangian, and ggh vertex (Equation C.15), we can write the amplitude of diagram 2.1e as,

$$i\mathcal{M}_4 = \frac{i\ C_{eff}\Lambda\ v\ g_s^2\ C_A}{(\hat{s}_{12}-m_h^2)+i\Gamma_h}\ \epsilon_\mu^a(p_1)\epsilon_\nu^a(p_2)\int\frac{d^dp}{(2\pi)^d}\ \frac{T^{\mu\nu}}{p^2\ (p_2-p)^2\ (p_1+p)^2} \tag{2.8}$$

where p is the loop momenta, the repeated indices are summed over. We used $f^{acd}f^{bcd} = C_A\delta^{ab}$ ($C_A = n$ for $SU(n)$) in the intermediate step.

$$p_1^\mu\epsilon_\mu(p_1) = p_2^\nu\epsilon_\nu(p_2) = 0 \tag{2.9}$$

The structure of the numerator is

$$\begin{aligned}T^{\mu\nu} =& -g_{\alpha\beta}\ g_{\delta\sigma}\ g_{\rho\lambda}\ [(p_2+p)^\delta g^{\nu\alpha} - 2p^\nu g^{\alpha\delta} + (p-2p_2)^\alpha g^{\delta\nu}]\ [(p_1-p)^\rho g^{\mu\beta} \\ & + 2p^\mu g^{\beta\rho} - (p+2p_1)^\beta g^{\rho\mu}]\ [(p_1+p)^\sigma(p_2-p)^\lambda - g^{\sigma\lambda}(p_1+p).(p_2-p)]\ .\end{aligned} \tag{2.10}$$

Note that the second term of the gluon propagators (Equation 2.2) is absent in the numerator structure since those terms do not contribute. The reason those terms do not contribute can also be understood from gauge invariance. If one calculates in the Feynman gauge instead of the axial gauge, the gluon propagator does not have that second term. However, in the Feynmann gauge, the ghost usually appears in the loop, but notice that the Higgs field does not couple to the ghost of the gluon, so the ghost can not appear in these diagrams. Therefore, the second term of the gluon propagator does not contribute in the axial gauge either.

Using Equations 2.4, 2.9, and $g^{\mu\nu}g_{\mu\nu} = d$, we can write the numerator after contraction of the indices as,

$$
\begin{aligned}
T^{\mu\nu} = g^{\mu\nu}\ [&\hat{s}_{12}^2 - \frac{9}{2}\hat{s}_{12}\ p^2 - 3\ \hat{s}_{12}\ p.p_1 + 3\ \hat{s}_{12}\ p.p_2 - 3\ p^2\ p.p_2 + 3\ p^2\ p.p_1 \\
&+ 2(p.p_1)^2 + 2(p.p_2)^2 + p^4] + p^\mu p^\nu\ [(8-2d)\hat{s}_{12} + p.p_1(4d-6) + p.p_2(6-4d) \\
&+ p^2(4d-5)] + p_2^\mu p^\nu\ [p^2 - 6p.p_1 - 2p.p_2] + p_2^\mu p_1^\nu\ [-2\hat{s}_{12} + 9p^2 + 6p.p_1 - 6p.p_2] \\
&+ p^\mu p_1^\nu\ [-2p.p_1 - 6p.p_2 - p^2]\ .
\end{aligned}
\tag{2.11}
$$

To decompose the tensor into scalar triangles, scalar bubbles, and tadpole integrals, we replace $p.p_1$ and $p.p_2$ with the following formulas.

$$
p.p_1 = \frac{(p+p_1)^2 - p^2}{2} \quad \text{and} \quad p.p_2 = \frac{p^2 - (p_2-p)^2}{2} \tag{2.12}
$$

After the replacement, we have the following.

$$
\begin{aligned}
&\frac{T^{\mu\nu}}{p^2\ (p_2-p)^2\ (p_1+p)^2} = \frac{1}{p^2\ (p_2-p)^2\ (p_1+p)^2}[g^{\mu\nu}\ \hat{s}_{12}^2 + (8-2d)\hat{s}_{12}\ p^\mu p^\nu \\
&- 2\hat{s}_{12}\ p_2^\mu p_1^\nu] + \frac{1}{p^2\ (p_2-p)^2}[(-\frac{3}{2}\hat{s}_{12} + p.p_1)g^{\mu\nu} + (2d-3)\ p^\mu p^\nu - 3p_2^\mu p^\nu + 3p_2^\mu p_1^\nu \\
&- p^\mu p_1^\nu] + \frac{1}{p^2\ (p_1+p)^2}[(-\frac{3}{2}\hat{s}_{12} - p.p_2)g^{\mu\nu} + (2d-3)\ p^\mu p^\nu + 3p_2^\mu p_1^\nu + 3p^\mu p_1^\nu \\
&+ p_2^\mu p^\nu] + \frac{1}{(p_2-p)^2(p_1+p)^2}[(-\frac{3}{2}\hat{s}_{12} - p.p_1 + p.p_2 - 2p^2)g^{\mu\nu} + p^\mu p^\nu + 3p_2^\mu p^\nu \\
&+ 3p_2^\mu p_1^\nu - 3p^\mu p_1^\nu]
\end{aligned}
\tag{2.13}
$$

In the intermediate state of the above equation, we discard the scaleless tadpole terms ($\frac{3}{2}\frac{g^{\mu\nu}}{(p_2-p)^2}$ and $\frac{3}{2}\frac{g^{\mu\nu}}{(p_1+p)^2}$) as those terms' integration over loop momentum in dimensional regularization is zero. For further simplification, we combine the denominators in a standard way.

$$\begin{aligned}\frac{1}{p^2\ (p_2-p)^2\ (p_1+p)^2} &= \int_0^1 dx\ \int_0^{1-x} dy\ \frac{2}{[x(p_1+p)^2+y(p_2-p)^2+(1-x-y)p^2]^3} \\ &= 2\int_0^1 dx\ \int_0^{1-x} dy\ \frac{1}{(l_1^2-\Delta_1)^3}\ ,\end{aligned} \tag{2.14}$$

where $l_1 = p + xp_1 - yp_2$ and $\Delta_1 = -xy\hat{s}_{12}$.

$$\begin{aligned}\frac{1}{p^2\ (p_2-p)^2} &= \int_0^1 dx\ \frac{1}{[x(p_2-p)^2+(1-x)p^2]^2} \\ &= \int_0^1 \frac{dx}{(l^2-0)^2}\ \text{(scaleless)}, \qquad (p_2^2=0, l=p-xp_2)\ .\end{aligned} \tag{2.15}$$

Similarly, $\frac{1}{p^2\ (p_1+p)^2}$ also gives the scaleless integral. Hence, the second and third lines of Equation 2.13 do not contribute.

$$\frac{1}{(p_2-p)^2(p_1+p)^2} = \int_0^1 dx\ \frac{1}{[x(p_2-p)^2+(1-x)(p_1+p)^2]^2} = \int_0^1 \frac{dx}{(l_4^2-\Delta_4)^2}\ , \tag{2.16}$$

where $l_4 = p+(1-x)p_1 - xp_2$ and $\Delta_4 = -x(1-x)\hat{s}_{12}$. The non-zero contribution of Equation 2.13 is provided below.

$$\begin{aligned}\frac{T^{\mu\nu}}{p^2(p_2-p)^2(p_1+p)^2} &= 2\int_0^1 dx \int_0^{1-x} dy \frac{1}{(l_1^2-\Delta_1)^3}[g^{\mu\nu}\ \hat{s}_{12}^2 + (8-2d)\hat{s}_{12}\ p^\mu p^\nu - 2\hat{s}_{12}\ p_2^\mu p_1^\nu] \\ &+ \int_0^1 \frac{dx}{(l_4^2-\Delta_4)^2}[(-\frac{3}{2}\hat{s}_{12} - p.p_1 + p.p_2 - 2p^2)g^{\mu\nu} + p^\mu p^\nu + 3p_2^\mu p^\nu + 3p_2^\mu p_1^\nu - 3p^\mu p_1^\nu]\end{aligned} \tag{2.17}$$

The numerator can be further simplified by eliminating loop momentum p in terms of l_1 and l_4. Using Equation 2.9, keeping the even power of l_1 and l_4 (since symmetric integral gives zero for the odd terms of l_1 and l_4), and replacing $l_\mu l_\nu$

with $l^2\ g_{\mu\nu}/d$, we will get

$$\int \frac{d^dp}{(2\pi)^d}\frac{T^{\mu\nu}}{p^2\ (p_2-p)^2\ (p_1+p)^2}$$
$$= 2\int_0^1 dx\ \int_0^{1-x} dy\ \frac{d^dl_1}{(2\pi)^d}\ \frac{1}{(l_1^2-\Delta_1)^3}\Big[l_1^2\ (\frac{8}{d}-2)\ \hat{s}_{12}\ g^{\mu\nu}+\Big(g^{\mu\nu}\ \hat{s}_{12}^2-2\hat{s}_{12}\ p_2^\mu p_1^\nu\Big)$$
$$+\Delta_1\ (8-2d)\ p_2^\mu p_1^\nu\Big]+\int_0^1 dx\ \frac{d^dl_4}{(2\pi)^d}\ \frac{1}{(l_4^2-\Delta_4)^2}\Big[l_4^2\Big(\frac{1}{d}-2\Big)g^{\mu\nu}-2\ \hat{s}_{12}\ g^{\mu\nu}$$
$$+\Delta_4\ \frac{1}{\hat{s}_{12}}\ p_2^\mu p_1^\nu-2\Delta_4\ g^{\mu\nu}\Big]\ . \tag{2.18}$$

We obtain the following by setting $d=4-2\epsilon$ and using Equations B.7-B.13.

$$\int \frac{d^dp}{(2\pi)^d}\frac{T^{\mu\nu}}{p^2\ (p_2-p)^2\ (p_1+p)^2}=\frac{i}{(4\pi)^2}\ (-\hat{s}_{12})^{-\epsilon}\ C_\Gamma\ \Big[\hat{s}_{12}\ g^{\mu\nu}\Big\{\frac{1}{(1-\epsilon)(1-2\epsilon)}$$
$$+\frac{1}{\epsilon^2}-\frac{1}{12}\ \frac{1}{\epsilon}\ \frac{(-7+4\epsilon)}{(1-\frac{2}{3}\epsilon)(1-2\epsilon)}-\frac{2}{\epsilon}\ \frac{1}{(1-2\epsilon)}+\frac{1}{3\ \epsilon}\ \frac{(1-\epsilon)}{(1-\frac{2}{3}\epsilon)(1-2\epsilon)}\Big\}$$
$$+p_2^\mu p_1^\nu\ \Big\{-\frac{2}{\epsilon^2}-\frac{2\epsilon}{(1-\epsilon)(1-2\epsilon)}-\frac{1}{6\ \epsilon}\ \frac{(1-\epsilon)}{(1-\frac{2}{3}\epsilon)(1-2\epsilon)}\Big\}\Big] \tag{2.19}$$

The interference of the virtual triangle diagram with LO amplitude gives the following after taking the average over color and polarization of the initial state gluons.

$$\overline{\mathcal{M}_4\ \mathcal{M}_0^*}=-i\ \frac{C_{eff}^2\Lambda^2v^2\ g_s^2\ C_A\ \mu^{2\epsilon}}{(\hat{s}_{12}-m_h^2)^2+\Gamma_h^2}\ \Big[\int \frac{d^dp}{(2\pi)^d}\ \frac{T^{\mu\nu}}{p^2\ (p_2-p)^2\ (p_1+p)^2}\Big]\times$$
$$\frac{1}{8\ 8}\Big[\frac{1}{d-2}\sum_{\text{pol}}\epsilon_\mu^a(p_1)\epsilon_\alpha^{*b}(p_1)\Big]\ \Big[\frac{1}{d-2}\sum_{\text{pol}}\epsilon_\nu^a(p_2)\epsilon_\beta^{*b}(p_2)\Big]\times\Big(p_1^\beta p_2^\alpha-g^{\alpha\beta}\frac{\hat{s}_{12}}{2}\Big) \tag{2.20}$$

As already said, $\mu^{2\epsilon}$ factor is because the cross section has mass dimension $[\sigma]=[M]^{2-d}$ in d dimensions. Substituting Equations 2.1 and 2.19 into Equation 2.20 and doing some simplifications, we obtain the following,

$$\overline{\mathcal{M}_4\ \mathcal{M}_0^*}=\overline{|\mathcal{M}_0|^2}\ \frac{\alpha_s}{2\pi}\ C_A\ (-1)^{-\epsilon}(\frac{\mu^2}{\hat{s}_{12}})^\epsilon\ (1-\epsilon)^{-1}C_\Gamma\Big[(\frac{2}{\epsilon}-\frac{2}{\epsilon^2})-\frac{13}{6}$$
$$-\frac{1}{6\ \epsilon}\ \frac{(1-\epsilon)}{(1-\frac{2}{3}\epsilon)(1-2\epsilon)}+\mathcal{O}(\epsilon)\Big] \tag{2.21}$$

where $\alpha_s=\frac{g_s^2}{4\pi}$.

Bubble diagram (Figure 2.1f): Using the four gluon vertex from the SM Lagrangian and ggh vertex (Equation C.15), we can write the amplitude of diagram 2.1f as follows.

$$i\mathcal{M}_5 = -\frac{i}{2}\frac{C_{eff}\Lambda\ v\ g_s^2\ C_A}{(\hat{s}_{12}-m_h^2)+i\Gamma_h}\ \epsilon_\mu^a(p_1)\epsilon_\nu^a(p_2)\int\frac{d^dp}{(2\pi)^d}\ \frac{\mathcal{N}^{\mu\nu}}{p^2\ (p+p_1+p_2)^2} \tag{2.22}$$

The numerator structure without p_1^μ and p_2^ν terms (since their contraction with the polarization gives zero):

$$\mathcal{N}^{\mu\nu} = g^{\mu\nu}\ (2d-4)\ (p+p_1+p_2).p + 2p^\mu\ p^\nu + p_2^\mu\ p^\nu + p_1^\nu\ p^\mu\ . \tag{2.23}$$

The denominator can be combined as

$$\frac{1}{p^2\ (p+p_1+p_2)^2} = \int_0^1 dx\ \frac{1}{[x(p+p_1+p_2)^2+(1-x)p^2]^2} = \int_0^1 \frac{dx}{[l^2-\Delta]^2}\ , \tag{2.24}$$

$$\text{where}\quad l = p + xp_1 + xp_2 \text{ and } \Delta = -x(1-x)\hat{s}_{12}. \tag{2.25}$$

Replacing loop momentum p in terms of l and keeping the even power of l, we have the following:

$$\int\frac{d^dp}{(2\pi)^d}\ \frac{\mathcal{N}^{\mu\nu}}{p^2\ (p+p_1+p_2)^2} = \int_0^1\ dx\int\frac{d^dl}{(2\pi)^d}\frac{1}{[l^2-\Delta]^2}\Big[l^2\ (2d-4+\frac{2}{d})g^{\mu\nu} + g^{\mu\nu}\ (2d-4)\ \Delta + p_2^\mu\ p_1^\nu\ \frac{2}{\hat{s}_{12}}\Delta\Big]\ . \tag{2.26}$$

With the help of Equations B.12 and B.13, we get the following.

$$\int\frac{d^dp}{(2\pi)^d}\ \frac{\mathcal{N}^{\mu\nu}}{p^2\ (p+p_1+p_2)^2} = -\frac{i}{(4\pi)^2}(-\hat{s}_{12})^{-\epsilon}\ C_\Gamma\ \frac{1}{3}\ \frac{1}{\epsilon}\ \frac{1}{(1-\frac{2}{3\epsilon})(1-2\epsilon)} \times\Big[\frac{4\epsilon^2-12\epsilon+9}{2}\ \hat{s}_{12}\ g^{\mu\nu} + (1-\epsilon)(2-2\epsilon)\ \hat{s}_{12}\ g^{\mu\nu} + (1-\epsilon)p_2^\mu\ p_1^\nu\Big] \tag{2.27}$$

Substituting Equation 2.27 into Equation 2.22, the interference of the bubble diagram with LO amplitude gives the following after taking the average over color and polarization of the initial state gluons.

$$\overline{\mathcal{M}_5\ \mathcal{M}_0^*} = \overline{|\mathcal{M}_0|^2}\ \frac{\alpha_s}{2\pi}C_A(-1)^{-\epsilon}\Big(\frac{\mu^2}{\hat{s}_{12}}\Big)^\epsilon(1-\epsilon)^{-1}C_\Gamma\Big[\frac{13}{6}+\frac{(1-\epsilon)}{(1-\frac{2}{3\epsilon})(1-2\epsilon)}\frac{1}{6\epsilon}+\mathcal{O}(\epsilon)\Big] \tag{2.28}$$

Total one loop amplitude can be written as,

$$\overline{\mathcal{M}_v \ \mathcal{M}_0^*} = \overline{\mathcal{M}_4 \ \mathcal{M}_0^*} + \overline{\mathcal{M}_5 \ \mathcal{M}_0^*} \ . \tag{2.29}$$

Thus from Equations 2.21 and 2.28, we have

$$\overline{\mathcal{M}_v \ \mathcal{M}_0^*} = \overline{|\mathcal{M}_0|^2} \ \frac{\alpha_s}{2\pi} \ C_A \ (-1)^{-\epsilon} (\frac{\mu^2}{\hat{s}_{12}})^{\epsilon} \ (1-\epsilon)^{-1} C_\Gamma \Big[\frac{2}{\epsilon} - \frac{2}{\epsilon^2} + \mathcal{O}(\epsilon)\Big] \ . \tag{2.30}$$

Real of the above equation will contribute to the loop amplitude, so taking the real of $(-1)^{-\epsilon}$ (Equation B.6), we obtain

$$Re[\ \overline{\mathcal{M}_v \ \mathcal{M}_0^*}\ ] = \overline{|\mathcal{M}_0|^2} \ \frac{\alpha_s}{2\pi} C_\Gamma \ (\frac{\mu^2}{\hat{s}_{12}})^{\epsilon} \ \Big[-\frac{6}{\epsilon^2} + 3\pi^2\Big] \ . \tag{2.31}$$

Taking the α_s correction of the C_{eff} (Equation C.14), we have the following

$$Re[\ \overline{\mathcal{M}_v \ \mathcal{M}_0^*}\ ] = \overline{|\mathcal{M}_0|^2} \ \frac{\alpha_s}{2\pi} C_\Gamma \ (\frac{\mu^2}{\hat{s}_{12}})^{\epsilon} \ \Big[-\frac{6}{\epsilon^2} + 3\pi^2 + 11\Big] \ , \tag{2.32}$$

where $\overline{|\mathcal{M}_0|^2}$ is given below

$$\overline{|\mathcal{M}_0|^2} = \frac{1}{128} \frac{C_0^2 \Lambda^2 v^2 \ \hat{s}_{12}^2}{(\hat{s}_{12} - m_h^2)^2 + \Gamma_h^2} \ (1-\epsilon)^{-1} \ . \tag{2.33}$$

The virtual amplitude is proportional to the Leading-order amplitude and has divergences in powers of $\frac{1}{\epsilon}$. The double pole $(\frac{1}{\epsilon^2})$ corresponds to the collinear and soft divergences. The above equation shows the absence of UV divergences $(\frac{1}{\epsilon})$. From naive power counting, one would expect the bubble diagram to have both IR and UV divergences. So, the absence of UV divergences is simply because of the accidental cancellation of IR divergences with UV divergences since we did not keep them separately. We renormalize the strong coupling constant, α_s, by replacing α_s with the following.

$$\alpha_s \to \alpha_s^{\overline{MS}}(\mu_R) = \alpha_s \Big[1 - \frac{\alpha_s}{2\pi} \ C_\Gamma \ \Big(\frac{\mu^2}{\mu_R^2}\Big)^{\epsilon} \ \frac{b_0}{\epsilon}\Big] \ , \tag{2.34}$$

where $b_0 = \frac{11}{6} \ C_A - \frac{2}{3} n_f T_f$, $T_f = \frac{1}{2}$, and n_f is the number of quark flavors of the SM under consideration. α_s on the left side of the above equation is the bare coupling, which is divergent, and α_s on the right is the renormalized coupling. The second term of the above equation is the divergent term. Bare coupling captures the UV divergence to give the renormalized coupling.

After the replacement (Equation 2.34), the LO amplitude becomes:

$$\begin{aligned}\overline{|\mathcal{M}_0|^2} &\rightarrow \overline{|\mathcal{M}_0|^2}\left[1-\frac{\alpha_s}{2\pi}\,C_\Gamma\left(\frac{\mu^2}{\mu_R^2}\right)^{\epsilon}\frac{2b_0}{\epsilon}+\mathcal{O}(\alpha_s^2)\right]\\ &\rightarrow \overline{|\mathcal{M}_0|^2}-\overline{|\mathcal{M}_0|^2}\,\frac{\alpha_s}{2\pi}\,C_\Gamma\left(\frac{\mu^2}{\mu_R^2}\right)^{\epsilon}\frac{2b_0}{\epsilon}+\text{higher order}\ .\end{aligned} \tag{2.35}$$

The Second term of the above equation has order α_s^3, so it must be added to the virtual correction term. Therefore, the total virtual amplitude, which includes soft and collinear divergences, is provided below.

$$Re[\ \overline{\mathcal{M}_v\ \mathcal{M}_0^*}\ ]=\overline{|\mathcal{M}_0|^2}\ \frac{\alpha_s}{2\pi}C_\Gamma\left(\frac{\mu^2}{\hat{s}_{12}}\right)^{\epsilon}\left[-\frac{6}{\epsilon^2}-\frac{2b_0}{\epsilon}+3\pi^2+11-2b_0\ \log\frac{\mu^2}{\mu_R^2}\right] \tag{2.36}$$

In the intermediate step, we use $\left(\frac{\mu^2}{\mu_R^2}\right)^{\epsilon}\equiv 1+\epsilon\log\frac{\mu^2}{\mu_R^2}$. Both single and double poles are IR in nature in the above equation. Another significant feature is the explicit appearance of the logs of the renormalization scale in the above expression, which reduces the scale dependence at NLO compared to LO.

2.2 Jets and Jet algorithms

Instead of a colored parton, we see a collimated spray of charged and neutral energetic hadrons, known as a jet. A jet algorithm groups many hadrons into jets. The algorithm also needs a recombination scheme specifying what momentum should be assigned to the resulting jets. In our thesis, we use the E-scheme, which adds the 4-momentum of the two combining particles to assign the 4-momentum of the resulting particle. Furthermore, the jet definition should meet several criteria; it should be simple in implementing the experimental analysis, and for the theoretical calculations, defined at all orders in the perturbation theory, it should be IRC safe, and the cross section should be insensitive to the hadronization. IRC safe means the cross section should not change by collinear splittings or soft radiations. Jet algorithms are widely classified into two types: cone algorithms and sequential algorithms.

Cone algorithms: The cone algorithms are iterative cones (IC). The iteration starts with a seed particle i, which sets the initial direction of the cone. The 4-momentum of all the particles (js) that lie within a circle of radius R (R is the jet's radius, $\Delta R_{ij}=\sqrt{(y_j-y_i)^2+(\phi_j-\phi_i)^2}<R$) around the y_i and ϕ_i of the seed particle are added. This momentum is set as a new seed direction.

The preceding iteration is then repeated using the new seed direction until the direction of the resultant cones becomes stable. That indicates that no particle is left in the region of the stable cone of radius R. The resulting stable cone is called the final jet.

This procedure has two issues: 1) what should one take as the initial seed, and 2) what to do if two stable cones originate from the iteration of two distinct seed particles overlap. Some different approaches to getting rid of these two issues are listed below.

The first one is progressive removal (IC-PR), where the initial seed particle (or calorimeter tower) is the one that has the largest transverse momentum. After identifying the corresponding stable cone, one labels it as a jet and removes it from the particle list of the event. The next seed is the hardest particle or calorimeter tower of the remaining list, and by doing the iteration, one can obtain the next stable cone. One has to repeat this iteration until no more particles are present (above some chosen energy or P_T threshold).

The second type is the split–merge approach (IC-SM) [90], where iteration starts from all particles (or calorimeter towers) of the list and obtains the stable cones. After obtaining stable cones, if two cones overlap, then merge these two cones if more than a fraction f (usually 0.5 or 0.75) of the transverse momentum of the softer cone contained in particles shared with the harder cone; otherwise, the shared particles are allocated to the cone to which they are closest (from the center of the circle of each cone) and treat them two separate cones.

Experimentalist prefers Cone algorithms because of the circular shape of the jets; however, it is found that it is IRC unsafe. Seedless cone algorithms are IRC-safe [91], but their execution consumes a lot of memory and processing resources. Therefore, in this thesis, we use sequential jet algorithms, which are IRC-safe, and their implementation in the FASTJET [92] package is speedy. Additionally, its clustering sequence of combining the particles (or calorimeter towers) has a close connection to the parton branching. Below we are describing a variety of sequential recombination algorithms commonly known as the generalized k_t family.

Generalized k_t algorithms: In the hadron collider, the generalized form of sequential algorithms is described in terms of dimensionful distance measures that longitudinally boost invariant,

$$d_{ij} = \min(p_{Ti}^{2p}, p_{Tj}^{2p}) \frac{\Delta R_{ij}^2}{R^2} , \quad d_{iB} = p_{Ti}^{2p} . \tag{2.37}$$

R is the jet radius, $\Delta R_{ij} = \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}$ is the angular separation between i and j particles, and p_{Ti}, and p_{Tj} are their transverse momentum. The event list of the final state particles (or calorimeter towers) is characterized by their 4-momentum $\{p_1, p_2, \cdots, p_N\}$. The formation of the jet using the generalized k_t algorithms goes as follows.

1. For each pair of particles i, j of the list, evaluate two distance measures, d_{ij} and d_{iB}.

2. If d_{ij} is the minimum among all possible d_{ij} and d_{iB}, combine i and j to form k via a recombination scheme. We use an E-scheme that adds the 4-momentum of these two particles $p_k = p_i + p_j$. After the recombination, remove p_i and p_j from the list, and add p_k to the list. Then again, go to step 1.

3. Form all possible d_{ij} and d_{iB}; if d_{iB} is the minimum, declare i as the final jet and remove it from the list.

4. Iterate again from step 1 until no more particles are in the list.

Since every particle in the list is assigned to a jet, a minimum transverse momentum of a jet is applied to ignore the soft particle's contribution.

k_t algorithm: $p = 1$ in Equation 2.37 gives the k_t algorithm [93, 94]. The soft radiation of a parton has divergences; hence, to get the finite cross section, the corresponding soft radiation of the parton should be added to the same jet. Hence, the k_t algorithm is motivated by that and clusters the softer particles first. However, the soft radiations of underlying or pile-up events are sensitive to this algorithm.

k_t-jet has an irregular shape in $\eta - \phi$ plane as it combines the soft particle first and large p_T particles later. Therefore the 4-momentum of the resulting particle after each combination changes rapidly, leading to an unstable jet axis.

Cambridge-Aachen (CA) algorithm: $p = 0$ in Equation 2.37 gives CA algorithm [95–97], where the two distance measures are $d_{ij} = \dfrac{\Delta R_{ij}^2}{R^2}$ and $d_{iB} = 1$. Therefore, the CA algorithm is oblivious to the transverse momentum of the particles and clusters the nearby particles first, gradually moving to the larger distance. It also produces irregular shapes, but less severely than the k_t algorithm.

anti-k_t algorithm: $p = -1$ in Equation 2.37 is known as an anti-k_t algorithm [98]. This algorithm clusters the largest p_T particle first and then the softer particles. As a result, after certain steps, the resulting 4-momentum does not change as it adds soft particles, leading to a stable jet axis. Therefore, the resulting jet is almost circular in the $\eta - \phi$ plane. However, the recombination sequence has no connection with the QCD. In case one also wants to know the underline QCD evolution of the anti-k_t jet, one can recluster the constituents of the anti-k_t jet using the CA or k_t algorithm.

2.3 Boosted jets and Jet Substructure variables

In many BSM scenarios, the particles in the final state, such as $W/Z/h$ and top quarks, can acquire a significant boost that all the daughter particles resulting from their hadronic decay become highly collimated. Consequently, due to the high degree of collimation among the decay products, it becomes challenging to distinguish or resolve two (or three) small-radius ordinary jets. Instead, all the decay products are typically encompassed within a single large-radius jet known as a fatjet. The hadronic decay of a boosted particle can be classified into two categories: two-pronged decays (W, Z, and h) or three-pronged decays (top quark).

One indication of the jet's origin is its mass. However, collinear branching of the massless partons can generate a large mass for the QCD jets at a large P_T. The fixed-order differential mass distribution (up to the leading log) of a QCD jet of radius R at its transverse momentum P_T is the following:

$$\frac{1}{\sigma}\frac{d\sigma}{dM_j} \sim \frac{1}{M_j}\frac{\alpha_s(P_T)C_i}{\pi}(\log\frac{P_T^2R^2}{M_j^2} + \mathcal{O}(1)) \ . \tag{2.38}$$

C_i is equal to $C_F = 4/3$ for quark and equal to $C_A = 3$ for gluon. The splitting function when a massless quark splits into a collinear gluon has the following form:

$$P_{g\leftarrow q}(z) = C_F\left[\frac{1+(1-z)^2}{z}\right] \ . \tag{2.39}$$

The integration of the divergence part of the splitting function $1/z$ provides the logarithm term in Equation 2.38, where the lower limit of z is set to $z > \dfrac{M_j^2}{P_T^2R^2}$ [a], and the upper limit is 1. For more detailed derivation, see [99]; for the corresponding resumed predictions in e^+e^- collision, see [100,101]. Therefore, the large

[a] Suppose a parton-a of transverse momentum p_T splits collinearly ($a \to i,j$) such that $p_{Tj} < p_{Ti}$, and define $z = p_{Tj}/p_T$, where $p_T = p_{Ti} + p_{Tj}$. Then, $m^2 = p_a^2 = (p_i+p_j)^2 \approx 2p_{Ti}p_{Tj}(1-\cos\theta)$. In the collinear limit($\theta \to 0$), one can write $m^2 \approx 2p_{Ti}p_{Tj}\dfrac{\theta^2}{2} = z(1-z)p_T^2\theta^2 \approx z(1-z)p_T^2\Delta R_{ij}^2$. Since the lower limit of the z is very small so, neglecting the z^2 term in the previous equation, one can get the approximate lowest limit of z for a jet of radius R.

radius QCD jets at larger P_T can have significant mass and mimic the $W/Z/h/t$ fatjets.

The QCD splitting functions have divergences corresponding to $z = 0$ and or $z = 1$, which means one of the split partons almost takes negligible energy, and the other has almost all, leading to asymmetric energy sharing. On the other hand, the fatjets originating from boosted W/Z bosons and top quarks, which exhibit two-pronged and three-pronged structures, often display a symmetric sharing of energy among their constituent particles. This idea served as the basis for jet substructure techniques. To reduce the QCD background that mimics the two-pronged and three-pronged fatjets, one can 1) demand a larger pruned jet mass and 2) looks inside the fatjet to identify how the energy of the fatjet is distributed. If the energy is distributed along only one axis, then it is a QCD jet, and if along two (three) axes, it can be a weak or Higgs boson (top quark). In this thesis, we used two jet substructure variables, pruned jet mass and N-subjettiness ratio, which we will discuss below.

2.3.1 Jet Grooming techniques

At the LHC, other than the showering and hadronizing of hard partons, many soft unassociated radiations come from the underline event, possible multi-parton interactions, and pile-up that degrade the jet mass resolution and the efficiency of many jet observables. Large radius fatjets are severely affected by those unassociated radiations. Therefore, to correctly predict the fatjet mass and other observables, we must clean up jets by removing those unassociated soft radiations. In the literature exist, many grooming techniques and a few of them are listed below.

Trimming: In this process [102], we begin by reclustering the elements of the fatjet into subjets of smaller radius $R_{sub} < R$ (R is the fatjet's radius). The k_t or CA algorithms are commonly used to cluster the subjets. Each subjet i whose transverse momentum $P_{Ti} > z_{\rm cut} P_T^{\rm fatjet}$ should keep. $z_{\rm cut}$ is the cutoff parameter of the algorithm. Any subjet that fails to satisfy the above threshold condition will be removed. All of the retained subjets constitute the final trimmed jet.

Pruning: In this thesis, we use the pruning technique [103,104] described below. Typically the last clusterings of the CA algorithms consist of soft radiations that are usually unassociated with the parent parton. Therefore, removing the soft and wide unassociated radiations gives the pruned jet. The procedure is as follows.

1. Collect the constituent particles of a given jet J (found by any jet algo-

rithm). After that, employ the CA algorithm to recluster its constituents. And then, unwind the cluster sequence sequentially.

2. At each unwinding step $P \to ij$, check the following two quantities:

$$z = \frac{\min(P_{Ti}, P_{Tj})}{P_{T_{i+j}}} \;, \qquad \text{and} \;, \qquad \Delta R_{ij} \;, \tag{2.40}$$

where ΔR_{ij} denotes the angular separation between i and j.

3. If both conditions $z < z_{cut}$ and $\Delta R_{ij} > R_{\text{fact}}$ are met, remove the softer of i, j, and continue unwinding.

4. Stop the unwinding if sufficient hard ($z > z_{cut}$) or collinear ($\Delta R_{ij} < R_{\text{fact}}$) splittings are achieved.

This technique has two parameters, the softness parameter z_{cut} and the radial separation R_{fact}, and both of these two parameters need to do optimized for every process under consideration. $z_{cut} = 0.1$ is the typical choice, and R_{fact} is generally chosen around the opening angle of a hard process, $R_{\text{fact}} \approx \dfrac{2M_J}{P_{TJ}} \times \dfrac{1}{2}$

Filtering: Filtering was initially proposed in reference [105] to sharpen the Higgs boson mass peak. However, it can be applied to any jet to eliminate the unassociated radiations. First, recluster the constituents of the given jet J into subjets of smaller radius R_{filt} using the CA algorithm. The hardest n_{filt} subjets are kept, and the rest are removed. All the remaining subjets form the final jet. The number of subjets n_{filt} depends on the prior knowledge of the signal jet; if the signal jet is N-pronged, n_{filt} is usually taken as $N + 1$ to accommodate one extra gluon radiation from the partons.

2.3.2 N-subjettiness ratio

The N-subjettiness ratio [106, 107] is a jet-shape observable that indicates the radiation pattern of a jet. If a jet exhibits N-body hadronic decay, it is called an N-pronged jet. The N-subjettiness ratio is derived from the inclusive variable N-jettiness [108] and denotes the number of subjet axes along which the jet's energy is spread. For defining N-subjettiness, we have to construct N axes within the jet. N subjets are obtained by reclustering the constituents of the jet using the k_t algorithm (one can use any other algorithm), and the resultant momentum of the subjets acts as the N-axes. Once we get the axes, N-jettiness is defined below,

$$\tau_N = \frac{1}{\mathcal{N}_0} \sum_i P_{Ti} \min\{\Delta R_{i,1}, \Delta R_{i,2}, \cdots \Delta R_{i,N}\}. \tag{2.41}$$

The normalization factor $\mathcal{N}_0$ is defined as $\mathcal{N}_0 = \sum_i P_{Ti} R$, where R is the radius of the jet. P_{Ti} is the transverse momentum of the i-th particle (or calorimeter tower) of the jet, and $\Delta R_{i,K} = \sqrt{(\eta_i - \eta_K)^2 + (\phi_i - \phi_K)^2}$ is its angular separation from the K^{th} subjet axis. The summation runs over all particles.

If $\tau_N \approx 0$, the jet's radiation is oriented along the N (or fewer) hard subjets. If $\tau_N >> 0$, the N subjet axes are unable to capture all of the jet's radiation; hence the jet must have at least N+1 hard subjet axes. The value of τ_n for a jet that has N-body hadronic decay is as follows:

$$\tau_n \rightarrow \text{ large if } n < N \quad , \quad \text{and } \tau_n \rightarrow \text{ small if } n \geq N \tag{2.42}$$

Since single N is insufficient to determine the exact behavior of the jet, we use N-subjettiness ratio $\tau_{N,N-1} = \frac{\tau_N}{\tau_{N-1}}$, which has more discriminating power in separating the signal jet from the SM background [106]. For example, we construct $\tau_{2,1}$ to discriminate the weak bosons, which have 2-body hadronic decay, from the QCD background. The value of $\tau_{2,1}$ for the W/Z jet is small since τ_2 is small and τ_1 is large, making the ratio even smaller. On the contrary, for the QCD jet, both τ_2 and τ_1 are small, making the ratio large (towards 1). Similarly, for a top fatjet, we construct $\tau_{3,2}$.

2.4 Multivariate analysis (MVA)

We adopt a multivariate analysis using the boosted decision tree (BDT) [109,110] to optimize the collider search. BDT is a sophisticated supervised machine learning technique. In this section, we describe the decision tree algorithm and then discuss its boosting procedure. Supervised means the classifier knows the features of the events (set of discriminating variables) and the class label. Although this algorithm can be applied for any number of classes, here we will talk about the binary classifier, which means the event belongs to either signal or background class.

In high-energy physics, simulated Monte Carlo signal and background events are typically assigned weights. These weights represent the ratio of the number of events generated to the product of the cross section and the integrated luminosity of the collected data. Since many different processes can contribute to the background (signal), the background (signal) class is simply the weighted combination of the different processes.

2.4.1 Decision Tree

A decision tree starts from the root (initial) node. Each node is then divided recursively into two daughter nodes/brunches until a stopping condition is met. Consider the signal and background events are described by a set of variables $\{X\}$, and signal and background events have weights w_i^S and w_j^B, respectively [b]. For the binary classifier, the decision tree algorithm goes as follows:

1. Sort all events (signal and background) by each variable in $\{X\}$.

2. For each variable in the list, find its optimal splitting value, which provides the best separation. The best means that after the splitting, among the two created daughter nodes, one will contain mainly signal events while the other is mostly background events.

3. Choose the variable (say x_k) and its optimal splitting value (say x_{k0}) that results in the best separation compared to the other variables in the list $\{X\}$. Then, divide the node into two daughter nodes based on the criteria $x_k < x_{k0}$. As a result, one daughter node holds events that satisfy the criteria, and the other contains events that fail the criteria. However, if the separation between the two classes can not improve by the above splitting, then do not split the corresponding node and call it a leaf.

4. This algorithm does not restrict using the same variable from the list in the multiple nodes if it gives the best separation.

5. Step 3 should be repeated recursively for each node unless any stopping criteria are met. After getting the stopping condition, declares the corresponding node as a leaf, and do not split it further.

6. Once no node is left other than the leaves, exit the algorithm.

The interesting point is that a decision tree is humanly readable, as one can easily track the variables and their optimal values that an event satisfies to reach an individual leaf. As a result, a tree can be interpreted in terms of some physics, defining selection rules. Because each leaf is either signal-like or background-like, the phase space comprises a lot of signal-like and background-like regions. Consequently, a non-linear boundary may be created to separate the signal from the background in this technique, which outperforms the cut-based analysis. The cut-based analysis performs a rectangular cut to each variable, labeling one side signal-like and the other as background. Therefore, a rectangular cut on several

[b] If more than one process contributes to the signal, the weight w_i^S for different signal events can be different (same for background).

variables in cut-based analysis selects one hypercube as a signal from the entire phase space.

The condition for the splitting of a node can be described in terms of its purity as below,

$$P = \frac{\sum_S w_i^S}{\sum_S w_i^S + \sum_B w_j^B} \tag{2.43}$$

Summation $\sum_S$ ($\sum_B$) represents one has to sum the weights of all the signal (background) events in that node. $P = 1,\ 0$ for the pure signal and pure background node, respectively, and hence $P(1-P) = 0$ for the pure signal or pure background node. The Gini index for a given node is as below,

$$\text{Gini} = \Big(\sum_{i=1}^{n} w_i\Big) P(1-P), \tag{2.44}$$

where n is the number of events present in that node. The condition of splitting is chosen to minimize $\text{Gini}_{\text{left, daughter}} + \text{Gini}_{\text{right, daughter}}$.

We use the parameter *MinNodeSize* as a stopping criterion. *MinNodeSize* is the minimum percentage of training events required for a leaf node. Additionally, we use the parameter *MaxDepth*, which is the maximum allowed depth of the decision tree. In the end, a leaf is referred to as a signal leaf if its purity is more than 0.5 (or whatever value is specified) or a background leaf if its purity is less than 0.5.

The usual approach of the multivariate (a set of discriminating variables need not be all independent) analysis randomly splits each signal and background dataset into two parts. Training is done on one part of the dataset; the other part, unseen during training, is reserved for testing. The decision trees have a reputation for being quite unstable due to overtraining. Therefore to get a stable model, one needs to minimize overtraining. Stability means a minor change in the training sample does not affect the performance of the testing sample. As a tree grows, each node has fewer and fewer events, increasing the statistical uncertainty with each subsequent split. Therefore the tree starts to learn the specific features of the events that may not be relevant to the desired result. If the tree becomes extremely specialized through learning almost every feature of the trained dataset, then the tree is overtrained. A first solution to mitigate overtraining is using stopping criteria like minimum node size, as described above. Other possible solutions also exist in the literature, like post-pruning.

If the distribution of the signal and background obtained from the test dataset fits well with the trained dataset, the network is not overtrained, and the obtained model is stable. The stable BDT model can be applied to unknown events as they do not depend on the training sample. After training, we also compute the

Kolmogorov-Smirnov (KS) probability for the training and testing samples to ensure that the network has not been overtrained. The KS probability measures the difference in the cumulative distribution functions of the training and testing datasets.

2.4.2 Boosted Decision Tree

The decision tree algorithm can be further improved using the boosted method. Start with unweighted events [c] and construct a tree as described above. If some training events are misclassified, such as when a signal event falls on a background leaf, or a background event falls on a signal leaf, then the weight of such events is raised (boosted). A second tree is constructed using the new weights of the misclassified events, while the weight remains the same for the correctly identified event (like when a signal event lands on a signal leaf). Since the weight increased, the newly constructed tree tried harder to identify the previously misclassified events. In that way, build a large number of trees by boosting the weights of the misclassified events of the previous tree, and the decision of the majority of trees is the desired output. In that way, the misclassification rate becomes less, which improves performance.

If the training sample consists of N number of events, then initially, each event has weight $1/N$. The first boosting technique is *AdaBoost* [111], which we use in the thesis, described below for N_{tree} number of trees.

- w_i^k is the weight of the ith event in the kth tree.
- Class label $y_i = +1$ if the ith event is a signal event, and -1 if it is a background event.
- Consider a function $\mathbb{I}(X)$ such that $\mathbb{I}(X) = 1$ if the statement X is true and 0 otherwise.
- $T_k(i) = +1$ if the ith event falls in the kth tree's signal leaf, and $T_k(i) = -1$ if found on a background leaf.
- Consider a misclassified function as below, equal to 1 if an event is misclassified and zero if not.

$$\text{isMisclassified}_k(i) = \mathbb{I}(y_i \times T_k(i) \leq 0) \tag{2.45}$$

[c] Supply the signal and background events with their corresponding weights to the machine, and the machine will unweight them.

- The misclassification rate is:

$$\varepsilon_k = \frac{\sum\limits_{i=1}^{N} w_i^k \times \text{isMisclassified}_k(i)}{\sum\limits_{i=1}^{N} w_i^k} \tag{2.46}$$

- Assign a weight to the kth tree T_k as follows:

$$\alpha_k = \beta \times \ln \frac{1-\varepsilon_k}{\varepsilon_k} \ , \tag{2.47}$$

  where the free parameter β is the strength of the boosting, it is also referred to as a learning rate or shrinkage coefficient in other machine learning algorithms. We set $\beta = 0.5$ in our analysis.

- The following step is the heart of the *AdaBoost* algorithm: build the next tree T_{k+1} by reweighting the events' weight. The weight of any event in tree T_{k+1} can be obtained from its weight in the previous tree T_k, as below.

$$w_i^{k+1} = w_i^k \times \exp[\alpha_k \times \text{isMisclassified}_k(i)] \tag{2.48}$$

  The above equation indicates that the weights of properly classified events remain unchanged, while the weights of misclassified events increase by a factor of e^{α_k} [d]. As a result, the next tree T_{k+1} will put more effort into classifying the challenging events that tree T_k failed to correctly identify while leaving those events that tree T_k correctly recognized.

- For event i, the final AdaBoost result is:

$$T(i) = \frac{1}{\sum\limits_{k=1}^{N_{tree}} \alpha_k} \sum_{k=1}^{N_{tree}} \alpha_k T_k(i) \ . \tag{2.49}$$

Therefore boosting helps to improve the result obtained by a single tree and reduces the error of the misclassification. If one builds a large number of trees, the error rate becomes zero or negligible, indicating that all the training events are correctly identified, leading to overtraining.

The error rate of the training and testing samples decreases with the number of trees in the boosted decision tree algorithm. However, after a certain number of trees, the error rate of the testing sample stops declining and remains constant

[d] Note that the misclassification rate ε_k should be less than 0.5; otherwise, $w_i^{k+1} < w_i^k$, and we will go in the wrong direction from our goal.

(known as a plateau), while the error rate of the training sample continues to decrease. In order to avoid overtraining and save resources without sacrificing performance, boosting can be stopped when such a plateau is achieved.

Now we will look at how to apply this technique to the collider analysis. After some basic cuts, we supply all the signal and background events with their corresponding weights for multivariate analysis in the TMVA framework [112]. We first determine the linear correlations among the different variables and their relative importance. The linear correlation coefficient between two variables, X and Y, is defined as,

$$\rho(X,Y) = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{< (X-\bar{X})(Y-\bar{Y}) >}{\sigma_X \sigma_Y} = \frac{< XY > - < X >< Y >}{\sigma_X \sigma_Y} \,, \tag{2.50}$$

where $< \cdot >$ denotes the expectation value. $\rho = 0$ if the two variables are uncorrelated, $\rho = +1$ if they are linearly correlated, and $\rho = -1$ if they are anti-correlated. σ_X is the standard deviation of the variable X. Linear correlation between the variables is important in determining whether or not the information carried by the variable is unique. If two variables are highly (anti)correlated, their simultaneous use in the MVA method does not improve the significance. Hence, we chose the less or moderately (anti)correlated variables.

Next, we determine the method unspecific relative importance of each variable: for a variable x_k, it is given below.

$$\Delta(x_k) = \int \frac{(\hat{y}_s(x_k) - \hat{y}_b(x_k))^2}{\hat{y}_s(x_k) + \hat{y}_b(x_k)} dx_k \,, \tag{2.51}$$

where, $\hat{y}_s$ and $\hat{y}_b$ are the probability distribution functions for the signal and background for a given observable x_k. The integration limit is the allowed range of the variable x_k. The separation power of the variable x_k equals 0, and 1 for a fully overlapping and non-overlapping probability distribution function, respectively. All the variables with substantial relative importance are kept. The data set is then split into two halves for training and testing. The training is done with the BDT algorithm, as discussed earlier. After ensuring that the network is not overtrained, we apply a cut to the BDT response. We optimized the cut that yields the highest statistical significance $\mathcal{N}_S/\sqrt{\mathcal{N}_S + \mathcal{N}_B}$. The number of signal and background events that survived after the optimum BDT cut are given by $\mathcal{N}_S$ and $\mathcal{N}_B$. We observe that the statistical significance of the MVA analysis is substantially higher than that of a cut-based analysis.

Chapter 3

Precise probing of the inert Higgs-doublet model at the LHC

3.1 Introduction

In this chapter, we consider the inert Higgs doublet model (IDM) [113, 114] as a prospective BSM scenario that gives a viable dark matter candidate and satisfies all the theoretical and experimental constraints. We explore this model in the context of the LHC with next-to-leading order QCD correction and match it to the parton shower for more accurate predictions. We find such corrections are very significant for the IDM model. In its simplest form, the present model can satisfy the whole amount of observed relic density of the DM in some particular parameter space, the so-called *resonance region* and *degenerate region*. In the former case, the relic density of the DM is produced thermally through the resonant Higgs portal annihilation. Hence, the DM mass is required to be nearly half of the Higgs boson mass and other BSM scalars carry larger masses. This region is also known as the *hierarchical mass region* as DM is the lightest, while others are quite heavy. On the other hand, DM and all other BSM scalars are nearly of equal mass ($\sim$ 500 GeV or more) in the degenerate region [115–117]. As expected, this region is harder to probe at the LHC because of the kinematic suppression due to heavy final state production, narrow mass gap and poor detection efficiency of the soft products coming from the decay of the BSM scalars.

We investigate the hierarchical mass spectrum, which cannot be analyzed using multi-jet plus missing transverse energy searches. The significant mass difference between BSM scalars and the DM leads to a very interesting signal topology due to the boosted vector boson created through heavy scalar decay. We focus on the associated production and pair production of heavy scalars. The pair production of the scalars also gets contributions from the Higgs boson-mediated S-channel diagrams. Gluon fusion is the main channel of the Higgs boson production in SM, which contains a loop at the leading order (LO). We work in the heavy top mass limit and that reduces the one loop diagram into an effective gluon-Higgs vertex. We consider $\mathcal{O}(\alpha_s)$ corrections to that effective term which is known to be as large as the LO alone. Therefore, the total Lagrangian is the sum of the IDM Lagrangian and the gluon-Higgs effective Lagrangian, and we consider $\mathcal{O}(\alpha_s)$ corrections to the total Lagrangian.

Since the decay of heavy BSM scalars produces boosted $W^{\pm}/Z$ boson, we analyzed di-fatjet plus $\not{E}_T$ signature, as this channel can explore its entire parameter space during the next phase of the LHC run. A sophisticated multivariate analysis (MVA) with jet-substructure variables is adopted in this analysis.

We organize the chapter as follows: Section 3.2 briefly describes the IDM model and the Higgs-gluon effective Lagrangian that we adopt in this computation. Section 3.3, points out various constraints on the IDM model and lists benchmark points accordingly in the hierarchical mass region. In Section 3.4, we mainly discuss the computational setup and show numerical results including the differential NLO K factor and scale uncertainties. Section 3.5 presents the distributions of different high-level kinematical variables involving jets at LO and NLO for the associated and pair production channels, demonstrating the importance of the QCD corrections. Section 3.6 explains the reason to consider $2J_V + \not{E}_T$ as the signal while dealing with a tiny IDM signal over an immense background. We also discuss here the MVA, which uses a highly non-linear cut, and use the full potential of NLO computation and jet-substructure variables to separate this tiny signal from the large background. Finally, we conclude in Section 3.7.

3.2 Theoretical framework

IDM has a new $SU(2)_L$ doublet Φ_2 in addition to the SM Higgs doublet, Φ_1, and a discrete $\mathbb{Z}_2$ symmetry is being imposed on it. All the fields of the SM are even under $\mathbb{Z}_2$ transformations. Φ_2 is odd under $\mathbb{Z}_2$ transformation and therefore the inert doublet can not acquire vacuum expectation value (vev), as vev can not change sign under any internal symmetry. As Φ_2 has no vev, we can write this doublet in terms of physical fields. $\mathbb{Z}_2$ symmetry also prevents the interaction between inert scalars and the SM fermions at any order in the perturbation series, aiding the lightest inert neutral scalar to act as a dark matter. The doublet, Φ_2, has hypercharge $Y = \frac{1}{2}$, which is equal to the hypercharge of Φ_1. These two doublets can be written in the unitary gauge as

$$\Phi_1 = \begin{pmatrix} G^+ \\ \frac{1}{\sqrt{2}}(v + h + iG^0) \end{pmatrix}, \quad \Phi_2 = \begin{pmatrix} H^+ \\ \frac{H + i\,A}{\sqrt{2}} \end{pmatrix}, \tag{3.1}$$

where G^+ and G^0 are the Goldstone bosons and the vev $v = 246\,\mathrm{GeV}$. H^+ is the charged BSM scalars. H and A are both neutral scalars; one is CP even, and the other is CP odd. Note that CP properties of the neutral scalars are basis-dependent. The most general potential [118] can be written as,

$$\begin{aligned} V_{IDM} =& \mu_1^2 \Phi_1^\dagger \Phi_1 + \mu_2^2 \Phi_2^\dagger \Phi_2 + \frac{\lambda_1}{2}(\Phi_1^\dagger \Phi_1)^2 + \frac{\lambda_2}{2}(\Phi_2^\dagger \Phi_2)^2 \\ &+ \lambda_3 (\Phi_1^\dagger \Phi_1)(\Phi_2^\dagger \Phi_2) + \lambda_4 (\Phi_2^\dagger \Phi_1)(\Phi_1^\dagger \Phi_2) \\ &+ \frac{\lambda_5}{2}\left[(\Phi_1^\dagger \Phi_2)^2 + (\Phi_2^\dagger \Phi_1)^2\right]. \end{aligned} \tag{3.2}$$

After electroweak symmetry breaking through the SM Higgs doublet, Φ_1, the masses of the BSM scalars at the tree level can be expressed as,

$$\begin{aligned} m_h^2 &= \lambda_1 v^2,\ m_{H^\pm}^2 = \mu_2^2 + \frac{1}{2}\lambda_3 v^2, \\ m_A^2 &= \mu_2^2 + \frac{1}{2}\lambda_c v^2,\ m_H^2 = \mu_2^2 + \frac{1}{2}\lambda_L v^2. \end{aligned} \tag{3.3}$$

All free parameters are real, so the scalar sector does not contain any CP violations and $\lambda_{L/c} = (\lambda_3 + \lambda_4 \pm \lambda_5)$. Higgs portal coupling λ_L, which can be positive or negative, plays an important role in the DM sector as it determines the annihilation rate of the DM in the hierarchical mass region. m_h is the SM Higgs boson mass, and $m_{H^\pm,A,H}$ are the masses of the BSM scalars. The parameters λ_1 and μ_1 can be written in terms of the mass of the Higgs boson and vev. So, IDM has five parameters – three masses of the inert scalars, self-coupling between inert scalars λ_2 and Higgs portal coupling λ_L. Self-coupling λ_2 does not affect the scalar masses and their phenomenology. In our study, we choose the inert scalar H as the dark matter candidate, but one can also choose the A as the dark matter without changing any phenomenology, just by flipping the sign of λ_5 preserving the CP properties of the DM candidate. The full IDM Lagrangian can be written as

$$\mathcal{L}_{IDM} = \mathcal{L}_{SM} + (\mathcal{D}_\mu \Phi_2)^\dagger (\mathcal{D}^\mu \Phi_2) + V_{IDM} \tag{3.4}$$

where the covariant derivative, $\mathcal{D}_\mu = (\partial_\mu - i g_Y Y B_\mu - i g \frac{\sigma^i}{2} W_\mu^i)$, and σ^i are the Pauli matrices; g and g_Y are the coupling strength of the weak and hypercharge interactions, respectively. In addition, we consider the following five-dimensional effective term to take into account Higgs interactions with gluons in the heavy top mass limit,

$$\mathcal{L}_{HEFT} = -\frac{1}{4} C_{eff}\, h\, G^a_{\mu\nu} G^{a\mu\nu}. \tag{3.5}$$

Here, $G^a_{\mu\nu}$ represents QCD field strength tensor and $C_{eff} = \frac{\alpha_s}{3\pi v}\,(1 + \frac{11}{4}\frac{\alpha_s}{\pi}) = C_0\,(1 + \frac{11}{4}\frac{\alpha_s}{\pi})$ contains terms up to $\mathcal{O}(\alpha_s^2)$, that basically takes part in the one loop corrected amplitude for Higgs boson mediated production channels.

3.3 Constraints and benchmark points

The parameter space of the IDM is very constrained from theoretical calculations, various experimental data and cosmological observations. We briefly demonstrate all these constraints and then set few benchmark points that will cover almost the entire hierarchical region of the IDM. Further details are provided in [118], [119].

The Potential must be bounded from below for any realistic model, and the vacuum should be neutral, which leads to the following constraint [119]:

$$\begin{aligned} &\lambda_1 > 0,\ \lambda_2 > 0,\ \lambda_3 + 2\sqrt{\lambda_1\lambda_2} > 0, \\ &\lambda_3 + \lambda_4 + \lambda_5 + 2\sqrt{\lambda_1\lambda_2} > 0. \end{aligned} \tag{3.6}$$

The condition $\lambda_4 + \lambda_5 < 0$ ensures the inert vacuum to be charge neutral. Generically, depending on the nature of additionally imposed symmetry, the electroweak symmetry breaking pattern has the following possibilities,

$$\begin{aligned} &v_1 = v,\ v_2 = 0, && \text{inert vacuum} \\ &v_1 = 0,\ v_2 = v, && \text{pseudo-inert vacuum} \\ &v_1 \neq 0,\ v_2 \neq 0, && \text{mixed vacuum} \end{aligned} \tag{3.7}$$

where v_1 denotes the vev of the doublet Φ_1, and v_2 is the vev of the Φ_2. v is the electroweak scale, $(G_F\sqrt{2})^{-1/2} = 246$ GeV. We want the inert vacuum as the global minima, which leads to the following constraint [120], [121]

$$\frac{\mu_1^2}{\sqrt{\lambda_1}} - \frac{\mu_2^2}{\sqrt{\lambda_2}} > 0. \tag{3.8}$$

The eigenvalues of the $2 \to 2$ scalar scattering processes [122] are given in Equation 3.9, and each eigenvalue ($|\Lambda_i|$) should be $\leq 8\pi$, coming from the perturbativity and unitarity constraints:

$$\begin{aligned}
&\Lambda_{1,2} = \lambda_3 \pm \lambda_4,\ \Lambda_{3,4} = \lambda_3 \pm \lambda_5,\ \Lambda_{5,6} = \lambda_3 + 2\lambda_4 \pm 3\lambda_5, \\
&\Lambda_{7,8} = -\lambda_1 - \lambda_2 \pm \sqrt{(\lambda_1 - \lambda_2)^2 + \lambda_4^2}, \\
&\Lambda_{9,10} = -3\lambda_1 - 3\lambda_2 \pm \sqrt{9(\lambda_1 - \lambda_2)^2 + (2\lambda_3 + \lambda_4)^2}, \\
&\Lambda_{11,12} = -\lambda_1 - \lambda_2 \pm \sqrt{(\lambda_1 - \lambda_2)^2 + \lambda_5^2}.
\end{aligned} \tag{3.9}$$

The contribution that navigates from the BSM physics to the electroweak radiative correction is parametrized by the S, T, U parameters [123], known as oblique parameters. The central values of the oblique parameters that we use in our analysis are [124]

$$S = 0.04 \pm 0.11,\ \ T = 0.09 \pm 0.14,\ \ U = -0.02 \pm 0.11\,. \tag{3.10}$$

The following parameter space of the IDM is ruled out from the neutralino search results at LEP-II [125], [126]:

$$m_H < 80 \text{ GeV},\ m_A < 100 \text{ GeV}\ ,\ \text{ and }\ (m_A - m_H) > 8 \text{ GeV}\,. \tag{3.11}$$

The charged Higgs mass gets the following bound from the chargino search results at LEP-II [127]:

$$m_{H^\pm} > 70 \text{ GeV}\,. \tag{3.12}$$

More recently, analyzing a pair of boosted hadronically decaying bosons together with MET from 13 TeV LHC data, ATLAS gave constraints on the masses of the charginos and neutralinos of the minimal supersymmetric model [128]. Based on a similar production mechanism from IDM, Reference [129] carried out a recasting analysis to show that the Higgs portal DM scenario and hierarchical heavy scalars of mass 123 GeV or above are allowed from this exclusion limit.

In the hierarchical region, the decay channels, $\Gamma(Z \to HA, H^+H^-)$ and $\Gamma(W^\pm \to H^\pm A, H^\pm H)$ are kinematically forbidden. The signal strength of the Higgs boson decay into the diphoton final state relative to the SM prediction is [130–132]:

$$\mu_{\gamma\gamma} = \frac{\sigma(pp \to h \to \gamma\gamma)}{\sigma(pp \to h \to \gamma\gamma)_{SM}} = 1.10^{+0.10}_{-0.09}\,. \tag{3.13}$$

The Higgs boson production rate is the same in both the SM and IDM models, dominated by the gluon gluon fusion channel, and so the signal strength turns out to be

$$\mu_{\gamma\gamma} = \frac{BR(h \to \gamma\gamma)_{IDM}}{BR(h \to \gamma\gamma)_{SM}}. \tag{3.14}$$

Input Parameters	BP1	BP2	BP3	BP4	BP5	BP6	BP7
$m_{H^\pm}$(GeV)	255.3	304.8	350.3	395.8	446.9	503.3	551.8
m_A(GeV)	253.9	302.9	347.4	395.1	442.4	500.7	549.63
λ_2	1.27	1.07	0.135	0.106	3.10	0.693	0.285

Table 3.1: Input parameters, masses of the BSM scalars ($m_{H^\pm}, m_A$), and the self-coupling constant (λ_2) between dark sector particles for several selected benchmark points that satisfy theoretical, DM relic density, DD data, and collider constraints listed in the text. Three other parameters are DM mass, $m_H = 53.71$ GeV, Higgs portal coupling, $\lambda_L = 5.4 \times 10^{-3}$ and Higgs boson mass $m_h = 125$ GeV.

A sufficiently large value of the λ_3 coupling and lighter charged Higgs mass can lead to enhanced decay of $h \to \gamma\gamma$, thereby pushing the ratio beyond the experimental limit and hence excluded. The upper limit of the Higgs invisible decay branching ratio measured by the ATLAS Collaboration [133] is 0.11 at 95% C.L. This measurement puts stringent constraints on the Higgs portal coupling (λ_L) and DM mass (m_H) in the region $m_H < \frac{m_h}{2}$. The Higgs invisible decay width in the IDM model is given by

$$\Gamma_{h\to HH} = \frac{\lambda_L^2 v^2}{64\pi m_h}\sqrt{1 - \frac{4m_H^2}{m_h^2}}\,. \tag{3.15}$$

$\Gamma_{h\to HH}/(\Gamma_{SM}+\Gamma_{h\to HH}) \leq 0.11$ must be satisfied in the kinematically allowed region of the decay of the Higgs boson into pair of the DM. Moreover, extremely precise measurements from WMAP [134] and PLANCK [21, 135, 136] have established that the relic abundance of the DM is $\Omega_{DM}h^2 = 0.120 \pm 0.001$ [21] with $h = \frac{\text{Hubble Parameter}}{(100\text{km s}^{-1}\text{Mpc}^{-1})}$. The dark matter annihilates into the SM particles and the relic density of the DM is inversely proportional to this annihilation rate. The observed relic density of DM sets a rigid constraint on the parameter spaces of the IDM so as not to overproduce the relic in the IDM. The spin-independent cross section of the DM-nucleon scattering processes at leading order mediated by the Higgs boson is given by [113]

$$\sigma = \frac{\lambda_L^2\, f^2}{4\pi}\,\frac{\mu^2\, m_n^2}{m_h^4\, m_{DM}^2}, \tag{3.16}$$

where m_n is the mass of the nucleon and $\mu = \frac{m_n\, m_{DM}}{m_n + m_{DM}}$. f is the Higgs-nucleon coupling strength and the allowed range of f is 0.26 - 0.63 [137]. However, the recent study suggests the value of f is 0.32 [138]. The upper bound of the DM-nucleon scattering cross section from the DM DD experiments like LUX [139] and

Xenon1T [140] poses a firm limit on the allowed values of λ_L. As already stated, we can divide the entire parameter space of the IDM into four distinct regions depending on the mass of the DM and the mass splitting between DM and other scalars - among these four, only the following two regions satisfy the observed relic density of the DM entirely.

The *hierarchical mass region* consists of a Higgs portal mass region with $m_{DM} \equiv m_H < 80$ GeV, and the mass gap with other BSM scalars as, $\Delta M \equiv \Delta M_{charged} \simeq \Delta M_{neutral} \sim 100$ GeV or more, where $\Delta M_{charged} = (m_{H^\pm} - m_{DM})$ and $\Delta M_{neutral} = (m_A - m_{DM})$. In this region, no bound on the DM mass comes from the LEP Z-boson width measurements. Since the DM mass is less than 80 GeV, the annihilation of the DM into the pair of weak gauge bosons is significantly suppressed. In this region, relic density of DM is achieved only through the Higgs portal annihilation channel. Since the mass differences between DM and other BSM scalars are significant, the co-annihilation effects are absent. As the annihilation cross section is proportional to λ_L, any small value of λ_L leads to overproduction of relic density. We get the total observed relic density of the DM in the range where the DM mass varies between 53 and 70 GeV for substantial λ_L values, constrained from DD of DM.

The *degenerate mass region* consists of high mass region, $m_{DM} \geq 500$ GeV, with rather tiny mass gap $\Delta M \sim 1$ GeV. In this regime, the following annihilation and co-annihilation processes open up:

$$
\begin{aligned}
&\text{annihilation} \begin{cases} H\,H \to W^+\,W^- \\ H\,H \to Z\,Z \end{cases} && \lambda_L \text{ sensitive} \\
&\text{co-annihilation} \begin{cases} H^+\,H^- \to W^+\,W^- \\ A\,A \to W^+\,W^- \\ A\,A \to Z\,Z \end{cases} && \lambda_L \text{ sensitive} \\
&\text{co-annihilation} \begin{cases} H^\pm\,H \to W^\pm\,\gamma & \text{gauge couplings} \\ H^\pm\,A \to W^\pm\,\gamma \end{cases}
\end{aligned}
\tag{3.17}
$$

The quartic coupling between DM and the longitudinal gauge bosons in the annihilation processes $H\,H \to W_L^+\,W_L^-$ and $H\,H \to Z_L\,Z_L$ is $(4\,m_{DM}\,\Delta M/v^2 + \lambda_L)$. In this degenerate mass spectrum $\Delta M \to 0$, and so this coupling remains sensitive to λ_L mostly. The relic density of DM increases with the DM mass and decreases with the annihilation cross section. Those combined effects set the correct relic density of DM in this region for $m_{DM} \geq 500$ GeV. Although this region is difficult to probe, with a charged long-lived Higgs boson, one can explore this region at the LHC with the charged track signal [141].

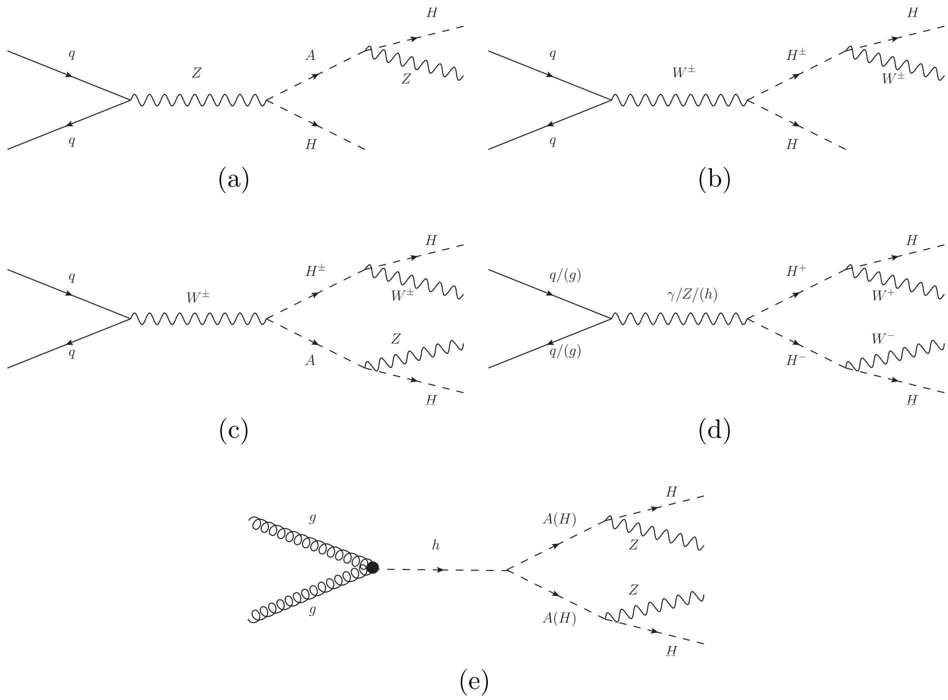


Figure 3.1: Parton level representative diagrams at LO of (a), (b) associate production of heavy scalar, and (c), (d), (e) pair production of heavy scalars. In our study, we consider one loop correction in α_S of all these diagrams.

The different benchmark points that we pick out for this study are given in Table 3.1, and all of them satisfy the constraints discussed above.

3.4 Computational setup and numerical results

We implement the Lagrangian given in Equation 3.4 together with the leading term of Equation 3.5 in FeynRules [26] and employ NLOCT [142] to generate UV and R_2 counter terms of the SM Lagrangian in order to have a NLO UFO model that we use under the MG5_aMC@NLO environment [69]. Inside this environment, real corrections are performed following the FKS subtraction method [143], whereas OPP technique [144] is the one that is being used to take care of the virtual contributions. Nevertheless, for AA, HH and H^+H^- pair production processes, gluon-gluon initiated processes mediated by Higgs propagator play a significant role and we insert the corresponding analytic form of the one loop amplitude in MadGraph5 virtual routine and that in $d = (4 - 2\epsilon)$

dimension reads as,

$$\begin{aligned}\overline{2\mathcal{R}(M_0M_v^\dagger)} &= \left(\frac{\alpha_s}{2\pi}\right)\frac{(4\pi)^\epsilon}{\Gamma(1-\epsilon)}\left(\frac{\mu^2}{s_{12}}\right)^\epsilon|\overline{M_0}|^2 \\ &\left[-\frac{6}{\epsilon^2}-\frac{2\,b_0}{\epsilon}+11+3\pi^2\right], \qquad (3.18)\end{aligned}$$

while setting the renormalization scale $\mu^2 = s_{12}$, partonic center-of-mass (CM) energy. $\mathcal{M}_0$ and $\mathcal{M}_v$ represent tree-level and one-loop amplitudes respectively. The leading term of the QCD β-function $b_0 = \frac{11}{6}C_A - \frac{2}{3}n_f T_F$, where n_f represents the number of active quark flavors and $C_A = 3, T_F = 1/2$. Note that the strong coupling is renormalized following the $\overline{\text{MS}}$ scheme and the $\mathcal{O}(\alpha_s^2)$ term of the Lagrangian given is Equation 3.5 is taken into account in the above expression. The color and spin averaged tree level squared amplitude in $d = (4-2\epsilon)$ dimension can be written as,

$$|\overline{M_0}|^2 = \frac{1}{128}(1+\epsilon+\epsilon^2)\frac{C_0^2\,\Lambda^2\,v^2\,s_{12}^2}{(s_{12}-m_h^2)^2+\Gamma_h^2}. \qquad (3.19)$$

Here $C_0 = \frac{\alpha_s}{3\pi v}$, Γ_h is the Higgs boson width, and Λ corresponds to $\Lambda_{L/c/3}$ as given in the Feynman rules furnished in Appendix C. Final state heavy scalar particles are decayed via MadSpin [145] which retains spin information at the tree level accuracy. NLO events thus obtained are then matched to Pythia8 [146,147] parton shower following the MC@NLO formalism [148] to avoid any double counting. For the signal, we use in-built NN23LO1 and NN23NLO PDF sets [149] for LO and NLO respectively. We use Delphes3 [70] to include the detector effects in our simulation, where we use the default card of the CMS. Jets are formed by clustering the particle-flow tower objects and particle-flow tracks. We employ anti-k$_T$ [98] clustering algorithm to form jets, where we have set radius parameter R=0.5. Using FastJet 3.2.2 [92] package, we reconstruct fatjets, utilizing Delphes tower objects as input for clustering. Cambridge-Achen (CA) [95] algorithm is hired for fatjets clustering where radius parameter is set to R=0.8. Fatjets are characterized by the radius parameter, $R \sim 2m_V/P_T$ ($V \equiv \{W^\pm, Z\}$), where P_T being the transverse momentum and m_V is the mass of the weak boson. We apply minimum P_T = 180 GeV for each fatjet formation. MVA analysis is done in the TMVA framework [112]. We implement the Boosted Decision Tree (BDT) algorithm in our MVA analysis. A decision tree splits the high-level input data recursively depending on a set of input features. The method that combines many trees (weak learners) into a strong classifier is called boosting. Figure 3.1 displays representative LO Feynman diagrams of the associated production of the heavy scalar and pair production of the heavy scalars, which we ultimately decay hadronically. Production cross sections for these channels before hadronic decay of the heavy scalars are given in Tables 3.2, and 3.3 at 14 TeV LHC. We choose the renormalization scale and the factorization scale as $\mu_R = \zeta_R\sqrt{s_{12}}$ and $\mu_F = \zeta_F\sqrt{s_{12}}$ respectively, where $\zeta_R = \zeta_F = 1$ represents the central scale choice. We vary $\zeta_R, \zeta_F = \{1/2, 1, 2\}$, which has a total of nine datasets. All the cross sections are given corresponding to the central scale where superscripts and subscripts denote the envelope of those nine scale choices. The Monte Carlo uncertainties are also given in those tables. We get reduced scale uncertainty in the total cross section at NLO than LO for both the associated and pair production processes, except in a few benchmark points for the associated production processes and the reason could be the cross-over of the envelopes around the maximum differential LO cross section, unlike NLO (see Figures 3.3a: bottom, 3.3c: bottom). Fractional scale uncertainty is defined as the envelope of the ratios of the differential cross sections at eight additional (ζ_R, ζ_F) choices to the central one. Dashed and solid lines in the fractional scale uncertainty subplot correspond to the lower and upper envelope respectively. Our study includes one order in α_S corrections to all these channels. Cross section of $pp \to HH$ channel at LO is 0.332 pb and at NLO it is 0.617 pb (*i.e,* K factor = 1.858) at 14 TeV LHC, independent of benchmark points since cross section depends only on m_H and λ_L, and both remain same for chosen benchmark points. This channel has a larger cross section than any other pair or associated production channels because of being less s-channel suppressed due to the presence of an on-shell Higgs boson mediator. Total transverse momentum distribution of the DM pair for BP2 of the channel, $pp \to HH$ is shown in Figure 3.2 for fixed order NLO (dashed blue) and NLO matched with parton-shower (solid red). It is clear from this figure that NLO+PS describes the low P_T region more vividly compared to a fixed order estimation. Note that, although such calculation is essential for traditional mono-jet search, possible contributions of $pp \to HH$ can only come in our di-fatjet study at the NNLO level. Characteristically, this process is background like and we find that much of the events will not pass the event selection criteria even while starting from a reasonably significant contribution. This channel is shown here for completeness, but we would not add such a contribution to our conservative estimate.

BP	$\sigma(pp \to AH)$ (fb)			$\sigma(pp \to H^{\pm}H)$ (fb)		
	NLO	LO	K-fac	NLO	LO	K-fac
BP1	$46.55^{+0.79(1.7\%)}_{-0.65(1.4\%)} \pm 8.0 \times 10^{-2}$	$35.12^{+0.04(0.1\%)}_{-0.14(0.4\%)} \pm 1.3 \times 10^{-1}$	1.33	$82.42^{+1.26(1.5\%)}_{-0.98(1.2\%)} \pm 2.0 \times 10^{-2}$	$61.67^{+0.12(0.2\%)}_{-0.37(0.6\%)} \pm 2.2 \times 10^{-1}$	1.34
BP2	$25.34^{+0.40(1.6\%)}_{-0.29(1.1\%)} \pm 4.0 \times 10^{-2}$	$19.01^{+0.17(0.9\%)}_{-0.23(1.2\%)} \pm 7.0 \times 10^{-2}$	1.33	$45.27^{+0.87(1.9\%)}_{-0.40(0.9\%)} \pm 8.0 \times 10^{-2}$	$33.77^{+0.37(1.1\%)}_{-0.47(1.4\%)} \pm 1.0 \times 10^{-2}$	1.34
BP3	$15.50^{+0.21(1.4\%)}_{-0.21(1.4\%)} \pm 3.0 \times 10^{-2}$	$11.60^{+0.17(1.5\%)}_{-0.21(1.8\%)} \pm 4.0 \times 10^{-2}$	1.34	$27.78^{+0.44(1.6\%)}_{-0.3(1.1\%)} \pm 5.0 \times 10^{-2}$	$20.69^{+0.37(1.8\%)}_{-0.41(2\%)} \pm 8.0 \times 10^{-2}$	1.34
BP4	$9.68^{+0.15(1.6\%)}_{-0.14(1.4\%)} \pm 2.0 \times 10^{-2}$	$7.19^{+0.16(2.2\%)}_{-0.17(2.3\%)} \pm 3.0 \times 10^{-2}$	1.35	$17.86^{+0.27(1.5\%)}_{-0.29(1.6\%)} \pm 3.0 \times 10^{-2}$	$13.23^{+0.32(2.4\%)}_{-0.33(2.5\%)} \pm 4.0 \times 10^{-2}$	1.35
BP5	$6.32^{+0.11(1.7\%)}_{-0.10(1.6\%)} \pm 1.0 \times 10^{-2}$	$4.67^{+0.13(2.8\%)}_{-0.13(2.8\%)} \pm 2.0 \times 10^{-2}$	1.35	$11.34^{+0.2(1.8\%)}_{-0.19(1.7\%)} \pm 2.0 \times 10^{-2}$	$8.39^{+0.25(3\%)}_{-0.25(3\%)} \pm 3.0 \times 10^{-2}$	1.35
BP6	$3.90^{+0.07(1.8\%)}_{-0.07(1.8\%)} \pm 6.5 \times 10^{-3}$	$2.87^{+0.10(3.5\%)}_{-0.10(3.5\%)} \pm 1.0 \times 10^{-2}$	1.36	$7.19^{+0.14(1.9\%)}_{-0.13(1.8\%)} \pm 1.0 \times 10^{-2}$	$5.33^{+0.19(3.6\%)}_{-0.19(3.6\%)} \pm 2.0 \times 10^{-2}$	1.35
BP7	$2.69^{+0.05(1.9\%)}_{-0.05(1.9\%)} \pm 4.6 \times 10^{-3}$	$1.97^{+0.08(4\%)}_{-0.07(3.6\%)} \pm 7.3 \times 10^{-3}$	1.37	$5.01^{+0.1(2\%)}_{-0.1(2\%)} \pm 8.9 \times 10^{-3}$	$3.68^{+0.15(4\%)}_{-0.15(4\%)} \pm 1.0 \times 10^{-2}$	1.36

Table 3.2: Cross sections for the associated production of heavy scalar at LO and NLO with integrated K-factor are given in this table at 14 TeV LHC before the decay of heavy scalars into DM and SM particles. The superscript and subscript denote the scale uncertainties in the total cross section (the percentages are given in bracket), while the last entry is the Monte Carlo uncertainty. AH and $H^{\pm}H$ channels are produced in five and four massless quark flavors, respectively.

BP	$\sigma(pp \to H^\pm A)$ (fb)			$\sigma(pp \to H^+H^-)$ (fb)		
	NLO	LO	K-fac	NLO	LO	K-fac
BP1	$16.93^{+0.28(1.7\%)}_{-0.25(1.5\%)} \pm 3.0 \times 10^{-2}$	$12.53^{+0.34(2.7\%)}_{-0.34(2.7\%)} \pm 5.0 \times 10^{-2}$	1.35	$11.01^{+0.43(3.9\%)}_{-0.42(3.8\%)} \pm 2.0 \times 10^{-2}$	$7.98^{+0.48(6\%)}_{-0.4(5\%)} \pm 2.0 \times 10^{-2}$	1.38
BP2	$8.41^{+0.14(1.7\%)}_{-0.18(2.1\%)} \pm 2.0 \times 10^{-2}$	$6.21^{+0.23(3.7\%)}_{-0.22(3.5\%)} \pm 3.0 \times 10^{-2}$	1.35	$6.01^{+0.37(6.2\%)}_{-0.32(5.3\%)} \pm 8.5 \times 10^{-3}$	$4.17^{+0.37(8.9\%)}_{-0.29(7\%)} \pm 1.0 \times 10^{-2}$	1.44
BP3	$4.78^{+0.1(2.1\%)}_{-0.1(2.1\%)} \pm 8.3 \times 10^{-3}$	$3.48^{+0.15(4.3\%)}_{-0.14(4\%)} \pm 1.0 \times 10^{-2}$	1.37	$3.76^{+0.29(7.7\%)}_{-0.24(6.4\%)} \pm 5.8 \times 10^{-3}$	$2.57^{+0.29(11.3\%)}_{-0.23(8.9\%)} \pm 7.6 \times 10^{-3}$	1.46
BP4	$2.81^{+0.06(2.1\%)}_{-0.06(2.1\%)} \pm 5.0 \times 10^{-3}$	$2.04^{+0.1(4.9\%)}_{-0.1(4.9\%)} \pm 7.8 \times 10^{-3}$	1.38	$2.50^{+0.23(9.2\%)}_{-0.19(7.6\%)} \pm 4.1 \times 10^{-3}$	$1.68^{+0.24(14.3\%)}_{-0.18(10.7\%)} \pm 5.3 \times 10^{-3}$	1.49
BP5	$1.69^{+0.04(2.4\%)}_{-0.04(2.4\%)} \pm 3.0 \times 10^{-3}$	$1.22^{+0.07(5.7\%)}_{-0.06(4.9\%)} \pm 4.7 \times 10^{-3}$	1.38	$1.68^{+0.18(10.7\%)}_{-0.15(8.9\%)} \pm 3.0 \times 10^{-3}$	$1.10^{+0.19(17.3\%)}_{-0.14(12.7\%)} \pm 3.1 \times 10^{-3}$	1.52
BP6	$0.97^{+0.02(2.1\%)}_{-0.02(2.1\%)} \pm 1.7 \times 10^{-3}$	$0.70^{+0.05(7.1\%)}_{-0.04(5.7\%)} \pm 2.4 \times 10^{-3}$	1.38	$1.14^{+0.14(12.3\%)}_{-0.12(10.5\%)} \pm 2.4 \times 10^{-3}$	$0.74^{+0.15(20.3\%)}_{-0.11(14.9\%)} \pm 2.2 \times 10^{-3}$	1.54
BP7	$0.63^{+0.018(2.8\%)}_{-0.018(2.8\%)} \pm 1.1 \times 10^{-3}$	$0.45^{+0.03(6.7\%)}_{-0.03(6.7\%)} \pm 1.6 \times 10^{-3}$	1.39	$0.85^{+0.11(12.9\%)}_{-0.09(10.6\%)} \pm 1.9 \times 10^{-3}$	$0.55^{+0.13(23.6\%)}_{-0.09(16.4\%)} \pm 1.7 \times 10^{-3}$	1.56

BP	$\sigma(pp \to AA)$ (fb)		
	NLO	LO	K-fac
BP1	$0.88^{+0.18(20.4\%)}_{-0.14(15.9\%)} \pm 3.3 \times 10^{-3}$	$0.46^{+0.15(32.6\%)}_{-0.11(23.9\%)} \pm 1.6 \times 10^{-3}$	1.92
BP2	$0.72^{+0.14(19.4\%)}_{-0.12(16.7\%)} \pm 2.6 \times 10^{-3}$	$0.38^{+0.13(34.2\%)}_{-0.09(23.7\%)} \pm 1.4 \times 10^{-3}$	1.87
BP3	$0.59^{+0.12(20.3\%)}_{-0.10(16.9\%)} \pm 2.2 \times 10^{-3}$	$0.32^{+0.11(34.3\%)}_{-0.08(25\%)} \pm 1.1 \times 10^{-3}$	1.86
BP4	$0.48^{+0.1(20.8\%)}_{-0.08(16.7\%)} \pm 1.7 \times 10^{-3}$	$0.27^{+0.09(33.3\%)}_{-0.07(25.9\%)} \pm 9.6 \times 10^{-4}$	1.80
BP5	$0.40^{+0.08(20\%)}_{-0.07(17.5)} \pm 1.7 \times 10^{-3}$	$0.22^{+0.08(36.4\%)}_{-0.05(22.7\%)} \pm 8 \times 10^{-4}$	1.78
BP6	$0.31^{+0.06(19.4\%)}_{-0.05(16.1\%)} \pm 1.1 \times 10^{-3}$	$0.18^{+0.06(33.3\%)}_{-0.05(27.8\%)} \pm 6.5 \times 10^{-4}$	1.75
BP7	$0.26^{+0.05(19.2\%)}_{-0.05(19.2\%)} \pm 9.3 \times 10^{-4}$	$0.15^{+0.05(33.3\%)}_{-0.04(26.7\%)} \pm 5.4 \times 10^{-4}$	1.70

Table 3.3: Same as Table 3.2 but for the pair production of the heavy scalars. $H^\pm A$ is produced in four massless quark flavors, while H^+H^- and AA channels are produced in five massless quark flavors.

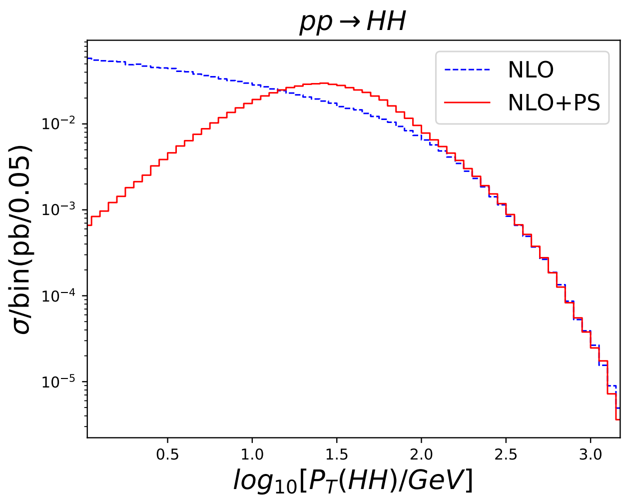
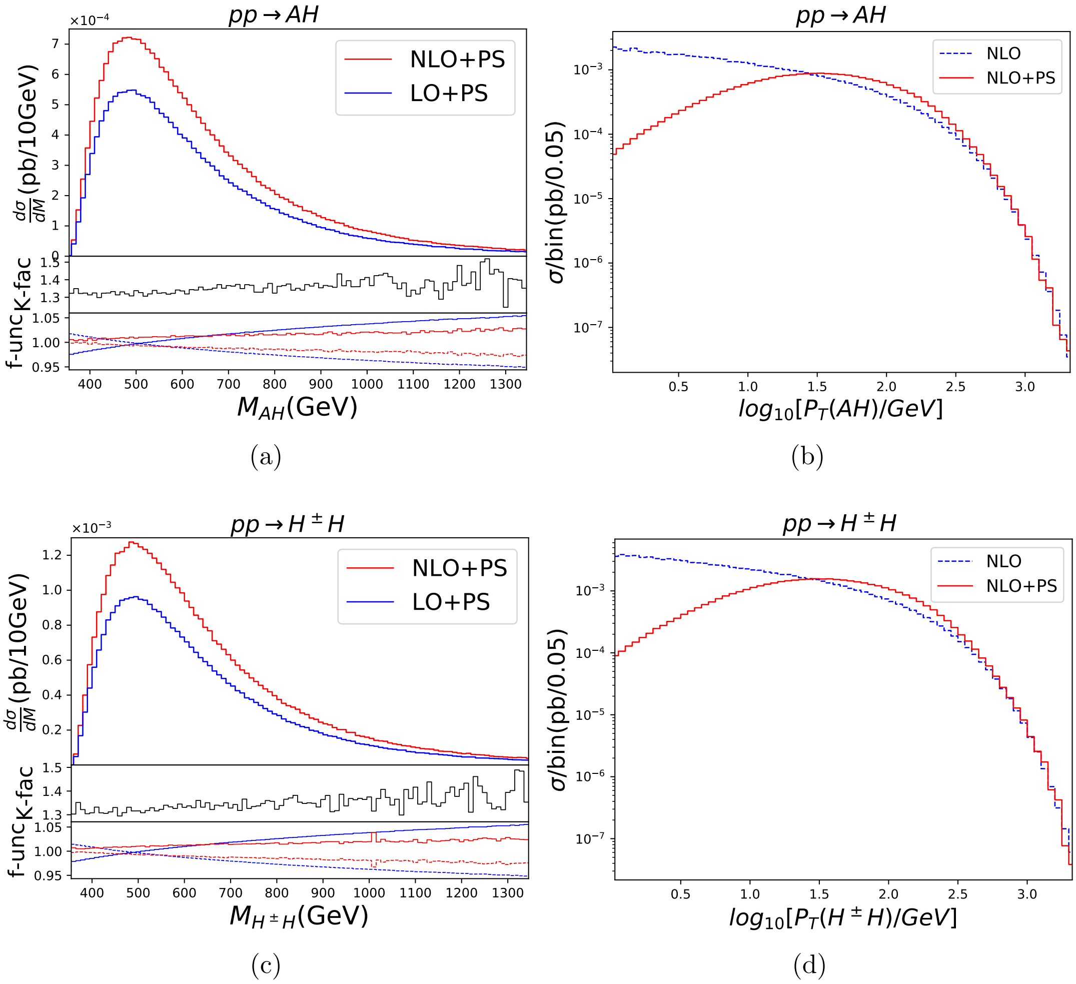


Figure 3.2: Differential distribution of the total transverse momentum of the DM pair for the channel $pp \to HH$ at fixed order NLO (dashed blue) and NLO+PS (solid red) accuracy.

In the subsequent figures, on the left panel, we show the improvement in NLO+PS results over the LO+PS ones on the invariant mass distribution (top) along with differential K-factor (middle) and fractional scale uncertainties (bottom) for all remaining production channels. Differential K-factor is vital in extracting correct signal efficiency, as most collider analyses usually do not cover the entire phase space and apply various kinematical cuts to distinguish signal from the background. Fractional scale uncertainty denotes how stable the NLO result is as compared to the LO under scale variation. On the right panel, Sudakov suppression due to NLO+PS computation is explicitly shown for each corresponding channel and that ensures re-summation of large logarithm terms in the low P_T region because of incorporating parton shower effect on top of the fixed order calculation. Note that, in these sets of representative figures, hadronic decays of final state heavy scalars are not considered for the time being. Figure 3.3 collects all the associated production channels of heavy scalars, whereas Figure 3.4 contains various pair production channels of heavy scalars. In all these figures, BP2 is considered as the representative benchmark point. The invariant mass distributions for the associated production channels peak around the same region, close to 485 GeV for both $pp \to AH$ (Figure 3.3a: top) and $pp \to H^{\pm}H$ (Figure 3.3c: top). However, among the pair production channels, vector boson mediated processes *viz.* $pp \to H^{\pm}A$ (Figure 3.4a: top) and $pp \to H^{+}H^{-}$ (Figure 3.4c: top) peak around 785 GeV and 730 GeV respectively, but the peak for the other one *i.e,* $pp \to AA$ (Figure 3.4e: top) occurs near to 650 GeV which is solely scalar mediated. This indicates that the final state particles coming from the associated production processes would be softer compared to the pair production processes.



(a) (b) (c) (d)

Figure 3.3: NLO effects on the associated production of heavy scalar channels, such as, $pp \to AH$ (subfigures [a], [b]), $pp \to H^{\pm}H$(subfigures [c], [d]). In each plot of the left panel, the top subplot shows the invariant mass distribution of the heavy scalar and DM pair at LO + PS (dashed blue) and NLO + PS (solid red) accuracy. The middle subplot displays the differential NLO K factor, the ratio of the NLO + PS cross section to the LO + PS one in each bin, while the bottom subplot presents the scale uncertainties for LO+PS (blue) and NLO+PS (red). The right panel shows the differential distribution of the total transverse momentum of the heavy scalar and DM pair for the respective channel at fixed order NLO (dashed blue) and NLO+PS (solid red) accuracy. All distributions are given for sample benchmark point BP2.

K-factor varies substantially, and in some kinematic regions, it indicates correction up to 90% . Nature of scale uncertainties for associated production processes are quite similar. Among pair production processes, fractional scale uncertainties for $pp \to H^+H^-$ (Figure 3.4c: bottom) and $pp \to AA$ (Figure 3.4e: bottom) are mostly stable in the high invariant mass region, whereas for $pp \to H^{\pm}A$ (Figure 3.4a: bottom) such uncertainties are monotonically increasing. Although these

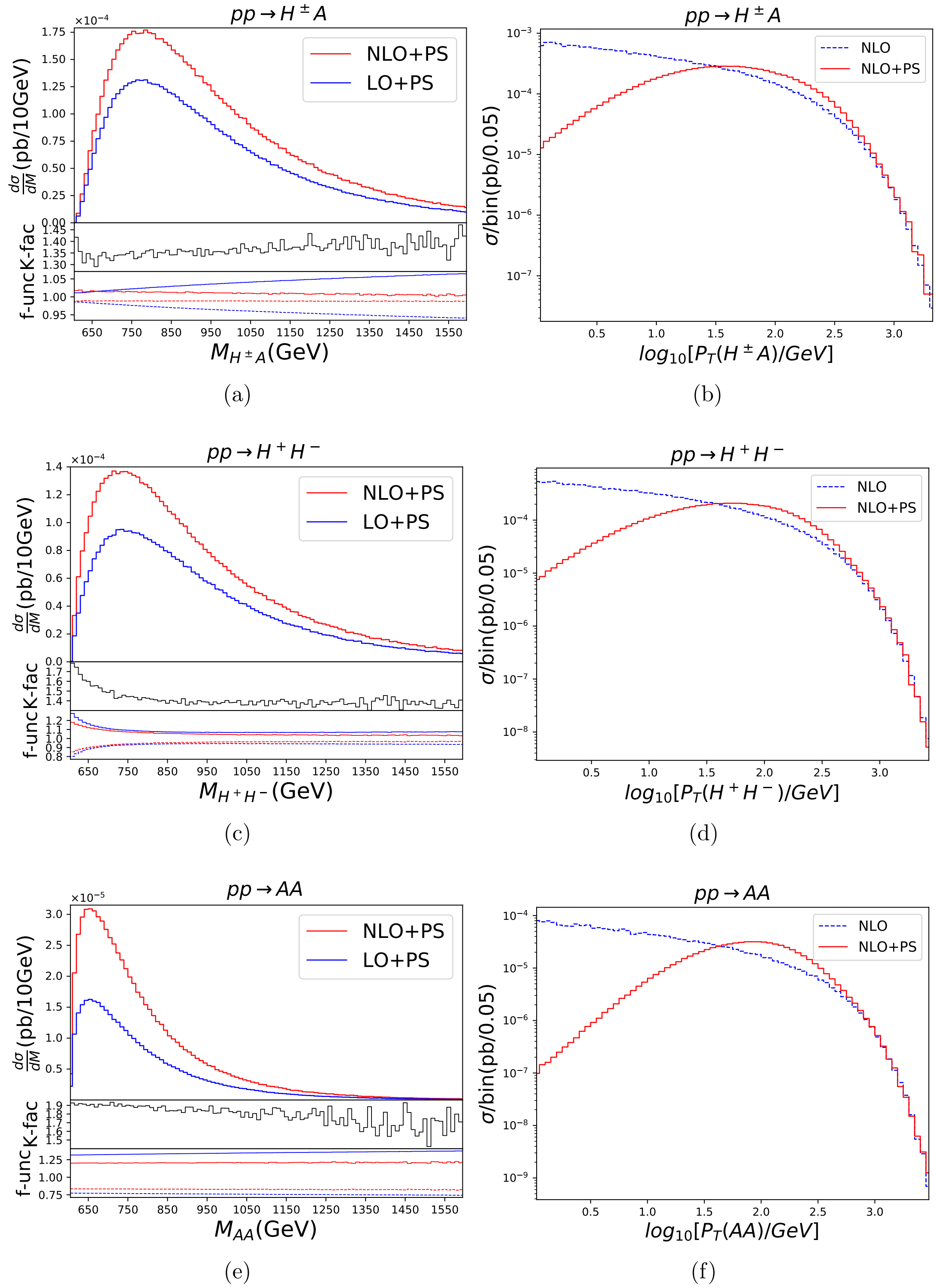
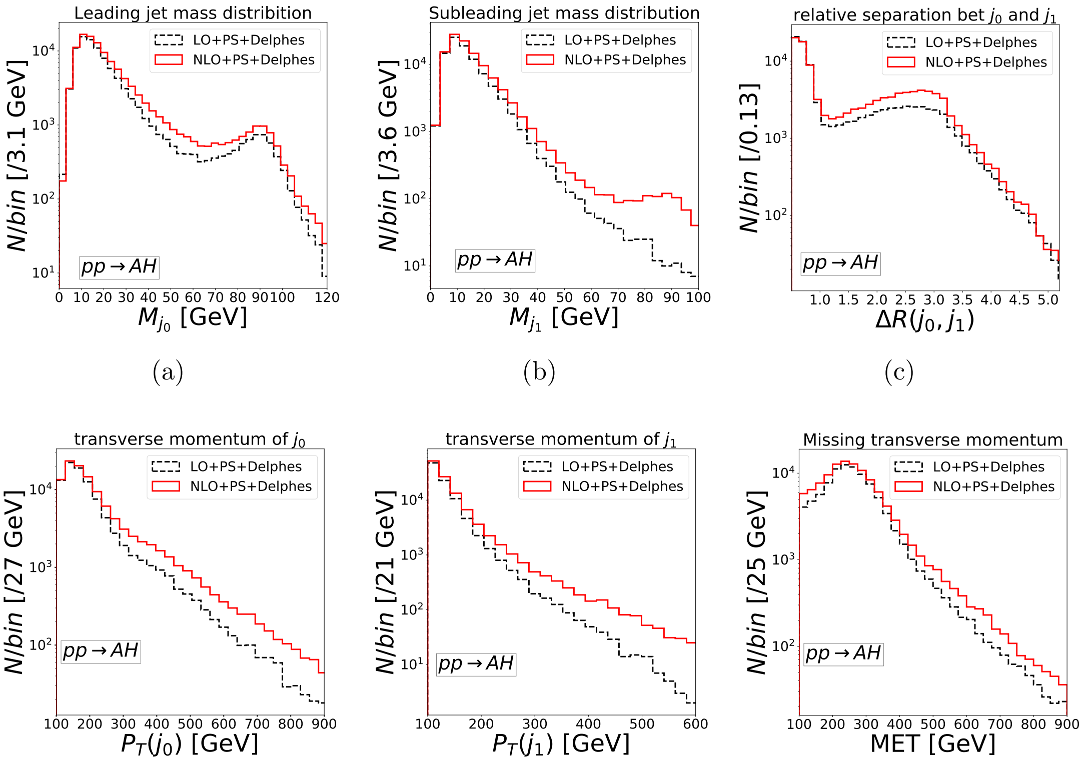


Figure 3.4: Kinematic variables in the left and right panels are same as in Figure 3.3, but these are for the pair production of the heavy scalar channels, such as, $pp \to H^\pm A$ (subfigures [a], [b]), $pp \to H^+H^-$ (subfigures [c], [d]), and $pp \to AA$ (subfigures [e], [f]).



(a) (b) (c)

(d) (e) (f)

Figure 3.5: Distributions of the various kinematic observables at LO (dashed black) and NLO (solid red) for the selected events with $\not{E}_T$, $P_T(j_0), P_T(j_1) > 100\,GeV$ from the channel $pp \to AH$, where A decay hadronically. This demonstration is for the benchmark point BP2. Plots (a) and (b) show distributions of the leading (j_0) and subleading (j_1) jet mass (M_{j_0}, M_{j_1}) respectively, (c) is the distribution of the relative separation between these two leading jets $\Delta R(j_0, j_1)$, while (d) and (e) are transverse momentum distribution of j_0 and j_1 respectively. Plot (f) shows the distribution of the total missing transverse energy; here the label MET represents $\not{E}_T$.

results are metaphorical as hadronic decay of the final state heavy scalars are not being considered here, they show the importance of doing $\mathcal{O}(\alpha_s)$ corrections to all the production channels to have better estimation of production rate and reduced scale uncertainty.

3.5 QCD jets from heavy scalar decay

Heavy scalars, after their creation through the associated channel along with DM candidate H, or from a pair production, primarily decay into H and a gauge boson, which is further decayed hadronically. It is imperative to look into their dominant hadronic decay channels as a possible probe for IDM. We select the

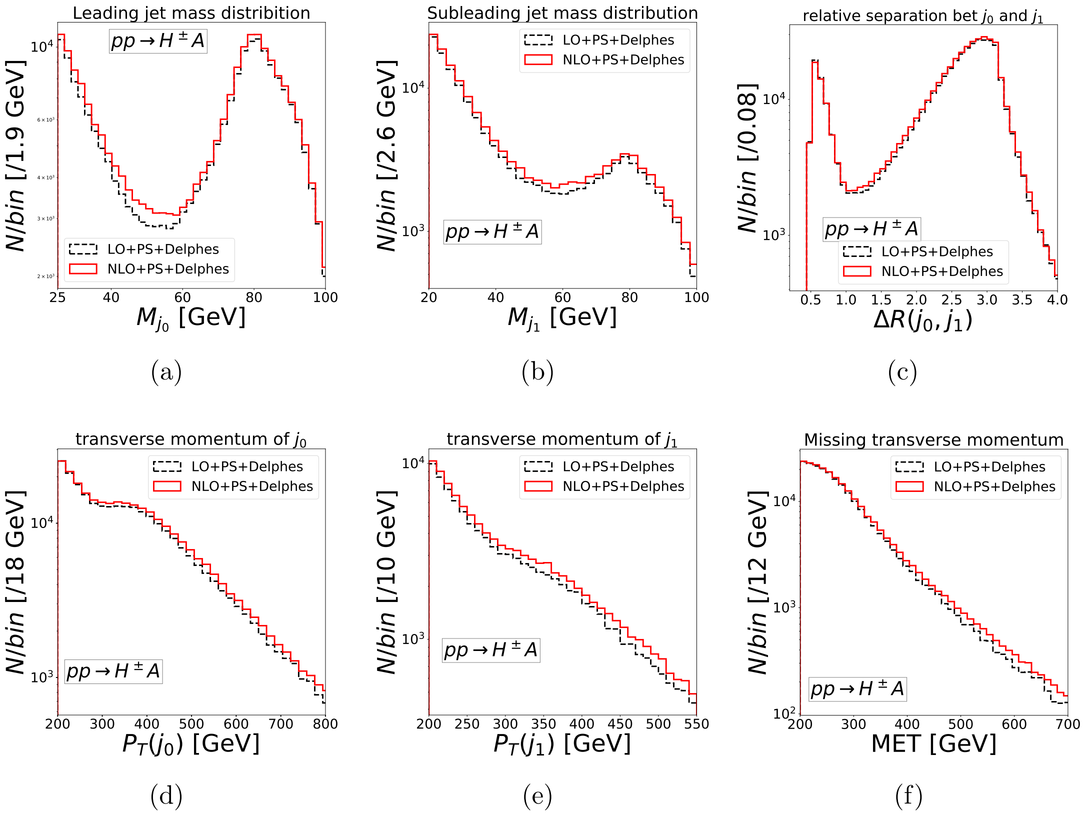


Figure 3.6: Panels are the same as in Figure 3.5, but for the pair production of the heavy scalars channel, such as $pp \to H^{\pm}A$, where both A, $H^{\pm}$ decay hadronically.

simulated events including the parton shower and detector effect with a minimum missing transverse energy, $\not{E}_T > 100$ GeV, and the minimum transverse momentum of the two leading jets $P_T(j_i) > 100$ GeV (for $i = 0, 1$). Particle-flow towers and particle-flow tracks are used as input to cluster the jets of radius parameter 0.5, where we use the anti-K$_T$ algorithm for clustering. The jet mass is defined by $M_j = (\sum_{i\in j} P_i)^2$, where P_i is the four-momentum of the i-th constituent within the jet. The missing transverse energy is defined as the negative sum of the transverse momentum of all the reconstructed constituents, $\not{E}_T = -\sum_i \vec{P}_{T,i}$. The angular distance between two jets in the transverse plane is denoted as $\Delta R(j_i, j_j)$. This section aims to examine relevant distributions of the jets from the signal to motivate the significance of NLO QCD calculation over the LO. In addition to upward shift, NLO corrections can change the shape of the distribution for a variety of kinematical variables. This has a profound effect in constructing the phenomenological study. These distributions also make a case for large-radius jets (fatjets) originated from boosted $Z/W^{\pm}$ boson decay which comes naturally in probing the hierarchical mass region of the IDM.

The distribution of the different high-level observables for one of the associ-

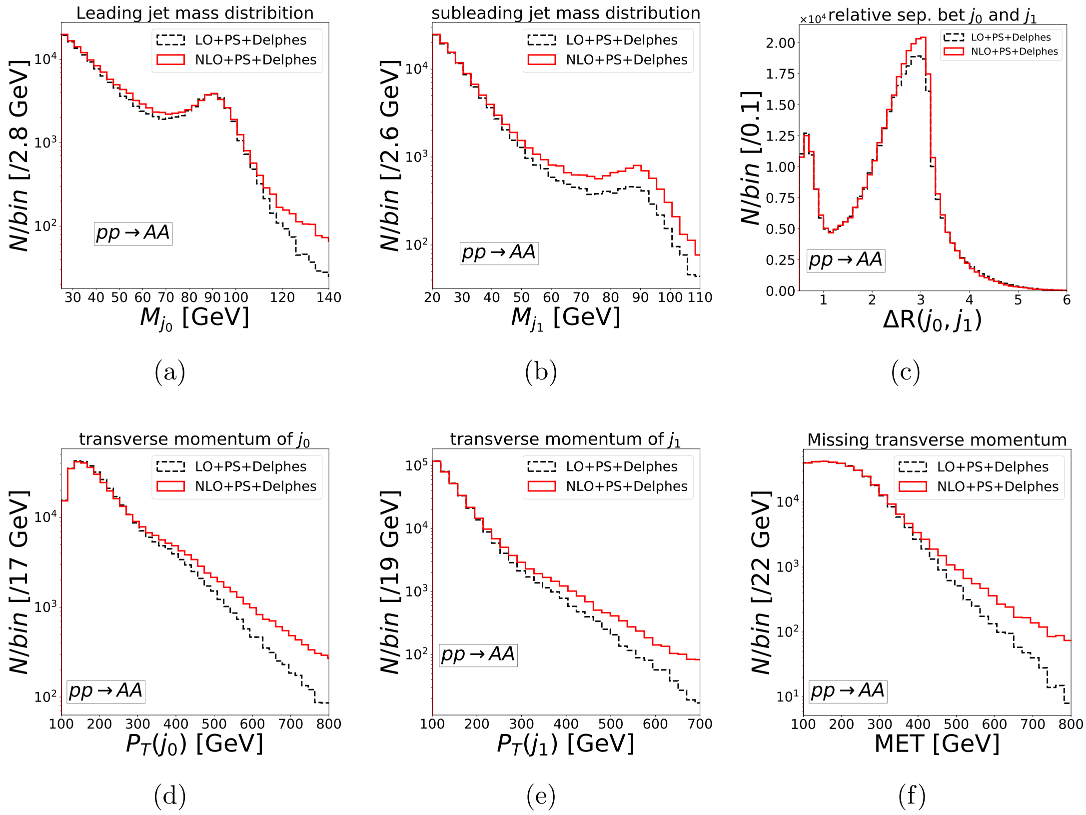


Figure 3.7: Panels are the same as in Figure 3.5, but for the pair production of the heavy scalars channel, such as $pp \to AA$, where both A decay hadronically.

ated production channels of heavy scalar [a], $pp \to AH$, and vector boson mediated pair production of the heavy scalars, $pp \to H^{\pm}A$ and scalar mediated pair production, $pp \to AA$ are shown in Figures 3.5, 3.6, and 3.7 respectively displaying the LO (dashed black) and NLO (solid red) contributions considering a sample benchmark point BP2. In each figure, the first two plots (a) and (b) present distributions of the leading (j_0) and subleading (j_1) jet mass, respectively. Plot (c) presents the distribution of the relative separation between these leading and subleading jets, whereas plots (d) and (e) exhibit their transverse momentum distributions, respectively. Finally, plot (f) shows the distribution of the total missing transverse energy from such production.

The channel $pp \to AH$ at the partonic level produces three hard jets, two from Z boson decay, and the other is the NLO radiation, while at LO, it has only two hard jets from Z boson decay. The first peak in the leading jet mass distribution (Figure 3.5a) is generated when a QCD hard parton forms a jet after PS and detector simulation. Interestingly, this same distribution points to a second peak both for LO and NLO results. This occurs when the Z boson is

[a]Both $pp \to AH$ and $pp \to H^{\pm}H$ channels follow similar distributions, as both A and $H^{\pm}$ masses are nearly degenerate and produced through vector mediator.

produced with enough boost to form a merged jet out of its full decay products, resulting into a peak at Z boson mass. The second hard parton from the Z boson forms the subleading jet, causing a peak near $M_{j_1} = 10$ GeV (Figure 3.5b) but no other peak in the LO j_1 mass distribution. However, the NLO distribution can have extra hard radiation. Occasionally when that carries enough transverse momentum to form a leading jet, Z boson decay still forms a merged subleading jet resulting in a second peak near Z boson mass (Figure 3.5b) deviating from a leading order estimate. Hence NLO estimate predicts an upward trend in the number of boosted di-jet events even from such associated production channels. One can also expect such abundance in boosted jets for other benchmark points with heavier scalars. Our previous argument is even more evident in the next distribution plot of the relative separation between two leading jets (Figure 3.5c) for the same channel $pp \to AH$. The number of events with smaller jet separation $\Delta R_{j_0,j_1} < 1.0$ is one order larger than in the other region. For a significant event sample, both leading and subleading jets come from the Z boson's decay and are closely separated. Naturally, the construction of large-radius jets embeds them together to form a single fatjet carrying properties of originating gauge boson. It is even more pronounced in larger masses of scalar. The distribution of the transverse momentum of the leading (Figure 3.5d) and subleading (Figure 3.5e) jets and the total missing transverse energy (Figure 3.5f) shows an upswing in NLO at larger PT. This is significant in view of the final selection of events (or, during multivariate analysis at boosted decision tree) comes with higher weightage from these distribution tails to deal with a tiny signal over an overwhelmingly large background.

Similarly, one requires to follow distributions from pair production channels of the heavy scalars. The leading and subleading jet mass distributions for vector boson mediated (Figures 3.6a, 3.6b) and scalar boson mediated (Figures 3.7a, 3.7b) channels in pair production of heavy scalars demonstrates two clear mass peaks both at LO and NLO. In this case, pairs of heavy scalars produce two boosted vector bosons, and as they have enough boost, it results into the second peak in both cases. Again, with the increase of scalar mass, the second peak rises, ensuring enhancement of di-fatjet events. The distributions of the relative separation for pair production of heavy scalars shown in Figures 3.6c and 3.7c contain two peaks. The second peak at $\Delta R \sim \pi$ appears when two jets originate from two different vector bosons. The first peak is when both the jets arrive from the same vector boson, which gradually diminishes for heavier mass. Pair production channel $pp \to AA$ has a significant shift between NLO and LO distributions in comparison to the $pp \to H^{\pm}A$ channel, as the former is Higgs mediated and has a larger K-factor. It is evident from this discussion that the tagging of large-radius jets originating from boosted vector bosons can significantly improve the efficiency of probing the hierarchical mass spectrum of the IDM. In the next section, we will describe the selection and properties of such boosted fatjets.

3.6 Boosted fatjet as a proxy for heavy scalar production

Our discussion in the previous section demonstrates that the multi-jet + $\not{E}_T$ search is not sufficient to explore the hierarchical mass region of IDM. Jet pair originated from the vector bosons, which comes out as boosted decay product of heavy scalar, is already collimated as a merged hadronic object. This process of getting a fatjet becomes more and more evident while probing a heavier scalar mass. A large radius fatjet can effectively identify this combined hadronic yield from the boosted vector boson. Moreover, it can carry a significant amount of information hidden inside the internal structure of jet formation through the orientation of fragmented hadrons and their energy deposits, revealing the properties and identity of the originating particle.

3.6.1 Signal and background processes

Representative LO Feynman diagrams both for associate production and pair production of heavy scalars are already depicted in Figure 3.1. Our primary focus is to analyze the NLO accurate di-fatjet signal arising from heavy IDM scalar decay using jet substructure variables. We do not discriminate W-jet or Z-jet and dub them as V-fatjet (J_V) since we consider a suitable mass window to accommodate both in our analysis. We will discuss the usefulness of the sophisticated multivariate analysis that can make the signature of $2J_V + \not{E}_T$ into the better discriminator in order to separate out tiny signal from an overwhelmingly large SM background. However, multivariate description creates a highly performant nonlinear cut at the cost of blurring the exact physical description of how different high-level variables affect our analysis. Hence, to better understand the kinematic variables that may affect LO and NLO computations, we would analyze them first with usual cut-based method before moving on to the MVA analysis. In passing, it is to be noted that the cross section of the di-Higgs production while one Higgs boson decay into a pair of bottom quarks ($h \to b\bar{b}$) and the other decays into pair of dark matter ($h \to HH$) is 1.05 fb. Although this channel has a sizable effect on the di-fatjet final state, we do not include this in our analysis since this process drops sharply after applying b-veto.

All the significant backgrounds that contribute to the $2J_V + \not{E}_T$ signal are included in our analysis. We do two to four additional jets merging using the MLM matching [150,151] scheme for different background processes, and normalize the cross section according to the available higher-order QCD corrections. Inclusive Z boson production is the principal background where Z boson decays invisibly ($pp \to Z + \text{jets} \to \nu\nu + \text{jets}$) and gives a large $\not{E}_T$ together with fatjets arising from QCD jets. This process is matched to four extra partons using the MLM scheme. Second, inclusive $W^{\pm}$ boson production has a significant contribution when the lepton from the leptonic decay of the W boson remains undetected ($pp \to W + \text{jets} \to l_{e,\mu}\nu + \text{jets}$). The neutrino from W-decay gives a substantial amount of $\not{E}_T$ and fatjets arise from QCD jets. This process is generated up to four extra partons with MLM matching. Note that the contribution from the above two background processes counts only when the missing transverse momentum is sufficiently large. We apply the generation level hard cut $\not{E}_T > 100$ GeV, as the region with lower missing transverse energy is of no interest for this present analysis. Additionally, di-boson production can offer a considerable amount of contribution in the background. The three different di-boson processes $pp \to WZ$, WW, and ZZ, are possible, where the WZ process gives the most significant contribution among these three. All three processes are generated and merged up to two extra partons. One of the vector bosons in all these processes decays hadronically, giving rise to a J_V. Other vector boson decaying invisibly ($Z \to \nu\nu$) or leptonically ($W \to l_{e,\mu}\nu$) with lepton being undetected, gives a large $\not{E}_T$. Another fatjet in all these di-boson processes arises from the QCD jets. Single top production is possible in the SM through three different types of process, S-channel ($p\,p \to t\,b$), t-channel ($p\,p \to t\,j$) and associated production ($p\,p \to t\,W$), where associated production gives a considerable amount of contribution to the background of our signal. This process is merged up to two extra partons using the MLM scheme. Finally, top pair production contributes to the background when one top decays leptonically and lepton is escaping the detection. Whereas the other top decays hadronically and that essentially gives rise to a vector-like fatjet J_V. Since such an event comes with a couple of b-jets, b-veto can effectively reduce this background. This process is generated to two extra partons with MLM matching. The other fatjet aries from the QCD jets or untagged b-jets. We found negligible contributions to the background from the QCD multi-jet and tri-boson processes compared to the processes mentioned above. Therefore we do not include these processes into our analysis. For our simulated backgrounds at 14 TeV LHC, we normalize their cross section according to the available higher-order QCD corrections, as tabulated in Table IV of Reference [119].

The associated production of heavy scalar with two jets merging and pair production of the heavy scalars are analyzed at LO [119] where it was found that the former processes contribute dominantly in the di-fatjet final state than the latter. A further estimate at NLO accuracy modifies the contribution in two vital directions. First, both for the associated production and pair production of heavy scalar processes, the differential NLO K-factor plays an important role, as already described in the previous section. Second, two jets merged associated production channels can mimic the Higgs mediated pair production of heavy scalar processes, and therefore may contribute to double-counting in a particular phase space region. NLO estimate eliminates such possibility giving non-overlapping contributions from all processes.

Now, along with both these effects, our estimate at NLO predicts reduced contribution from associated production, thereby enhancing the part from the pair production. This has a profound significance in setting up the phenomenological analysis. On contrary to a more complex mixed-signal region analysis by taking into account the admixture of $1J_V$ and $2J_V$, that has been carried out in Reference [119], it is tempting to concentrate only on the $2J_V$ identification for a significant gain. Demanding that both the fatjets have V-jet like characteristics, one finds a more effective background control and, as a result, a higher statistical significance.

3.6.2 Construction of high level variables

The total energy of the fatjet originated from the hadronic decay of boosted W, Z is distributed around two subjet axes. N-subjettiness ratio (τ_{21}) and the jet-mass (M_J) are two potent variables to classify such fatjets J_V from those that arise from the fragmentation of QCD parton. The jet-mass is defined by $M_J = (\sum_{i\in J} P_i)^2$, where P_i is the four-momentum of the i-th hit in the calorimeter. Large-radius jets are prone to attract additional soft contributions from underlying QCD radiation, which needs to be eliminated to get reliable estimates from the different high-level variables. Pruning, filtering, and trimming [102–105] are different grooming techniques prescribed to remove those soft and wide-angle radiations. We consider pruned jet in our analysis as discussed in refs. [103, 104].

We run the pruning algorithm repeatedly to remove the soft and wide-angle emission and veto such recombinations. One has to estimate two variables, the angular separation of the two proto-jets, ΔR_{ij} and softness parameter $Z = \min(P_{Ti}, P_{Tj})/P_{T(i+j)}$, at every recombination step. The recombination between i-th and j-th proto jets is not performed dropping the softer one, if $\Delta R_{ij} > R_{fact}$ and $Z < Z_{cut}$. We choose standard default values of $R_{fact} = 0.5$ and $Z_{cut} = 0.1$

BP	Pre-selection cuts + $\not{E}_T > 200\,GeV$, b-veto, $65\,GeV < M(J_0), M(J_1) < 105\,GeV$, $\tau_{21}(J_0), \tau_{21}(J_1) < 0.35$					
	$H^\pm A$			H^+H^-		
	N_S^{NLO}	$N_S^{LO\times K}$	relative change%	N_S^{NLO}	$N_S^{LO\times K}$	relative change%
BP1	$168.2^{+2.8}_{-2.5}$	$119.5^{+3.3}_{-3.2}$	40.75%	$121.2^{+4.6}_{-4.6}$	$82.7^{+5.0}_{-4.1}$	46.55%
BP2	$190.7^{+3.1}_{-4.1}$	$155.6^{+5.7}_{-5.5}$	22.56%	$150.4^{+9.2}_{-8.0}$	$111.1^{+9.8}_{-7.7}$	35.37%
BP3	$202.8^{+4.2}_{-4.2}$	$162.8^{+7.0}_{-6.5}$	24.57%	$153.8^{+11.8}_{-9.8}$	$122.5^{+13.8}_{-10.9}$	25.55%

Table 3.4: N_S^{NLO} and $N_S^{LO\times K}$ are shown to represent the expected number of the $2J_V + \not{E}_T$ final state events for different pair production of heavy scalars channels, $pp \to H^\pm A$, and $pp \to H^+H^-$. These two numbers coming from NLO computation and LO with integrated NLO K-factor multiplication, respectively, at 14 TeV HL-LHC. Superscripts and subscripts are the change in the corresponding number of events due to the envelope of eight different (μ_R, μ_F) scale choices. Additional cuts over existing pre-selection (see text) are $\not{E}_T > 200\,GeV$, b-veto, $65\,GeV < M(J_0), M(J_1) < 105\,GeV, and\, \tau_{21}(J_0), \tau_{21}(J_1) < 0.35$. Relative change, $(N_S^{NLO} - N_S^{LO\times K})/N_S^{LO\times K}$ reflects the importance of differential K-factor in the present computation. Relative changes are given corresponding to the central scale.

Topology	BP1	BP2	BP3	BP4	BP5	BP6	BP7
Associated production	452.29	377.73	327.56	266.9	217.53	176.9	138.11
Pair production	1677.13	1432.67	1184.16	969.0	785.63	622.99	516.62
Z+jets	W+jets	tW+jets	tt+jets	WZ+jets	ZZ+jets	WW+jets	Total
652519	527312	46011.8	54635	36126.5	3689.51	12002.4	1.3323×10^6

Table 3.5: Expected number of events from different signal and background processes at 14 TeV HL-LHC corresponding to the central scale after applying the Pre-selection cuts with leading and subleading fatjet mass $M_{J_0}, M_{J_1} > 40$ GeV and b-veto.

[103]. The N-subjettiness determines the jet shape of hadronically-decaying boosted V-bosons. Considering that N number of subjets exist within the jet, N-subjettiness (τ_N) is defined by the angular separation between constituents of the jet from their nearest sub-jet axis as given below [106, 107].

$$\tau_N = \frac{1}{\mathcal{N}_0} \sum_i P_{T,i} \min\{\Delta R_{i,1}, \Delta R_{i,2}, ..., \Delta R_{i,N}\} \tag{3.20}$$

The summation runs over all the constituents of the jet, and $P_{T,i}$ is the transverse momentum of the i-th constituent. $\mathcal{N}_0 = \sum_i P_{T,i}\, R$ is the normalization factor, and R is the jet radius. τ_{21} denotes the ratio of τ_2 and τ_1, which is an excellent variable to tag a hadronically-decaying boosted V-boson as it tends to zero (far from zero) for a correctly identified two-prong (one-prong) jet.

To proceed further, we define the following *pre-selection criteria* based on which signal and background event samples are prepared: (i) each event has to have at least two fatjets constructed by the Cambridge-Aachen (CA) jet clustering algorithm with radius parameter $R = 0.8$, and the minimum transverse momentum of each fatjet $P_T(J_i) > 180\,GeV$, (ii) since pair of DM particles are produced in the signal, a minimum missing transverse energy $\not{E}_T > 100\,GeV$ is applied to select the events, (iii) we also impose a minimum azimuthal angle separation between the identified fatjet and missing transverse momentum direction, so that, $|\Delta\phi(J_i, \not{E}_T)| > 0.2$. This minimizes any jet mismeasurement effect contributing to $\not{E}_T$, (iv) since no leptons are expected in signal region, backgrounds can be further suppressed by vetoing a lepton tag. So, events are vetoed if they contain leptons that have pseudo-rapidity $|\eta(l)| < 2.4$ and transverse momentum $P_T(l) > 10\,GeV$.

It is clear from our previous discussion on boosted fatjet that several interesting variables can contribute to strengthen the signal efficiency. We would demonstrate the distribution of all such variables, but before that we point out some of the significant changes that appeared due to NLO computation in the signal region. In Table 3.4, we show the expected number of $2J_V + \not{E}_T$ final state events corresponding to the central scale, originated from different pair production of heavy scalar processes. Such numbers for $pp \to H^{\pm}A$, and $pp \to H^+H^-$ at NLO (N_S^{NLO}) level are given for three sample benchmark points, together with LO level numbers multiplied by overall NLO K-factor ($N_S^{LO\times K}$) for 3000 fb^{-1} integrated luminosity at 14 TeV LHC. Superscripts and subscripts are the change in the corresponding number of events due to the envelope of eight different (μ_R, μ_F) scale choices. In both cases, that makes the overall cross section normalized to the NLO value. Signal region criteria in conjunction with pre-selection cuts are described in the table [b]. Relative change, defined as $(N_S^{NLO} - N_S^{LO\times K})/N_S^{LO\times K}$, is given for the central scale. Relative change is independent of the luminosity and ascertains the necessity of considering actual NLO events instead of using LO events multiplied by a flat K-factor. It is evident that NLO and LO computations have different efficiencies for the given kinematic cuts. Relative change between these two estimations exhibits the role of the differential NLO K-factor by changing the LO estimation up to 40% for the process $pp \to H^{\pm}A$, and 46% for $pp \to H^+H^-$ for the given kinematic cuts mentioned at the top of the Table 3.4.

In addition to the pre-selection cuts described above, final event selection criteria for multivariate analysis includes a very relaxed cut on pruned jet mass. All other variables are kept free to provide the multivariate analysis with enough scope to optimize the nonlinear cut based on suitable variables. We select the

[b] One can, in principle, use such stiff event selection criteria for a realistic cut based analysis. Our purpose is purely for demonstration, as we would finally employ multivariate analysis to construct the suitable optimization based on rather loosely set criteria.

signal and background events after applying the following cuts: (i) both leading and subleading fatjets have to have a minimum pruned jet mass of 40 GeV to reduce the contribution of fatjets originated from QCD, (ii) b-veto is applied on the jets that are formed using anti-k$_t$ algorithm with radius parameter $R = 0.5$ and this significantly reduces $t\bar{t}$ background.

3.6.3 Multivariate analysis (MVA)

In Table 3.5, we present the expected number of signal events coming from the associated production and pair production of the heavy scalar channels together with all background processes at 14 TeV LHC with integrated luminosity 3000 fb^{-1}. From this table, we see contribution in the $2J_V + \not{E}_T$ final state coming from the pair production of the heavy scalars is always more prominent than the associated production after these cuts. We construct two independent event samples for our multivariate analysis, one for the signal and another for the background. The entire dataset is splitted randomly – 50% for the training and the remaining for testing purposes for both samples. We employ an adaptive BDT algorithm for MVA. We generate different signal processes separately at NLO and combine them according to their weights to get the kinematic distributions of the combined signal. Similarly, the different background processes are generated separately at LO with two to four extra jet MLM matching and combined thereafter according to their weights to get the kinematic distributions of the combined background. A set of kinematic variables is chosen from a bigger group of variables employed in the MVA analysis depending on their relative importance while discriminating the signal class from the background class. We present in Figure 3.8 the normalized kinematic distributions of all nine input variables that are used in MVA. We obtain the signal distributions using sample benchmark point BP2, including all the associated production and pair production of the heavy scalars at NLO. We do not include the process $pp \to HH$ in our analysis although it has a larger cross section than any other associated or pair-production channels, as b-veto and cuts on the fatjet mass and N-subjettiness ratio τ_{21} weaken its effect and the remaining events reside well away from the maximum *BDT response* region. The background comprises of all the processes discussed in Subsection 3.6.1 after applying the cuts $M_{J_0}, M_{J_1} > 40$ GeV and b-veto along with the pre-selection cuts mentioned in Subsection 3.6.2 at 14 TeV LHC. The distributions of the pruned jet mass $M_{J_{0,1}}$ of the leading (Figure 3.8a) and subleading (Figure 3.8b) fatjets, have a peak near 80-90 GeV for the signal close to the vector boson mass, however no such peak for the background reflects that fatjets are predominantly formed from QCD jets. The distributions of the N-subjettiness ratio, $\tau_{21}(J_{0,1})$ of the leading (Figure 3.8c) and subleading (Figure 3.8d) fatjets establish that both the fatjets of the signal have a two-prong structure as they peak at a smaller value of τ_{21}. In contrast, both the fatjets in the background has a characteristic one-prong structure producing a larger value for this variable. Hence these four jet substructure variables are crucial in discriminating the signal from the background. The relative separation between the leading (J_0) and subleading

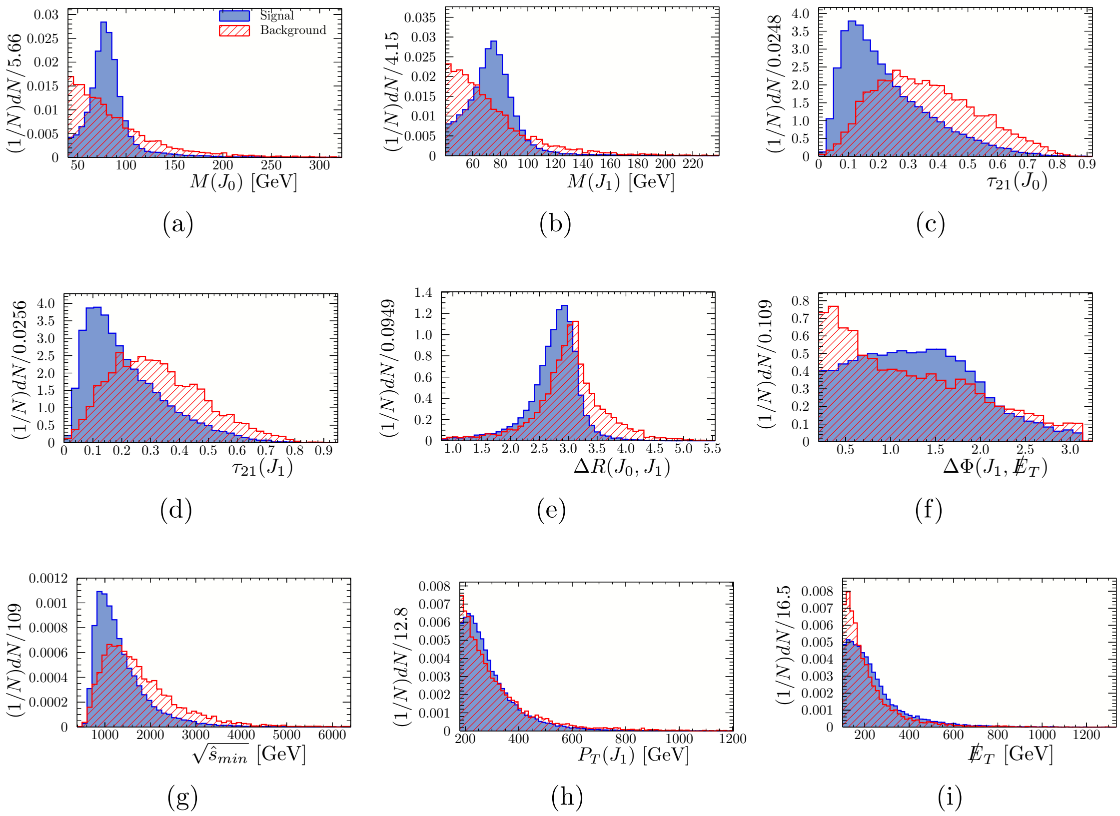


Figure 3.8: Normalized kinematic distributions of the different input variables used in MVA for the background (red) and the signal (blue). Plot (a) and (b) represent the distribution of pruned jet mass of the leading and subleading fatjets, respectively, whereas plot (c) and (d) are the distributions of the N-subjettiness ratio of the leading and subleading fatjets, respectively. Plot (e) shows distribution of the relative separation between the two leading fatjets. Azimuthal separation distribution of the subleading fatjet from the missing energy direction is depicted in plot (f). Plot (g) shows distribution of the global inclusive variable $\sqrt{\hat{S}_{min}}$ and distribution of the transverse momentum of the subleading fatjet and total missing transverse momentum are presented in plot (h) and (i), respectively. We display the signal distributions for BP2, including all contributions from associated production of the heavy scalar and pair production of the heavy scalars at NLO. The background comprises all the processes discussed in Subsection 3.6.1 after applying the cuts $M_{J_{0,1}} > 40$ GeV and b-veto together with the pre-selection criteria mentioned in Subsection 3.6.2.

Variable	$\tau_{21}(J_0)$	$M(J_0)$	$\tau_{21}(J_1)$	$M(J_1)$	$\Delta R(J_0, J_1)$	$\sqrt{\hat{S}_{min}}$	$\Delta\phi(J_1, \not{E}_T)$	$\not{E}_T$	$P_T(J_1)$
Separation	16.58	15.71	13.71	11.57	11.27	9.039	3.011	2.451	1.324

Table 3.6: Method unspecific relative importance (or separation power) of the different variables according to their rank before using at MVA. We obtain the numbers for BP2 from the TMVA package during MVA. Those numbers can change modestly for different benchmark points and different algorithms.

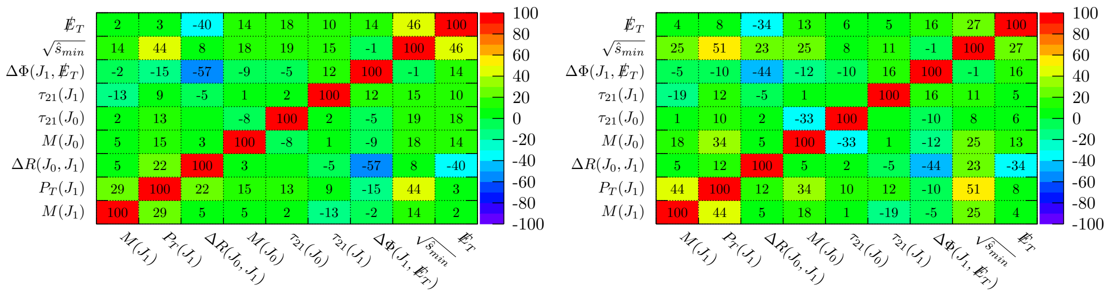
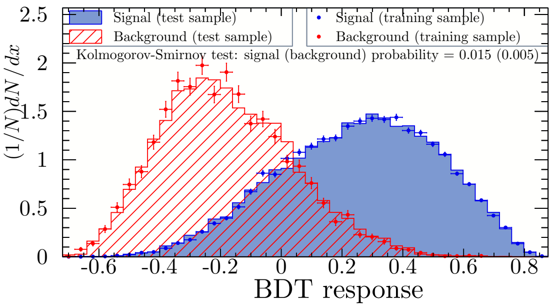


Figure 3.9: The linear correlation coefficients among different kinematic variables used in MVA (in percentage) for the signal (left panel, BP2) and background (right panel). The positive and negative signs signify the positive and negative correlations (anti-correlated) among the two variables.

(J_1) fatjets $\Delta R(J_0, J_1)$ (Figure 3.8e), azimuthal separation between J_1 and $\not{E}_T$ is represented as $\Delta\phi(J_1, \not{E}_T)$ (Figure 3.8f), and the inclusive global variable $\sqrt{\hat{S}_{min}}$ (Figure 3.8g) are effective observables to separate the signal from the background. The inclusive variable $\sqrt{\hat{S}_{min}}$, defined as the minimum CM energy required to satisfy all observed objects and $\not{E}_T$ was proposed in [152–154] to find the new physics mass scale for the signals containing invisible particles like ours. All the reconstructed objects of the detectors are used to construct the reconstructed object-level $\sqrt{\hat{S}_{min}}$ that demonstrate better efficiency than the other inclusive variables H_T, $\not{H}_T$ etc.

A variable is considered to be more powerful discriminator, if it possesses a larger separation between the signal and background. For different kinematic variables, the method unspecific relative importance is shown in Table 3.6, where we principally keep the variables that have less (anti-)correlation among themselves both for the signal and background. We notice four jet substructure variables $M_{J_{0,1}}$ and $\tau_{21}(J_{0,1})$ are very good discriminators. The relative importance of the different kinematic variables can change modestly for different benchmark points. Although very high P_T for both fatjets and large $\not{E}_T$ are considered during event selection, transverse momentum of the subleading fatjet, $P_T(J_1)$ (Figure 3.8h) and $\not{E}_T$ (Figure 3.8i) still can take a role in discriminating the signal from the background in MVA. Note that $P_T(J_0)$ and $P_T(J_1)$ are highly correlated (pos-



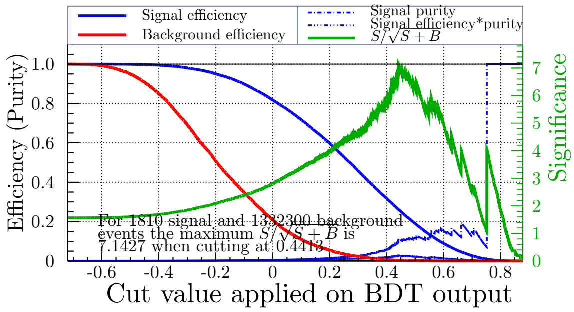
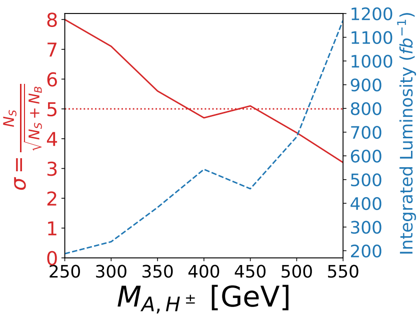


Figure 3.10: The left panel shows the normalized BDT response for the training and testing samples for both signal (BP2) and background classes. The right panel contains the cut efficiencies for the background (red) and the signal (blue) and the statistical significance of the signal over the background (green) as a function of the cut value applied on the BDT response.

itively) both in signal and background classes, so we keep only $P_T(J_1)$ in the analysis as it has more relative importance than $P_T(J_0)$. Similarly, $\Delta\phi(J_0, \not{E}_T)$ and $\Delta\phi(J_1, \not{E}_T)$ are highly anti-correlated, but we keep $\Delta\phi(J_1, \not{E}_T)$ because of its larger relative importance. The linear correlation coefficients among different kinematic variables used in MVA (in%) for the signal and background are shown in Figure 3.9. The positive (negative) signs signify the positive (negative) correlation (anti-correlation) among the two variables. Modestly large anti-correlation between $\Delta\phi(J_1, \not{E}_T)$ and $\Delta R(J_0, J_1)$ is present, although we kept them both as they have large relative importance.

Finally, we present the normalized BDT response for the training and testing samples for both signal and background classes in the left panel of Figure 3.10. The signal distribution is presented for BP2. The distributions of the BDT response get well separated for the signal and background. Cut efficiencies can be estimated by applying a cut $\text{BDT}_{res} > \text{BDT}_{cut}$ on the BDT response. In the right panel of Figure 3.10, such cut efficiencies are demonstrated for the background (red) and signal (blue), along with the statistical significance of the signal over the background (green) as a function of the cut value applied on the BDT response. We use the prescription $\sigma = \frac{N_S}{\sqrt{N_S+N_B}}$ for computing the statistical significance. N_S and N_B are respectively the expected number of signal and background events after using the optimal cut BDT_{opt} at 3000 fb^{-1} luminosity at 14 TeV LHC. N_S, N_B, and σ are shown in the right panel of Figure 3.11 for different benchmark points. We find more than 5σ discovery potential for four different benchmark points. In the left panel of Figure 3.11 we summarize the result in terms of statistical significance of the signal as a function of the masses of the heavy BSM scalars (solid red) at 14 Tev LHC with integrated luminosity 3000 fb^{-1}. At the



Model	N_S^{bc}	BDT$_{opt}$	N_S	N_B	σ
BP1	2129	0.6927	357	1650	8.0
BP2	1810	0.4413	474	3930	7.1
BP3	1511	0.6003	390	4516	5.6
BP4	1236	0.5270	207	1738	4.7
BP5	1003	0.6499	178	1041	5.1
BP6	800	0.6098	102	485	4.2
BP7	654	0.6462	97	820	3.2
N_{SM}	1332300				

Figure 3.11: The left panel shows the statistical significance of the signal over the background as a function of masses of the heavy BSM scalars (solid red line) at 14 TeV HL-LHC. The dashed blue curve on the same plot exhibits the required luminosity for two sigmas (2σ) exclusion for different benchmark points– the horizontal dotted red line to mark 5σ discovery potential. The right table demonstrates the corresponding expected number of signal events (N_S^{bc}) at NLO and background events (N_{SM}) before applying the BDT cut, where N_S and N_B are the expected number of signal and background events that survive after applying the optimum BDT$_{opt}$ cut, respectively.

same time, the dashed blue line exhibits the required luminosity for 2σ exclusion for different benchmark points.

3.7 Conclusions

IDM is a simple extension of the SM where a new $SU(2)_L$ scalar doublet owning a discrete $\mathbb{Z}_2$ symmetry provides a viable DM candidate together with additional heavy BSM scalars. This model offers two distinct parameter spaces, consisting of hierarchical mass spectrum and degenerate mass spectrum of these scalars, that satisfy the observed relic density of the dark matter and other theoretical and experimental constraints.

Despite of several studies being performed in exploring this viable dark matter model at the LHC, in this chapter we initiate the effort of looking into a promising channel with NLO QCD precision. This study focuses on the hierarchical mass region and considers NLO QCD corrections on the associated and pair production channels of heavy scalars. We find that the effect of QCD correction is significant for encrypting the correct search strategy at the LHC. Table 3.2, and Table 3.3 encapsulate the correction factors for different benchmark points. We get an overall correction of about 33%-39% for the associated production processes and for a gauge boson mediated pair production channel, $pp \to H^\pm A$. Similarly, the $pp \to H^+H^-$ process, which encompasses both gauge boson and

Higgs mediator, has the correction factor in between 38% and 56%. In contrast, $pp \to AA$ being scalar mediated, receives a correction factor in the range of 70%-92%. Nevertheless, notable improvement on scale uncertainties is achieved due to the inclusion of NLO corrections. We also take into account the parton shower effect and demonstrate its practicality at the low transverse momentum region.

After jet clustering and detector simulation, we compare distributions of various crucial kinematic observables at LO and NLO. Noted shifts in the shape of these distributions over the LO computation can significantly influence the construction of phenomenological analysis. We notice a substantial relative change in the number of survived signal events as an effect of the differential NLO K-factor. For example, this change is up to 46% for the gauge mediated pair production of heavy scalar processes. We also emphasize that gauge boson mediated decay products of hadronically decayed heavy scalars are highly collimated in this signal region and therefore large-radius fatjets come naturally in probing the hierarchical mass region. The internal structure and properties of the fatjet are key ingredients to know about their genesis. Fatjets originated from the QCD radiation of partons pose different characteristics compared to the fatjets generated from boosted vector boson. The jet substructure is a powerful tool to get control over the colossal SM background and identify the signal correctly. We find jet-substructure observables $M_{J_{0,1}}$ and $\tau_{21}(J_{0,1})$ are excellent discriminators in discriminating fatjets originated from the boosted vector boson and the QCD jets. We work with parton shower matched NLO QCD corrected signal and employ sophisticated multivariate analysis to distinguish the signal using these powerful jet-substructure variables. We discuss the set of nine variables that are used in the MVA analysis and their linear correlation coefficients are presented for the signal at a sample benchmark point and for the background.

We observe that the discovery potential for different benchmark points nearly up to 350 GeV of heavy scalar mass in the hierarchical mass region has a statistical significance above 5σ at the HL-LHC. Hence this parameter space of the hierarchical mass spectrum which is well motivated having a dark matter candidate of mass $m_{\rm DM} \sim m_h/2$, would be quite interesting to look into. We also notice through this study that the heavy BSM scalar mass falling in the range of 250-550 GeV can be excluded with 1200 fb^{-1} integrated luminosity at the 14 TeV LHC.

Chapter 4

Top-philic Dark Matter in a Hybrid KSVZ axion framework

4.1 Introduction

As discussed in the introduction of the thesis, an extension of the SM with a global Peccei-Quinn (PQ) symmetry [155, 156] provides solutions for two of the critical shortcomings of the SM in one go, and they are the *Strong CP problem* and the existence of *dark matter*. This global symmetry is expected to be broken at a scale much larger than the Electroweak (EW) scale. The breaking of $U(1)_{PQ}$ predicts a pseudo-Goldstone particle, popularly known as the *QCD axion*, that is not absolutely stable but can have a lifetime much greater than the age of the Universe [157–160] to play the role of DM. There are primarily three different QCD axion models that can simultaneously explain the presence of the DM in the Universe and solve the Strong CP problem. The (i) Peccei-Quinn-Weinberg-Wilczek (PQWW) [155, 161, 162] model introduces an additional singlet scalar that also obtains a non-zero vacuum expectation value (vev) at the time of EW phase transition. This setup is already ruled out from the experiments. The (ii) Kim-Shifman-Vainshtein-Zakharov (KSVZ) [56, 57] model introduces an extra colored particle together with a complex scalar that breaks the PQ symmetry. Anomaly-free condition is ensured by the introduction of vector-like quarks (VLQ). Finally, the (iii) Dine-Fischler-Srednicki-Zhitnitsky (DFSZ) [163, 164] model incorporates an additional Higgs field apart from the PQ breaking scalar. It is also interesting to point out that the breaking of PQ symmetry in these models also leaves a remnant Z_2 symmetry that remains intact. If such a setup is extended with an extra particle that also carries a non-trivial Z_2, then this unbroken symmetry can naturally ensure its stability. This motivates us to study two-component DM scenarios in these models.

In this chapter and the next chapter, we aim to explore a hybrid KSVZ scenario, where an extra complex scalar singlet extends the particle spectrum of the KSVZ setup in addition to the usual complex scalar that breaks the PQ symmetry. Axion can provide the correct relic density of DM, but after the corresponding breaking scale is fine-tuned. We analyze an extended KSVZ model which circumvents such fine-tuning by adding another complex scalar S, a singlet under SM. Under the residual Z_2 symmetry of the KSVZ model, VLQ is odd in this setup, and S is likewise Z_2-odd. Therefore the lightest component of S serves as the second dark matter candidate. VLQ interacts with the SM quarks and the scalar S in the present configuration. The hypercharge of VLQ is determined by the kind (up or down) of SM quarks considered.

Given that we are considering up-type quarks, the hypercharge of VLQ is $\frac{2}{3}$. VLQ plays a critical role in dark matter phenomenology because it opens up new co-annihilation and annihilation channels, such as co-annihilation between scalar DM and VLQ and annihilation of VLQs into the SM particles, which has a significant impact on relic density calculations. Since DM interacts with the SM quarks through VLQ, additional direct detection channels open up, such as VLQ-mediated t-channel elastic scattering between the SM-quark and scalar DM. Moreover, the VLQ and its interaction with the SM quarks also affect the LHC phenomenology. VLQ decays into an SM-quark and a missing DM particle after being produced at the LHC. As a result, multijet plus missing transverse momentum may be employed as a possible probe. If the mass difference between VLQ and DM is more than the top quark mass, VLQ can be probed from its decay into top quarks, along with a sizeable missing transverse energy from dark matter in the final state.

The Yukawa interaction takes the form $f_i S\overline{\Psi}_L u_{iR} + \ h.c$, where u_R denotes right-handed up-type SM-quarks with $i = u,\ c,\ t$. In this chapter, we do the dark matter and collider phenomenology of the extended KSVZ model. Our signal at the LHC comprises two boosted top-like fatjets and missing transverse energy. We consider the top-philic dark matter scenario, where f_u and f_c are tiny, and $f_t \sim 1$. We will show the vast parameter space of this model that gives the correct relic density of DM and is also allowed from the direct, indirect, and many other constraints, which can be discovered at the LHC with 139 fb^{-1} integrated luminosity.

The chapter is organized as follows. We introduce our model in Section 4.2 where the particle spectrum together with their charges under different symmetry groups have been discussed. Various theoretical and experimental constraints in our model are presented in Section 4.3. In Section 4.4 we discuss the dark matter phenomenology of the model. The collider analysis and the result based

	η	S	Ψ
$SU(3)_C$	1	1	3
$SU(2)_L$	1	1	1
$U(1)_Y$	0	0	2/3
$U(1)_{PQ}$	2	1	1

Table 4.1: Particle contents and their respective charge assignments under different symmetry groups.

on multivariate analysis are presented in Section 4.5. Finally, we summarize our findings in Section 4.6.

4.2 The Model

As stated in the introduction, the present chapter aims to study dark matter and collider phenomenology in a hybrid KSVZ framework of QCD axion. As is well known, the vanilla KSVZ model requires a complex scalar singlet η that breaks a global symmetry, popularly known as $U(1)_{\rm PQ}$. In addition, this model also demands a $SU(2)_L$ singlet colored fermion Ψ with a $+1$ unit of $U(1)_{\rm PQ}$ charge. This extra quark is vector-like and hence does not introduce any chiral anomaly. In addition, the hybrid KSVZ model also introduces an additional complex singlet scalar S charged under the $U(1)_{\rm PQ}$. The BSM fermion and scalar content of the model and their respective charges are listed in Table 4.1. The most general renormalizable and gauge-invariant Lagrangian for the present setup can be written as,

$$-\mathcal{L}^{\rm VLQ} = f_i S \overline{\Psi}_L u_{iR} + f_\Psi \eta \overline{\Psi}_L \Psi_R + h.c., \tag{4.1}$$

where, u_R represents right-handed up-type quarks in the SM with $i = u, c, t$. Here, L and R denote left- and right-handed projections. Note that the hypercharge of the newly introduced VLQ depends on its interaction with the SM quarks. The relevance of introducing an up-type VLQ in this setup will be clear once we discuss the DM and collider phenomenologies in Sections 4.4 and 4.5 respectively.

Moving on to the scalar part of the Lagrangian, the most general renormalizable scalar potential of our model, $V(H, \eta, S)$ can be written as,

$$\begin{aligned} V(H,\eta,S) &= \lambda_H(|H|^2 - v_H^2/2)^2 + \lambda_\eta(|\eta|^2 - F_a^2/2)^2 + \lambda_{\eta H}(|H|^2 - v_H^2/2)(|\eta|^2 - F_a^2/2) \\ &+ \mu_S^2|S|^2 + \lambda_S|S|^4 + \lambda_{SH}|H|^2|S|^2 + \lambda_{S\eta}|\eta|^2|S|^2 + [\epsilon_S \eta^* S^2 + h.c]. \end{aligned} \tag{4.2}$$

After the breaking of both $U(1)_{\rm PQ}$ and the SM gauge symmetry, the different scalars involved in the present setup take the following form,

$$H = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(v_H + h_0) \end{pmatrix}, \quad \eta = e^{\frac{ia}{F_a}} \frac{(F_a + \sigma_0)}{\sqrt{2}}, \quad S = \frac{S_1 + iS_2}{\sqrt{2}}, \tag{4.3}$$

where v_H denotes the vacuum expectation value (vev) of H obtained after the electroweak symmetry breaking (EWSB) and F_a represents the $U(1)_{\rm PQ}$ breaking scale. It is to be noted that, after the breaking of both symmetries, a non-zero $h_0 - \sigma_0$ mixing leads to the following mass terms:

$$M^2 \equiv \begin{pmatrix} 2v_H^2\lambda_H & F_a v_H \lambda_{\eta H} \\ F_a v_H \lambda_{\eta H} & 2F_a^2\lambda_\eta \end{pmatrix}. \tag{4.4}$$

The mass matrix can be diagonalised using

$$\begin{pmatrix} h_0 \\ \sigma_0 \end{pmatrix} = \begin{pmatrix} \cos\theta_m & \sin\theta_m \\ -\sin\theta_m & \cos\theta_m \end{pmatrix} \begin{pmatrix} h \\ \sigma \end{pmatrix} \tag{4.5}$$

where the mixing angle is given by,

$$\tan(2\theta_m) = \frac{F_a v \lambda_{\eta H}}{F_a^2\lambda_\eta - v^2\lambda_H}. \tag{4.6}$$

Finally, after diagonalization, the physical masses of the h and σ are given as,

$$M_{h,\sigma}^2 = (\lambda_H v^2 + \lambda_\eta F_a^2) \pm \sqrt{(\lambda_H v^2 - \lambda_\eta F_a^2)^2 + F_a^2 v^2 \lambda_{\eta H}^2}. \tag{4.7}$$

Next, as an artifact of two different symmetry breakings, the masses of the different components of the S can be expressed as,

$$M_{S_{1,2}}^2 = \frac{1}{2}(2\mu_S^2 + v_H^2\lambda_{SH} + F_a^2\lambda_\eta \mp 2\sqrt{2}\epsilon_s F_a). \tag{4.8}$$

Notice that the presence of the term proportional to ϵ_S in Equation 4.2 plays a crucial role in generating the mass splitting among the components of S. Subsequently, the mass of the VLQ is given as,

$$M_\Psi = f_\Psi \frac{F_a}{\sqrt{2}}. \tag{4.9}$$

At this stage, it is interesting to point out that, even after the breaking of both the symmetries, there still exists a remnant Z_2 symmetry under which both the Lagrangian as well as the scalar potential remains invariant. This remnant Z_2 can remain intact if S does not acquire a non-zero vev. Under such a scenario, the lightest neutral component of S can provide a vital DM candidate.

Finally, the setup also contains a pseudo-Nambu Goldstone boson a, associated with scalar η, popularly known as *axion*. The axion obtains a mass as a result of non-perturbative QCD effects given as [157, 159],

$$m_a \simeq 0.6 \text{ meV} \times \left(\frac{10^{10} \text{ GeV}}{F_a}\right). \tag{4.10}$$

Note that a suitable choice of decay constant F_a can adjust the fraction of which QCD axion can contribute toward the relic density of the dark matter. That makes the preset setup a tunable two-component dark matter scenario. The role of QCD axion as a DM candidate and its constraints are elaborated in Section 4.4. Now with the knowledge of all the particles and their interactions in this hybrid KSVZ setup, we are in a position to list the set of independent parameters important for the dark matter and collider phenomenology:

$$\{M_\Psi, M_{S_1}, M_{S_2}, M_\sigma, F_a, \lambda_{SH}, \lambda_{S\eta}, f_i\}.$$

4.3 Experimental and Theoretical Constraints

The extended KSVZ model under consideration is subjected to various theoretical as well as experimental constraints. In this section, we summarize all the relevant ones.

- **Stability and Perturbativity:** The scalar sector is extended over the vanilla model. Hence, different scalars in the present setup can help stabilize the electroweak vacuum. The stability of the electroweak vacuum also demands that the scalar potential should be bounded from below in all the field directions of the field space. On the other hand, a perturbative theory demands that the model parameters should obey:

$$|\lambda_i| < 4\pi \text{ and } |g_i|, |y|, |f_i| < \sqrt{4\pi}. \tag{4.11}$$

  where g_i and y are the SM gauge and Yukawa couplings, whereas f_i are Yukawa couplings involving different BSM fields, respectively.

- **Relic density, Direct and Indirect detection of DM:** For any dark matter model, it is essential to satisfy the observed abundance of DM relics from the precision measurement in the Planck experiment [25],

$$\Omega_{\rm DM} h^2 = 0.120 \pm 0.001. \tag{4.12}$$

Apart from DM relic density, the DM-nucleon scattering cross section is also constrained by various direct search experiments like LUX [165], PandaX-II [166, 167], and XEXON1T [168]. Finally, the DM annihilation to the SM particles are also subjected to the constraints coming from the indirect search experiments like PAMELA [46], Fermi-LAT [48], MAGIC [47] etc. Nonetheless, in all these cases, one also needs to take care of the multi-component nature of DM in our extended scenario, which is further discussed in Section 4.4.

- **Flavor constraints:** The Yukawa interactions of the complex singlet scalar S with VLQ and the SM right-handed quarks like u and c in the present setup can contribute towards the $D^0 - \bar{D}^0$ mixing [169]. The measured value of the $D-$meson mass splitting significantly constrained this mixing. The Feynmann diagrams that contribute to this mixing are shown in Figure D.5; each diagram has four possible configurations with a total of sixteen diagrams. Effective operator contributing to this mixing in the present setup can be expressed as

$$\mathcal{L}_{\text{eff}} = \frac{\tilde{z}}{M_\Psi^2}\bar{u}_R^\alpha \gamma^\mu c_R^\alpha \bar{u}_R^\beta \gamma_\mu c_R^\beta. \tag{4.13}$$

where

$$\tilde{z} = -\frac{f_u^2 f_c^2}{96\pi^2}[g_\psi(M_{S_1}^2/M_\Psi^2) + g_\psi(M_{S_2}^2/M_\Psi^2) - 2g_\psi(M_{S_1}M_{S_2}/M_\Psi^2)]. \tag{4.14}$$

Here $g_\Psi(x) = 24x f_6(x) + 12\tilde{f}_6(x)$ where the expressions of f_6 and $\tilde{f}_6$ can be found in [170]. The measurement of the $D-$meson mass splitting demands [169, 170]

$$|\tilde{z}| \lesssim 5.7 \times 10^{-7}(M_\Psi/\text{TeV})^2\ . \tag{4.15}$$

- **LHC diphoton searches**: As a result of mixing between h and σ, all the tree level interactions with the SM Higgs get modified. In such a case, the signal strength in the di-photon channel takes a form:

$$\mu_{\gamma\gamma} = c_\theta^2 \frac{BR_{h\to\gamma\gamma}}{BR_{h\to\gamma\gamma}^{\text{SM}}} \simeq c_\theta^2 \frac{\Gamma_{h\to\gamma\gamma}}{\Gamma_{h\to\gamma\gamma}^{\text{SM}}}. \tag{4.16}$$

The LHC sets a limit on this new mixing angle as $|\sin\theta| \leq 0.36$ [171].

- **Invisible Higgs decay**: Involvement of the new interactions of the SM Higgs with various BSM particles in the present setup can lead to its new decay modes if kinematically allowed. These extra decays of Higgs can contribute toward invisible Higgs decay. In such a situation, we need to employ the bound on the invisible Higgs decay width as [172]:

$$Br(h \to \text{Invisible}) \equiv \frac{\Gamma(\text{h} \to \text{Invisible})}{\Gamma(\text{h} \to \text{SM}) + \Gamma(\text{h} \to \text{Invisible})} < 0.11. \quad (4.17\text{a})$$

In the case of light DM, the Higgs can decay to a pair of it when kinematically allowed. However, in our present analysis, we primarily focus on the parameter space where $m_i > \frac{m_h}{2}$ so the above constraint is not applicable.

- **Direct collider constraints:** Due to the presence of colored vector-like quarks, the present model is subjected to various collider constraints. Being non-trivially charged under the $U(1)_{PQ}$ allows the VLQ to couple with the complex scalar and the SM up type quarks. If kinematically allowed, the heavier states can always decay into the DM and an SM quark. Therefore a generic collider signature of this model contains a considerable amount of missing (transverse) energy from the escape of final DM particles from detection at the detector.

  Vector-like fermion can be pair produced through electroweak interaction performed at CERN's Large Electron Positron Collider (LEP):

$$e^+e^- \to \gamma^*, Z \to \Psi\bar{\Psi} \quad (4.18)$$

  The interaction between vector-like fermion, light SM quarks, and the DM can lead to the decay of Ψ to a light quark associated with DM at LEP if kinematically allowed. The reinterpreted LEPII results of squark search [173, 174] exclude the mass of Ψ up to 100 GeV. Such constraint is incorporated in our final exclusion plots. Please follow the brown region in Figure 4.10. Similar searches were also carried out at the LHC. In a recent ATLAS search, the vector-like mediator is searched while it decays into an invisible particle and light quark up (charm) when the mass difference between the mediator and DM is less than the top quark mass. The green region in Figure 4.10 is excluded from the reinterpreted result [173] of the ATLAS search [175] for multijet (2-6 jets) plus missing transverse momentum at center-of-mass (CM) energy $\sqrt{s} = 8$ TeV with an integrated luminosity of 20.3 fb^{-1}. Exploring a larger mass difference between the mediator and DM candidate, top-antitop plus missing transverse momentum signal has been extensively studied by both CMS and ATLAS collaborations, particularly superpartners searches of the top quark [176–184] and some dedicated dark matter searches [185]. The vector-like mediator can be pair produced at the LHC mainly through strong interaction and then decay into an up-type quark and invisible particle. So, the search for a top pair along with the missing transverse momentum signature by ATLAS and CMS can be reinterpreted to exclude some of the parameter spaces of this model. The CMS analysis [184] is reinterpreted in Ref [186] at 13 TeV LHC for an integrated luminosity of 35.9 fb^{-1}, assuming vector-like mediator decay with 100% branching fraction into the top and invisible particle. In their analysis, the signal consists of two oppositely charged isolated leptons from leptonic decays of both top and anti-top. The signal also consists of at least two hard jets; one of them is b-tagged and a large missing transverse momentum. The olive region in Figure 4.10 is the exclusion region (2σ) obtained from this analysis.

The existing LHC search relies on finding the top pair based on two hard leptons and a b-tagged jet. It is evident that the sensitivity of such detection deteriorates when these tops are boosted, especially while decaying from a heavy mother particle. We propose an alternative search strategy in this chapter by recognizing these boosted double top jets with a large missing transverse energy signature using jet substructure variables and multivariate analysis. We are examining the spectrum where the mass difference between vector-like mediator and DM is larger than the mass of the top quark such that on-shell decay into the top is possible. Our search strategy helps to explore the significant parameter space of this model that gives observed relic density of DM and also allowed from the direct-detection experiment with the current luminosity of the LHC.

4.4 Dark Matter Phenomenology

In this section, we aim to elaborate on the DM phenomenology of the model under consideration. As discussed earlier, the setup is a hybrid of the KSVZ model that includes an extra complex scalar (S) whose lightest component (S_1) plays the role of one of the DM while the part of the second DM is played by the QCD axion of the KSVZ setup. The involvement of the two DMs in this extended KSVZ scenario makes the layout a two-component DM system. Besides the QCD axion, the KSVZ setup naturally demands a presence of an extra colored fermionic $SU(2)_L$ singlet. This fermion plays a non-trivial role in the DM phenomenology and the collider searches of the DM as it talks directly to it through the Yukawa interaction given in Equation 4.1. Next, we discuss the DM phenomenology of both the DM candidates of the present model.

4.4.1 Relic density and DM detection

Apart from providing a solution to the *strong CP problem*, another interesting consequence of introducing a PQ symmetry is the emergence of the Nambu Goldstone boson, popularly known as *axion*. If the breaking scale (F_a) of the PQ symmetry is chosen appropriately, the resulting axion can be light as well as stable. This QCD axion can be an excellent DM candidate in such a scenario. Axions can be produced non-thermally as a result of the misalignment mechanism. Here, the axion field begins to coherently oscillate around the minimum of the PQ vacuum when its mass becomes comparable to the Hubble parameter. This coherent oscillation of the axion field behaves like a cold matter in the Universe. The relic density of the axion in such a case is approximately given by [187–189],

$$\Omega_a h^2 \simeq 0.18\,\theta_a^2 \left(\frac{F_a}{10^{12}\text{GeV}}\right)^{1.19}. \tag{4.19}$$

Here, θ_a represents the initial misalignment angle of the axion.

For the case of the scalar DM, we consider the mass hierarchy $M_{S_2} > M_{S_1}$ such that the lightest scalar component represents the second DM candidate. Its interactions with the SM Higgs and the VLQ keep it in equilibrium with the thermal bath in the early Universe. As the temperature of the Universe drops below the DM mass, its production from the thermal bath stops while its annihilation of the SM particle continues. Once the Universe's expansion rate becomes larger than the interaction rate of the DM, its annihilation to the SM bath also stops, and its abundance freezes out. DM can annihilate to the SM particles through: (a) its contact interactions, (b) Higgs-mediated channels [a] and (c) VLQ mediated channels (as a result of Yukawa interaction given in Equation 4.1). The presence of the Yukawa interaction also allows the DM to co-annihilate if the mass-splitting between the DM and newly introduced quark is sufficiently small. Note that as the VLQ and S_2 share the same Z_2 charge similar to the DM, their annihilations would also be important for evaluating the effective annihilation cross section. In Appendix D, we present all the important annihilation and co-annihilation channels of the DM that are crucial in determining its final relic abundance [b]. Once all the important annihilation and co-annihilation channels are identified, one can use them to determine the final relic density of the DM, which can be

[a] From Equation 4.7, it is evident that until and unless λ_η is very small, M_σ will remain much heavier than M_h and hence the σ mediated annihilation channels be very much suppressed.

[b] Just for completeness we have also shown the DM annihilation to the axion final states. These annihilations are highly suppressed and do not contribute towards the relic density of scalar dark matter. This is because most of the vertices involved in these annihilation cross sections are either proportional to $1/F_a$ or $\frac{\sin\theta}{F_a}$ or $\frac{\epsilon_S}{F_a}$ or ϵ_S.

expressed as [190],

$$\Omega_{S_1} h^2 = \frac{1.09\times 10^9 \text{ GeV}^{-1}}{g_*^{1/2}\ M_{Pl}} \frac{1}{J(x_f)}, \tag{4.20}$$

where $J(x_f)$ is given by,

$$J(x_f) = \int_{x_f}^{\infty} \frac{\langle \sigma|v|\rangle_{\text{eff}}}{x^2}\ \text{dx}. \tag{4.21}$$

$\langle \sigma|v|\rangle_{\text{eff}}$ in Equation (4.21) is the effective thermal average DM annihilation cross sections including contributions from the co-annihilations and is given by,

$$\begin{aligned}\langle \sigma|v|\rangle_{\text{eff}} = &\frac{g_{s_1}^2}{g_{\text{eff}}^2}\sigma(\overline{S_1}S_1) + 2\frac{g_{s_1}g_{s_2}}{g_{\text{eff}}^2}\sigma(\overline{S_1}S_2)(1+\Delta_{12})^{3/2}\ \exp[-x\Delta_{12}] + 2\frac{g_{s_1}g_\Psi}{g_{\text{eff}}^2}\sigma(\overline{S_1}\Psi)\\ &(1+\Delta_\Psi)^{3/2}\ \exp[-x\Delta_\Psi] + \frac{g_{S_2}^2}{g_{\text{eff}}^2}\sigma(\overline{S_2}S_2)(1+\Delta_{12})^3\ \exp[-2x\Delta_{12}]\\ &+ \frac{g_\Psi^2}{g_{\text{eff}}^2}\sigma(\overline{\Psi}\Psi)(1+\Delta_\Psi)^3\ \exp[-2x\Delta_\Psi].\end{aligned} \tag{4.22}$$

In the equation above, g_{s_1}, g_{s_2} and g_Ψ are the spin degrees of freedom for S_1, S_2 and Ψ. Here, $x = \frac{M_{S_1}}{T}$ representing dimensionless parameter with inverse of temperature, while Δ_Ψ and Δ_{12} are two dimensionless parameters qualifying mass splittings from dark scalar candidate:

$$\Delta_\Psi = \frac{M_\Psi - M_{S_1}}{M_{S_1}};\ \Delta_{12} = \frac{M_{S_2} - M_{S_1}}{M_{S_1}}. \tag{4.23}$$

The effective degrees of freedom in Equation 4.22 is given by,

$$g_{\text{eff}} = g_{s_1} + g_{s_2}(1+\Delta_{12})^{3/2}\ \exp[-x\Delta_{12}] + g_\Psi(1+\Delta_\Psi)^{3/2}\ \exp[-x\Delta_\Psi]. \tag{4.24}$$

In the following analysis, we first generate the model using FeynRules [26] and then implement it in `micrOMEGAs` -v5 [27] to find the region of parameter space that corresponds to correct relic abundance for our scalar DM candidate in accordance with the relation,

$$\Omega_{\text{T}} h^2 = \Omega_a h^2 + \Omega_{S_1} h^2, \tag{4.25}$$

where $\Omega_{\text{T}} h^2$ corresponds to the total relic density of the DM satisfying PLANCK constraints [25].

Next, the present model is subjected to the constraints coming from the direct search experiments for the dark matter. Experiments like LUX [165], PandaX-II [166, 167] looks for the DM recoil in the DM-nucleon scattering and subsequently provides a bound on the DM-nucleon scattering cross section. Being a two-component DM system, the direct detection cross section of the scalar DM should be rescaled as,

$$\sigma^{\rm SI}_{S_{1,\rm eff}} = \frac{\Omega_{S_1}}{\Omega_{\rm T}}\sigma^{\rm SI}_{S_1} \tag{4.26}$$

As mentioned earlier, due to the direct Yukawa interaction of the scalar DM with the up-quark, two other scattering processes contribute to the direct detection cross section of the scalar apart from the usual SM Higgs-mediated scattering. In Appendix D we listed all the scattering processes of the DM S_1 with the detector nucleon.

Finally, the model is also subjected to the constraints coming from the indirect search experiments. Indirect search experiments looking for an excess of gamma rays can help in probing the WIMP dark matter. DM particles can annihilate and produce the SM particles, out of which photons (and also neutrinos), being electromagnetically neutral, have better chances of reaching the detector from the source without getting deflected. Experiments like PAMELA [46], Fermi-LAT [48], MAGIC [47] etc. look for such excess in order to confirm the particle nature of the DM. The present set up being a two-component DM scenario, the indirect detection cross section of the scalar DM should be rescaled as well,

$$\sigma^{\rm ID}_{S_{1,\rm eff}} = \left(\frac{\Omega_{S_1}}{\Omega_{\rm T}}\right)^2\sigma^{\rm ID}_{S_1}. \tag{4.27}$$

At this stage, it is worth commenting on some of the detection possibilities of the axion as a DM candidate. Several ongoing and proposed experiments rely on axion being a DM. All these experiments lean on different detection techniques. For example, ADMX [191] searches for DM-photon conversion in the presence of the magnetic field. CASPEr [192] uses nuclear magnetic resonance to hunt for the axion DM; it is known that if the axion exists, it will modify Maxwell's equation. ABRACADABRA [193] utilizes this by using a toroidal magnet to source an effective electric current, and finally, MADMAX [194] is a proposed experiment that uses dielectrics haloscopes.

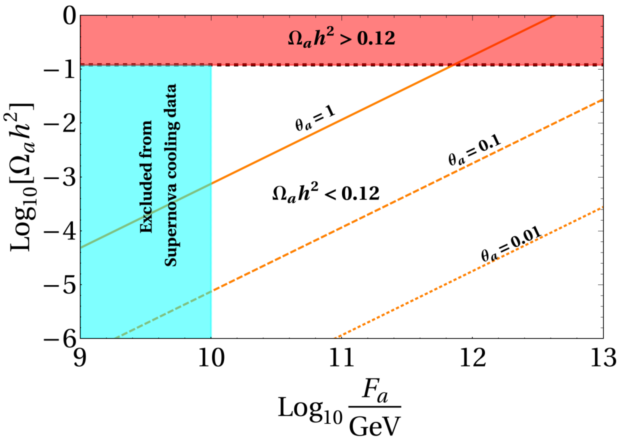


Figure 4.1: Variation of QCD axion relic density with the decay constant F_a for three different values of misalignment angles: $\theta_a = 1.0$ (solid), $\theta_a = 0.1$ (dashed), and $\theta_a = 0.01$ (dotted). Black thick dashed line corresponds to observed relic $\Omega_{\mathrm{DM}}h^2 = 0.12$. The cyan region is disallowed from the Supernova cooling data. The light pink region corresponds to the parameter space where the DM relic density remains overabundant.

4.4.2 Parameter Space of Hybrid KSVZ Axion Framework

It is well known that the KSVZ model provides a DM in the form of QCD axion. For this axion to play the role of the DM or contributes sufficiently towards the relic density of the DM, the decay constant F_a should lie in the range,

$$10^{10}\ \mathrm{GeV} \leq F_a \leq 10^{12}\ \mathrm{GeV}. \tag{4.28}$$

The lower bound on F_a comes from the supernova cooling data [58] whereas the upper bound results from the overproduction of the axion or, in other words, the relic density of the axion become overabundant. To understand this, in Figure 4.1 we study the variation of the axion relic density ($\Omega_a h^2$) with the decay constant for three different values of the misalignment angles *i.e.* $\theta_a = 1.0$ (solid), $\theta_a = 0.1$ (dashed), and $\theta_a = 0.01$ (dotted). The region in cyan is ruled out from the supernova cooling data, whereas the light pink region corresponds to the overproduced DM relic density. As can be seen from Figure 4.1, for $\theta_a = 1.0$ and $F_a \simeq 10^{12}$ GeV, the axion alone can contribute 100% towards the relic density of the dark matter. Finally, the white region corresponds to the parameter space where QCD axion as a DM remains under-abundant.

The present setup is an extended version of the KSVZ scenario, which contains an additional DM candidate as a singlet scalar. The presence of this extra DM here demands us to choose a parameter space for axion from Figure 4.1 where the relic density of the axion remains under-abundant so that the relic density of the axion together with the scalar can satisfy the Planck limit. For a demonstrative purpose, we fix $F_a = 10^{11}$ GeV and choose the misalignment angle as $\theta_a = 1$ for the rest of the analysis. This choice of F_a and θ_a corresponds to $\Omega_a h^2 = 0.012$. Without losing generality in our analysis, we set a heavier M_σ at 50 TeV (as the setup requires it to be quite heavy). At this stage, we would like to point out that the DM matter couples to σ through $S_1 - S_1 - \sigma$ interaction. This interaction can also help DM to annihilate into the SM particles through scalar mixing. Although these annihilations will have suppression coming from the mass of σ, they might still not be that small as these annihilations are also proportional to the $F_a \lambda_{S\eta}$. With $F_a = 10^{11}$ GeV and not so small value of $\lambda_{S\eta}$, the DM can still have significant annihilation cross sections and such cross sections might violate perturbative unitarity [195]. This demands $\lambda_{S\eta}$ to be extremely tiny. For simplicity, we set $\lambda_{S\eta} = 0$ throughout our analysis. Next, for the analysis purpose, we also define a mass-splitting, $\Delta M = M_{S_2} - M_{S_1}$ and consider it to be a free parameter rather than M_{S_2}. It is interesting to point out that once ΔM and F_a are fixed, the parameter ϵ_S automatically gets fixed, as can be seen from Equation 4.8. Before diving into the detailed analysis of the second DM candidate, we will like to mention the set of parameters that are relevant for the analysis of the DM phenomenology of the second DM candidate:

$$\{M_\Psi, M_{S_1}, \Delta M, F_a, \lambda_{SH}, f_i\}.$$

To demonstrate the above discussions, we display the variation of the relic density of S_1 with its mass in all the left panel plots of Figure 4.2. In the right panel, we also exhibit the variation of the effective direct detection cross section with M_{S_1} for different choices of parameters. In the top left panel of Figure 4.2, we project the importance of the Yukawa coupling f_t while choosing fixed values of $\Delta M = 100$ GeV, $M_\Psi = 500$ GeV and $\lambda_{SH} = 0.01$. It is interesting to point out that for $\lambda_{SH} = 0.01$, the DM does not satisfy the correct relic density in a pure scalar singlet DM scenario. In these plots, we also set both Yukawa couplings $f_u = f_c = 0.01$ to highlight the importance of the top Yukawa coupling f_t for three values of f_t: 0.1 (blue), 0.5 (red) and 1.0 (green). Notice that for $f_t = 0.1$, with the increase in the DM mass, we first observe a resonance dip at $M_{S_1} = M_h/2$ [c], next, a fall is observed at $M_{S_1} = 80$ GeV where the annihilation of the DM to

[c] In a lower DM mass regime, DM annihilating to the three body final states $q\bar{q}g$ can also contribute significantly towards the relic density, where chirality suppression in the lower order process is lifted by final state radiation. We do not consider this three-body final state in our analysis as the entire low mass regime of the DM is already ruled out from the DD searches, as can be seen from the right panels of Figure 4.2.

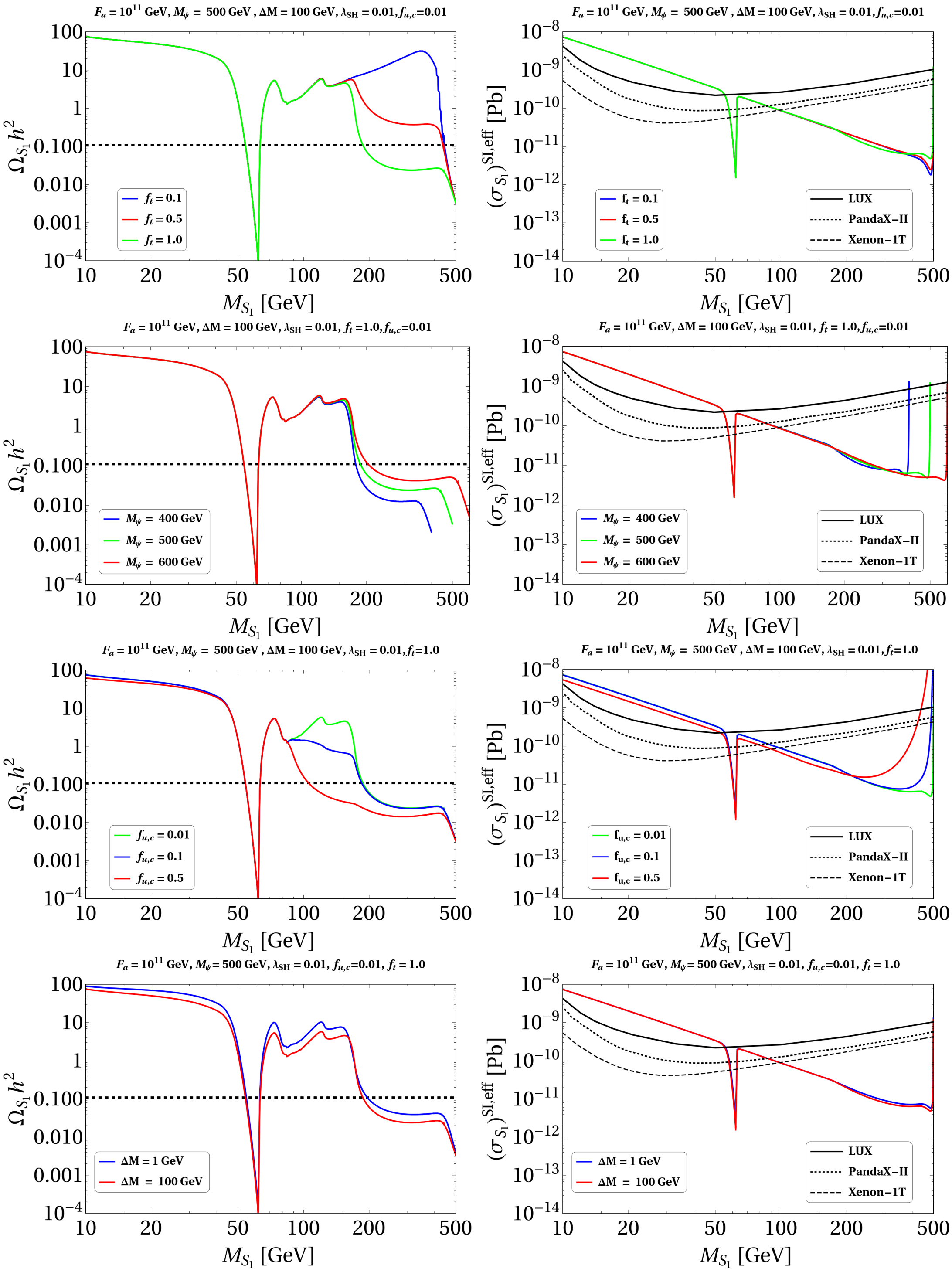


Figure 4.2: Variation of $\Omega_{S_1}h^2$ (left panel) and $\sigma^{\rm SI}_{S_{1,\rm eff}}$ (right panel) versus dark matter mass M_{S_1}. In all the plots we fix $F_a = 10^{11}$ GeV, $\lambda_{SH} = 0.01$. The Black dashed line in all the left plots corresponds to $0.120 - \Omega_a h^2$.

the $W^\pm$ boson opens up after which the relic density increases with the increase in the DM mass ($\langle \sigma v \rangle \propto 1/M_{S_1}$) and again drops at $M_{S_1} = 125$ GeV when

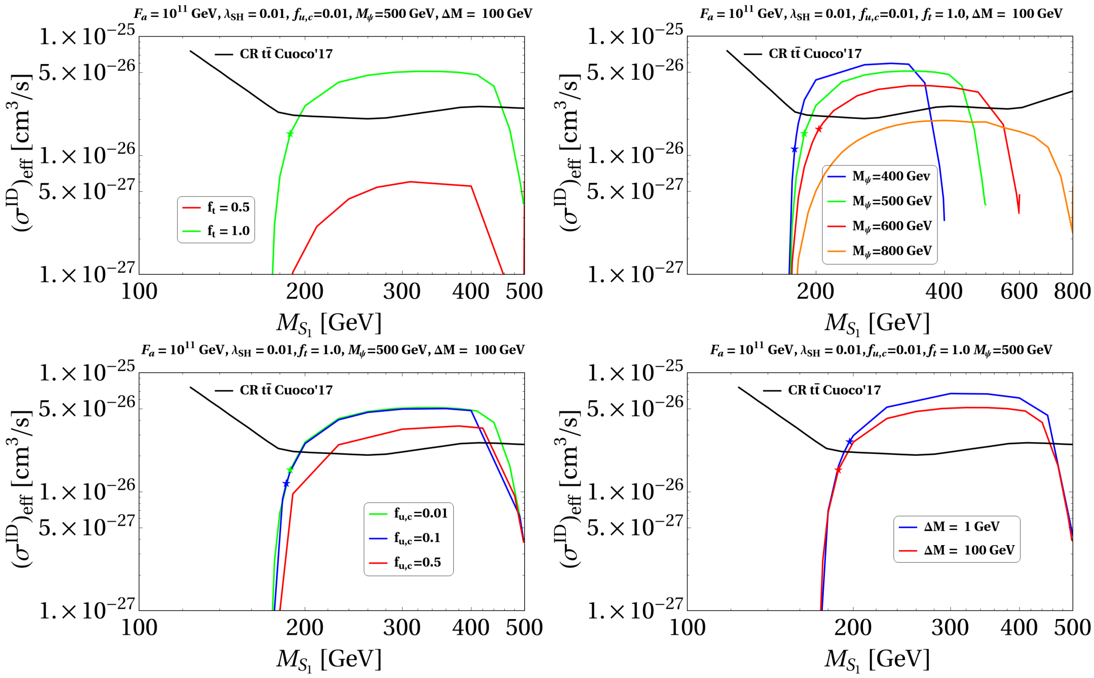


Figure 4.3: Variation of effective indirect detection cross section versus DM mass. Variation for different values of f_t, the mass of VLQ, $f_{u,c}$, and ΔM are shown in top left, top right, bottom left and bottom right panels respectively. In all the plots we fix $F_a = 10^{11}$ GeV, $\lambda_{SH} = 0.01$. The solid black line shows the experimental upper bound in $t\bar{t}$ final state.

the DM starts annihilating into the Higgs boson [d]. Finally, at a larger value of M_{S_1} ($M_{S_1} = 345$ GeV), when the mass difference between M_Ψ and M_{S_1} becomes relatively small, and the effect of DM co-annihilation with the VLQ comes into the picture, and a sharp drop in DM relic density is observed. In this region, although the DM co-annihilates with M_Ψ, the annihilation of Ψ with $\bar{\Psi}$ to the gluons dominates the effective annihilation cross section (see Equation 4.22) of the DM. Due to these co-annihilations, the relic density dark matter finally satisfies the condition $0.120 - \Omega_a h^2$ (denoted by the black dashed line) at $M_{S_1} = 444$ GeV. Further increasing f_t to the higher values like 0.5 and 1.0, one notices that the DM annihilations to top quarks mediated by M_Ψ (see Appendix D) start to dominate once the threshold of $M_{S_1} = M_{top}$ is crossed and the relic density can also satisfy the condition $0.120 - \Omega_a h^2$ near about $M_{S_1} \simeq 200$ GeV for $f_t = 1.0$. This figure illustrates the importance of VLQ in the DM phenomenology of the present setup as it makes huge parameter space allowed from the relic density, which was originally disallowed in the pure scalar singlet DM scenario.

[d] We would also like to point out that for DM mass $M_{S_1} < M_t$, DM annihilations to gg (via a box diagram) [196,197] or three-body final states like tWb [196] can also contribute towards the relic density of the DM for a significantly large Yukawa coupling f_t. We do not consider these processes in our analysis as we found that these processes remain suppressed for the choice of f_t we are interested in for our analysis.

In the top-right panel, we plot the spin-independent effective direct detection cross section of S_1 with the DM mass and compare it with the experimental results. Notice that only the coupling f_u enters the direct detection cross section of DM apart from the Higgs-portal coupling λ_{SH}. Note that the reduced values in the effective direct detection cross section are the result of the rescaling in the two-component DM system (see Equation 4.27). Additionally, Higgs resonance dip at $M_{S_1} = M_h/2$ and rise at $M_{S_1} \simeq 500$ GeV because of the interference among the direct detection diagrams (see Appendix D). Here, one notices that the near resonance region remains discarded from the experimental bounds. Still, the other regions where relic density can be satisfied remain allowed from the direct detection searches.

In the left panel of the second row in Figure 4.2 we fix $f_t = 1.0$ and then study the effect of varying M_Ψ in $\Omega_{S_1}h^2 - M_{S_1}$ plane. As expected, the final fall in the relic density pattern happens at three different positions corresponding to the three different values of M_Ψ. With the heavier propagator mass *i.e.* $M_\Psi =$ 600 GeV, the effective annihilation cross section of the DM to the top quark remains smaller in comparison to what is observed for $M_\Psi = 400$ GeV and hence relatively larger relic density is observed for $M_\Psi = 600$ GeV than for $M_\Psi = 400$ GeV. Similarly, in the left panel of the third row, we depict the effect of varying f_u and f_c. For simplicity, here we also assume $f_u = f_c$. As expected, for a large value of $f_{u,c}$ the annihilation of DM to top and up (charm) quark final state also becomes dominant the moment the threshold $2M_{S_1} = M_{top} + M_{u(c)}$ is achieved. This leads to an increase in the DM annihilation cross section, and consequently, a decrease in the relic density is observed. Next, in the left panel of the fourth row we show the effect of varying ΔM on the relic density of the DM. As expected, a smaller ΔM results in a larger effective DM annihilation cross section and hence a smaller relic, so in order to satisfy the correct relic density a heavier DM mass is required. On the other hand, a larger ΔM requires a smaller DM mass to satisfy the observed relic density and hence the plot shifts towards the lower DM mass. Finally, the middle and bottom right panel of Figure 4.2 can be followed from the one observed in the top right panel.

As can be seen from Figure 4.2, the correct relic density is mostly satisfied in the parameter space where $M_{s_1} > M_t$. Hence, one needs to check the prospects of indirect detection of DM in our model specifically focusing on $t\bar{t}$ final states from the DM annihilations. We display our findings in Figure 4.3 where we plot the effective indirect detection cross section with DM mass and show the experimental upper bound (black solid line) of DM annihilating to $t\bar{t}$ final states that can be obtained from antiproton cosmic ray data [186]. In the top left panel of Figure 4.3, we find that the DM mass for which the observed relic density satisfied (shown

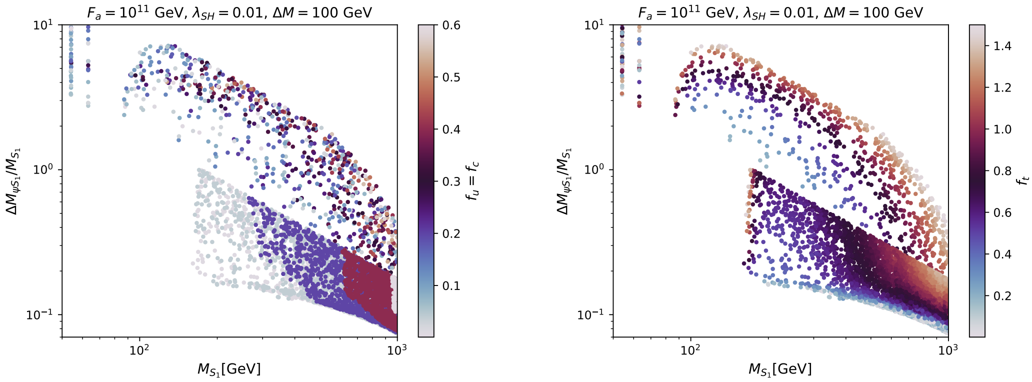


Figure 4.4: Parameter space satisfying observed DM abundance and also allowed by the direct search experiments in the bi-dimensional plane of $\frac{\Delta M_{\Psi S_1}}{M_{S_1}} - M_{S_1}$, where the color coding is done with respect to: Left Panel: the Yukawa couplings $f_u = f_c$ and Right Panel: the Yukawa coupling f_t. In both the plots we fix $F_a = 10^{11}$ GeV, $\lambda_{SH} = 0.01$, $\Delta M = 100$ GeV while we vary f_t in the range 0 - 1.5 and $f_u = f_c$ in the range 0 - 1.5.

by green $\star$) in the top left panel of Figure 4.2 is also allowed from the constraints coming from the indirect search bound. A similar situation is also observed in the top right and bottom left panels. On the other hand, in the bottom right panel where the variation with ΔM is studied, it found that a larger $\Delta M \sim 100$ GeV is prefered if one also considers the constraints coming from indirect search experiments. For this reason, we fix $\Delta M = 100$ GeV throughout our analysis.

In Figure 4.4, we show the parameter space that remains consistent with the DM constraints and is also allowed by the constraints that come from the flavor observable like $D^0 - \bar{D}^0$ mixings in the bi-dimensional plane of $\Delta M_{\Psi S_1}/M_{S_1}$ Vs M_{S_1}, where $\Delta M_{\Psi S_1}$ is the mass difference between VLQ and DM, $M_\Psi - M_{S_1}$. Here, the dependence upon different Yukawa couplings ($f_u = f_c$ in the left panel and f_t in the right panel of the figure) is spotted with a continuous color map. Two discrete narrow slices at the top-left corner due to Higgs resonance. We are primarily interested in the non-resonant continuous region extended over vast parameter space. At the lower M_{S_1} end, this continuous region opens up when the DM pair annihilate into a top quark and up (charm) quark (see Appendix D). Eventually, for a choice of heavier mass, the DM pair starts annihilating into the top pair.

This allowed region can be categorized into two distinct parts as upper and lower regions separated by a line where the mass difference between VLQ and DM equates to the top mass. Hence the upper region can be probed at the collider with on-shell production of top quark from VLQ decay, while the lower region is sensitive to a probe with light quark search. We will further demonstrate in the next section how top quark searches from the boosted top jet can improve the search strategy in this region.

As a consequence of the narrow mass gap between scalar DM and VLQ, co-annihilation takes a leading role in most parts of the lower region. Precisely because of the same reason, this region is also susceptible to the direct detection probe. Variations of color contours for different f_u values are evident in the lower region of the left plot in Figure 4.4. This reflects the gradually larger parameter space excluded due to direct detection constrain for a choice of larger f_u values. The lower value of mediator mass increases the direct detection cross section. In order to keep this cross section below the current direct detection bounds, a smaller f_u is required. On the contrary, the direct search experiments allow the upper region irrespective of the choice of f_u, and hence a uniform distribution of the colors is observed. This is because even for a large f_u, the DM-nucleon scattering cross section still remains small due to the presence of a heavier mediator *i.e* M_Ψ.

At the right part of the same plot, one finds that with an increasing DM mass, a relatively smaller $\Delta M_{\Psi S_1}/M_{S_1}$ is required in order to satisfy the correct relic density, while the interplay between the DM mass, mediator's mass and the f_u makes these points allowed from the direct detection constraints. In this region, the DM dominantly annihilates into the top-quark pair and sub-dominantly into the top quark and up (charm) quark final states. Next, in the right panel of Figure 4.4, we show the color coding with respect to f_t in order to highlight its significance. One observes the correct relic density in the top-left region of the plot due to the involvement of a large f_t as is also evident from the top left panel of Figure 4.2. As expected for a lighter mediator mass, a relatively smaller f_t is required to satisfy the correct relic density, as is also observed while moving downward in the plot. Finally, the role of f_t becomes more prominent once the $M_{S_1} = M_{top}$ threshold is opened, as can also be seen from the right side of the plot.

4.5 Collider Analysis and Results

The involvement of VLQ (Ψ) in the present setup opens up interesting collider prospects as they can be produced either in pair or associated with a scalar in the proton-proton collision at LHC. Among these production channels, $pp \to \Psi\bar{\Psi}$ and $pp \to \overset{(-)}{\Psi} S_{1,2}$, the cross section of the second process strongly depends on the Yukawa coupling $f_{u,c}$, while the pair of VLQs is produced primarily by the strong interaction and hence model-independent. Once produced, the VLQ can decay preferably into scalar DM candidate or its heavier pair along with one of

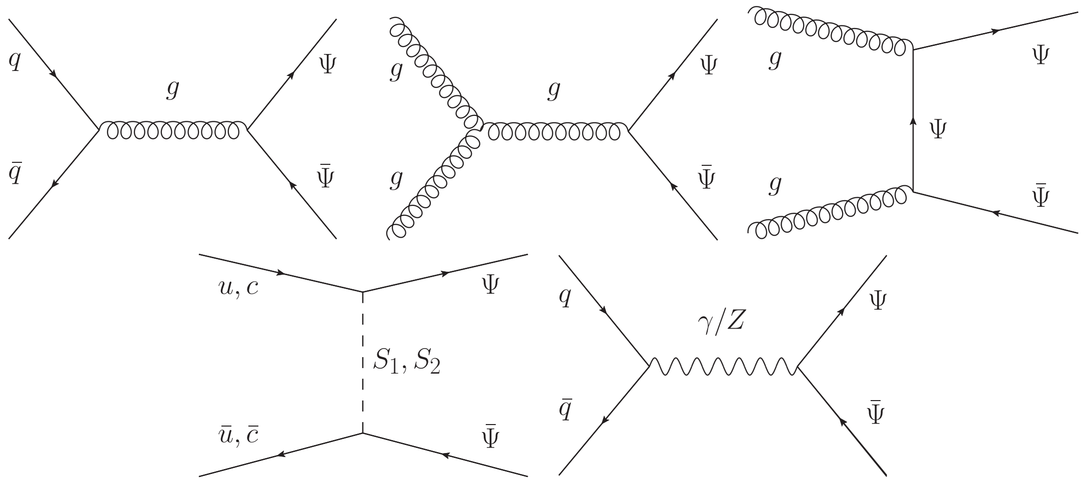


Figure 4.5: Representative Feynman diagrams for leading order partonic processes contributing VLQ pair production $pp \to \Psi\bar{\Psi}$ at LHC.

the up-type quarks as allowed kinematically. Hence primary LHC searches rely on identifying such quark jets along with missing transverse energy (MET or $\not{E}_T$) from DM production, as discussed in Section 4.3.

It is noteworthy that a substantial parameter space exists in this model where the mass difference between VLQ and DM is significantly larger than the top quark mass while providing correct dark matter relic density and also allowed from the direct detection experiments. Here produced top quarks are expected to be fairly boosted by production from the decay of heavy mother (VLQ) particles. Such a prospect motivates us to look at this signal possessing a unique topology where hadronic decay of the top retains its collimated structure as a boosted fatjet [e] and is identified as a top-like-fatjet (J_t).

To probe these regions at the LHC, we consider pair production of VLQs and each of those further decay into the top quark associated with the scalar (DM or S_2). Here we adopt a significantly smaller Yukawas $f_u = f_c$ $(= 0.01)$ so that the primary branching fraction of the decay of Ψ into the top quark is close to 100%. The signal topology is below, where we identify two final top fatjets associated with significant missing transverse momentum from dark matter.

$$pp \to \Psi\bar{\Psi} \to (t, S_{1,2}), (\bar{t}, S_{1,2}) \equiv 2J_t + \not{E}_T \tag{4.29}$$

Note, S_2 can decay through two-body ($S_2 \to S_1 a$), three-body ($S_2 \to S_1 j j$),

[e]We encountered a similar feature in the succession of different BSM scenarios [119,198–200], where boosted fatjet is probed in association with MET. Fatjets, in these searches, still harbor the intrinsic footprint of their root and manifest such features inside the jet substructure. Exploring this can provide additional tools to deal with a significant background involving QCD jets.

Benchmark points	M_{S_1} (GeV)	$\Delta M_{\Psi S_1}$ (GeV)	ΔM (GeV)	f_t	$\Omega_{S_1} h^2$	$\sigma^{SI}_{S_1,eff}$ (pb)	$\sigma(pp \to \Psi\bar{\Psi})$ (fb)
BP1	301	305	100	0.8	0.108	9.24×10^{-12}	966
BP2	302	475	100	1	0.104	9.96×10^{-12}	223
BP3	403	405	100	1	0.109	5.77×10^{-12}	175
BP4	358	448	100	1	0.109	6.48×10^{-12}	177
BP5	433	364	100	1	0.107	4.6×10^{-12}	188
BP6	459	326	100	1	0.109	4.2×10^{-12}	208
BP7	494	273	100	1	0.107	3.79×10^{-12}	239
BP8	510	238	100	1	0.099	3.69×10^{-12}	278
BP9	527	224	200	1	0.103	3.56×10^{-12}	272
BP10	542	188	100	0.98	0.106	3.57×10^{-12}	321
BP11	678	349	100	1.3	0.109	2.14×10^{-12}	37

Table 4.2: Different benchmark points satisfy the observed relic density of DM, direct and indirect detection (not shown in the table) bounds, along with the constraints coming from the theoretical and the LHC data, as listed in the text. M_{S_1} is the mass of the DM, S_1. f_t is the coupling strength of the interaction between top quark, VLQ, and the scalar S (see: Equation 4.1). $\Delta M_{\Psi S_1} = M_\Psi - M_{S_1}$ and $\Delta M = M_{S_2} - M_{S_1}$. $\Omega_{S_1} h^2$ (see: Equation 4.25) and $\sigma^{SI}_{S_1,eff}$ (see: Equation 4.27) are the relic density of DM, S_1 and effective direct detection cross section, respectively. Other parameters are $F_a = 10^{11}$ GeV, $\lambda_{SH} = 0.01$, and $f_u = f_c = 0.01$. The production cross section of the partonic process $pp \to \Psi\bar{\Psi}$ at LO for different benchmark points before decaying into the SM quark and scalar at 14 TeV LHC is given at the last column.

and four-body ($S_2 \to S_1\, j\, b\, W$) decay modes, where suppressed multi-body decay occurs through off-shell VLQ. Partonic level Feynman diagrams of the production of VLQ pair are shown in Figure 4.5. Although the main contribution comes from the strong interaction, we keep all the diagrams for completeness. Few representative benchmark points (BPs) are listed in Table 4.2; those provide observed relic density of DM and allowed from the direct and indirect detection experiments along the constraints coming from the theoretical and the LHC data as listed in Section 4.3. Also, the production cross section of the partonic process $pp \to \Psi\bar{\Psi}$ at LO for different benchmark points before decaying into the SM quark and scalar at 14 TeV LHC is shown at the last column Table 4.2. For our analysis, we have used an NLO QCD $K-$factor of 1.33 for the $pp \to \Psi\bar{\Psi}$ production [f].

[f] We estimate an approximate NLO (QCD) $K-$factor for the process $pp \to \Psi\bar{\Psi}$ by replacing Ψ with the top quark of mass m_Ψ at the MadGraph5_aMC@NLO and took the most conservative value over this mass range.

4.5.1 Simulation Details with Signal and Backgrounds

In preparation for our investigation of this Hybrid KSVZ framework through VLQ pair production at the LHC, we require a realistic setup to simulate both the signal processes as well as a careful selection of background processes that can mimic the signal.

We implement this Hybrid KSVZ framework in FeynRules [26] to generate the UFO model file required for matrix element generation for Monte-Carlo event generator. Parton level events are generated in the MadGraph5_aMC@NLO environment [69] and further pass through Pythia8 [146,147] for showering, fragmentation and hadronization. Background events are generated along with two to four additional jets MLM matching [150, 151] with virtually-ordered Pythia showers to avoid any double counting. We include higher-order corrections for different processes by multiplying the appropriate K factor. An in-built NN23LO1 pdf set is adopted for the parton distribution functions (PDF), and a default dynamical factorization scale is used for events generation. The showered events are further passed through Delphes3 [70] to include detector effects with the default CMS card. Jets (j) of radius parameter 0.5 are constructed with the anti-k$_T$ [98] clustering algorithm, where we used the particle-flow towers and particle-flow tracks as input. We implement the Cambridge-Achen (CA) [95] algorithm to construct large radius fatjets - J. Fastjet 3.2.2 [92] is used for clustering fatjets of radius parameter $R = 1.5$. A boosted top gives a fatjet whose radius parameter is approximately govorned by $R \sim 2m_t/P_T$, where m_t (P_T) is the mass (transverse momentum) of the top quark. Hence, the minimum transverse momentum required by each top to form such a fatjet is $P_T \gtrsim 200$ GeV. Finaly, we implement the adaptive Boosted Decision Tree (BDT) algorithm to perform the multivariate analysis (MVA) in the TMVA [112] framework.

Our analysis considers all the backgrounds that significantly contribute to the two boosted top fatjets with large missing transverse momentum, as listed below.

<u>$t\bar{t}+$ jets:</u> Top pair production with the semi-leptonic top decays is the most dominant background for our signal process. Although pure hadronic decay of tops can offer two boosted top jets, the requirement of a considerable amount of missing energy reduces this background by a significant factor of 100, where mismeasurement of hadronic activities acts as a source of MET. In the semi-leptonic decay, one top decay hadronically and is reconstructed as boosted top jet, and the other top decay leptonically gives a significant amount of missing energy when the lepton escapes detection. Other boosted jet comes from the QCD radiation. This background is matched with the MLM matching scheme up to two extra jets.

QCD background: QCD background is enormous at the LHC but can be reduced to a negligible contribution [201]. Even after the requirements of two boosted fatjets, we are left with a remarkably large number of events from this background. We further require at least one b tag within the leading or sub-leading fatjet. Contribute negligibly after additional suppression of 100 comes from fake MET from hardons and another 50 from the requirement of b tag fatjet. We do not include this background in our analysis.

$tW+$ jets: Single top production associated with W boson significantly contributes to the SM background. The top is reconstructed as the boosted top where the b quark is tagged within it, and the W boson decays leptonically to give rise to missing transverse momentum. In contrast, another boosted fatjet arises from QCD jets. MLM matching up to two extra jets is done for this process.

$V+$ jets: (Simi-)invisible decay of W/Z vector boson in addition to QCD radiation that emulates the fatjet can contribute sizably even with a requirement of sizeable reconstructed mass of the fatjet. We do MLM matching up to four extra jets for both processes. A generation level cut $\not{E}_T > 100$ GeV is applied for both processes to obtain statistically significant background events.

di-boson + jets: Minor contribution can come from Di-boson + jets. We retain all the three di-boson background processes ($pp \to WZ$, WW, ZZ) in our analysis. Among the three, WZ + jets contribute the most. All three processes are matched up to two extra jets with an MLM matching scheme. In all the cases, one of the vector boson decay invisibly ($Z \to \nu\nu$) or leptonically ($W \to l\nu$) to give $\not{E}_T$. One of the boosted fatjet comes from QCD jets, and another fatjet comes from either hadronically decaying vector boson or the QCD jets.

$t\bar{t}V$: Such processes have three body phase spaces and have less cross section than other background processes mentioned above. Both the tops can be reconstructed as boosted fatjets, while $\not{E}_T$ comes from the invisible or leptonically decay of the Z and W boson, respectively. Among these two, $t\bar{t}Z$ contributes the most because of the larger cross section and more significant efficiency when applying $\not{E}_T$.

We consider all contributions generating those events at leading order and normalize with the NLO (QCD) cross section. Higher-order QCD corrected production cross section at the 14 TeV LHC for different background processes accounted in this analysis are listed in Table 4.3.

Background		σ (pb)
top pair [202]	$t\bar{t}$+ jets	988.57 [N^3LO]
single top [203]	tW	83.1 [N^2LO]
mono-V boson [204, 205]	Z+ jets	6.33×10^4 [N^2LO]
	W+ jets	1.95×10^5 [NLO]
di boson [206]	ZZ+ jets	17.72 [NLO]
	WW+ jets	124.31 [NLO]
	WZ+ jets	51.82 [NLO]
mono-V + $t\bar{t}$	$t\bar{t}Z$	0.911 [NLO]
	$t\bar{t}W^{\pm}$	0.636 [NLO]

Table 4.3: Higher-order QCD corrected cross section at the 14 TeV LHC of different background processes considered in our study. The order of QCD correction is given in brackets. For the final process, higher-order QCD corrected cross section in five massless quark flavors at 14 TeV LHC obtained from MG5_aMC@NLO. Default factorization and renormalization scales and an in-built NN23NLO pdf set are used.

4.5.2 Construction of High-Level Variables and Cut-Based Analysis

Once we have generated our signal and background processes after the realistic detector-level simulation, the next task is constructing high-level event variables sensitive to kinematic configuration signal and background processes. For example, the unique point of this collider study counts on the fatjet characteristic and its different properties related to the mass-energy distribution within these fatjets. We categorize some of the useful variables for our analysis in the following bulleted points:

N-subjettiness ratio: In the case of a highly boosted top quark, one can capture all three hadronically decayed constituents of the top quark within a single large-radius jet (fatjet). The whole energy of a reconstructed top-fatjet is distributed around three subjet axes. Assuming N number of subjets belong to the fatjet, N-subjettiness is defined by the angular distance in the transverse plane of constituents of the fatjet from the nearest subjet axis and weighted by the transverse momentum of the constituents as below [106, 107]:

$$\tau_N = \frac{1}{\mathcal{N}_0} \sum_i P_{T,i} \min\{\Delta R_{i,1}, \Delta R_{i,2}, ..., \Delta R_{i,N}\}. \tag{4.30}$$

Here, the summation goes over all the particles inside the jet. The denominator is $\mathcal{N}_0 = \sum_i P_{T,i} R$, where $P_{T,i}$ and R are the transverse momentum of the i-th constituent and radius of the jet, respectively. Since N-subjettiness determines the jet shape, the N-subjettiness ratios, such as τ_{31} and τ_{32} are good observables in signal background analysis. τ_{32} effectively distinguishes the top signal from two-prong fatjets arising from the boosted W or Z boson in the background. In contrast, τ_{31} is also effective for separating the top signal from the one-prong QCD fatjets that contribute significantly to the background.

Pruned jet mass: Jet-mass is a good variable for classifying a boosted top-fatjet from the two-prong fatjets from the boosted W/Z boson or one-prong QCD fatjets. The jet mass, $M_J = (\sum_{i\in J} P_i)^2$, where four-momentum of i-th energy hit in the calorimeter is P_i. Since large radius jets pick additional soft contributions from underlying QCD radiations, we must remove these soft and wide-angle radiations for more realistic predictions. Different jet grooming techniques, pruning, filtering, and trimming [102–105] are available to remove those softer and wider angle radiations while we consider pruning in our analysis. In the first step of pruning, we define fatjet using the CA algorithm, and in the second step, we pruned its constituents in each recombination step.

$$Z = \min(P_{Ti}, P_{Tj})/P_{T(i+j)} < Z_{\rm cut} \qquad \text{and} \qquad \Delta R_{ij} > R_{\rm fact}\ . \tag{4.31}$$

The merging $i, j \to J$ is vetoed when both the conditions are satisfied. Pruning is parametrized by two parameters, the softness parameter, Z, and the angular distance of the constituents, ΔR_{ij}. We chose $Z_{cut} = 0.1$ [103] and $R_{\rm fact} = 0.86$ ($\sim m_t/P_{T,\rm top}$) [104] in our analysis.

Primary event selection criteria : Based on our previous discussion and construction of high-level variables, we identify two large-radius jets, leptons, and missing transverse energy as per the following event selection criteria both for the signal and background events alike:

1. Each event should contain at least two fatjets constructed by CA algorithm with radius parameter $R = 1.5$, and each of them has transverse momentum, $P_T(J_0), P_T(J_1) > 200$ GeV. Here, J_0 and J_1 represent the leading and subleading fatjet.

2. Each event is selected with a minimum missing transverse energy $\not{E}_T > 100$ GeV.

3. Since our signal does not contain lepton, we veto any event if it contains any lepton with transverse momentum, $P_T(l) > 10$ GeV within pseudo-rapidity $|\eta(l)| < 2.4$.

4. To minimize jet mismeasurement contribution to $\not{E}_T$, we keep an azimuthal separation between each fatjet and $\not{E}_T$, $|\Delta\Phi(J_{0,1}, \not{E}_T)| > 0.2$.

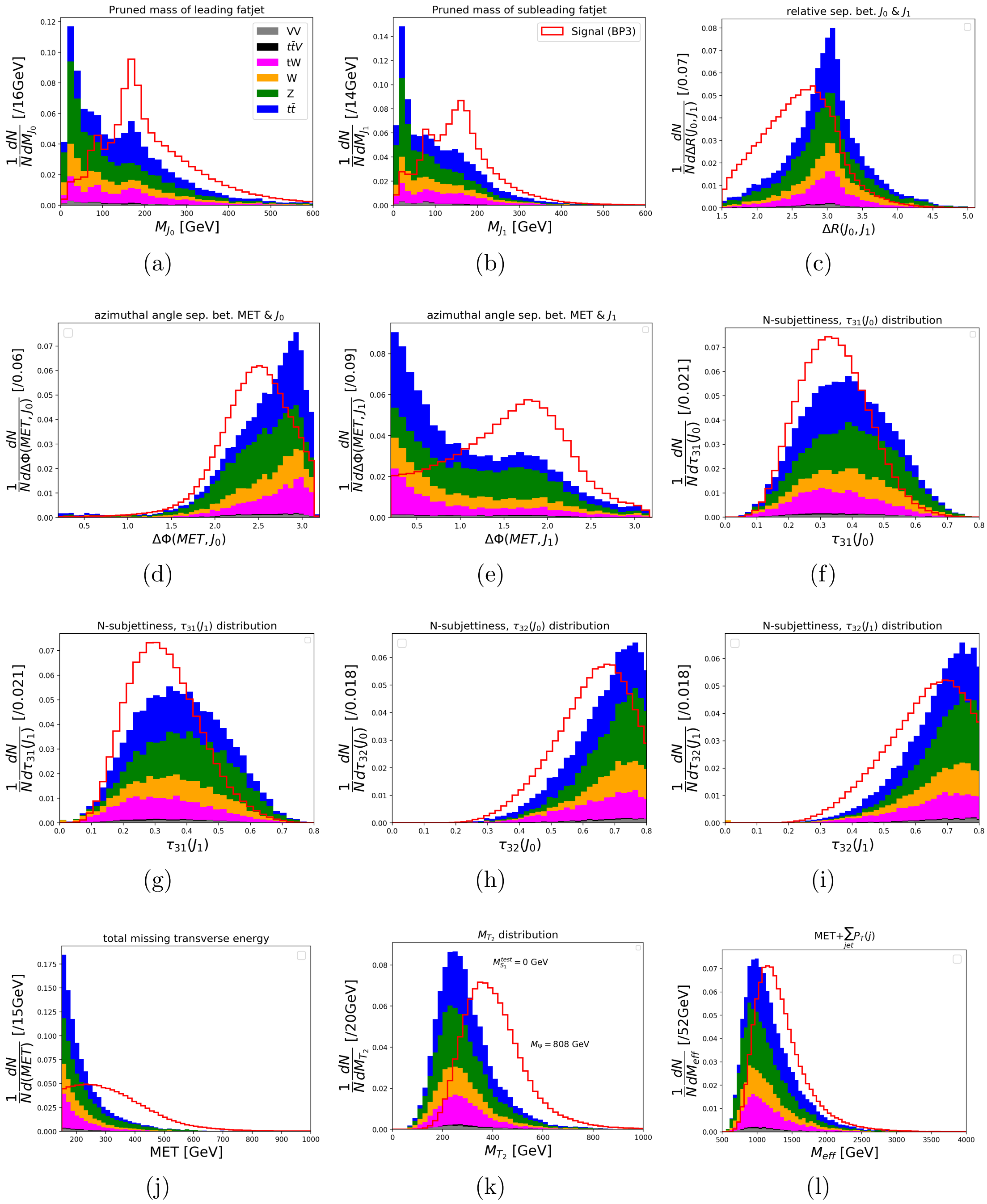


Figure 4.6: Distributions of different kinematical variables for the signal (BP3) and all the backgrounds contributing to the fatjets $+\not{E}_T$ final state after imposing the b tag within leading or subleading fatjet, $\not{E}_T > 150$ GeV, and the primary event selection criteria (described in the text) for 14 TeV LHC. The normalized distribution for the signal is given by the solid red line. The events of each background process have been weighted by their cross section and the cut efficiency after applying the previously mentioned cuts. Each background process is then normalized to the sum of individual cross section times cut efficiency. Colors show the contribution of the individual background process.

The normalized distribution of different observables of a sample signal benchmark point, BP3, and bin-wise stacked histogram of all the backgrounds are shown in Figure 4.6. These plots are shown after demanding at least one the b tag within leading or subleading fatjet, enhanced $\not{E}_T > 150$ GeV, over the preselection cuts already described for 14 TeV LHC.

The prime background $t\bar{t}+$ jets, where one of the top decay hadronically and the other decays leptonically, is shown by the top most blue shade, while the solid red line indicates the sample signal, BP3. The distributions of the pruned jet mass of the leading (M_{J_0}) and subleading (M_{J_1}) fatjets are given in Figure 4.6a and Figure 4.6b, respectively. At LO, Ψ and $\bar{\Psi}$ produce back to back, each one followed by decay into an (anti)top quark and $S_{1,2}$. In most events in this benchmark point, these tops are boosted as they are produced from the decay of heavy particles. When the top is sufficiently boosted, all three constituents of the top quark fall within a single large-radius jet, giving a three-prong jet substructure and pruned jet mass very close to the top quark mass. For the signal, we get a sharp peak around the top quark mass for both the leading and subleading fatjet. These large radius jets sometimes misses some of the constituent sub-jets, especially when the boost of the top quark is relatively low, causing a secondary peak near the W/Z boson mass for both the fatjets of the signal. For semi-leptonic $t\bar{t}+$ jets background, the top which decays hadronically gives the leading fatjet for a significant number of events and causes a sharp peak near top mass in the leading fatjet mass distribution. From the demand for a very high missing transverse momentum, $t\bar{t}+$ background contributes to a phase space region where the b-jet from the leptonically decaying top quark generates the subleading fatjet predominantly. Consequently, subleading fatjet mass generates its peak near 20 GeV from QCD radiation.

The total missing transverse energy distribution is shown as another interesting variable in Figure 4.6j. In the case of signal, we have two missing DM particles coming from the decay of Ψ pair, where they primarily produce back to back, so the $\not{E}_T$ has uniform distribution as two missing particles can avail entire phase space. In contrast, the background drops sharply for large $\not{E}_T$. Distributions of the azimuthal separation of the leading and subleading fatjets from the $\not{E}_T$ are presented in Figure 4.6d and Figure 4.6e, respectively. As stated earlier, two missing particles can avail the entire phase space for the signal, so both $\Delta\Phi(MET, J_{0,1})$ have a uniform distribution. For a significant amount of events of the $t\bar{t}+$ jets background, the b-jet from the leptonically decaying top quark behaves as a subleading fatjet (J_1), and the neutrino gives the $\not{E}_T$, where we select the events that have large $\not{E}_T$. Hence, the azimuthal separation of J_1 from $\not{E}_T$ gets a maximum at a lower value. In contrast, the azimuthal separa-

	BP3	$t\bar{t}$+jets	tW+jets	ttZ	ttW	Z+jets	W+jets	WZ+j	ZZ+j	WW+j
C1	5969 [100%]	9.6×10^4 [100%]	5.1×10^4 [100%]	1048 [100%]	111 [100%]	3.5×10^5 [100%]	1.9×10^5 [100%]	1.3×10^4 [100%]	1.6×10^3 [100%]	3.6×10^3 [100%]
C2	5296 [88.73%]	4.2×10^4 [43.89%]	2.12×10^4 [41.96%]	793 [75.71%]	64 [57.53%]	2.28×10^5 [65.06%]	1.06×10^5 [53.78%]	8.11×10^3 [64.34%]	969.2 [61.73%]	1.6×10^3 [43.97%]
C3	4424 [74.11%]	3.21×10^4 [33.60%]	1.59×10^4 [31.42%]	656 [62.63%]	54.1 [48.73%]	3.36×10^4 [9.57%]	1.64×10^4 [8.32%]	1.5×10^3 [11.89%]	267 [17.0%]	341.1 [9.37%]
C4	1005 [16.85%]	4.02×10^3 [4.20%]	1.72×10^3 [3.39%]	185 [17.66%]	16.7 [15.07%]	1.54×10^3 [0.44%]	926 [0.47%]	72 [0.57%]	10.4 [0.66%]	26 [0.71%]
C5	666 [11.16%]	2.46×10^3 [2.57%]	1.07×10^3 [2.12%]	132.5 [12.64%]	12 [10.84%]	842 [0.24%]	493 [0.25%]	42.5 [0.337%]	7.1 [0.45%]	15.7 [0.43%]
C6	432 [7.24%]	411 [0.43%]	197 [0.39%]	54 [5.12%]	3.1 [2.78%]	260 [0.074%]	132 [0.067%]	17.5 [0.139%]	4.3 [0.272%]	1.7 [0.047%]

Table 4.4: The cut efficiency and expected number of events after the corresponding cuts for the signal and all the backgrounds contribute to the fatjets $+\not{E}_T$ final state at the 14 TeV LHC and 139 fb^{-1} integrated luminosity. The effectiveness of different selection cuts can be followed in the form a cut flow from top to bottom after applying (C1) Preselection cuts, (C2) $\not{E}_T > 150$ GeV, (C3) requiring at least one b-tag within J_0 or J_1, (C4) 120 GeV$< M_{J_0}, M_{J_1} < 230$ GeV, (C5) $\tau_{31}(J_0), \tau_{31}(J_1) < 0.4$ and finally, (C6) $M_{T_2} > 320$ GeV. A sample benchmark point, BP3, is presented in this table.

tion of the leading fatjet (J_0) from $\not{E}_T$ peaks near $\sim \pi$ rad. The distribution of $\Delta R(J_0, J_1)$, angular distance between J_0 and J_1 in the transverse plane is given in Figure 4.6c.

The distribution of the kinematic variable τ_{31} for both leading and subleading fatjet are shown in Figures 4.6f and 4.6g, respectively. In both distributions, as expected, the signal has a peak for a smaller value of τ_{31} representing that signal fatjets have a three-prong structure. Similarly, the distribution of the kinematic variable τ_{32}, which separates the three-prong fatjet from the two-prong fatjet, are presented in Figures 4.6h and 4.6i. $\tau_{32}(J_0)$ has a peak near 0.6 and 0.75 for signal and background, respectively. Note that we do not apply any mass window in generating these distributions, but if we do, the peaks of τ_{32} move towards a lower value. So, in the final event selection in the cut-based analysis, we apply a mass window to discriminate the signal from the background better.

The distribution of kinematic transverse mass variable M_{T2} [154, 207, 208] is given in Figure 4.6k. Assuming DM mass is unknown to us, we construct M_{T2} after setting trial DM mass as zero in this construction. The SM particles have a smaller mass compared to the mass of Ψ, so the M_{T2} distribution of signal and background are well separated. Since we do not want to find the correct mass of the mother particle (Ψ), this variable is used to discriminate the signal from the background efficiently. The distribution of M_{eff} is given in Figure 4.6l. Effective mass is defined as

$$M_{\text{eff}} = \not{E}_T + H_T, \tag{4.32}$$

BP	BP1	BP2	BP3	BP4	BP5	BP6	BP7	BP8	BP9	BP10	BP11
σ	11.4	12.8	11.1	9.9	6.9	5.3	4.0	3.6	2.5	1.2	2.8
$\frac{S}{B}$	0.41	0.47	0.40	0.35	0.23	0.17	0.13	0.12	0.08	0.04	0.09

Table 4.5: Statistical significance (σ) and the signal-to-background ratio ($\frac{S}{B}$) are shown for the signal corresponding to different benchmark points contributing to the fatjets $+\not{E}_T$ final state at 14 TeV LHC and 139 fb^{-1} integrated luminosity.

where $H_T \equiv \sum_{i=1}^{N_J} P_{iT}$ (N_J is the number of visible jets) is the scalar sum of the transverse momentum of the jets. The above distributions show that all the variables are very good at distinguishing the signal from the background.

We apply the following selection cuts to demonstrate a cut-based analysis (CBA) over the preselection cuts (described before) to increase the signal-to-background ratio. Note that our final results are based on sophisticated multi-variate analysis with improved statistics. So, the next part is for demonstration purposes without putting much effort into optimizing all the selection criteria. Here, we offer a cut-flow in cut-based analysis to better understand the signal and background differential distributions.

Final selection cuts:

5. We increase $\not{E}_T$ from 100 GeV to 150 GeV since it reduces the background sharply than the signal.

6. Demand an additional b-tag within either leading or subleading fatjet is applied. The b-tag efficiency for the signal within leading or subleading fatjet is 84%. This requirement reduces $Z+$ jets and $W+$ jets backgrounds substantially below $t\bar{t}+$ jets background.

7. We select the events for which the pruned mass of the leading and sub-leading fatjets falls within 120 GeV$< M(J_0), M(J_1) <$ 230 GeV [g]. The lower threshold helps us reduce different backgrounds where one or both the fatjets originated from QCD radiation or W/Z boson.

8. To discriminate further the fatjets from QCD jets, we use N-subjettines and collect the events that satisfy $\tau_{31}(J_0)$ and $\tau_{31}(J_1) < 0.4$ [h].

9. We impose $M_{T_2} >$ 320 GeV. This requirement increases the signal-to-background ratio ($\frac{S}{B}$). For example, in the case of BP3, $\frac{S}{B}$ changes from

[g]Note that, in the MVA next section, we retain only the lower mass threshold and let the framework select the non-linear cuts to get the optimal signal-to-background ratio.

[h]One may use the N-subjettiness variables $\tau_{32}(J_0)$ and $\tau_{32}(J_1)$ to discriminate the fatjets from two-prong fatjets originated from boosted W/Z bosons. Since we analyze the same signal using MVA in the next section, we do not check τ_{32} variables in CBA.

Signal	BP1	BP2	BP3	BP4	BP5	BP6	BP7	BP8	BP9	BP10	BP11
	6625	3525	2341	2711	2176	1924	1424	1081	915	552	385

SM BG	tt+jets	tW+jets	ttZ	ttW	Z+jets	W+jets	WZ+jets	ZZ+jets	WW+jets
	8928.06	3815.42	294.35	25.93	3527.96	2408.37	172.35	27.54	46.49

Table 4.6: The expected number of signal and the SM background events after applying $M_{J_0} > 120$ GeV, $M_{J_1} > 120$ GeV and b-tag (within leading or subleading fatjet), $\not{E}_T > 150$ GeV in addition to preselection cuts at 14 TeV LHC for 139 fb^{-1} integrated luminosity.

0.13 to 0.4 (Table 4.4).

The expected number of signal (for a sample benchmark point, BP3) and background events and cut efficiency after imposing the preselection cuts and final selection cuts at 14 TeV LHC for 139 fb^{-1} integrated luminosity are shown in Table 4.4. Statistical significance and the signal-to-background ratio for different benchmark points are shown in Table 4.5. $\sigma = \frac{N_S}{\sqrt{N_S+N_B}}$ defines the statistical significance, where N_S and N_B are the expected signal and background events after the cuts, respectively. The statistical significance for the signal of different benchmark points is above the discovery potential for an integrated luminosity of 139 fb^{-1}. We also have good statistics indicating that extracting the VLQ pair from the Standard Model background is not tough.

4.5.3 Analysis based on the Multivariate Gradient Boosting Technique

In the previous section, we constructed high-level variables and demonstrated their potential in a CBA. This section extends that idea to perform a more sophisticated MVA. In these analyses, MVA generally gives better sensitivity than CBA if appropriate kinematic variables are utilized, where we may get significance above the discovery limit for the benchmark points that is unable through CBA. The M_{J_0} and M_{J_1} distribution (Figure 4.6a, Figure 4.6b) have the largest peak around the top mass, and the signal is much harder than the background for $M_{J_{0,1}} > 120$ GeV. Instead of both lower and upper mass thresholds to set an allowed window, we retain only a lower mass threshold of 120 GeV for both the fatjets for event selection in MVA for a higher number of events. We expect the MVA framework to select nonlinear variable space to get the optimal signal-to-background ratio. The 120 GeV cut on both fatjets reduces the backgrounds drastically compared to the signal for which fatjets arrises from the QCD jet (one-prong) or boosted W/Z boson (two-prong). We also demand at least one b tag within the leading or subleading fatjet for event selection in MVA, reducing

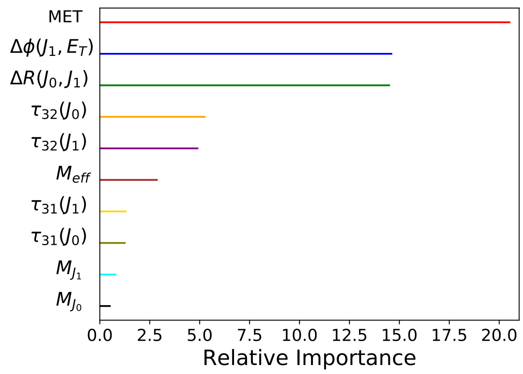


Figure 4.7: Relative importance (Method unspecific) of the different kinematic variables used in MVA. We get those numbers for BP3 from the TMVA package. Those numbers can change a little bit if one chooses a different algorithm.

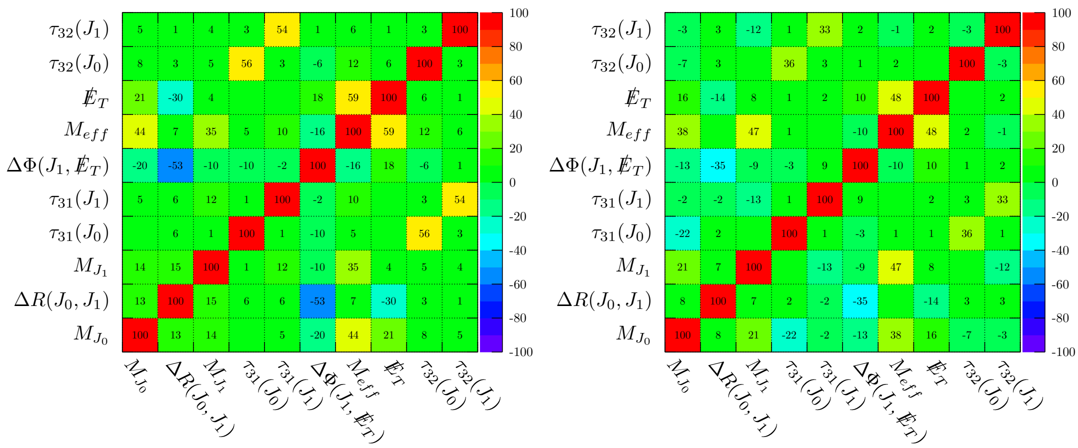


Figure 4.8: Linear correlation coefficients (in percentage) between different kinematical variables for the signal (BP3, left panel) and background (right panel). Missing entries correspond to a negligible correlation smaller than one. Positive and negative coefficients indicate that two variables are correlated or anti-correlated, respectively.

the background much more than the signal. From the missing energy distribution (Figure 4.6j), we see most background events exist in low missing energy, so after demanding large missing energy, we reduce the background significantly compared to the signal. So we apply $\not{E}_T > 150$ GeV for the event selection in MVA.

With these selection criteria, we keep all other variables unrestrained, giving enough scope to the multivariate analysis to find an optimal nonlinear cut based on the suitable variables. The expected number of signal and background events after applying MVA selection cuts at 14 TeV LHC for 139 fb^{-1} integrated luminosity is given in Table 4.6. We apply the adaptive Boosted Decision Tree (BDT)

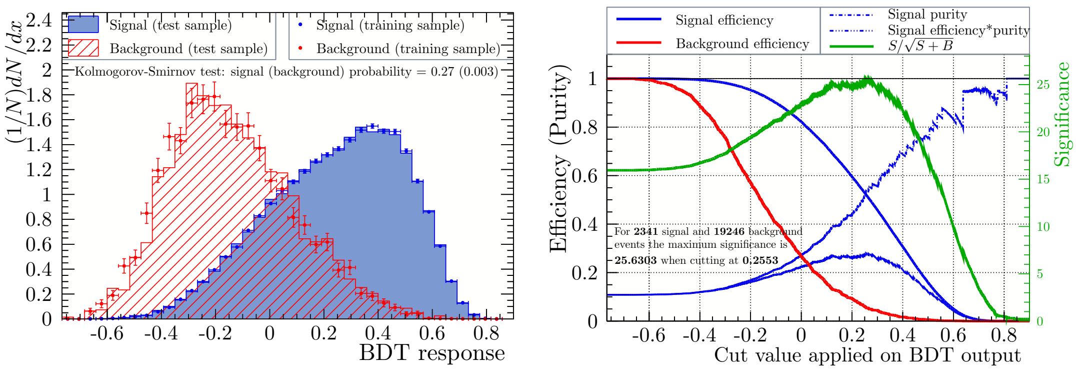


Figure 4.9: Left panel: normalized distribution of the BDT response for both signal (blue, BP3) and background (red) classes (both training and testing samples of both classes). Right panel: signal (blue) and background (red) efficiencies and the statistical significance of the signal (green) as a function of cut applied on BDT output.

algorithm in our analysis and construct statistically independent signal and background event samples. Each event sample is split randomly for training and testing purposes. Since multiple processes contribute to the total background, we generate them with two to four extra jets MLM matching separately and combine them in proportion to their proper weight to get a combined background sample. For multivariate analysis, a final set of kinematic variables are accepted from a larger set, where we retain only those variables that are less (anti) correlated in both signal and background and have larger relative importance. Even before implementing any model, a variable can have more relative importance than another when it has larger discriminating power separating the signal class from the background class. We find $P_T(J_0)$, $P_T(J_1)$, and $\sqrt{\hat{S}_{min}}$ [152, 153] are highly correlated with M_{eff} in both signal and background. However, we keep M_{eff} as it has larger relative importance than other variables. $\sqrt{\hat{S}_{min}}$ is defined as

$$\sqrt{\hat{S}_{min}} = \sqrt{(\sum_j E_j)^2 - (\sum_j P_{z,j})^2} + \not{E}_T \tag{4.33}$$

where summation runs over all the visible jets. From Equations 4.32 and 4.33, the above correlations are expected. We also observe that $\Delta\Phi(J_0, \not{E}_T)$ and $\Delta\Phi(J_1, \not{E}_T)$ are moderately anti-correlated in signal but highly anti-correlated in background. The moderate anti-correlation of the signal is because of the total availability of phase space of the two missing particles. In the case of background, for example, the principle $t\bar{t}+$ jets background, the only allowed phase space is when both top and anti-top are highly boosted and move almost in the opposite direction, where one of the reconstructed tops gives the leading fatjet, and another one gives both subleading fatjet and large missing transverse momentum. As a result, these two variables are highly anti-correlated in the background. We keep $\Delta\Phi(J_1, \not{E}_T)$ as it has larger relative importance than $\Delta\Phi(J_0, \not{E}_T)$. We notice that M_{T2} and $\not{E}_T$ are also highly correlated in signal and moderately in background. Since $\not{E}_T$ has the largest relative importance, we choose $\not{E}_T$ over M_{T2} for MVA analysis. The relative importance of the different kinematic variables used in MVA is presented in Figure 4.7 for sample benchmark point BP3. From the normalized distributions in the previous section, we notice that all variables used in MVA are outstanding in distinguishing the signal from the background. However, the $\not{E}_T$, $\Delta R(J_0, J_1)$, and $\Delta\Phi(J_1, \not{E}_T)$ are the finest among all these useful variables. The linear correlation coefficients (in percentage) between different kinematical variables for the signal (BP3, left panel) and background (right panel) is presented in Figure 4.8. BDT algorithm may lead to overtraining for wrong choices of different (BDT specific) parameters during training. Such overtraining can be avoided if one checks the Kolmogorov-Smirnov probability during training. We train the algorithm separately for every benchmark point and confirm that no overtraining exists in our analysis.

The normalized distribution of the BDT response of the signal (BP3) and the background classes (both training and testing samples of both classes) is shown in Figure 4.9. We notice both the classes are well separated. We present the variation of the signal and background efficiencies and the statistical significance of the signal (BP3) with the cut applied on the BDT response in the right panel of Figure 4.9. Statistical significance is defined as $\sigma = \frac{N_S}{\sqrt{N_S+N_B}}$. The number of events that survive after applying the $\text{BDT}_{\text{res}} > \text{BDT}_{opt}$ cut for signal and background is N_S and N_B, respectively. BDT_{opt} is the optimal cut for which the significance is maximum. In Table 4.7 (upper) N_S, N_B, σ, and $\dfrac{N_S}{N_B}$ are presented for different benchmark points at 14 TeV LHC with integrated luminosity 139 fb^{-1}. We find that for a few of the chosen BPs, the number of signal events is larger than the background events after the $\text{BDT}_{\text{res}} > \text{BDT}_{opt}$ cut, and for all eleven benchmark points, we reach the discovery potential with integrated luminosity 139 fb^{-1}.

Our next interest would be to verify how significance varies with the mass of the scalar S_2. For that purpose, we generated the event samples separately for different masses of the S_2 with the same DM mass $M_{S_1} = 403$ GeV, VLQ mass $M_\Psi = 808$ GeV, and the coupling constant $f_t = 1$. We train the algorithm separately for different samples of different S_2 masses and confirm that no overtraining exists in our analysis and perform the MVA.

There are two possible hierarchies possible: $M_{S_2} > M_\Psi > M_{S_1}$ and $M_\Psi >$

BP	N_S^{bc}	BDT$_{opt}$	N_S	N_B	σ	$\frac{N_S}{N_B}$ (MVA)	$\frac{N_S}{N_B}$ (CBA)
BP1	6625	-0.0259	5643	8584	47.3	0.66	0.41
BP2	3525	0.1584	2325	2047	35.1	1.14	0.47
BP3	2341	0.2553	1222	1048	25.6	1.17	0.40
BP4	2711	0.1975	1556	1277	29.2	1.22	0.35
BP5	2176	0.1446	1366	2502	21.9	0.55	0.23
BP6	1924	0.1325	1170	2727	18.7	0.43	0.17
BP7	1424	0.1398	821	2845	13.6	0.29	0.13
BP8	1081	0.0942	624	2951	10.4	0.21	0.12
BP9	915	0.0785	627	4820	8.5	0.13	0.08
BP10	552	0.0506	311	3469	5.0	0.09	0.04
BP11	385	0.3875	134	346	6.1	0.39	0.09
N_{SM}	19246						

ΔM (GeV)	N_S^{bc}	BDT$_{opt}$	N_S	N_B	$\sigma = \frac{N_S}{\sqrt{N_S+N_B}}$
75.1	2315	0.2645	1252	1240	25.1
100 (BP3)	2341	0.2553	1222	1048	25.6
153.6	2258	0.2129	1246	1176	25.3
200.4	2035	0.1899	1284	1384	24.9
244.5	2019	0.2963	1159	948	25.2
302.7	2036	0.208	1254	1240	25.1
351	2034	0.2271	1261	1199	25.4
402.7	2023	0.2778	1133	882	25.2
N_{SM}	19246				

Table 4.7: The upper table demonstrates the effectiveness of the current search in terms of statistical significance (σ) for different benchmark points conceived for this study. The lower table illustrates the variation of this potential for one benchmark point, changing the mass of the heavy scalar, S_2, and shows that this mass does not have much impact in exploring the parameter space. N_S^{bc} and N_{SM} are the total number of events for different signal benchmark points and the combined background before applying any cut on BDT output (as shown in Table 4.6). After using an optimal selection on the BDT response (BDT$_{opt}$) surviving number of signal and background events are given by N_S and N_B, respectively for 14 TeV LHC for an integrated luminosity 139 fb^{-1}. Corresponding statistical significance and the signal-to-background ratio are also presented for ready reference. To better compare the sensitivities between the different analysis methods, we add the $\frac{N_S}{N_B}$ ratio of CBA from Table 4.5 in the last column of the upper table.

$M_{S_2} > M_{S_1}$. We consider two boosted tops associated with missing transverse momentum as our signal. It is interesting to note that the significance of the former hierarchy is always greater or equal to the second. In the case of the former, S_2 can decay into $\Psi(\to S_1 j)j$ or $S_1 jj$ (through off-shell Ψ), where j is the up-type SM quark. If at least one of these jets is the top quark, then the signal efficiency increases and hence the significance. So the hierarchy $M_\Psi > M_{S_2} > M_{S_1}$ gives a lower statistical significance, and we consider this scenario throughout our result for a conservative estimation. The total number of events coming from the signal topology for different masses of S_2 and background events after applying an optimal cut (BDT$_{opt}$) is given in Table 4.7 (lower) for the hierarchy

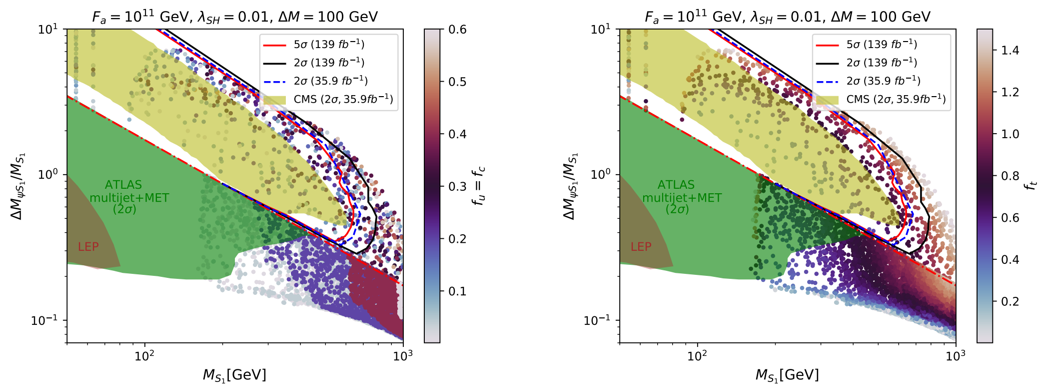


Figure 4.10: The solid red line is the 5σ discovery contour, and any point inside the red line has a statistical significance $> 5\sigma$ at 14 TeV LHC for an integrated luminosity 139 fb^{-1}. The dashed red line corresponds to $\Delta M_{\Psi S_1} = M_\Psi - M_{S_1} = M_{\rm top}$. Below the dashed red line, we can not probe with the boosted tops plus missing energy signal, as we can not get any on-shell top from the decay of Ψ. The solid black and blue dashed lines are the exclusion contour (2σ) of our analysis for an integrated luminosity of 139 fb^{-1} and 35.9 fb^{-1}, respectively. The exclusion region (2σ) from the existing LEP, ATLAS (multijet + MET), and CMS ($t\bar{t}$ + MET) analysis are shown by brown, green, and olive color, respectively.

$M_\Psi > M_{S_2} > M_{S_1}$. The statistical significance variation with the S_2 mass is also shown here for a given mass of Ψ and couplings. The mass of S_2 has no effect on the statistical significance of the boosted top fatjets plus a large missing momentum signature. However, if one of the decay products of S_2 ($S_2 \to S_1 jj$) is at least a top quark, then it can increase the significance.

Finally, we present the discovery (5σ) and exclusion (2σ) contours from our analysis at 14 TeV LHC for an integrated luminosity 139 fb^{-1} in the bi-dimensional plane of $\dfrac{\Delta M_{\Psi S_1}}{M_{S_1}} - M_{S_1}$ in Figure 4.10 by solid red and solid black lines, respectively. Our analysis is effective when the on-shell top is produced from Ψ decay. Hence the region below the dashed red line can not be probed in the present channel. Considering the 100% branching fraction of the decay of Ψ into the top quark associated with the scalar, the existing search [186] can exclude vector-like quark masses up to 1 TeV. We find the masses of the vector-like quark ranging up to 1.41 TeV can be excluded, while the masses extent to 1.28 TeV can be discovered at 14 TeV LHC with an integrated luminosity of 139 fb^{-1}. In the region below the dashed red line, the mass difference between a vector-like quark and the scalar DM is less than the top quark's mass, and the vector-like quark fully decays into a light quark associated with a scalar when kinematically allowed. So, one can probe those regions using multi jets plus missing transverse momentum signature, which is beyond the scope of our present analysis.

4.6 Summary and Conclusion

In this chapter, we analyze a hybrid KSVZ setup, where the model is extended by an extra complex scalar singlet whose lightest component plays the role of dark matter. We highlight the fact that the presence of a colored vector-like quark that occurs naturally in the KSVZ model plays a crucial role both in the dark sector and collider phenomenology of the setup. Being charged under $U(1)_{PQ}$ allows the VLQ to couple with all up-type quarks and the DM through the Yukawa interactions. When appropriately tuned, this coupling can enhance the DM parameter space in comparison to what is observed in a pure scalar singlet DM scenario. In this chapter, we demonstrate that the Yukawa couplings play a non-trivial role in obtaining the observed relic density. Moreover, the same couplings also allow the parameter space from the direct search bounds by entering into extra Feynman diagrams that contribute toward the direct detection cross section of the dark matter.

A search of vector-like quarks in events with two boosted top fatjets with large missing transverse momentum is presented. The analysis is done for 139 fb^{-1} integrated luminosity at 14 TeV LHC. We discuss all the significant backgrounds that can potentially mimic the signal. Jet substructure variables and various other variables are used in our analysis. Sophisticated multivariate analysis is performed to increase the sensitivity over cut-based one. Different jet substructure variables, $\Delta R(J_0, J_1)$, N-subjettiness ratios, and M_{eff} are outstanding in distinguishing the signal from the background and take a central role in getting very high significance. However, the missing transverse momentum distribution and the azimuthal separation between the subleading fatjet and the missing transverse energy direction have the uppermost importance in separating the signal from the background. With a conservative estimation, we give discovery and exclusion contours in Figure 4.10 in the region where the mass difference between the vector-like quark and the scalar dark matter is larger than the top quark mass.

Chapter 5

Precision prediction of a democratic up-family philic KSVZ axion model at the LHC

5.1 Introduction

In this chapter, we explore the extended KSVZ model described in the previous chapter but from different perspectives. Here we take NLO QCD corrections of the VLQ pair production at the LHC and do the multivariate analysis considering NLO+PS signal. Unlike the previous chapter, we consider equal (democratic) coupling strengths $f_u = f_c = f_t$ at all three generations and try to find the parameter spaces that yield the correct relic density and are permissible from other experimental observations such as direct detection (DD), collider data, etc. Interestingly, one would find that the flavor constraint strongly disfavors this democratic option, although such models can be allowed from observed relic density and all other constraints.

The correct relic density is achieved through $S_1S_1 \to t\bar{q},\ \bar{t}q\ (q = u, c)$ or $S_1S_1 \to t\bar{t}$ annihilation processes in parameter spaces where S_1 and VLQ are not degenerate and apart from the Higgs resonance. The parameter spaces that give correct relic density and are also allowed from DD experiments for the democratic choice of all equal coupling strengths $f_u = f_c = f_t$ are practically forbidden by the preceding flavor restriction (Equation 4.15 in the previous chapter). In contrast, in the previous chapter, we set $f_u = f_c$, which remains tiny, but f_t was free to take any large value and showed that such a combination could yield the correct relic density and is enabled by this flavor constraint. *The flavor constraint requires either or both lighter flavor couplings (f_u, f_c) tiny to be allowed.* Instead of making two of these three couplings negligibly small, another interesting scenario emerges if we choose one of f_u or f_c vanishingly small (or zero) while the other remains democratically as large as f_t. We set $f_c = 0$ and $f_u = f_t$ such that all parameter spaces that generate correct relic density are concurrently allowed by the flavor constraint. The direct detection experiment yields more limited parameter spaces with this arrangement since the nucleon comprises the light quarks and the gluon, $f_u \neq 0$ results in a tree-level direct detection scattering diagram, $S_1 u(\bar{u}) \to S_1 u(\bar{u})$, via VLQ exchange (see direct detection diagrams E.2).

The present study investigates the reach of this compelling parameter space at the 14 TeV LHC. We especially employ next-to-leading order (NLO) correction for VLQ pair production for precise computation and match the NLO fixed-order result to the Pythia8 parton shower. The partonic leading-order (LO) cross section has the order $\sigma(pp \to \Psi\bar{\Psi})_{\text{LO}} = \mathcal{O}(\alpha_S^2) + \mathcal{O}(f^4) + \mathcal{O}(f^2\alpha_S)$. Although the dominant contribution comes from the pure QCD sector ($\mathcal{O}(\alpha_S^2)$), we also keep subleading terms ($\mathcal{O}(f^4)$, $\mathcal{O}(f^2\alpha_S)$) at LO for more accurate results. The integrated NLO-K factor for pure QCD couplings is around 1.3, which means 30% enhancement over the LO cross section. We also observe that the differential distributions change significantly, and the theoretical scale uncertainties reduce considerably.

Another interesting point is that the scalar DM parameter spaces that provide the correct relic density while simultaneously being allowed by direct detection constraints differ dramatically. In previous chapter, for example, when the mass of the scalar DM is more than the mass of the top quark (m_t), DM annihilates into $t\bar{t}$ through VLQ exchange t-channel, giving the correct relic density when $f_t \sim 1$ while other two couplings f_u, f_c are tiny. A tiny f_u is required since it has to be allowed from the direct detection experimental constraint. In the present chapter, when the DM is heavier than the top quark, DM annihilation into $t\bar{u}$, $\bar{t}u$ contributes the most in relic density, followed by the annihilation into $t\bar{t}$ final state. Likewise, allowed parameter space can neither support arbitrarily large coupling f ($= f_u = f_t$) from direct detection nor the too-small value of it to obtain correct relic density. Therefore, their interplay remains vital for selecting the available parameter spaces. Contrary to the previous chapter, only a minuscule model space is left when the mass difference between the scalar DM and the VLQ ($\Delta M_{\Psi S_1}$) is smaller than m_t since DD constraints prohibit such spaces due to having a high f value, although having the correct relic density.

The majority of parameter spaces are present when $\Delta M_{\Psi S_1} > m_t$. After the pair production of colored VLQ at the LHC, each VLQ can decay into an on-shell top quark associated with the scalar particle. The branching ratio $\text{BR}(\Psi \to tS) < 0.5$ and counts on the coupling f, in contrast to the previous chapter where

VLQ entirely decays into the top quark while kinematically allowed. Our signal comprises two boosted top-like fatjets and missing transverse energy (MET). We consider all the SM background processes that mimic the signal. We do a sophisticated multivariate analysis of two top fatjets plus a significant MET signal using the Boosted Decision Tree (BDT) algorithm. The available higher-order QCD cross section is used to normalize all the background processes. The parameter spaces of this model are shown to be well within the scope of the 14 TeV LHC with 139 fb^{-1} luminosity.

The extended KSVZ model and its constraints were presented in the previous chapter. The current chapter is organized as follows. The dark matter phenomenology for democratic choice ($f_u = f_t,\ f_c = 0$) is presented in Section 5.2. Section 5.3 demonstrates the impact of NLO+PS calculations, the differential k-factor, and the scale uncertainty of NLO+PS compared to LO+PS. Section 5.4 displays our collider analysis technique using relevant high-level observables, including jet substructure variables, with multivariate analysis (MVA). Finally, we summarize our findings in Section 5.5.

5.2 Dark Matter Phenomenology

This section examines the dark matter phenomenology due to the scalar component. Before we get into the details, let's look at the relevant free parameters. Since axion couplings with scalar DM or the SM particles are inversely proportional to F_a, such couplings are severely suppressed and have practically no role in scalar DM phenomenology. The radial excitation of the field η is decoupled from the rest of the particle spectrum of the model because of its enormous mass, $M_\sigma \sim F_a$ for $\lambda_\eta \sim 1$. As a result, σ does not have any impact on DM phenomenology. The relevant parameters are $\{M_\Psi, M_{S_1}, \Delta M, f\}$. As previously stated, one way to bypass the prohibitory flavor constraint is by setting $f_c = 0$.

Relic density of DM: To estimate the component of relic density offered by scalar DM, we solve the Boltzmann equation using `micrOMEGAs` -v5 [27]. We first construct our model in Feynrules [26]. The variation of the scalar DM relic density with its mass is displayed in Figure 5.1 while we fix $F_a = 10^{11}$ GeV, $\lambda_{SH} = 0.01$, $f_c = 0$, and $\Delta M = 100$ GeV. We present three solid lines for three distinct values of democratic coupling $f = 0.1,\ 0.5$, and 1.0 for the 500 GeV mass of the VLQ. In these variation curves, the first sharp dip ensues due to the Higgs resonance, in which pair of DM annihilate into the SM particles through the resonant Higgs boson when $M_{S_1} \sim \frac{m_h}{2}$, while the second dip occurs when $M_{S_1} \sim M_W$, in which pair of S_1 annihilate into a W boson pair through s-channel Higgs-mediated

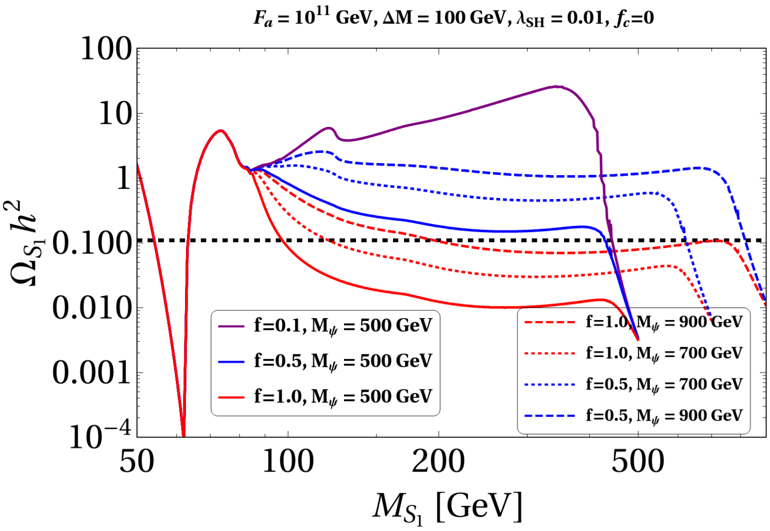


Figure 5.1: Variations of the scalar DM relic density with its mass (M_{S_1}) for different values of f and M_Ψ are shown. Here, we fix $F_a = 10^{11}$ GeV, $\lambda_{SH} = 0.01$, $f_c = 0$, and $\Delta M = 100$ GeV. The Black dashed line corresponds to $0.120 - \Omega_a h^2$.

diagram, see Figure E.1.

For $f = 0.1$ (solid purple line), after the second dip, the relic density increases along with the increase in the DM mass, and a third dip is observed at $M_{S_1} = m_h$. The pair of S_1 begin to annihilate into the Higgs bosons via contact interaction, Higgs-mediated s-channel, and S_1-mediated t-channel diagrams (Figure E.1) and produce the third dip. Ultimately, when the mass difference between VLQ and DM becomes smaller, the impact of DM co-annihilation with the VLQ and annihilation of the VLQ pair into gluons becomes apparent, and a final decline in DM relic density is observed. Further increasing f (0.5 with solid blue line and 1.0 with solid red line) reveals that relic density declines just after the second dip due to the significant contribution of $S_1 S_1 \to t\bar{u}$, $\bar{t}u$ annihilation channels via the VLQ-exchange t-channel processes. The correct relic density is achieved for $f = 1.0$ when DM mass is around 96 GeV.

Blue (red) dotted and dashed lines correspond to the same values of f as in solid lines, except with a heavier choice of mediator Ψ. Because the annihilation cross section decreases as propagator mass increases, and relic density is inversely proportional to the annihilation cross section, the dotted and dashed lines move to higher relic density than the solid line. One clearly follows from these variations that significant parameter space for heavier dark matter masses can open up for different choices of these parameters (over and above the typical Higgs portal). Interestingly, in the case of a pure scalar singlet DM scenario, the DM does not satisfy the correct relic density for $\lambda_{SH} = 0.01$. However, the interaction of the DM with the SM top quark in the present model affords many parameter spaces

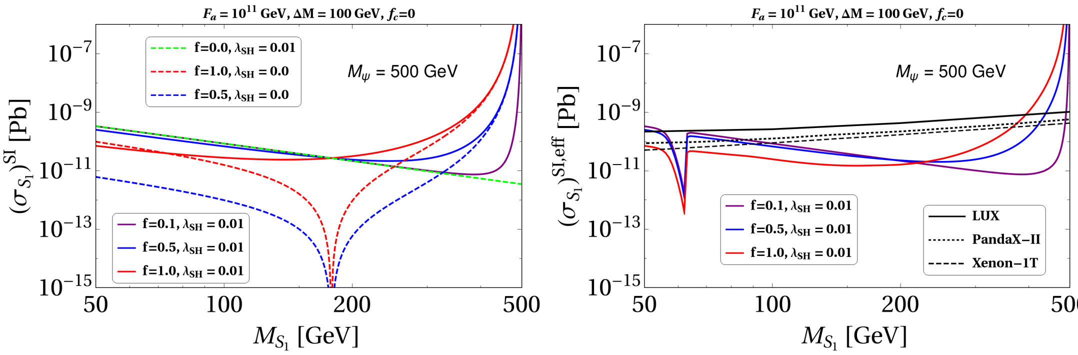


Figure 5.2: **Left panel:** Spin-independent cross section of scattering between scalar DM and nucleon as a function of dark matter mass M_{S_1}. We set $f = 0, \lambda_{SH} = 0.01$ (green dashed line), and $\lambda_{SH} = 0$ for the blue-dashed ($f = 0.5$) and red-dashed ($f = 1.0$) lines to illustrate the individual contributions. **Right panel:** Effective spin-independent scattering cross section (Equation 5.1) vs dark matter mass. All of the plots on the left and right are for $M_\Psi = 500$ GeV. The solid colors purple, blue, and red in both panels stand for $f = 0.1,\ 0.5$, and 1.0, respectively, and $\lambda_{SH} = 0.01$ for all solid lines.

that satisfy the Planck limit.

Direct and indirect detection of DM: WIMPs may also scatter off nuclei, depositing energy that can be detected by detectors like LUX [165], PandaX-II [166,167], and XEXON1T [168]. These experiments can set strong constraints on the scattering cross section and DM mass. All Direct detection channels and the square amplitude of these diagrams are shown in Appendix E. For demonstration purposes, we present a spin-independent direct detection cross section of S_1 with its mass shown in the left panel of Figure 5.2. All solid lines correspond to $\lambda_{SH} = 0.01$ but for different values of $f = 0.1$ (solid purple), 0.5 (solid blue), and 1.0 (solid red). Because f and λ_{SH} are both non-zero, the Higgs-mediated and VLQ-mediated channels and their interference diagrams contribute. It is instructive to note how individual channel contributes. One can first set $f = 0,\ \lambda_{SH} = 0.01$ (dashed green line) so only the Higgs-mediated channel contributes. Subsequently, setting $\lambda_{SH} = 0$ for two choices of $f = 1.0$ (0.5) in the dashed-red (blue) line demonstrates the contribution from pure VLQ-mediated s and t-channels and their interference diagram. The Higgs-mediated diagram does not contribute here.

The amplitude square of the Higgs-mediated diagram does not rely on the mass of S_1 (Equation E.5); nevertheless, the cross section of the dashed green line decreases with the DM mass, which comes from the phase space part of the integral. We see dashed red and blue lines strongly depend on the M_{S_1} since the amplitude square of the VLQ mediated s and t- channels and their interference explicitly depends on M_{S_1} (see Equations E.1- E.4), and the cross section is minimum when $M_{S_1} = \dfrac{M_\Psi}{2\sqrt{2}}$. When comparing dashed-green (only Higgs-mediated channel contributes), dashed-red (only VLQ-mediated channels contribute), and solid-red (total cross section) lines, one can witness a substantial negative (positive) interference between Higgs and VLQ-mediated diagrams when $M_{S_1} < \dfrac{M_\Psi}{2\sqrt{2}}$ ($M_{S_1} > \dfrac{M_\Psi}{2\sqrt{2}}$). Finally, when DM mass is large, we see a sharp rise because of the on-shell production of VLQ (Figure E.2a).

In a two-component DM scenario, the direct detection cross section of the scalar DM should be rescaled as

$$\sigma^{\mathrm{SI,eff}}_{S_1} = \left(\frac{\Omega_{S_1} h^2}{\Omega_{\mathrm{T}} h^2}\right) \sigma^{\mathrm{SI}}_{S_1}, \tag{5.1}$$

where $\Omega_{\mathrm{T}} h^2$ is the total relic density, comprising the scalar and axion DM. The spin-independent effective direct detection cross section for three distinct values of $f = 0.1,\ 0.5$, and 1.0 are presented in the right panel of Figure 5.2 (same as solid lines in the left panel). The black lines show the experimental upper bounds. A dip around $M_{S_1} \sim \frac{m_h}{2}$ is found because of rescaling, as given in Equation 5.1. Here, one notices that the DD experiments disallow the region with a considerable mass difference between the VLQ and DM for significantly large f values. For example, the regions when $M_{S_1} > 390$ GeV for $M_\Psi = 500$ GeV and $f = 1$ (solid red line) are disallowed by DD. Additionally, Figures 5.1 and 5.2, for $f = 1$, suggest that the parameter spaces that offer correct relic density are likewise allowed from the DD.

WIMPs may self-annihilate, emitting a significant amount of gamma and cosmic rays while looking at the dense DM regions at the galactic center. Indirect detection experiments like PAMELA [46], Fermi-LAT [48], MAGIC [47] etc. can constrain model parameter spaces substantially. Because this is a two-component DM scenario, the scalar DM's indirect detection cross section should also be rescaled as,

$$\sigma^{ID}_{S_1,\mathrm{eff}} = \left(\frac{\Omega_{S_1} h^2}{\Omega_{\mathrm{T}} h^2}\right)^2 \sigma^{ID}_{S_1}\ . \tag{5.2}$$

We note that most indirect detection experiments enable our model because the necessity for the axion keeps S_1 under abundance, which lowers the indirect detection constraints since it depends on the fractional scalar DM relic density squared. For very small $\Delta M = (M_{S_2} - M_{S_1})$, S_1 and S_2 are nearly degenerate and can open an additional channel where S_1 and S_2 co-annihilate into top anti-top pair via VLQ and contribute to the indirect detection cross section, which may be disallowed by anti-proton cosmic ray data [186], so we set $\Delta M = 100$ GeV throughout our analysis, ensuring that no co-annihilation channel exists.

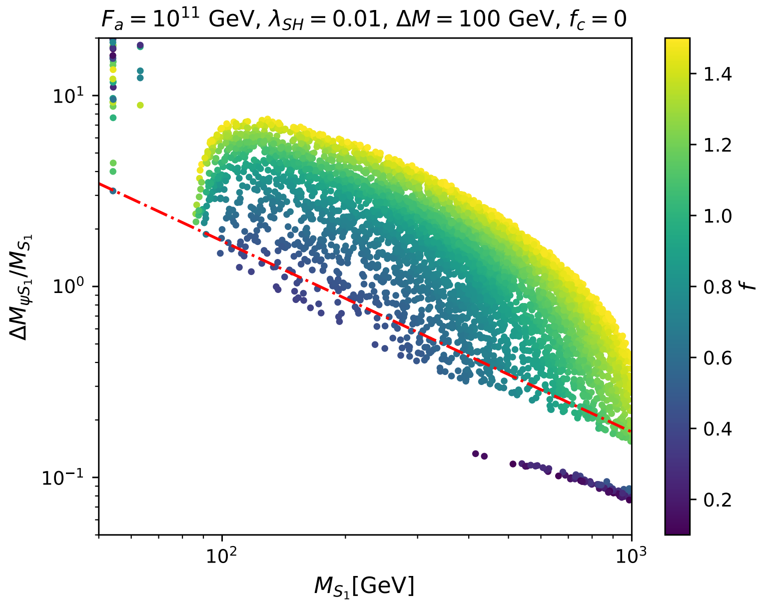


Figure 5.3: On the $\frac{\Delta M_{\Psi S_1}}{M_{S_1}} - M_{S_1}$ plane, the parameter spaces that satisfy the measured DM abundance, permitted by the direct search experiments, and comply with other restrictions as stated in the text are displayed. The color coding is done with respect to f, with f varying from 0.1 to 1.5. Here, we fix $F_a = 10^{11}$ GeV, $\lambda_{SH} = 0.01$, $\Delta M = 100$ GeV, and $f_c = 0$.

	M_{S_1} (GeV)	$\Delta M_{\Psi S_1}$ (GeV)	ΔM (GeV)	f	$\Omega_{S_1} h^2$	$\sigma^{SI}_{S_1,eff}$ (pb)	Processes (percentage)	BR($\Psi \to t S_{1,2}$)
BP1	332	500	100	0.83	0.109	7.98×10^{-12}	$S_1 S_1 \to t\bar{u}, \bar{t}u$ (60%) $S_1 S_1 \to t\bar{t}$ (40%)	0.4907
BP2	402	407	100	0.82	0.109	7.65×10^{-12}	$S_1 S_1 \to t\bar{u}, \bar{t}u$ (56%) $S_1 S_1 \to t\bar{t}$ (44%)	0.4875
BP3	450	300	100	0.79	0.107	1.14×10^{-11}	$S_1 S_1 \to t\bar{u}, \bar{t}u$ (54%) $S_1 S_1 \to t\bar{t}$ (45%)	0.435

Table 5.1: A few representative benchmark points (BPs) from the scan plot are presented; these BPs satisfy the correct relic density and are permissible under all constraints. $\Omega_{S_1} h^2$ and $\sigma^{SI}_{S_1,eff}$ (Equation 5.1) are the relic density and the effective direct detection cross section of the scalar DM, S_1, respectively. $\Delta M_{\Psi S_1} = M_\Psi - M_{S_1}$ and $\Delta M = M_{S_2} - M_{S_1}$. Other parameters are $F_a = 10^{11}$ GeV, $\lambda_{SH} = 0.01$, and $f_c = 0$. The second last column shows the different processes that contribute to the relic density with the percentage contributions in the bracket. The branching fraction of VLQ decays into the top quark associated with the scalar is shown in the last column.

Parameter scan and benchmark points: To demonstrate the relevant parameter space that offers correct relic abundance while being allowed by direct detection and all other constraints as specified in the last section, we identify the three most important parameters, that is, the masses M_{S_1}, $\Delta M_{\Psi S_1}$ and the

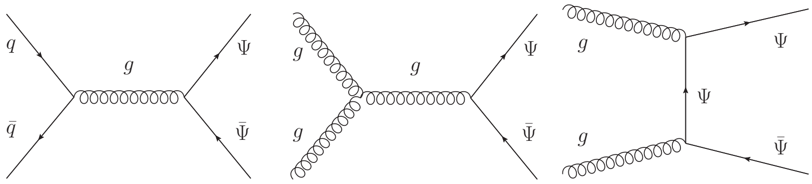
BP	$\sigma^{\rm ID}_{S_1}$ (cm^3/s)	$S_1S_1 \to t\bar{u}, \bar{t}u$ (in %)	$S_1S_1 \to t\bar{t}$ (in %)	$\sigma^{ID}_{S_1,{\rm eff}}$ (in $t\bar{t}$) (cm^3/s)	σ^{ID}_{exp} (in $t\bar{t}$) (cm^3/s)
BP1	2.30×10^{-26}	59.8	40.0	7.59×10^{-27}	2.37×10^{-26}
BP2	2.36×10^{-26}	56.0	43.7	8.51×10^{-27}	2.54×10^{-26}
BP3	2.43×10^{-26}	54.4	45.4	8.77×10^{-27}	2.60×10^{-26}

Table 5.2: Total indirect detection cross section, $\sigma^{\rm ID}_{S_1}$ and the percentage of the different processes that contribute to indirect detection are presented in the third and fourth columns. Effective indirect detection cross section in the $t\bar{t}$ final state for different benchmark points is shown in the fifth column, which is defined as (% contribution in $t\bar{t}$ final state) $\times\ \sigma^{ID}_{S_1,{\rm eff}}$, where $\sigma^{ID}_{S_1,{\rm eff}}$ is given in Equation 5.2. The last column is the experimental upper bound in the $t\bar{t}$ final state [186].

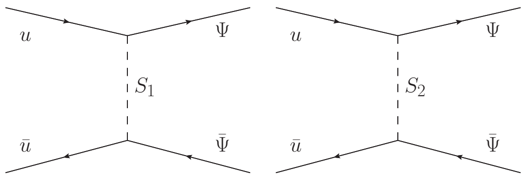
Yukawa coupling f. Figure 5.3 display such points on the plane of M_{S_1} vs $\frac{\Delta M_{\Psi S_1}}{M_{S_1}}$, while the Yukawa coupling f is color coded, with f ranging from 0.1 to 1.5. The red dash-dot line corresponds to $\Delta M_{\Psi S_1} = m_t$. Hence, the upper portions of this line can be investigated at the LHC with top quark on-shell production as VLQ decays into a top quark and invisible DM, while the lower area can be probed with jets + MET as VLQ decays into an u-quark associated with DM. Points along two vertical lines at the top left region correspond to the part satisfied by Higgs resonance.

It is enlightening to note that the lower sections of the plot, which correspond to the small mediator mass M_Ψ, typically generate an increased DD cross section; therefore, those regions are excluded from the DD bounds despite having the correct relic density. Only a few points exist at the lower right corner when f is tiny. Those regions have correct relic density because being nearly degenerate, the VLQ and DM co-annihilate and the pair of VLQ annihilates into gluons. Note that for larger f values, those co-annihilation regions are ruled out by the DD experiments, as we already see in Figure 5.2 (right panel, red, blue lines). Interestingly, non-perturbative effects like Sommerfeld enhancement and bound state formation can significantly affect relic density in those co-annihilation regions. Further study of this region is beyond the scope of the present discussion. Such points are challenging to probe at the LHC as the DM mass is quite large and VLQ is degenerate to the DM, so the partonic cross section of VLQ production will be small, and VLQ will emit a soft jet that is very difficult to detect.

A few representative benchmark points (BPs) from the scan plot are listed in Table 5.1, which are allowed from all the constraints and provide correct relic density. The scalar DM relic density, spin-independent DD scattering cross section of S_1, the percentage contribution of each process to the relic density, and the branching ratio of VLQ decay into the top quark are also given. Table 5.2



(a) $\mathcal{M}^a_{\text{LO}} = \mathcal{O}(\alpha_S)$



(b) $\mathcal{M}^b_{\text{LO}} = \mathcal{O}(f^2)$

Figure 5.4: The Feynman diagrams for the pair production of VLQ at Leading order.

shows the total cross section of indirect detection (ID) and the percentage contribution of the various processes to the indirect detection. The theoretical ID cross section in the final state of $t\bar{t}$ and the experimental upper bound are given in the last two columns, where we find that all of those BPs are well inside the experimental upper bound.

5.3 Pair production of vector-like quark at NLO+PS accuracy

We implement the model Lagrangian in FeynRules [26] and employ the NLOCT [142] package to generate UV and R_2 counter terms of the virtual contribution in NLO UFO model that we finally use under the MadGraph5_aMC@NLO [69] environment. Inside this, the real corrections are performed using the FKS subtraction method [143, 209], whereas the OPP technique [144] takes care of the virtual contributions. Showering of the events is done using Pythia8 [146, 147]. For leading order (LO) and next-to-leading order (NLO) event generation, we use NN23LO and NN23NLO PDF sets, respectively.

All the tree-level diagrams in the pair production of VLQ at the LHC are shown in Figure 5.4. The three Feynman diagrams in the upper row depend only on the QCD coupling and are the dominant production channels. The bottom two diagrams depend on the BSM Yukawa coupling, f. The LO cross section has the order $\sigma_{\text{LO}} = \mathcal{O}(\alpha_S^2) + \mathcal{O}(f^4) + \mathcal{O}(f^2\alpha_S)$. The term $\mathcal{O}(f^2\alpha_S)$ comes from the interference between the bottom two Feynman diagrams and the subset of the first diagram in the upper row. It is important to note that the gluon-initiated diagrams do not interfere with the bottom two diagrams.

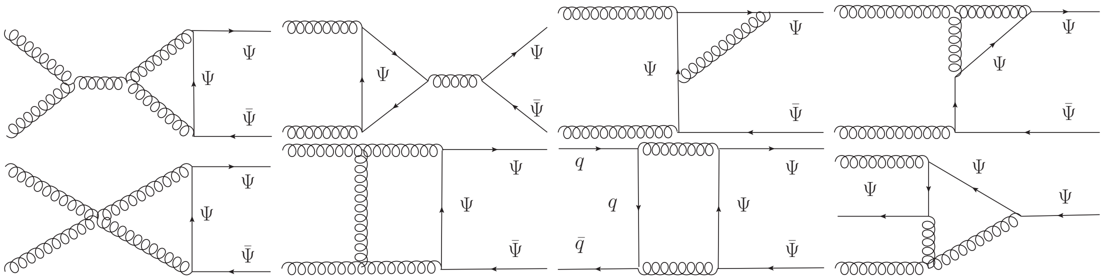


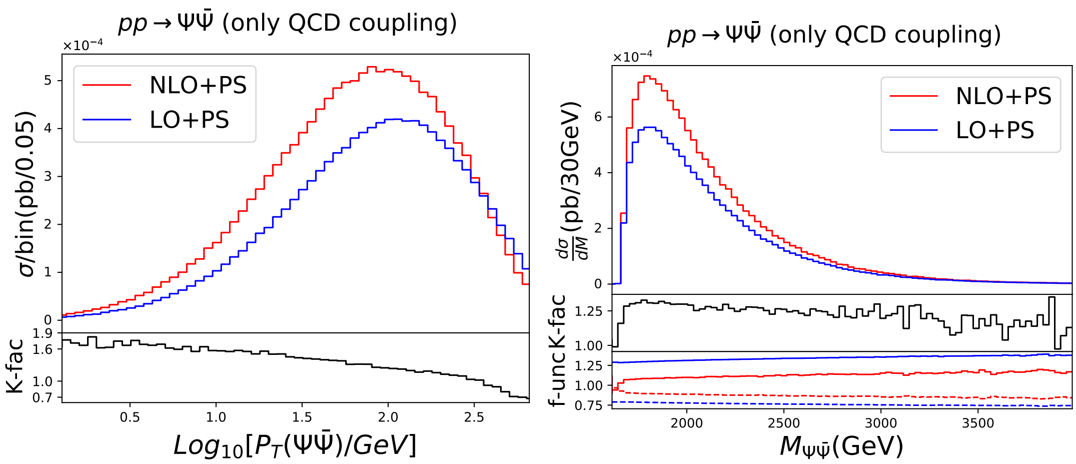
Figure 5.5: Representative Feynman diagrams for the pair production of VLQ at NLO-QCD for the processes where the tree-level diagrams only have QCD coupling (Figure 5.4a). $\sigma^a_{\rm NLO} \propto \mathcal{M}^\dagger_V \mathcal{M}^a_{\rm LO} = \mathcal{O}(\alpha_S^3)$.

BP	$\sigma(pp \to \Psi\bar{\Psi})$ (fb) LO $\sigma_{\rm LO} = \mathcal{O}(\alpha_S^2) + \mathcal{O}(f^4) + \mathcal{O}(f^2\alpha_S)$	$\sigma(pp \to \Psi\bar{\Psi})$ (fb) for leading production processes at LO and NLO		
		LO, $\mathcal{O}(\alpha_S^2)$	NLO, $\mathcal{O}(\alpha_S^3)$	K-fac
BP1	$96.39^{+31.5\%}_{-22.5\%}$	$105.8^{+31.3\%}_{-22.2\%}$	$138.5^{+9.6\%}_{-11.3\%}$	1.31
BP2	$114.0^{+31.9\%}_{-22.5\%}$	$125.7^{+31.4\%}_{-22.4\%}$	$162.1^{+10.1\%}_{-11.5\%}$	1.29
BP3	$181.5^{+32.1\%}_{-22.7\%}$	$201.6^{+31.3\%}_{-22.3\%}$	$257.3^{+9.8\%}_{-11.4\%}$	1.28

Table 5.3: Total leading-order cross section, including QCD and BSM coupling, and their interference in the pair production of VLQ at 14 TeV LHC before their decay is given in the left panel. Right panel: Leading contribution of the tree-level VLQ pair production process ($\mathcal{O}(\alpha_S^2)$) and its next-to-leading order cross section, along with the integrated K-factor, are given. The superscript and subscript denote the scale uncertainties (in percentage) of the total cross section. Five massless quark flavors are used for computation.

The leading production channels (three diagrams at the top row) form a gauge invariant subset, and we do one-loop QCD correction of those processes. As a result, the NLO cross section has the order $\mathcal{O}(\alpha_S^3)$. Few representative Feynman diagrams at NLO-QCD are shown in Figure 5.5. The total LO cross section is given in the left panel of Table 5.3. The leading contribution of the VLQ pair production at the tree level and its next-to-leading order cross section, along with the integrated K-factor, are given in the right panel of Table 5.3. The K-factor is defined as the ratio of NLO to LO cross section. We find a significant enhancement of about 30% in the NLO-QCD cross section over LO. Table 5.3 shows that the interference term has a small negative contribution to the total LO cross section. The vertices $u\Psi S_1$ and $u\Psi S_2$ differ by a i-factor due to the definition of the scalar in Equation 4.3. Therefore a relative minus sign is present when the bottom two diagrams interfere with the first diagram in the top row.

We designate the partonic center-of-mass energy of the event as the central choice for both the factorization and renormalization scales. To compute the scale variance, we vary both the factorization and renormalization scales from a



(a) (b)

Figure 5.6: (a) Distribution of $\log_{10}[P_T(\Psi\bar{\Psi})/GeV]$ at LO+PS and NLO+PS, and the differential K-factor for VLQ pair production at the LHC. (b) The distribution of invariant mass of the VLQ pair is shown in the upper panel, and the differential K-factor and the scale uncertainties are shown in the middle and bottom panels, respectively. The plots correspond to BP1, and LO consists only of QCD coupling.

factor of two to half of this central scale, resulting in nine different data sets. The superscripts and subscripts in the tables indicate the envelopes of the nine data sets, although all of the cross sections shown in Table 5.3 correspond to the central scale. At NLO, the uncertainties associated with the renormalization and factorization scales are significantly smaller than at LO.

LO+PS and NLO+PS distributions of $\log_{10}[P_T(\Psi\bar{\Psi})/GeV]$ (upper panel) and the differential K-factor (lower panel) are given in Figure 5.6a. $P_T(\Psi\bar{\Psi})$ is the transverse momentum of the VLQ pair. For $\log_{10}[P_T(\Psi\bar{\Psi})/GeV] < 1.6$, the left plot shows that K-factor is more than 1, but for $\log_{10}[P_T(\Psi\bar{\Psi})/GeV] > 1.6$, the K-factor is less than 1, indicating that the NLO cross section is less than the LO cross section. The differential K-factor is not flat everywhere. It is almost flat at the lower values and then starts to go down, so scaling the LO events by a constant K-factor would not give accurate results.

The invariant mass distribution of the VLQ pair is shown on the top panels of Figure 5.6b for BP1. The differential K-factor is shown in the middle panel. Invariant mass distribution peaks around 1800 GeV, and the differential K-factor is almost flat around the peak. The bottom panel shows the envelope of the factorization and renormalization scale uncertainties. The red solid and dashed lines show the width of the scale uncertainty for NLO+PS, while the blue solid and dashed lines show the LO+PS scale uncertainties. We can see that both LO+PS and NLO+PS results are stable, but the NLO+PS result has much-reduced scale uncertainty. Although these are figurative findings because the decay of Ψ is not considered, they demonstrate the need to do $\mathcal{O}(\alpha_s)$ corrections on the pair production channels for more accurate prediction of the total cross section, differential distribution of various variables, and lower scale uncertainty.

5.4 Multivariate Analysis (MVA)

For completeness, we further carried out the collider analysis on this model using the $t\bar{t}$+MET final state, as in our previous chapter. However, we use the NLO+PS accurate events in this chapter to generate the signal. Moreover, the previous analysis assumed the BR($\Psi \to tS_{1,2}$) = 1.0, based on the top-philic coupling, which is invalid for a democratic coupling with other generations. Hence, based on calculated BR($\Psi \to tS_{1,2}$) < 0.5 (Table 5.1) for all the BPs when the pair of Ψ decay into the top quarks, the cross section is reduced by a factor of at least 0.25 than the previous. Representative benchmark points in Table 5.1 are allowed by 139 fb^{-1} projected exclusion contour of recent ATLAS analysis [176] of the stop pair production. In contrast to the prior leading-order estimate, the QCD corrections in this chapter significantly increase the overall cross section by around 30% over LO, and the factorization and renormalization scale uncertainties reduce considerably.

The signal topology is given by [a],

$$pp \to \Psi\bar{\Psi}\,[\mathrm{QCD}] \to (t\,S_{1,2})(\bar{t}\,S_{1,2})\,j \Rightarrow 2J_t + \not{E}_T + X \tag{5.3}$$

The advantages of studying the hadronic final state are as follows. The large hadronic branching ratio and significant mass difference between VLQ and the scalar significantly boost the top quark. As a result, reconstructing the top quark as a boosted fatjet additionally provides the inherent properties of the jet. Furthermore, we can have additional handel using jet substructure variables.

The expected number of signal (BP1) and background events (in fb, expected event numbers are obtained by multiplying them with the luminosity) is listed as cut flow, along with the cut efficiencies, after each set of event selection criteria is shown in Table 5.4. In preselection cut (C1) we demand at least two fatjets of radius $R = 1.5$, each with a transverse momentum $P_T(J_0), P_T(J_1) > 200$ GeV, missing transverse momentum $\not{E}_T > 100$ GeV, a lepton-veto, and $|\Delta\Phi(J_{0,1}, \not{E}_T)| > 0.2$ (to minimize jet mismeasurement contribution to $\not{E}_T$). The other cuts are: (C2)

[a] Since we are considering two top-like fatjet (J_t) without measuring the jet charge, $uu \to \Psi\Psi$ (through the t-channel scalar mediator) can also contribute to the same signature followed by the decay of the VLQ into the top. Interestingly, since scalars and Ψ have the same PQ charge, this type of t-channels exchange is impossible unless PQ symmetry is spontaneously broken finally to contribute negligibly. To give some perspective in our benchmark point BP1, we find $\sigma_{\mathrm{LO}}(uu \to \Psi\Psi) = 0.3$ fb, and $\sigma_{\mathrm{LO}}(\bar{u}\bar{u} \to \bar{\Psi}\bar{\Psi}) = 0$, so we safely ignore those processes.

	Signal (BP1) (fb)	Z**+jets** (fb)	W**+jets** (fb)	$t\bar{t}$**+jets** (fb)	tW**+jets** (fb)	WZ**+j** (fb)	WW**+j** (fb)	ZZ**+j** (fb)	$t\bar{t}\ V$ (fb)	**tot BG** (fb)
C1	5.99 [100%]	2517.99 [100%]	1366.91 [100%]	690.65 [100%]	366.91 [100%]	93.53 [100%]	25.90 [100%]	11.51 [100%]	8.34 [100%]	5081.74 [100%]
C2	5.49 [91.65%]	1640.29 [65.14%]	762.59 [55.79%]	302.16 [43.75%]	152.52 [41.57%]	58.35 [62.39%]	11.51 [44.44%]	6.973 [60.58%]	6.17 [73.98%]	2940.56 [57.87%]
C3	4.58 [76.46%]	241.73 [9.60%]	117.99 [8.63%]	230.94 [33.44%]	114.39 [31.18%]	10.79 [11.54%]	2.45 [9.46%]	1.92 [16.69%]	5.11 [61.27%]	725.32 [14.27%]
C4	2.23 [37.23%]	25.38 [1.01%]	17.33 [1.27%]	64.23 [9.30%]	27.45 [7.48%]	1.24 [1.33%]	0.33 [1.27%]	0.2 [1.74%]	2.30 [28.13%]	138.46 [2.72%]

Table 5.4: After applying various kinematic event selection cuts, signal and background events (in fb) indicate the efficiency for each set of cuts to reduce the backgrounds. The kinematic cuts (C1-C4) are described in the text. After applying the C4 cut, the remaining events are passed for the multivariate analysis.

	$\not{E}_T$	$\Delta\phi(J_1, \not{E}_T)$	$\Delta R(J_0, J_1)$	$\tau_{32}(J_1)$	$\tau_{32}(J_0)$	$\Delta\phi(J_0, \not{E}_T)$	$\mathrm{M_{eff}}$	$M(J_1)$	$M(J_0)$	$\tau_{31}(J_0)$	$\tau_{31}(J_1)$
BP1	31.29	20.19	17.82	8.61	8.49	8.38	3.29	2.10	1.48	1.10	0.9
BP2	19.39	16.74	17.39	6.75	6.99	8.04	2.26	0.67	0.72	1.11	0.73
BP3	9.25	11.30	12.11	6.52	5.74	6.55	1.29	0.95	0.38	0.52	0.72

Table 5.5: Method unspecific relative separation power of different kinematic variables in separating the signal and background classes.

$\not{E}_T > 150$ GeV, (C3) a b-tag within the leading or subleading fatjet, and (C4) pruned mass of the two leading jets $M_{J_0}, M_{J_1} > 120$ GeV. After applying the preselection cut (C1), we find V+jets ($V = Z, W$) are the principal background while $t\bar{t}$+jets is the sub-dominant background. However, after a b-tag within J_0 or J_1 and demanding large fatjet masses, we found $t\bar{t}$+jets becomes the primary background, while V+jets are the sub-dominant. Applying all those cuts, we still retain a substantial number of signal events while the background reduces significantly. All the signal and background processes are passed through all these event selection criteria up to C4 before passing events to MVA. We create two separate signal and background classes. The combined background is the weighted combination of all the different background processes. Each signal and background class is randomly divided into 50% for training and the rest 50% for testing. We use boosted decision tree (BDT) algorithm and choose a set of kinematic variables from a wider collection of variables for MVA. The variables with high relative importance distinguishing the signal class from the background class are preferable. Table 5.5 lists the relative importance of the various kinematic variables involved in the MVA. The left (signal) and right (background) tables of Figure 5.7 show the linear correlation coefficients among the variables employed in MVA for BP1.

In the previous chapter, we provide the normalized distributions of all background processes after performing all event selections up to C4. We avoid demonstrating these distributions since the shapes are qualitatively similar for physics

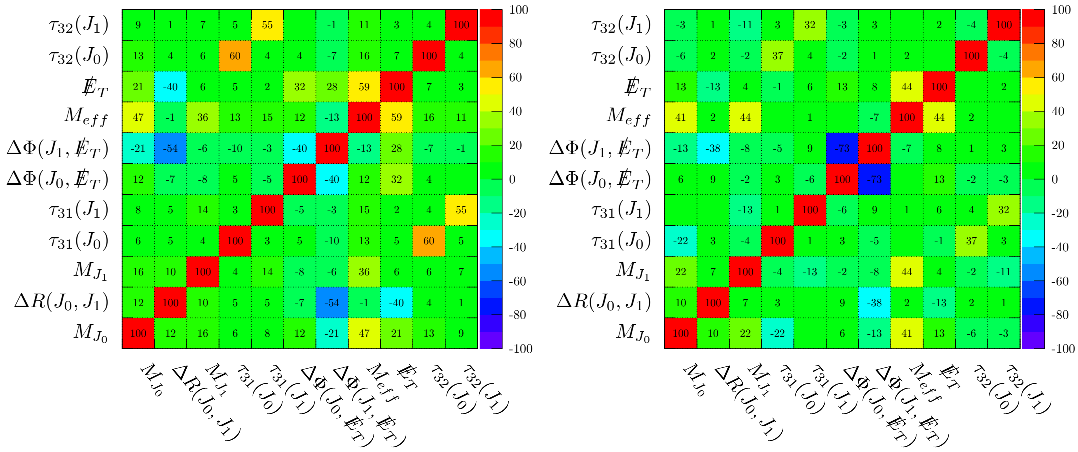


Figure 5.7: Coefficients of linear correlation (in percentage) between various kinematical variables for the signal (BP1, left panel) and background (right panel) are presented. Missing entries have an insignificant correlation of less than one. Two variables are correlated or anti-correlated based on positive and negative coefficients.

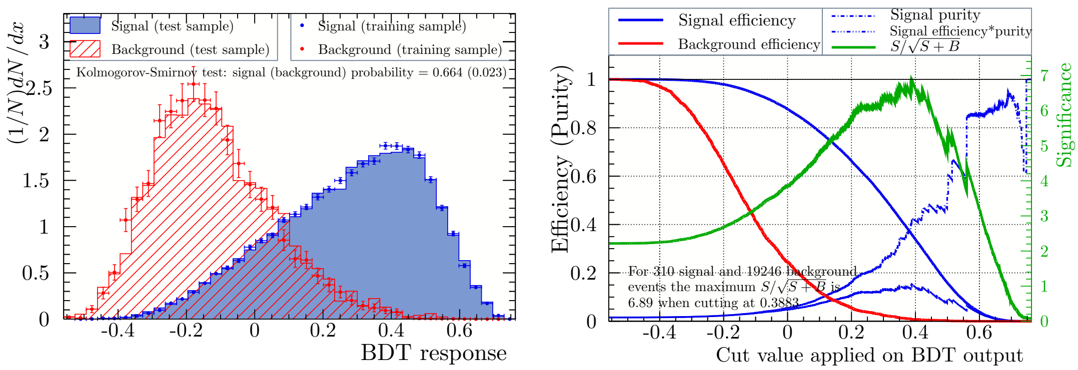


Figure 5.8: The left panel shows the normalized distribution of the BDT output for training and testing samples of both signal and background. The statistical significance of the signal with the cut applied to the BDT output is shown in the right panel, along with signal and background efficiency.

understanding. The normalized distribution of the BDT response for test and train samples of both signal (BP1) and background classes is plotted on the left side of Figure 5.8. We find signal and background are well separated. With the cut applied to the BDT output, the signal and background efficiency, as well as the statistical significance ($\frac{N_S}{\sqrt{N_S+N_B}}$) for 139 fb^{-1} data, are presented in the right plot of Figure 5.8. Before applying any cuts to the BDT output, Table 5.6 shows the number of signals (N_S^{bc}) and background (N_{SM}) events for various BPs. It also shows the expected number of signal events (N_S) and background events

	N_S^{bc} (fb)	BDT_{opt}	N_S (fb)	N_B (fb)	$\frac{N_S}{\sqrt{N_S+N_B}}$ for 139 fb^{-1}	$\frac{N_S}{N_B}$
BP1	2.23	0.3883	0.8012	1.0783	6.89	0.743
BP2	2.53	0.2582	1.2207	6.1302	5.31	0.199
BP3	1.67	0.2961	0.4252	2.3529	3.0	0.180
N_{SM}	138.46					

Table 5.6: The table shows the efficacy of the current search in terms of statistical significance for various benchmarks. N_S^{bc} and N_{SM} are the total numbers of signal and background events before performing MVA (see Table 5.4), while N_S and N_B, respectively, provide those following BDT analysis. BDT_{opt} is the optimal BDT cut. The second-to-last column provides the statistical significance of the signal for 139 fb^{-1} luminosity.

(N_B) that remain after applying an optimal cut (BDT_{opt}) to the BDT output. The last two columns show the statistical significance of the signal at 139 fb^{-1} luminosity and the signal-to-background ratio. We optimize each of the three BPs separately.

Table 5.6 shows that the statistical significance of BP3 is lower than that of the other two benchmark points even though it has the most significant partonic cross section of VLQ pair production since the mass of the VLQ is the smallest for BP3. This is attributed to a smaller mass difference between VLQ and DM than the other two BPs, which results in a relatively less boosted top quark and a smaller signal efficiency. Table 5.6 also demonstrate that a significant parameter space of this model can be explored with more than 5σ significance using the 139 fb^{-1} data at the 14 TeV LHC.

5.5 Conclusions

We explore a complex scalar extended KSVZ axion framework, where the scalar is singlet under the SM gauge groups but only has the Peccei-Quinn charge. This model has the capability to solve two of the most outstanding problems of SM, that is, the strong-CP problem and a natural candidate for dark matter in the form of QCD axion having a lifetime comparable to the age of the Universe. Axion can satisfy the correct dark matter relic density, measured by the Planck collaboration, but at the expense of fine-tuning the corresponding breaking scale. The residual $\mathbb{Z}_2$ symmetry in this model ensures that the lightest component of the complex scalar is stable and thus plays the role of a second dark matter, removing the need for any such fine-tuning.

KSVZ axion framework also provides a rich phenomenology by introducing a vector-like quark which can be explored at a hadron collider like LHC. In the extended scenario, VLQ interacts with the scalar (DM) candidate and the SM quarks (up or down) based on its hypercharge. Hence the VLQ now plays a critical role in dark matter phenomenology because it opens up new annihilation and coannihilation channels.

Here, we explore the possibility of democratic Yukawa interaction of the vector-like quark with all up-type quarks and scalar dark matter candidate. One must find the allowed parameter spaces that provide the correct relic density and agree with other experimental observations such as direct detection (DD), collider data, etc. It is found that the flavor constraint strongly disfavors this democratic option, which requires either one or both lighter flavor couplings (f_u, f_c) needs to be tiny. For simplicity, we consider $f_c = 0$ while keeping the other two democratic. It is interesting to note that the allowed parameter space can neither support arbitrarily large coupling $f(= f_u = f_t)$ from direct detection nor the too-small value of it to obtain the correct relic density. Therefore, their interplay remains vital for selecting the available parameter spaces.

We employ NLO-QCD correction on dominant production channels of colored VLQ pair production at the LHC. The total NLO cross section increases by approximately 30% compared to LO. Additionally, the differential distributions of various observables exhibit significant changes when considering NLO+PS compared to LO+PS. A notable reduction in scale uncertainty is also observed at NLO+PS, leading to a more precise and accurate result than at LO+PS. Following pair production at the LHC, each VLQ undergoes decay into a top quark accompanied by a scalar DM. The VLQ and scalar DM exhibit a substantial mass difference, considerably boosting both top quarks. Boosted top-like fatjets generated from the hadronic decay of top quark still carry different characteristics of it, which are primarily captured in a dedicated jet analysis and different substructure variables. Multivariate analysis with these variables and attributes of event topology is demonstrated to establish a strong ability to explore a significant parameter space of this model at the 14 TeV LHC with 139 fb^{-1} integrated luminosity.

Chapter 6

Precise probing and discrimination of third-generation scalar leptoquarks

6.1 Introduction

In the previous three chapters, we explored several BSM models motivated by dark matter. We conducted a multivariate analysis at the LHC and studied the phenomenology of dark matter within these models. However, this chapter will focus on studying scalar leptoquarks, which are exotic particles present at the LHC. We will investigate the discovery potential of leptoquarks at the LHC, considering NLO QCD corrections. Furthermore, we will employ polarization variables to differentiate between various scalar leptoquark models.

Leptoquark (LQ) is a hypothetical particle that couples to quark and lepton together. It carries both baryon number and lepton number and provides a means to unify quarks and leptons. It can appear in many interesting scenarios beyond the Standard Model, for example, Pati-Salam model [61,62], Grand unified theory [59,60], Composite model [63] *etc.*, and therefore it remains as a very active area in experimental searches. In some of these models, the baryon number gets violated and that allows protons to decay. But the strong constraints from the non-observation of proton decay so far have pushed the masses of the leptoquark to a very high scale, typically around 10^{16} GeV. However, imposing baryon number or lepton number conservation one gets a set of leptoquarks in the Buchmüller-Rückl-Wyler (BRW) framework [211], which allows leptoquark masses to be in a range accessible to the collider searches. Also, such leptoquarks are favorable to explain anomalies observed in the B-meson decays in BaBar [212], Belle [213–215] and LHCb experiments [216, 217].

In this chapter, we focus on the third-generation scalar leptoquarks. Phenomenology of such leptoquarks are studied widely in different channels [218–221] and they are also searched by the ATLAS and CMS collaborations [222–225]. In a recent analysis [226], the ATLAS collaboration did a cut-based analysis and extracted the limit for the up-type third-generation scalar leptoquark, assuming LQ decaying into a top quark and neutrino with a 100% branching ratio. Their analysis put a lower limit of 1240 GeV on the LQ mass at 95% confidence level for an integrated luminosity of 139 fb^{-1} at the 13 TeV LHC. This chapter presents an alternative search strategy considering two top-like fatjets plus significant missing energy in the final state with a sophisticated multivariate analysis of the NLO+PS signal events including jet substructure variables. Given the already constrained parameter space, a relatively heavy leptoquark would naturally produce top quark at the boosted region once produced from its decay. Thus, it is prudent to identify such top quarks as a top-like fatjet from its hadronic decay. Note that the corresponding leptonic decay mode not only suffers from branching ratio suppression, but also identifying such leptons inside a jetty signature is a challenging task and therefore it affects the efficiency significantly. We observe that our result is consistent with the existing search and find that the third-generation LQ can be discovered with a significance of $\geq 5\sigma$ for masses below 1380 GeV with 3000 fb^{-1} data at the 14 TeV LHC (HL-LHC).

After discussing discovery potential, we move to distinguish different scalar leptoquarks of the same electromagnetic charge. Therefore we also analyze distinguishing different scalar leptoquark models based on the same final state signature at the LHC. One proposal has been made to determine different leptoquark types of the same spin and different electromagnetic charges by measuring jet charge [227]. We show that in the context of third-generation up-type leptoquark, measuring the polarization of the top quark resulting from the leptoquark decay can be an efficient way to distinguish scalar leptoquark models of the same electromagnetic charge without requiring the measurement of jet charge. In this chapter, for the first time, we use polarization variables to distinguish two scalar leptoquark models, considering all the backgrounds.

As the top quark decays before it hadronizes, its spin information can be obtained from its decay products[a] [228]. Top quark polarization has been studied for more than last thirty years [229–241]. Determination of the polarization of boosted top-quark is studied in [242]. The possibility of distinguishing two models in the $t\bar{t}\tau\bar{\tau}$ channel was explored before for scalar leptoquark in Reference [243] without signal-to-background study.

We set our probe strategy based on two chosen leptoquark models, S_3 and

[a]Other quarks form bound states before their decay and hence lose their spin information.

R_2. The cross section of pair production of leptoquarks at the LHC in these two models is the same, but the top quark originates from the leptoquark decay in these two models are different. The S_3 model produces a left-chiral top quark, while the R_2 model gives a right-chiral top quark. The polarization variables, like the ratio of the b-jet energy to the reconstructed top fatjet energy, can be used to distinguish two models at 14 TeV LHC and a futuristic 27 TeV collider (HE-LHC).

The rest of the chapter is organized as follows. In Section 6.2, we describe the third-generation scalar leptoquark models. In Section 6.3, we show the effect of NLO calculations. We study the impact of parton shower over the fixed-order (FO) NLO calculation, k-factor variation in differential distributions, and reduction of scale uncertainties at the NLO+PS accuracy. In Section 6.4, we describe our search strategy and provide details on multivariate analysis used to discriminate the signal and the background. In Section 6.5, we discuss how the polarization observables can be instrumental in distinguishing two above mentioned models. Finally, we summarize and conclude in Section 6.6.

6.2 The models

Under the Standard Model (SM) gauge group $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$, there are total six species of scalar leptoquarks, namely S_3, R_2, $\tilde{R}_2$, $\tilde{S}_1$, S_1 and $\bar{S}_1$. Since a quark transforms as a triplet of $SU(3)_c$, a leptoquark should also transform as the same multiplet of $SU(3)_c$ in order to form gauge invariant interaction terms. In Table 6.1, we show the SM quantum numbers of all scalar leptoquarks. The subscripts on the model name denote their $SU(2)_L$ quantum numbers. If two or more models have same $SU(2)_L$ quantum number but different hypercharges, a tilde or bar is used to identify them. The component fields of the electroweak multiplets are written in the third column of the table with superscripts denoting their electric charges. In this chapter we are interested in studying the third generation scalar leptoquarks only. Various decay channels of the component fields for third generation leptoquarks are written inside parentheses.

It is interesting to notice that in only two models fields transform as **3** and for the rest of the models they transform as $\bar{\mathbf{3}}$ under $SU(3)_c$. Let us, for example, consider the S_3 model in which fields transform as $\bar{\mathbf{3}}$. The reason behind this transformation is that the fields in S_3 should couple to a quark doublet Q and lepton doublet L, as it transforms as **3** under $SU(2)_L$[b]. But it can couple to only $\bar{Q^C} L$, not with $\bar{Q}L$, since the latter is zero . As $\bar{Q^C}$ transforms as **3**, S_3 would

[b] $2 \otimes 2 = 3 \oplus 1$

Models	$(SU(3)_c, SU(2)_L, U(1)_Y)$	Components & Decay
S_3	$(\bar{3}, 3, \frac{1}{3})$	$S_3^{\frac{4}{3}}(\tilde{b}, \tau^+)$, $S_3^{\frac{1}{3}}((\tilde{t}, \tau^+), (\tilde{b}, \tilde{\nu}_\tau))$, $S_3^{-\frac{2}{3}}(\tilde{t}, \tilde{\nu}_\tau)$
R_2	$(3, 2, \frac{7}{6})$	$R_2^{\frac{5}{3}}(t, \tau^+)$, $R_2^{\frac{2}{3}}((t, \tilde{\nu}_\tau), (b, \tau^+))$
$\tilde{R}_2$	$(3, 2, \frac{1}{6})$	$\tilde{R}_2^{\frac{2}{3}}((t, \tilde{N}_\tau), (b, \tau^+))$, $\tilde{R}_2^{-\frac{1}{3}}((b, \tilde{\nu}_\tau), (b, \tilde{N}_\tau))$
$\tilde{S}_1$	$(\bar{3}, 1, \frac{4}{3})$	$\tilde{S}_1^{\frac{4}{3}}(\tilde{b}, \tau^+)$
S_1	$(\bar{3}, 1, \frac{1}{3})$	$S_1^{\frac{1}{3}}((\tilde{t}, \tau^+), (\tilde{b}, \tilde{\nu}_\tau), (\tilde{b}, \tilde{N}_\tau))$
$\bar{S}_1$	$(\bar{3}, 1, -\frac{2}{3})$	$\bar{S}_1^{-\frac{2}{3}}(\tilde{t}, \tilde{N}_\tau)$

Table 6.1: All the possible scalar leptoquark models which give gauge invariant terms in the Lagrangian under the SM gauge group transformations. To learn about the naming convention used for the models, see the text.

transform as $\bar{\mathbf{3}}$ [c]. Obviously the conjugate of S_3 transforms as $\mathbf{3}$, however for that lepton doublet precedes the quark doublet in the Lagrangian. Here we label a leptoquark as field (as opposed to the conjugate field) if in the interaction term quark precedes lepton. Transformation properties of the other leptoquarks under $SU(3)_c$ can be understood in a similar way.

One might be interested to probe third generation up-type scalar leptoquark component fields which have $\frac{2}{3}e$ electric charge. There are four such component fields, namely $S_3^{-\frac{2}{3}}$, $R_2^{\frac{2}{3}}$, $\tilde{R}_2^{\frac{2}{3}}$, and $\bar{S}_1^{-\frac{2}{3}}$. At the LHC, the first two fields can give two top fatjets plus missing energy as the signature, whereas the last two, depending on the right handed heavy neutrino decay mechanisms, will give more complicated and model dependent signatures. In this chapter, we are interested to study the phenomenology of the $S_3^{-\frac{2}{3}}$ and $R_2^{\frac{2}{3}}$ fields only.

The kinetic term for the generic scalar Leptoquark (S) can be written as,

$$\mathcal{L}_{\rm kin} = (D_\mu S)^\dagger (D^\mu S) - M_S^2 S^\dagger S \, . \tag{6.1}$$

Here the covariant derivative D_μ is given as,

$$D_\mu = \partial_\mu - i g_s \lambda^a G^a_\mu \, , \tag{6.2}$$

where g_s is the strong coupling, λ^a and G^a $(a = 1, ..., 8)$ denote the Gellman matrices and gluon fields respectively. The above Lagrangian gives rise to the

[c]For R_2 model, fields transform as $\mathbf{2}$ under $SU(2)_L$ and therefore one fermion needs to be doublet, while the other one needs to be singlet. The doublet and singlet are left and right handed respectively. Hence, in this case, interaction with charge conjugated quark field vanishes, while the one without charge conjugation survives. This explains why R_2 transforms as $\mathbf{3}$ under $SU(3)_c$.

following two vertices – (i) gluon-LQ-LQ, (ii) gluon-gluon-LQ-LQ. The Feynman rules for these vertices are independent of the type of leptoquarks.

The quantum numbers of leptoquark S_3 is such that it can allow diquark coupling. However, without baryon or lepton number conservation, for TeV-scale leptoquark, this coupling has to be too tiny as otherwise it would lead to proton decay [d]. As this coupling is too constrained, in our analysis they do not play any role[e]. The interaction terms for the third generation scalar leptoquarks S_3 and R_2 of charge $\frac{2}{3}e$, with a quark and a lepton, are given by [244],

$$\mathcal{L}_{\text{Int}}^{S_3^{\frac{2}{3}}} = y_{S_{LL}} * \bar{t_L^C}\ v_\tau\ S_3^{-\frac{2}{3}} + h.c., \tag{6.3}$$

$$\mathcal{L}_{\text{Int}}^{R_2^{\frac{2}{3}}} = y_{R_{RL}} * \bar{t_R} v_\tau\ R_2^{\frac{2}{3}} + y_{R_{LR}} * \bar{b_L}\tau_R\ R_2^{\frac{2}{3}} + h.c., \tag{6.4}$$

where "RL" in $y_{R_{RL}}$ signifies that the chiralities of the quark and lepton are right-handed and left-handed, respectively. Other subscripts also carry the same convention. As $S_3^{\frac{2}{3}}$ has only one decay channel ($i.e.$, $S_3^{\frac{2}{3}} \to t_L\nu_\tau$), it has 100% branching fraction for it. Although $R_2^{\frac{2}{3}}$ has two decay channels, in our present analysis we shall assume 100% branching fraction to its $t_R\tilde{\nu}_\tau$ decay mode, which can easily be scaled to other values as required[f].

6.3 Pair production at NLO+PS accuracy

We consider signal events at the NLO in QCD matched to parton shower. The production of events at the NLO(FO) QCD accuracy requires calculating amplitudes of LO, virtual and real-emission Feynman diagrams. We show all the LO and a few virtual Feynman diagrams for the pair production of scalar leptoquarks in Figure 6.1. The real-emission diagrams are not shown which are tree level diagrams with an extra gluon or light quark. The diagrams are drawn using JaxoDraw package [246]. The events are produced using MadGraph5_aMC@NLO [69]. For the signal, we first write the model in FeynRules [26] and use NLOCT [142] package[g] to produce the UFO model [247]. This UFO model is then used in MadGraph5_aMC@NLO to generate events at the NLO(FO) accuracy. To account

[d] The models R_2 and $\tilde{R}_2$ which do not allow any diquark coupling are called genuine leptoquark. The rest four scalar leptoquark models allow diquark couplings.

[e] The diquark coupling can also be forbidden by demanding either baryon number or lepton number conservation.

[f] For other branching fraction, the production cross section of leptoquark pair will also depend on $y_{R_{LR}}$ in the five flavor scheme, since a t-channel production diagram will appear when $y_{R_{LR}}$ is non-zero. However, in Ref [245] it has been shown that the dependence of the cross section on this parameter is quite small.

[g] The NLOCT package calculates the UV and R2 terms of the OPP method [144].

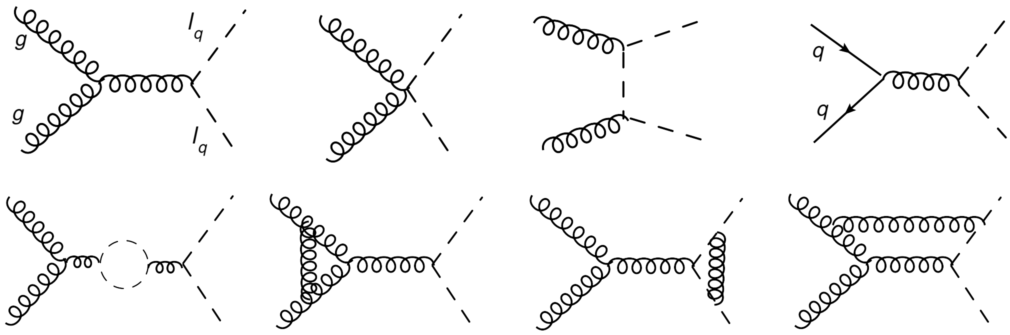


Figure 6.1: In the upper row, all possible prototype born diagrams are shown. In the lower row, only few prototype virtual diagrams are shown.

order / model	LO (fb)	NLO(FO) (fb)
$S_3^{\frac{2}{3}}$	$0.6621^{+37.8\%}_{-25.8\%}$	$0.7229^{+14.5\%}_{-14.7\%}$
$R_2^{\frac{2}{3}}$	$0.6631^{+37.8\%}_{-25.8\%}$	$0.7163^{+14.9\%}_{-14.8\%}$

Table 6.2: Cross sections for the pair production of scalar leptoquarks of mass $M_{\text{LQ}} = 1300\ GeV$ at the 14 TeV LHC. The scale variation are shown in subscript and superscript.

for the infrared divergence in real emission processes, MadGraph5_aMC@NLO uses the FKS subtraction method [143, 209].

In Table 6.2, we show the cross sections for the production of pair of scalar leptoquarks of mass $M_{\text{LQ}} = 1300\ GeV$ at 14 TeV LHC at LO and NLO(FO). The cross sections for both the models $S_3^{\frac{2}{3}}$ and $R_2^{\frac{2}{3}}$ are same up to Monte Carlo uncertainty, as expected from the discussion in the previous section. Corrections due to the NLO QCD effects are around 10%. We have used NNPDF23_lo_as_0119_qed and NNPDF23_nlo_as_0119_qed parton distribution functions, respectively, for the LO and NLO calculations. The partonic center of mass energy is used as the central choice for the renormalization and factorization scales. For the scale variation study, we vary the renormalization and factorization scales up and down by a factor of two, resulting in total nine points including the central choice. The upper and lower envelopes of the variations of the cross section due to these different choices of scales are shown as the percentage change from the central cross section in the superscript and subscript, respectively. From the table, we see the NLO QCD correction here reduces the scale uncertainty by around a factor of two.

The NLO(FO) results discussed in the above two paragraphs can give distri-

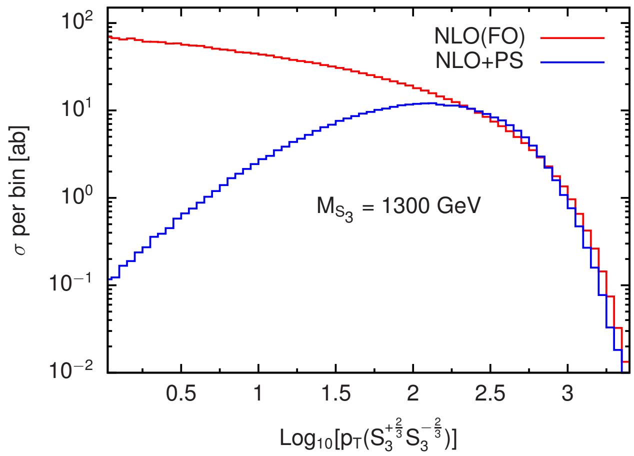


Figure 6.2: Distributions of $\mathrm{Log}_{10}\,[p_T(S_3^{+\frac{2}{3}}S_3^{-\frac{2}{3}})]$ for NLO(FO) and NLO matched to parton shower.

butions of different kinematic variables using weighted events, but unweighting of the these events cannot be done as the matrix elements are not bounded in this case [69]. Also in this case, result is not physical for low p_T region. However, it can produce unweighted events while matched to the parton shower making use of the MC@NLO formalism [148]. Results at the NLO+PS accuracy give correct description of the low p_T region. For showering of events, we use PYTHIA8 [248]. In Figure 6.2, we see that NLO+PS calculation over the fixed order one reduces the cross section at the lower transverse momentum region of the leptoquark pair system $p_T(S_3^{+\frac{2}{3}}S_3^{-\frac{2}{3}})$ due to the Sudakov suppression.

In Figure 6.3, we show LO+PS and NLO+PS normalized distributions of MET and $Log_{10}\,[p_T(S_3^{+\frac{2}{3}}S_3^{-\frac{2}{3}})]$ on the upper panels of two subfigures. The shapes of MET distributions for LO+PS and NLO+PS are identical and they peak around 700 GeV. For $p_T(S_3^{+\frac{2}{3}}S_3^{-\frac{2}{3}})$ distribution in the right figure, the peak for NLO+PS is slightly shifted towards left of LO+PS one and they peak in the range of 100-300 GeV. On the lower panels, we show the k-factor for differential distribution, i.e. the ratio of differential NLO+PS cross section to LO+PS one. In the left figure for MET, we see that for the shown range the k-factor at different bins stays nearly same and takes a value around 1.1. In the right figure, the differential k-factor is not flat for $\mathrm{Log}_{10}\,[p_T(S_3^{+\frac{2}{3}}S_3^{-\frac{2}{3}})]$ and therefore scaling the leading order events by a constant k-factor would not give precise results.

On the upper panel of Figure 6.4, we show differential distribution of cross section with respect to the top transverse momentum at the LO+PS and NLO+PS level for the central scale choice. We see that the NLO+PS corrections lead to increased cross section at every bin. In the lower panel, the effect of scale variation is shown as red and blue bands, where a band is drawn between the upper and

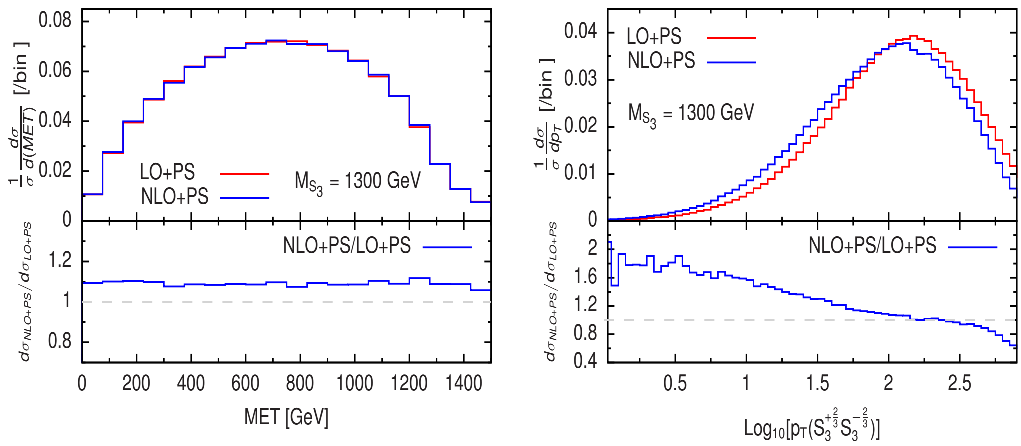


Figure 6.3: The distributions of MET and Log_{10} $[p_T(S_3^{+\frac{2}{3}}S_3^{-\frac{2}{3}})]$ at LO+PS and NLO+PS.

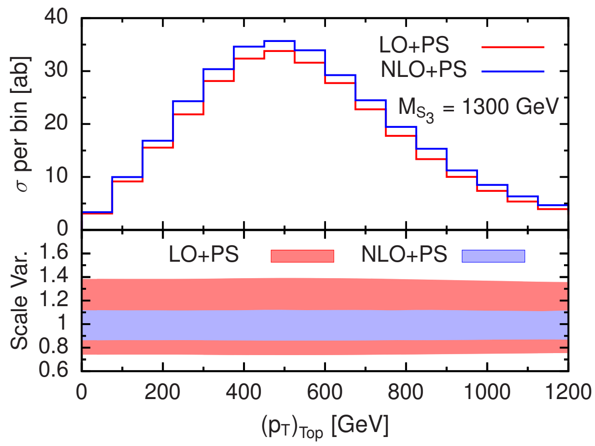


Figure 6.4: In the upper panel, we show distribution of $(p_T)_{\text{Top}}$ at LO+PS and NLO+PS accuracies. The bands in the lower panel show the scale variation of the distribution with respect to central value. The bands are drawn between the envelopes of the different distributions arising from the different scale choices.

lower envelopes of different results for different scale choices. It can be seen that the scale variation of NLO+PS result is significantly smaller compared to the LO+PS one, confirming that the NLO QCD correction leads to more accurate result in addition to the enhancement in the cross section.

6.4 Collider Analysis

We consider pair production of $\frac{2}{3}e$-charged third-generation scalar leptoquarks ($S_3^{2/3}$ and $R_2^{2/3}$) and try to probe them at the 14 TeV LHC with two top-like fatjets plus large missing transverse momentum. Third-generation scalar leptoquark pair production is possible only through gluon fusion and $q\bar{q}$ annihilation, and hence the cross section is independent of any model-dependent coupling and depends only on the leptoquark mass. We consider NLO QCD corrections matched to parton shower of the LQ pair production channel and few representative diagrams are already shown in Figure 6.1. Equations 6.3 and 6.4 show the decay modes of $S_3^{2/3}$and $R_2^{2/3}$, respectively. We consider decay of $R_2^{2/3}$ fully into a top quark and a neutrino. Since the current ATLAS study [226] excludes the third-generation LQ of mass lower than 1.24 TeV, the top quark originating from the decay of heavy LQ will have a high boost. The top quark will decay further, and all the decay components will start collimated resulting into a boosted large-radius jet, called top fatjet (J_t). We consider the hadronic decay of the top quarks. So, in the final state, we have two boosted top-like fatjets and a significant missing transverse momentum. We use jet substructure variables, missing energy, and other high-level observables to distinguish the signal from the SM background. The signal topology is given below,

$$\begin{aligned} pp \to S_3^{2/3} S_3^{-2/3}\,[\mathrm{QCD}] \to (t\nu_\tau)(\bar{t}\bar{\nu}_\tau)j \Rightarrow 2J_t + \not{E}_T + X\,, \\ pp \to R_2^{2/3} R_2^{-2/3}\,[\mathrm{QCD}] \to (t\bar{\nu}_\tau)(\bar{t}\nu_\tau)j \Rightarrow 2J_t + \not{E}_T + X\,, \end{aligned} \tag{6.5}$$

where the top quarks coming from the $S_3^{2/3}$ and $R_2^{2/3}$ decay are respectively left and right chiral.

6.4.1 Background simulation

All the background processes that can potentially mimic the signal are included in our analysis. Each background process is generated with two to four additional QCD jets and matched according to the MLM scheme [150, 151] with virtually-ordered Pythia shower. PDF sets, renormalization, and factorization scales that are used in our analysis remain same as described in Section 6.3. The showered events are then passed through DELPHES3 [70] for detector simulation purpose, and we use the default CMS card provided there. Particle-flow towers and tracks are clustered to form anti-kT jets of radius parameter 0.5. Fatjets (J or J_t) of radius 1.5 are constructed with the Cambridge-Achen (CA) algorithm [95] using FASTJET 3.2.2 [92].

$t\bar{t}+$ jets: One of the main backgrounds for our signal process is the pair production of top quarks when one of the top quarks decays hadronically and the other decay leptonically. The top quark that decays hadronically is reconstructed as top-fatjet. The neutrino from the leptonic decay of the other top quark and the lepton that escapes detection provide missing energy (MET or $\not{E}_T$), while another fatjet comes from the QCD radiation or b-jet. Hadronic decay of both top quarks can give two boosted top-fatjets; however, the requirement of significant missing energy reduces this background compared to the previous setup by a factor of 100, since the MET comes from the mis-measurement of the hadronic activities. This background is produced with two additional radiations and matched with the MLM matching scheme.

$Z+$ jets: Another main background of our signal is the inclusive Z-boson production, where the Z-boson decays invisibly. This process is generated with four extra partons, and the MLM matching is used. Two fatjets essentially originate from the QCD jets.

$W+$ jets: It contributes considerably but is smaller than $Z+$ jets background. When the W boson decays leptonically, the missing energy comes from the neutrino and the lepton that escape detection. This background is also generated with four partons following MLM matching and here also the fatjets come from the extra radiations.

Since our analysis requires large missing energy, we generate $Z+$ jets and $W+$ jets backgrounds with a generation-level hard-cut $\not\!\!E_T > 100$ GeV for better statistics.

$tW+$ jets: Single top quark production at the LHC in association with the W boson, contributes considerably as a background, which is generated with two extra parton using MLM matching. Top quark decays hadronically to give rise to a boosted top-like fatjet, while another fatjet comes from the QCD radiation. The neutrino with the missing lepton from W decay is the source of the missing energy.

$VV+$ jets: A small contribution can come from the diboson production, which can be classified into three different categories, WZ, WW, and ZZ, where all of these are matched with two extra partons applying MLM matching scheme. WZ contributes the most among these three, where Z boson decays invisibly to produce missing energy and hadronic decay of the W boson gives one fatjet. Even though WW and ZZ contribute almost negligibly we keep these backgrounds in our analysis. In either case, one of them decays hadronically and the other one decays leptonically (W) or invisibly (Z). In all these three processes, another fatjet comes due to the QCD radiation.

$t\bar{t}Z$: The cross section of $t\bar{t}Z$ is smaller than any of the above mentioned background processes, but we keep this too in our analysis. This process becomes

Background	Ref	σ (pb)
$t\bar{t}+$ jets	[202]	988.57 [N^3LO]
$tW+$ jets	[203]	83.1 [N^2LO]
$Z+$ jets	[204, 205]	6.33×10^4 [N^2LO]
$W+$ jets		1.95×10^5 [NLO]
$ZZ+$ jets	[206]	17.72 [NLO]
$WW+$ jets		124.31 [NLO]
$WZ+$ jets		51.82 [NLO]

Table 6.3: Higher-order QCD corrected production cross sections of different background processes at the 14 TeV LHC used in our analysis, where the order of QCD correction is presented in brackets.

signal like when Z-boson decays invisibly and two tops are reconstructed as top-like fatjets. This process gives almost negligible contribution compared to $Z+$ jets and $t\bar{t}+$ jets backgrounds. We omit $t\bar{t}W$ background since its contribution is found to be even more suppressed.

QCD background: The di-jet production cross section is vast at the LHC; even after constructing two fatjets, huge events remain from this background. The requirement of large missing energy gives additional suppression of order 100 since MET here can only occur due to the mis-measurement of hadronic activities. An additional suppression of order 50 comes from the requirement of b-tagged fatjet. So, QCD backgrounds are found to be negligible compared to the other backgrounds and therefore we do not include this in our analysis.

The background processes considered in our analysis are normalized with the available higher-order QCD corrected production cross section, as presented in Table 6.3.

6.4.2 Construction of Jet Substructure Variables

Jet substructure variables provide good efficiencies when analyzing boosted topologies. The substructure variables that we use in our analysis are listed below.

Pruned Jet Mass: Jet mass is a good variable in separating a boosted top-like fatjet from the boosted W/Z boson or the QCD fatjets. Additional soft and wide angle radiations from the underlying QCD interactions can contribute to the fatjet mass. So for realistic predictions, one needs to remove those contributions. Pruning, filtering, and trimming [102–105] are different jet grooming techniques and we use pruning in our analysis. The fatjet mass is defined as

$M_J = (\sum_{i\epsilon J} p_i)^2$, where the four-momentum of the i-th constituent is denoted as p_i. After clustering a fatjet using the CA algorithm, we de-cluster its constituents in each recombination step and remove the soft and wide-angle radiations from the fatjet. The merging of i-th and j-th proto-jets into the fatjet is vetoed, and the softer one is removed, if the following conditions are achieved,

$$Z = \min(P_{Ti}, P_{Tj})/(P_{Ti} + P_{Tj}) < Z_{\text{cut}}, \quad \text{and} \quad \Delta R_{ij} > R_{\text{fact}}. \tag{6.6}$$

The angular separation between two proto-jets is ΔR_{ij}, and we choose $R_{\text{fact}} = 0.86 \sim \frac{m_{\text{top}}}{P_{T,\text{top}}}$ [104]. Z and P_{Ti} are the softness parameter and the transverse momentum of the i-th proto-jet respectively. We set $Z_{\text{cut}} = 0.1$ [103] in our analysis.

N-subjettiness ratio: N-subjettiness is a jet shape variable that measures how the energy of a fatjet is distributed around different subjet axes and is defined as follows [106, 107],

$$\tau_N = \frac{1}{\mathcal{N}_0} \sum_i P_{T,i} \min\{\Delta R_{i,1}, \Delta R_{i,2}, \cdots \Delta R_{i,N}\}. \tag{6.7}$$

The summation runs over all the constituent particles of the jet. $\mathcal{N}_0$ is the normalization factor, defined as $\mathcal{N}_0 = \sum\limits_i P_{T,i} R$, where $P_{T,i}$ is the transverse momentum of the i-th constituent of the jet of radius R. $\Delta R_{i,K} = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ is the angular separation of the i-th constituent of the jet from its K^{th}-subjet axis in the pseudo-rapidity-azimuthal angle, *i.e.*, $\eta - \phi$ plane. Rather than τ_N, the ratio $\frac{\tau_N}{\tau_{N-1}}$ is a more effective discriminating variable between N-prong fatjets and the SM background [106]. Our analysis uses $\tau_{32} = \frac{\tau_3}{\tau_2}$ and $\tau_{31} = \frac{\tau_3}{\tau_1}$ to differentiate top-fatjets from the SM background.

6.4.3 Event Selection

<u>Baseline-Selection Criteria:</u> We apply the following pre-selection cuts (C1) to select events for further analysis.

- The radius parameter of the top fatjet is $R \sim \dfrac{2m_t}{P_T}$, where P_T and m_t are the transverse momenta and top quark's mass, respectively. For each event, we reconstruct at least two fatjets using CA algorithm of radius parameter 1.5 with minimum transverse momentum $P_T(J_0), P_T(J_1) > 200$ GeV

- The missing energy of each event should be greater than 100 GeV

- Since lepton is not present in the final state of our signal, we veto the events which contain any lepton of transverse momentum $P_T(l) > 10$ GeV and pseudo-rapidity $|\eta(l)| < 2.4$

- A minimal cut on the azimuthal separation between any fatjet and the missing momentum $\Delta\phi(J_i, \not{E}_T) > 0.2$ is applied to minimize the hadronic mis-measurement contribution

Cuts	S_3 (fb)	R_2 (fb)	Z +jets (fb)	W +jets (fb)	$t\bar{t}$ +jets (fb)	tW +jets (fb)	WZ +jets (fb)	WW +jets (fb)	ZZ +jets (fb)	$t\bar{t}Z$ (fb)	tot BG (fb)
C1	0.2315 [100%]	0.232 [100%]	2517.99 [100%]	1366.91 [100%]	690.65 [100%]	366.91 [100%]	93.53 [100%]	25.90 [100%]	11.51 [100%]	5.24 [100%]	5078.64 [100%]
C2	0.2258 [97.54%]	0.2262 [97.5%]	1640.29 [65.14%]	762.59 [55.79%]	302.16 [43.75%]	152.52 [41.57%]	58.35 [62.39%]	11.51 [44.44%]	6.973 [60.58%]	3.96 [75.57%]	2938.36 [57.86%]
C3	0.1810 [78.19%]	0.1801 [77.63%]	241.73 [9.60%]	117.99 [8.63%]	230.94 [33.44%]	114.39 [31.18%]	10.79 [11.54%]	2.45 [9.46%]	1.92 [16.69%]	3.28 [62.60%]	723.48 [14.25%]
C4	0.1047 [45.23%]	0.1033 [44.53%]	25.38 [1.01%]	17.33 [1.27%]	64.23 [9.30%]	27.45 [7.48%]	1.24 [1.33%]	0.33 [1.27%]	0.2 [1.74%]	1.474 [28.13%]	137.634 [2.71%]

Table 6.4: The expected number of events (in fb, multiplying with the luminosity gives the expected event numbers) and cut efficiency for the signal S_3 and R_2 (1.3 TeV mass of leptoquark for both models) and all the background processes that contribute to the fatjets $+\not{E}_T$ final state after implementing the corresponding cuts at the 14 TeV LHC are shown. The effectiveness of different kinematic cuts can be followed from top to bottom after applying (C1) Preselection cuts, (C2) $\not{E}_T > 150$ GeV, (C3) requiring at least one b-tag within J_0 or J_1, and finally (C4) $M_{J_0}, M_{J_1} > 120$ GeV. After applying C4 cut, the remaining events are passed for the multivariate analysis.

Final selection cuts: After the primary selection, we apply the following cuts before passing events for multivariate analysis (MVA).

(C2) Missing energy cut is raised from 100 GeV to 150 GeV, which reduces the background sharply

(C3) We do additional b-tagging inside the leading (J_0) or subleading (J_1) fatjets

(C4) We demand pruned mass of both the leading M_{J_0} and subleading M_{J_1} fatjets to be greater than 120 GeV

Table 6.4 displays the cut flow along with the cut efficiencies, anticipated number of events (in fb, multiplying with the luminosity gives the expected event numbers) for the signal and the background processes for the 14 TeV LHC. One can see that the higher missing energy cut, b-tagging within a fatjet, and the pruned fatjet masses are very effective in significantly reducing backgrounds while maintaining good signal acceptance. The principal backgrounds $Z+$ jets and $W+$ jets are drastically reduced when a b-jet is tagged within the leading or subleading fatjet, and their effects are nearly identical to that of the $t\bar{t} + jets$ background (see the rows up to C3 in Table 6.4).

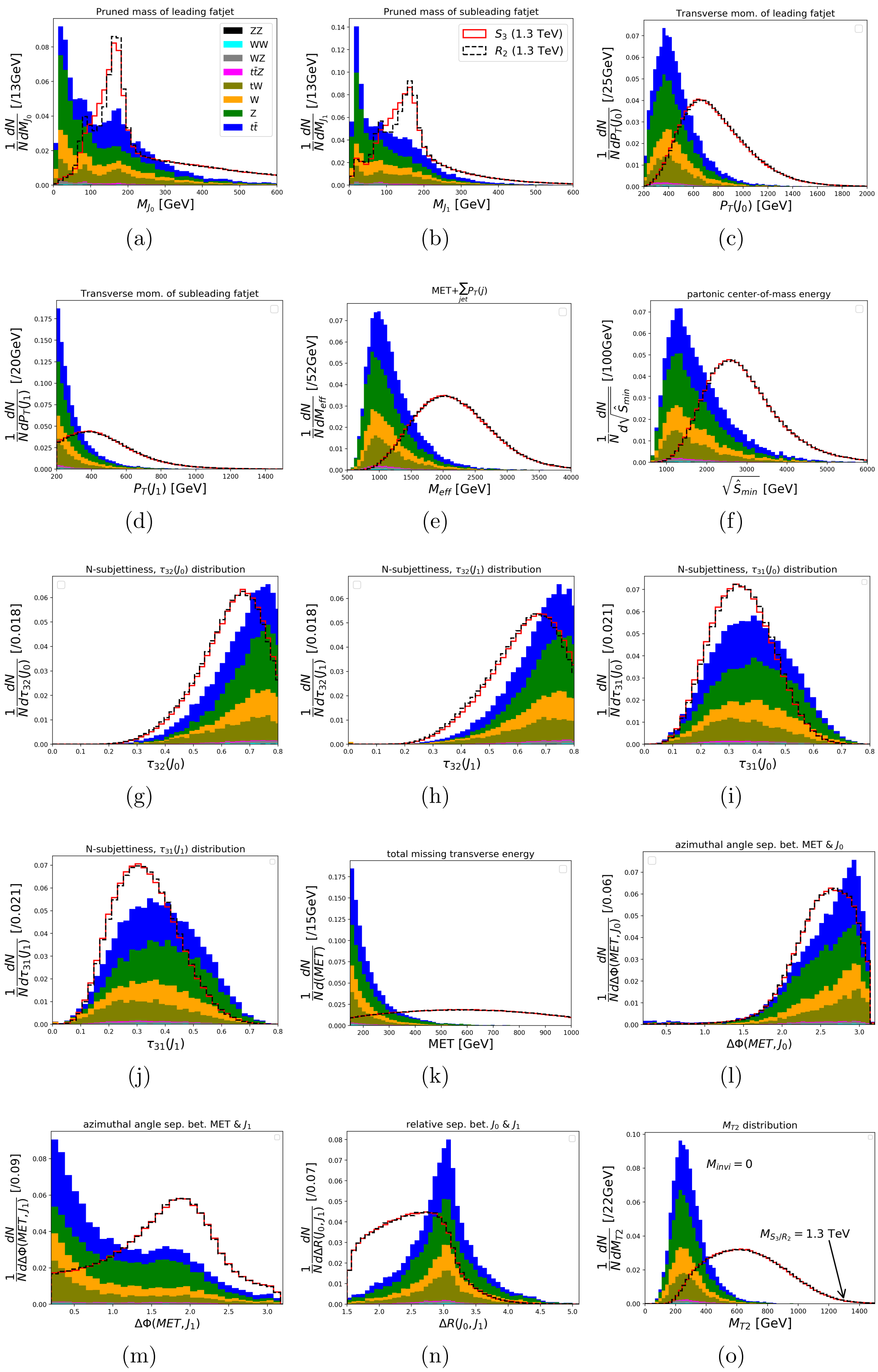


Figure 6.5: After imposing $\not{E}_T > 150$ GeV and b-tagging inside J_0 or J_1, together with preselection cuts as indicated in the text, the normalized distribution of kinematic variables of the signal S_3 (solid red), R_2 (dashed black), and bin-wise stacked histogram of all the background processes are shown.

The normalized distributions of various kinematic variables of the signal S_3 and R_2, as well as bin-wise stacked histograms of all background processes after imposing $\not{E}_T > 150$ GeV and b-tagging inside J_0 or J_1, together with preselection cuts, are shown in Figure 6.5, where leptoquark mass is set at 1.3 TeV. The contributions of individual background processes are represented by different colors: blue, green, orange, olive, and magenta, for $t\bar{t}$+jets, Z+jets, W+jets, tW+jets, and $t\bar{t}Z$, respectively. Each background process is weighted by its effective cross section after applying the cuts listed and normalized to the total cross section.

The distributions of the leading and subleading fatjets' pruned masses are depicted in Figure 6.5a and Figure 6.5b respectively. $Z+$ jets background demonstrates no peak in the M_{J_0} and M_{J_1} distributions, near the Z-boson mass or around the top mass as expected, since the fatjets originate from the QCD radiations. One of the tops in the $t\bar{t}+$ jets background decays hadronically and is reconstructed as a top-fatjet, while the other fatjet comes because of the QCD radiation. As a result, the M_{J_0} distribution exhibits a peak near the top mass, but the M_{J_1} distribution does not exhibit a peak near the top mass.

It is also interesting to note that fatjet mass distributions are slightly different for the two signals while other kinematic variables remain similar. This is a direct implication of two different polarization. The bottom quark and the W-boson travel in the opposite direction in the top quark's rest frame to conserve the linear momentum. As in the S_3 model the top quark is left chiral, the majority of the b-quark in the top quark's rest frame lie in the same direction of the boost (this will be further discussed in the next section). This means that the majority of the W boson emerges at an angle greater than 90 degrees to the boost. However, in the R_2 model (top quark is right chiral), most of the b-quarks are found in the direction opposite to the boost in the top quark's rest frame. This suggests that most of the W bosons exist around the boost direction. As a result, in the lab frame, the quarks from the hadronic decay of the W boson are more collimated in the R_2 model compared to the S_3 model. When W and the b-quark get combined to form a single large radius three-prong fatjet, the S_3 model produces fewer events than the R_2. Because the W boson is heavier than the b-quark, S_3 needs more boost to bring back all the W bosons along the boost direction compared to R_2. As a result, R_2 model exhibits larger peaks in both the leading and subleading fatjet mass distributions around the top quark mass than the S_3 model. Moreover, for the R_2 model, we observe also a distinct peak at the W-boson mass in either of the fatjet mass distribution. This is because most R_2 events carry W bosons along the boost direction in the top quark's rest frame, and in lab frame decay products of W boson, are more collimated compared to S_3 events. However, we see more S_3 events than R_2 between the W boson and top quark mass because the overall cross section is the same for both models.

Figures 6.5c and 6.5d respectively depict the transverse momentum of J_0 and J_1. From these distributions, we can observe that the signal is substantially harder than the background. Figure 6.5e displays the M_{eff} distribution, where M_{eff} is the scalar sum of the total transverse momentum of the visible jets plus MET.

$$M_{\text{eff}} = \not{E}_T + \sum |\overrightarrow{P}_{iT}|\,, \tag{6.8}$$

where the summation runs over all the visible jets and $\overrightarrow{P}_{iT}$ is the transverse momentum of the i-th jet. Global and inclusive quantities are used to define $\sqrt{\hat{s}_{\text{min}}}$ [152], the minimum partonic center-of-mass energy, and its distribution is shown in Figure 6.5f. Neutrinos are the missing particles in our system, and the definition of $\sqrt{\hat{s}_{\text{min}}}$ is given by

$$\sqrt{\hat{s}_{\text{min}}} = \sqrt{E^2 - P_Z^2} + \not{E}_T \tag{6.9}$$

where E and P_Z are the total energy and longitudinal component of the total visible momentum in the event, respectively. Here visible means all the visible objects in the detector, e.g., jets, electrons, photons, and muons. The signal has a peak towards a larger value of $\sqrt{\hat{s}_{\text{min}}}$ compared to the background since the signal requires more partonic center-of-mass energy to produce two heavy LQs that subsequently decay into the top quark and neutrino.

The N-subjettiness variables, τ_{32}, for both the leading and subleading fatjets are shown in Figures 6.5g and 6.5h. τ_N tries to quantify the number of subjets inside the fatjet. One would anticipate a smaller value of τ_{32} for a boosted top-fatjet since the value of τ_3 for a three-prong fatjet is small and the value of τ_2 is large, therefore their ratio produces a smaller value. In contrast, backgrounds are mostly QCD dominated (1-prong) or coming from the weak bosons (2-prong), so the value of τ_2 is small for both QCD jets and fatjets originating from weak bosons, giving larger τ_{32}. The distributions show that the signal has considerably lower τ_{32} values [h] than the backgrounds, indicating that the signal has a more three-prong structure than the background. Different chirality of the top quarks accounts for the slight difference in these distributions for S_3 and R_2 models. The distributions of τ_{31} for J_0 and J_1 are shown in Figures 6.5i and 6.5j. The distributions show that both the signal and the background peak at a lower value of τ_{31}, indicating that it is not as good as τ_{32} for distinguishing the signal from

[h] Although the signal peaks at a lower value of τ_{32} than the background, the peak emerges at roughly 0.6, which is rather substantial. The three subjets of the top quark are highly collimated, therefore the τ_2 value is also small for the top fatjets, which causes the three-prong top fatjets peak to arise for the signal at a significantly large value of τ_{32}.

the background.

The distribution of missing transverse momentum is shown in Figure 6.5k, where the background can be seen to drop sharply for large MET. In the case of signal, both the neutrinos from the decay of LQs, have equal access to the phase space, resulting in a nearly uniform distribution of the missing transverse momentum. Figures 6.5l, 6.5m, and 6.5n show, respectively, the distributions of the azimuthal separation of the leading and subleading fatjets from the $\not{E}_T$ and the relative separation between the fatjets in the $\eta - \phi$ plane. The distribution of M_{T2} [154, 207] is shown in Figure 6.5o. M_{T2} is useful in measuring the mass of the parent particle, which is pair-produced at the collider, and subsequently decays into one visible object and one missing particle from the end-point of the distribution, and it is defined as follows

$$M_{T2} = \min_{\overrightarrow{p_{1T}}^{\rm invi} + \overrightarrow{p_{2T}}^{\rm invi} = \not{E}_T} [\max\{M_T^{(1)}, M_T^{(2)}\}]. \tag{6.10}$$

$M_T^{(i)}$ $(i = 1, 2)$ are the transverse masses of the LQ and anti-LQ as defined below,

$$(M_T^{(i)})^2 = m_i^2 + M_{\rm invi}^2 + 2(E_{iT} E_{iT}^{\rm invi} - \overrightarrow{p_{iT}} \cdot \overrightarrow{p_{iT}}^{\rm invi}), \qquad \{i = 1, 2\}\,. \tag{6.11}$$

Since LQ decays into a top quark and massless neutrino, we set $M_{\rm invi}^2 = M_\nu^2 = 0$ and $E_{iT}^{\rm invi} = |\overrightarrow{p_{iT}}^{\rm invi}|$, where $\overrightarrow{p_{iT}}^{\rm invi}$ is the transverse momentum of an individual neutrino. $\overrightarrow{p_{iT}}^{\rm invi}$ is constrained by the measured missing transverse momentum,

$$\overrightarrow{p_{1T}}^{\rm invi} + \overrightarrow{p_{2T}}^{\rm invi} = \overrightarrow{\not{E}_T}. \tag{6.12}$$

m_i, and $\overrightarrow{P_{iT}}$ $(i = 1, 2)$ are the reconstructed mass and the transverse momentum of the (sub)leading top-fatjets, respectively. E_{iT} is the transverse energy of the fatjets defined as $E_{iT} = \sqrt{m_i^2 + \overrightarrow{P_{iT}}^2}$. One can observe from the distribution Figure 6.5o that its end point correctly predicts the mass of the LQ (1.3 TeV). Since the SM particles have masses that are significantly less than the LQ mass, the background and signal distributions are quite well separated. So this variable not only predicts LQ mass, but also helps in background reduction.

6.4.4 Multivariate Analysis

In the previous subsection, distribution of several observables (without C4 cut), that can be used as input variables for sophisticated multivariate analysis using the gradient boosting technique, are described. For MVA input, we use a loose-cut (up to C4), as mentioned in the preceding subsection. The last row of Table 6.4 shows the estimated amount of signals (in fb) from two models, the contribution

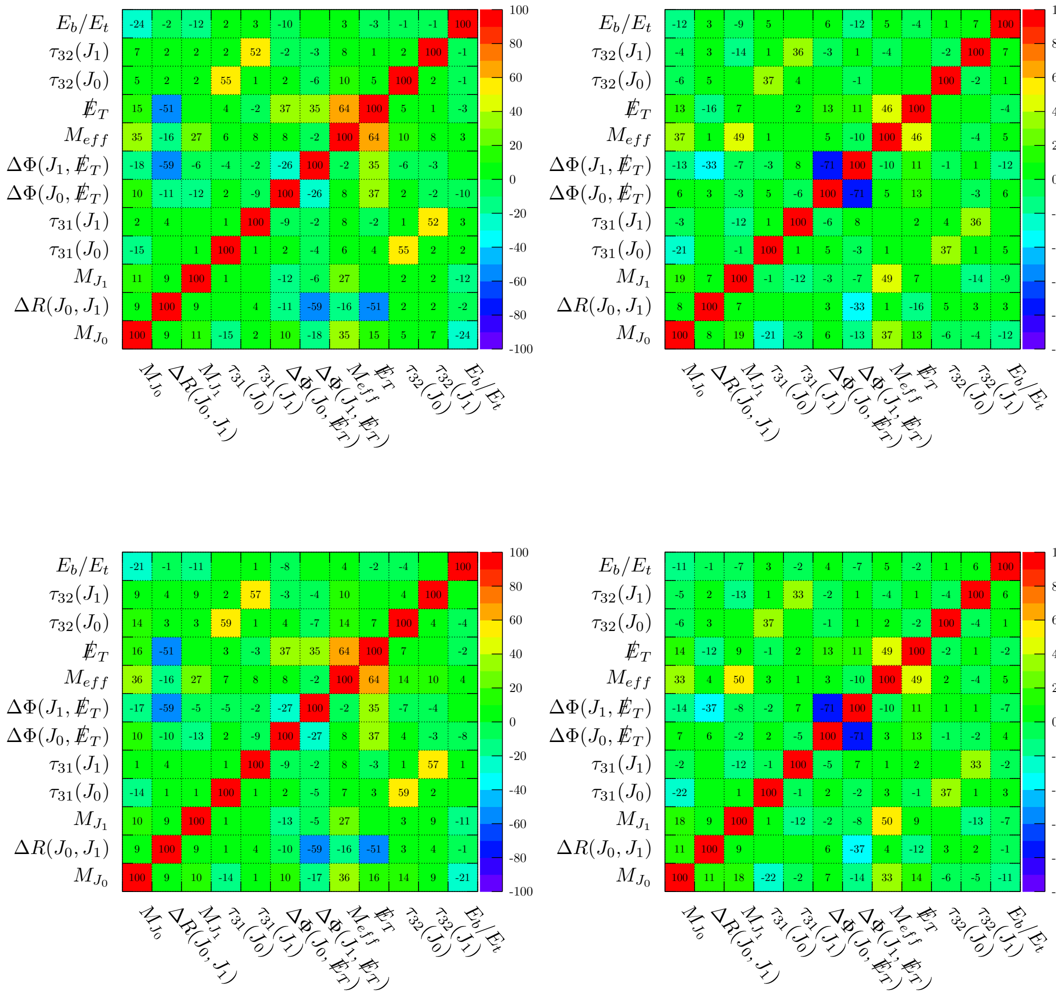


Figure 6.6: Linear correlation coefficients (%) between different variables for signal S_3 (top left panel) and corresponding background (top right panel); same for signal R_2 (bottom left panel) and corresponding background (bottom right panel). Positive and negative coefficients show that two variables are correlated or anti-correlated, respectively. Missing entries indicate an insignificant correlation of less than one.

of different background processes, and the total background at the 14 TeV LHC after applying MVA selection cut (C4). For MVA, we use the adaptive Boosted Decision Tree (BDT) algorithm and construct two statistically independent signal and background event samples. The background is the weighted sum of individual SM background processes. MVA picks a subset of kinematic variables from a larger collection based on the linear correlation among the variables and their relative importance in distinguishing the signal from the background. As expected by Equation 6.8, we notice that $P_T(J_0)$ and $P_T(J_1)$ have large correlations with $M_{\rm eff}$, and $\sqrt{\hat{s}_{\rm min}}$ also exhibits high correlations with $M_{\rm eff}$ due to their linear dependence on MET, as shown by Equations 6.8 and 6.9. We keep M_{eff} because of its high relative importance compared to $P_T(J_0)$, $P_T(J_1)$, and $\sqrt{\hat{s}_{\text{min}}}$. A high correlation exists between M_{T2} and MET; however, we retain MET because it has the highest relative importance than any other variables in separating the signal from the background. Although M_{eff} and MET exhibit a significant correlation in both the signal and background (as predicted by Equation 6.8), we keep them both in our study since they have exceptionally high separation powers to distinguish the signal from the background. Figure 6.6 exhibits the linear correlation coefficients between different variables for signal S_3 (top left panel) and the corresponding background (top right panel). The bottom left and bottom right panels depict the signal R_2 and its corresponding background. Positive and negative coefficients indicate whether two variables are correlated or anti-correlated. In the TMVA package [112], the linear correlation coefficient is calculated using the following formula,

$$\rho(x, y) = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y}, \tag{6.13}$$

where the covariance between x and y is $\text{cov}(x, y) = \langle xy \rangle - \langle x \rangle \langle y \rangle$ and σ_x, σ_y are the standard deviation of these variables.

Variable	$\not{E}_T$	M_{eff}	$\Delta R(J_0, J_1)$	$\Delta\phi(J_1, \not{E}_T)$	$M(J_0)$	$\tau_{32}(J_1)$	$\tau_{32}(J_0)$	$\Delta\phi(J_0, \not{E}_T)$	$M(J_1)$	$\tau_{31}(J_1)$	$\tau_{31}(J_0)$
S_3	59.98	49.43	23.44	21.42	5.99	4.15	3.99	3.83	2.33	0.99	0.95
R_2	59.33	50.63	21.97	20.87	6.42	5.85	4.82	4.36	2.49	1.49	1.11

Table 6.5: Before employing at MVA, the method unspecific relative importance (separation power) of the individual variables.

The separation power of different kinematic variables for the two models used in MVA, is presented in Table 6.5. This table shows that the order of the variables in for distinguishing the leptoquark signal from the overwhelming background are the MET, M_{eff}, relative separation between the fatjets in $\eta - \phi$ plane, and azimuthal separation between the subleading fatjet and MET. Due to improper selection of various (BDT-specific) parameters during training, the BDT method may result in overtraining. Overtraining can be prevented if the Kolmogorov-Smirnov probability is checked throughout training. We train the algorithm separately for the S_3 and R_2 models and ensure that there is no overtraining in our analysis. The top left panel of Figure 6.7 shows the normalized distribution of the BDT output for the signal S_3 (blue) and its background (red) for both training and testing samples, whereas the bottom left plot shows the same for the $R2$ model. We observe that for both models, signal and background are well separated. In the same figure, top right plot illustrates the signal S_3 (blue) and background (red) efficiencies, as well as statistical significance (green) as a function of the cut applied to BDT output, while the bottom right plot depicts the same for the R_2 model.

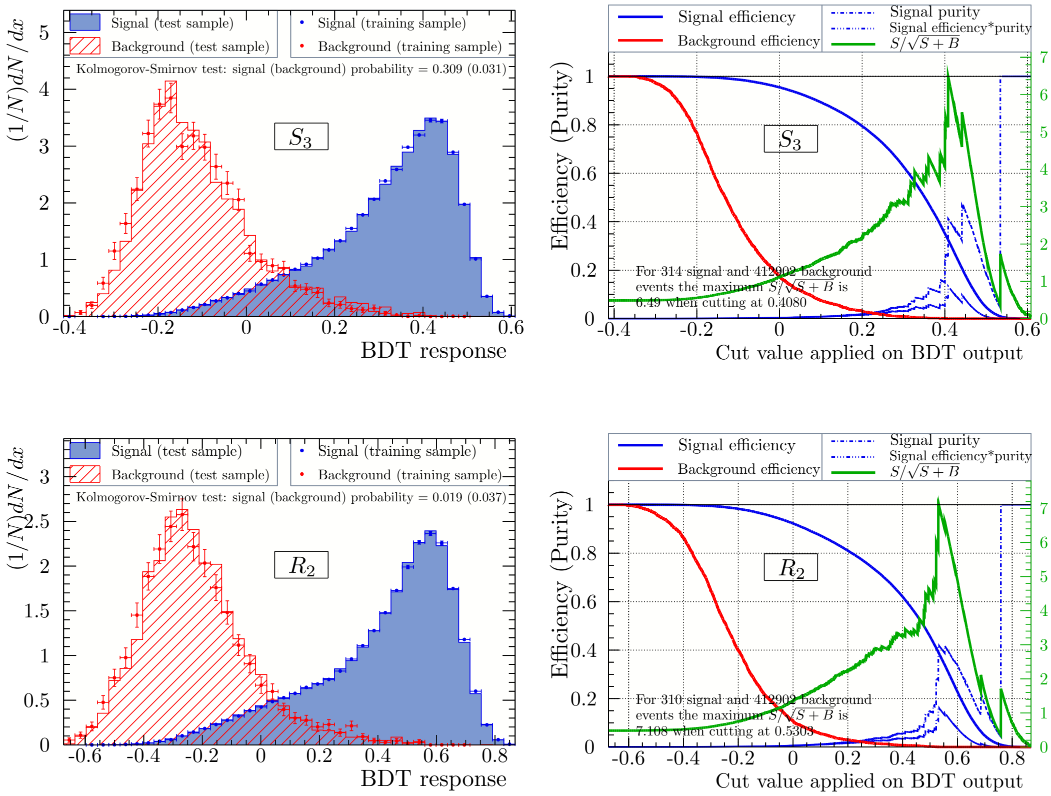


Figure 6.7: The top-left plot depicts the distribution (normalized) of the BDT output for the training and testing samples for both the signal S_3 (blue) and background (red) classes. The right plot depicts signal S_3 (blue) and background (red) efficiencies, as well as statistical significance ($\frac{N_S}{\sqrt{N_S+N_B}}$) as a function of the cut applied to BDT output. The same for the R_2 model is shown in the bottom left and bottom right plots.

	N_S^{bc} (fb)	BDT$_{\text{opt}}$	N_S (fb)	N_B (fb)	$\frac{N_S}{\sqrt{N_S+N_B}}$ ($\frac{N_S}{\sqrt{N_B}}$), 3 ab^{-1}	$\frac{N_S}{N_B}$
S_3	0.1047	0.4080	0.03403	0.04850	6.5 (8.5)	0.702
R_2	0.1033	0.5303	0.04047	0.05677	7.1 (9.3)	0.713
N_{SM}	137.634					

Table 6.6: The table shows the effectiveness of the present search in terms of statistical significance for S_3 and R_2 models. Before applying any cuts to the BDT output, the total number of events for different models and the combined background are N_S^{bc} and N_{SM}, respectively (as shown in Table 6.4). For the 14 TeV LHC, after employing an optimum cut (BDT$_{opt}$) on the BDT response, the surviving number of signal and background events are provided by N_S and N_B (in fb), respectively. For quick access, the statistical significance corresponding to 3 ab^{-1} luminosity are also shown.

The statistical significance of the two models at 3 ab^{-1} integrated luminosity at the 14 TeV LHC and the signal-to-background ratio are shown in Table 6.6. There, N_S^{bc} [i] and N_{SM} represent the total number of events for the signal and background before applying any cut to the BDT output, while N_S and N_B represent the same after applying an optimal cut $\mathrm{BDT_{opt}}$ to the BDT response. We observe that both models at the HL-LHC have discovery potential for the 1.3 TeV scalar leptoquark. The 5σ-discovery and 2σ-exclusion limits of these two models at the HL-LHC are presented in Table 6.7. There is a slight difference in the discovery potential of these two models because of their polarization. Note that explicitly the polarization variables have a negligible role compared to other variables such as MET, $\Delta R(J_0, J_1)$, etc., as given in Table 6.5 for discovering the leptoquark signal at the LHC. However, we see significant differences in the J_0 and J_1 mass distribution and slight in N-subjettiness distributions because of the different chirality of the top quarks of the two models. So, once the LQ signal is discovered at the LHC, we can use the polarization variables to distinguish these two models, which are described in detail in the next section. In our analysis, we find that a scalar LQ of mass 1270 GeV or smaller can be rejected with 2σ with an integrated luminosity of 140 fb^{-1}, which is compatible with the existing ATLAS search and analysis. We also find that a luminosity around 1600 fb^{-1} is required for the 5σ discovery of 1.3 TeV scalar LQ.

$\mathcal{L} = 3ab^{-1}$	$S_3^{\frac{2}{3}}$	$R_2^{\frac{2}{3}}$
5σ discovery	1380 GeV	1370 GeV
2σ exclusion	1520 GeV	1520 GeV

Table 6.7: Discovery and exclusion reach at 14 TeV LHC for 3 ab^{-1} luminosity.

6.5 Distinguishing two models

If a leptoquark signature is observed at the collider in some particular final state, the next goal will be to distinguish different models in order to probe its genesis. In the above section, we have seen that pair production in both $S_3^{\frac{2}{3}}$ and $R_2^{\frac{2}{3}}$ models can finally give two fatjets plus large missing energy signature. The leptoquarks in these two models decay to top quarks of different helicities. Top quark's polarization can be probed by studying the distribution of some particular kinematic variables of its decay products, which can in turn allow us to probe the type of the leptoquark. In the following subsection, we discuss some such polarization variables that can address the leptoquark identity.

[i] Although we use full NLO events, if one uses LO events but normalizes with the total NLO cross section, N_S^{bc} number for both models decreases by around 2%.

Daughters	b	W^+
k_i	-0.41	+0.41

Table 6.8: Spin analyzing power of bottom quark and W^+ coming from top decay.

6.5.1 Polarization Variables

There are different variables which can exhibit dependence on top quark polarization. In the following, we discuss a few of them.

6.5.1.1 Angular variable in the rest frame of (anti-)top

In the rest frame of top quark, if θ_i be the angle between the decay particle i and the direction of boost of the top quark , the differential distribution of the decay width Γ with respect to the angular variable $\cos\theta_{\rm i}$ is given by,

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta_i} = \frac{1}{2}(1 + \rm P_t\ k_i\ \cos\theta_i), \tag{6.14}$$

where $\rm P_t$ is the top quark polarization, which is +1 for the right-handed top and -1 for the left-handed top. k_i is the spin analyzing power of i-th decay particle. In Table 6.8, we show the spin analyzing power of different decay particles. In Appendix F.0.1, the spin analyzing power of bottom quark is derived. Similar distribution in the anti-top rest frame can be written as,

$$\frac{1}{\bar{\Gamma}}\frac{d\bar{\Gamma}}{d\cos\bar{\theta}_{\bar{i}}} = \frac{1}{2}(1 + \rm \bar{P}_{\bar{t}}\ \bar{k}_{\bar{i}}\ \cos\bar{\theta}_{\bar{i}}), \tag{6.15}$$

where the entities with bar are the corresponding quantities for the anti-top quark. Here as well, $\bar{P_{\bar{t}}}$ is +1 for right-handed anti-top and -1 for left-handed anti-top. $\bar{k_{\bar{i}}}$ is given by $\bar{k_{\bar{i}}} = -k_i$. So it is evident that the distribution of the i-th decay particle coming for the right-handed top will be same as the distribution of $\bar{i}$-th decay product of the left-handed anti-top. As we are producing leptoquark pair which will decay to top and anti-top with opposite helicities, this feature will ensure the distributions of b and $\bar{b}$ for a model are the same.

The decay of top quark gives rise to mostly left-handed ($\lambda_b = -1$) b-quark and the other component, *i.e.* the right-handed one, is heavily suppressed because of small mass of b-quark[j]. It is known that the top quark decays 70% of the time to longitudinal ($\lambda_{W^+} = 0$) and 30% of the time to one of the transverse ($\lambda_{W^+} = -1$) component of the W boson [249,250][k]. So for top quark, essentially only two decay

[j]This happens as the decay is governed by weak interaction, which couples to only left-handed fermions in the massless limit.

[k]The top quark decay to other transverse component ($\lambda_{W^+} = +1$) is almost negligible,

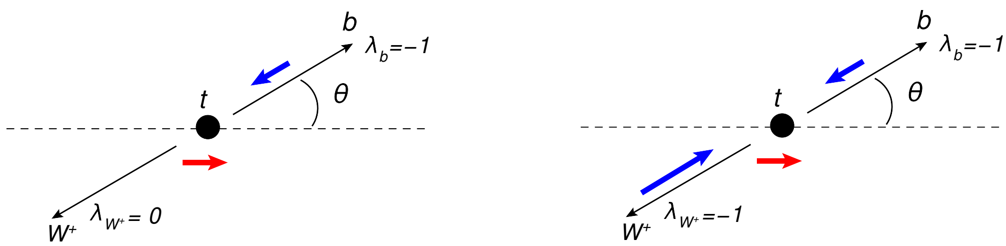


Figure 6.8: Decay diagram of right-handed top quark in its rest frame. Black dot represents top quark. Thick colored arrows denote spin of the particles. For b-quark, essentially $\lambda_b = -1$ component gets produced and the other component $\lambda_b = +1$ is heavily suppressed, because of its small mass. The top decays to $\lambda_{W^+} = 0$ and $\lambda_{W^+} = -1$ helicity components of W^+ 70% and 30% times, respectively. As the other transverse component W boson, i.e. $\lambda_{W^+} = 1$, requires right-handed b-quark to conserve spin, it is also suppressed. So in effect only the two diagrams shown here contribute to the right-handed top quark decay.

configurations exist. In Figure 6.8, to illustrate, we show these two configurations for decay of a right-handed top quark in its frame. To conserve the total spin in the decay process, the total spin of the b-quark and W boson system must be equal to $\frac{1}{2}$. Moreover, we can write the spin state of the b-quark and W boson system in the basis of $|+\rangle_{\hat{z}}$ and $|-\rangle_{\hat{z}}$ states[l] (with positive z-axis along the top boost direction). So to conserve third component of spin, only $|+\rangle_{\hat{z}}$ component can contribute, as the top quark spin is along the the boost direction. For the left diagram, the total spin of b-quark and W boson system makes an angle $(180-\theta)$ with the boost direction, whereas for the right diagram it makes angle θ. So the left diagram follows a $\sin^2\frac{\theta}{2}$ distribution, whereas the right diagram follows a $\cos^2\frac{\theta}{2}$ distribution[m]. Obviously, the weighted sum of these two distributions should lead to Equation 6.14 [n].

6.5.1.2 Energy variables in the Lab frame

In the literature [235, 240, 243, 251], two most discussed energy variables for the polarization study are $z = \frac{E_b}{E_t}$ and $u = \frac{E_l}{(E_l+E_b)}$. However, the variable $z = \frac{E_b}{E_t}$, which is the fraction of energy of the top quark carried by the b quark in the lab frame, is only the relevant one here as the W boson originating from top quark decays hadronically in our study. The variable "z" and cos θ_b are fully correlated

as this requires right-handed b-quark (which is heavily suppressed) to conserve spin angular momentum.

[l] $|+\rangle_{\hat{n}} = \cos\frac{\Theta}{2}|+\rangle_{\hat{z}} + \sin\frac{\Theta}{2}e^{i\Phi}|-\rangle_{\hat{z}}$, where $\hat{n}$ is a unit vector along (Θ,Φ) direction.

[m] $(\cos\frac{\Theta}{2})^2|_{\Theta=180-\theta} = \sin^2\frac{\theta}{2}$

[n] $0.7\sin^2\frac{\theta}{2} + 0.3\cos^2\frac{\theta}{2} = 0.5 - 0.2\ \cos\theta$

and they are related by the following relation [235] (see Appendix F.0.2),

$$\cos\theta_{\rm b} = \frac{1}{\beta_{\rm t}}\Big(\frac{2{\rm m_t^2}}{{\rm m_t^2}-{\rm m_W^2}}{\rm z}-1\Big), \tag{6.16}$$

where β_t represents the boost of the top quark in the lab frame. The distribution of decay width with respect to z (using Equation 6.14 and Equation 6.16) can be given as [243],

$$\frac{1}{\Gamma}\frac{d\Gamma}{dz} = \frac{1}{\beta_t}\frac{m_t^2}{m_t^2-m_W^2}\left(1 - P_t\, k_f\frac{1}{\beta_t} + P_t\, k_f\frac{1}{\beta_t}\frac{2m_t^2}{m_t^2-m_W^2}z\right). \tag{6.17}$$

The similar expression will hold for anti-top particle with every element replaced by their corresponding barred element.

6.5.1.3 Distributions of polarization variables

In Figure 6.9, we show truth level normalized distributions of $\cos\theta_{\rm b}$ and $\dfrac{E_b}{E_t}$ at LO at the left and right subfigures, respectively[o]. The distribution with respect to $\cos\theta_{\rm b}$ can be understood from Equation 6.14. Therefore in S_3 model, for most of the events in the rest frame of the top quark, the b-quark moves in the same direction as the boost of the top quark. Obviously, the opposite happens for the R_2 model. For the $z = \frac{E_b}{E_t}$ variable, we see for the S_3 and R_2 models, the distribution peak near the right and left end of the plots, respectively. This can also be understood from the cos θ_b distribution. As for the R_2 model, in the rest frame, for majority of events, the b-quarks move in the direction opposite to the boost and their energy E_b will be less. Therefore the distribution in this case peaks towards the left. The reverse happens for the S_3 model. Another interesting thing to observe in the right figure is that the cross section is zero after $z = 0.8$. This happens because all the top quark energy cannot be carried by the b-quark only, as the W boson needs at least its rest mass energy, M_W. In Figure 6.10, we show these distributions after including NLO calculation, showering effect, and applying various cuts up to C4 (discussed in Subsection 6.4.3) in Delphes simulation. Here the distribution of b-jet is found to be different from that of b-quark because of showering effects and formation of jets. Near the boost direction, i.e. near cos $\theta_b \sim 1$, the difference between the b-quark and b-jet distributions is striking as there b-jet gets contaminated with the particles originating from W boson because of very large boost of top quark.

[o] We discussed in the Section 6.5.1.1, b and $\bar{b}$ jets have the same distributions for a model. For an event, now onwards by b we will mean either b or $\bar{b}$ jet and t will mean corresponding top or anti-top fatjet.

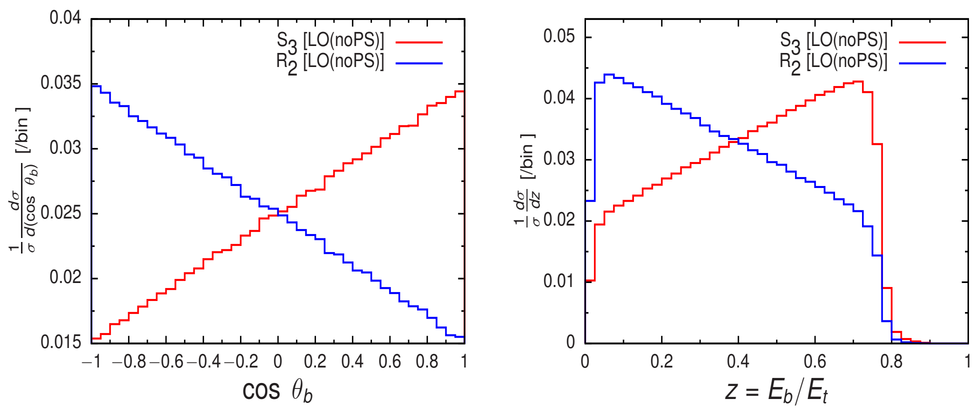


Figure 6.9: The distributions of cos θ_b and $\dfrac{E_b}{E_t}$ at LO without parton shower. The mass of the leptoquark has been taken to be 1300GeV.

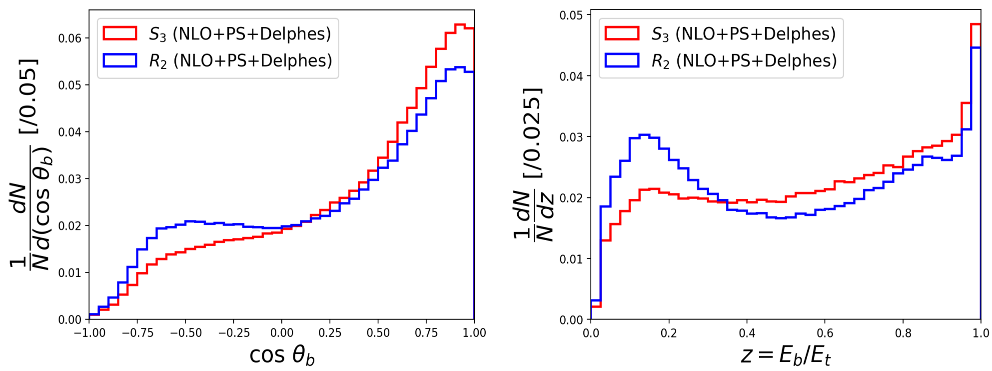


Figure 6.10: The distributions of cos θ_b and $\dfrac{E_b}{E_t}$ after Delphes simulation and applying cuts up to C4 mentioned in Subsection. 6.4.3. The effect of radiation causes significant changes in the distribution compared to truth level results. For cos $\theta_b \sim 1$ and for z around 0.8 and more, the distributions are strikingly different from the truth level results because of the contamination in the b-jet from W decay products, owing to very large boost of top quark.

6.5.2 Log-likelihood ratio test

In this subsection, we study the prospect of distinguishing two models, if in the future, a scalar leptoquark of mass 1300 GeV is observed. It will take around 1600 fb^{-1} of data for a 5σ discovery. At this mass, for $\mathcal{L} = 3000\text{fb}^{-1}$, with the optimized cuts chosen by BDT, the number of signal and background events are found to be (102,145) for the S_3 model and (121,170) for the R_2 model [p]. For these number of events we find the distribution of events with respect to $\frac{E_b}{E_t}$. We use log-likelihood ratio (LLR) hypothesis test for distinguishing two models[q]. The likelihood function is given by the product of Poisson distribution functions

[p]multiplying luminosity with the cross sections given in Table 6.6 gives these event numbers.
[q]We have also checked with χ^2 hypothesis test and got similar kind of results.

at all bins. That is, for O_i being the observed data and E_i being the expected data, the likelihood function $\mathcal{L}$ is given as,

$$\mathcal{L}(\mathrm{E}|\mathrm{O}) = \prod_{i=1}^{n} e^{-E_i} {E_i}^{O_i} / \Gamma(O_i + 1)\,. \tag{6.18}$$

The exclusion significance of a model M1, when another model M2 is observed, is given as

$$Z_{M1|M2} = \sqrt{-2ln\frac{\mathcal{L}(\mathrm{M1}|\mathrm{M2})}{\mathcal{L}(\mathrm{M2}|\mathrm{M2})}}\,. \tag{6.19}$$

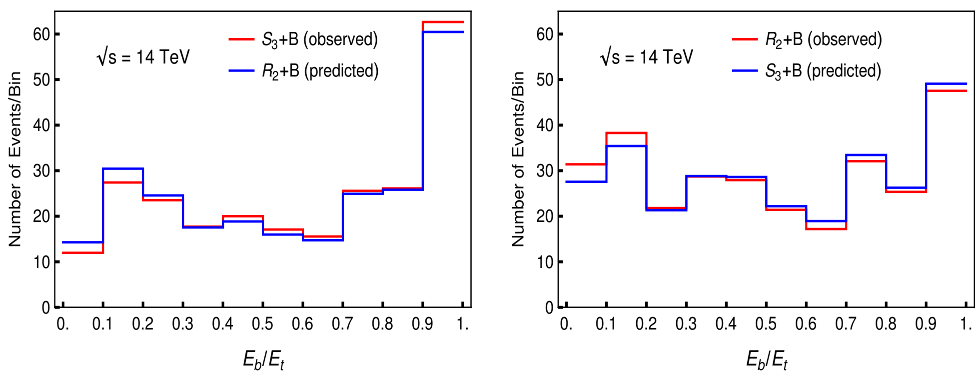


Figure 6.11: The signal+background event distributions in $\frac{E_b}{E_t}$ for observed and predicted models data after applying an optimal BDT cut (given in Table 6.6) with 3000 fb^{-1}. To find the events for the predicted model, the signal events of it are passed through the same BDT model used for finding the event numbers of the observed model.

We have considered both the scenarios when either of the models is observed and the other one is predicted for which we want to find the exclusion significance. To find distribution for events numbers for the predicted model, the signal events of it are scanned through the same BDT-model used for the observed model. In Figure 6.11, we show $\frac{E_b}{E_t}$ distribution for event numbers for observed and predicted models at 14 TeV LHC with 3 ab^{-1} of data. For the analysis, we have taken first 8 bins, starting from the left, of the $\frac{E_b}{E_t}$ distribution[r], given in Figure 6.11. We obtain an exclusion significance (Z) of 0.98 σ, when $S3 + B$ is taken as observed at the LHC and $R2 + B$ is considered as the predicted one. For the reverse case, we obtain Z value as 1.01 σ, see Table 6.9. As the exclusion significance is quite

[r]For the bins around z=0.8 and above, the b-jet energy is not very well measured. In this region, because of very high boost of top quark, b-jet gets contaminated with the other two light jets, originated from the hadronic decay of top quark.

low, it shows that two models can not be distinguished well at the LHC. However, it is prompting to see whether these two models can be distinguished at 27 TeV (HE-LHC) collider for the same mass of the leptoquark. To do this study, we assume that the shape of the signal and individual background distributions will remain same at the 27 TeV LHC as that of the 14 TeV collider. We then scale the distributions by overall factors after calculating their total cross sections at these two different center of mass energy colliders. In Figure 6.12, we show the plot for exclusion significance vs. required luminosity at the HE-LHC. We find that with moderate amount of luminosity (around 1800 fb^{-1}) at this collider, either of the models can be excluded at 5σ significance when the other one appears as observed. In the last column of Table 6.9, we show the exclusion significance for 3 ab^{-1} data at this collider.

$\mathcal{L}$	predicted	observed	Rejection Prob. (Z) (14 TeV)	Rejection Prob. (Z) (27 TeV)
$3ab^{-1}$	$R_2 + B$	$S_3 + B$	0.98 σ	6.45 σ
	$S_3 + B$	$R_2 + B$	1.01 σ	6.59 σ

Table 6.9: Probability of excluding one model when other model is the observed model at 14 TeV LHC and 27 TeV HE-LHC with $\mathcal{L} = 3\ ab^{-1}$.

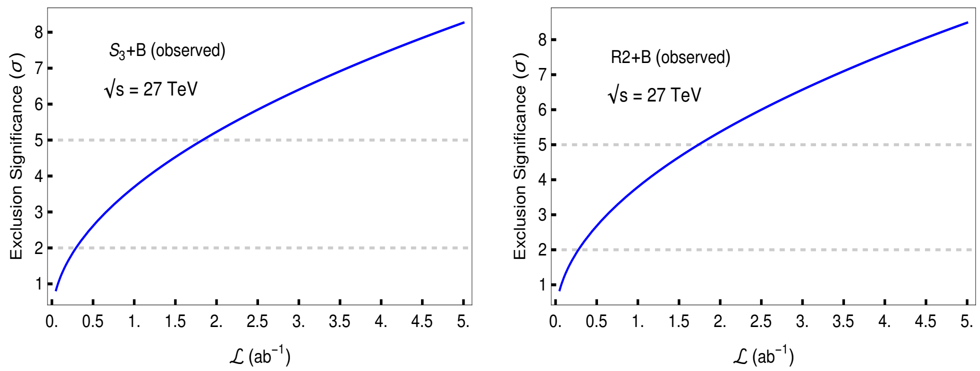


Figure 6.12: The exclusion significance vs. required luminosity at 27 TeV collider by projecting the distributions at 14 TeV collider to 27 TeV collider. The mass of the leptoquark has been taken to be $M_{\rm LQ} = 1300$ GeV.

6.6 Conclusions

TeV-scale leptoquarks that can emerge from various models are well-motivated and phenomenologically interesting to be searched at high-energy collider experiments. Present chapter investigates the pair production of third-generation

$\frac{2}{3}e$-charged scalar leptoquark at the LHC using NLO QCD accuracy, matched to parton shower for precise probing. Among different potential scalar leptoquark models, two primary interests - S_3 and R_2 can be probed by looking at their decay into a top with a tau neutrino, thus producing a compelling signature of a pair of top-like fatjets along with substantial missing transverse energy. Here tops, created from heavy leptoquarks, are naturally boosted and therefore considering them as boosted jets is quite meaningful.

With a precise understanding of jet physics, it is now possible to study the intrinsic substructure and properties of such jets, thereby pointing out the origin of these jets with a high degree of accuracy. Therefore the considered channel has excellent potential for separating the tiny signal from the overwhelming SM background. Parton shower effects are included in our study and its usefulness in the low transverse momentum region is seen in Figure 6.2. We also demonstrate that the factorization and renormalization scale uncertainties for the NLO+PS events are much lower than that of LO+PS events (see Figure 6.4 and Table 6.2).

For accurate prediction, we include all the relevant background processes with two to four extra QCD radiations and normalize them using the available higher-order QCD-corrected production cross section. Different high-level variables, such as MET, M_{eff}, $\Delta R(J_0, J_1)$, $\Delta\phi(J_i, \not{E}_T)$, jet substructure based pruned jet mass and N-subjettiness are proved to be efficacious to pinpoint the signal. Multivariate analysis is carried out individually for these two models and we show that at the 14 TeV LHC with an integrated luminosity of 3000 fb^{-1}, the leptoquarks of mass 1380 GeV can be discovered (5σ), and up to 1520 GeV can be excluded (2σ).

Among the two scalar leptoquark models considered here, it is interesting to note that the top quarks resulting from the decay of leptoquarks possess different chiralities. Most of the high-level variables utilized for multivariate analysis are not sensitive to this polarization. Only the jet mass variables acquire some minor effect due to the modified distribution pattern in the decay process. However, these are insignificant enough, thereby providing almost equivalent mass constraints for both models.

We further construct different polarization sensitive variables to distinguish these scalar leptoquark models of the same charge. We exhibit the effectiveness of such variables in terms of (i) an angular variable in the top quark's rest frame, (ii) the ratio of the energy variables $\frac{E_b}{E_t}$. Such effects are demonstrated at the truth level and after including parton shower and (fat)jet formation (see Figures 6.9, 6.10). Significant distortion is noticeable following detector simulation and (fat)jet formation. This is primarily attributed to the contamination and poor measurement efficiency of the b-jet momenta within a highly collimated top-like fatjet. The log-likelihood-ratio (LLR) hypothesis test is used to distinguish the models in the presence of combined background events. We find that the statistical exclusion significance remains low at around 1σ confidence level at the LHC. However, it is shown that the 27 TeV collider can play a promising role and it is estimated that the required luminosity would be around 300 fb^{-1} (1800 fb^{-1}) to distinguish these two models with 2σ (5σ) significance.

Chapter 7

Summary and future directions

The Standard Model (SM) successfully explains various experimental observations but fails to address phenomena such as the strong CP problem, dark matter (DM), and others. This thesis explores extensions of the SM to address some of these issues and proposes potential solutions within the framework of particle physics.

The ongoing study of the Higgs boson and the search for new physics at the LHC also requires high precision on the theoretical side. Computing higher-order corrections of physical observables within perturbation theory is crucial for achieving this precision. The strong coupling constant (α_S) varies with the energy scale and is small at the LHC's hard interaction energy scale, enabling a perturbative expansion of physical observables in terms of it. The leading-order (LO) predictions only provide approximate estimates and suffer from a large unphysical scale dependence. We include next-to-leading order (NLO) in QCD corrections in our analysis to get a more accurate cross section. NLO-QCD corrections have additional radiation, which alters differential distributions compared to LO predictions. Furthermore, at NLO, the dependence on unphysical renormalization and factorization scales diminishes, leading to more precise estimates of theoretical uncertainties. By incorporating these higher-order corrections, the precision of theoretical predictions is enhanced, improving our understanding of the underlying BSM physics at the LHC.

The total NLO fixed-order cross section is finite, but the differential NLO distributions can contain large logarithmic terms in specific phase spaces. To ensure reliable predictions, it is crucial to resum the large logarithmic terms and exponentiate them. Therefore we match the NLO fixed-order results with parton shower, which numerically resum the logarithmic terms and provide finite and reliable differential distributions across the entire phase space. By combining the NLO fixed-order calculation with parton shower, more accurate and precise predictions can be obtained for various processes at the LHC.

The non-observation of new physics signatures at the LHC has imposed constraints on the masses and couplings of BSM particles, confining their existence to the TeV to a few TeV mass ranges. Exploring new physics at these TeV scales presents significant challenges and necessitates using innovative techniques. In many BSM scenarios, heavy particles decay into Standard Model gauge bosons or top quarks. The subsequent hadronic decay of these particles predominantly leads to hadronic final states composed of jets. Investigating these final states offers several advantages, including a large hadronic branching ratio and the ability to extract the properties of the original jet through the reconstructed boosted jets. Indeed, hadronic final states face substantial challenges due to the overwhelming SM background, and QCD jets mimic the fatjets. Nonetheless, recent developments in jet substructure variables have proven valuable in extracting subtle signals from the overwhelming SM background by analyzing the internal structure of jets. One example is the N-subjettiness ratio, which provides insights into the radiation pattern of a jet and helps determine whether it exhibits a one-, two-, or three-prong structure, aiding in the discrimination of signal from the background.

Additionally, we use multivariate analysis (MVA), which outperforms the traditional cut-based analysis. MVA effectively combines many observables and defines a non-linear boundary to extract the signal from the background more efficiently.

In the first study (Chapter 3), we explore the inert Higgs-doublet model (IDM), an extension of the SM that provides a viable Higgs portal scalar dark matter candidate and heavier scalars with masses of 100 GeV or more (hierarchical region). The investigation focuses on the intriguing hierarchical mass spectrum of the IDM, which successfully accounts for the observed relic abundance and satisfies various theoretical, collider, and astrophysical constraints. We incorporate the NLO-QCD corrections of heavy scalar pair and associated production processes. $\mathcal{O}(\alpha_s)$ corrections to the gluon-gluon-Higgs effective coupling have been taken into account in this study wherever appropriate. We find that the effect of QCD correction is significant for encrypting the correct search strategy at the LHC. We also observed that LO+PS (parton shower) events normalized with the NLO cross section exhibit notable differences from the full NLO+PS events due to changes in differential distributions. We propose a novel signal process that involves two large-radii boosted jets and substantial missing transverse momentum (MET). A robust investigation of the hierarchical region of IDM is accomplished through the MVA using the boosted decision tree (BDT) algorithm. This analysis brings almost all of the parameter spaces of the hierarchical region of IDM well within reach of the 14 TeV LHC.

KSVZ axion model offers a natural solution to the strong CP problem. The QCD axion of the KSVZ model behaves as a dark matter (DM) candidate when the axion decay constant is appropriately tuned. We extend the model with an additional complex singlet scalar. This model offers a two-component dark matter scenario without fine-tuning. We explore this extended KSVZ model in Chapters 4 and 5. The colored vector-like quark (VLQ) present in the model plays a crucial role in dark matter and collider phenomenology, mediating interactions between the scalar DM and up-type SM quarks. We investigate various Yukawa couplings and identify parameter spaces that satisfy the observed relic density of DM and other experimental constraints. We study the NLO-QCD corrections for VLQ pair production at the LHC, finding a 30% increase in the total cross section compared to LO. Using two boosted top fatjets with large MET, we perform a multivariate analysis and obtain discovery and exclusion potential for a significant parameter space of the model with 139 fb^{-1} integrated luminosity at the 14 TeV LHC.

The search for leptoquarks is an active research area due to their potential to explain observed anomalies such as W-mass, muon g-2, $R_D(*)$, and various others [252]. In Chapter 6, we studied the pair production of third-generation scalar leptoquarks at the 14 TeV LHC. By considering next-to-leading order corrections and matching them with parton shower, we obtained precise predictions for the pair production process. The pair production of leptoquarks is independent of any specific BSM coupling, making it a model-independent search. We focused on S_3 and R_2 scalar leptoquark models and their decay into a top quark and a neutrino. We proposed a distinctive signature of top-like fatjets with missing transverse energy, offering a discovery potential of leptoquark mass up to 1380 GeV at the 14 TeV LHC through multivariate analysis. The top quark appearing from the decay of the S_3 leptoquark is left-chiral, while when it appears from R_2 is right-chiral. After obtaining the discovery potential of these two leptoquarks at the LHC, we use polarization variables sensitive to the top quark polarization to distinguish these two models.

This thesis explores BSM scenarios and incorporates NLO-QCD corrections for more accurate predictions. The analysis techniques have broad applicability to other BSM models and SM analyses, enhancing the discovery potential at the 14 TeV LHC. In future studies, we will explore other leptoquark scenarios in the NLO-QCD precision and study their collider phenomenology. Additionally, we plan to investigate the QCD and/or electroweak corrections of dark matter annihilation processes. These corrections will play a crucial role in gaining insights into the evolution of dark matter in the universe.

Appendix A

Altarelli-Parisi splitting function

Altarelli-Parisi splitting function at 4-dimension is given below.

$$P_{q\leftarrow q}(z) = C_F\Big[\frac{1+z^2}{(1-z)_+} + \frac{3}{2}\delta(1-z)\Big] \tag{A.1}$$

$$P_{g\leftarrow q}(z) = C_F\Big[\frac{1+(1-z)^2}{z}\Big] \tag{A.2}$$

$$P_{q\leftarrow g}(z) = T_R\Big[z^2 + (1-z)^2\Big] \tag{A.3}$$

$$P_{g\leftarrow g}(z) = 2C_A\Big[\frac{z}{(1-z)_+} + \frac{1-z}{z} + z(1-z)\Big] + b_0\delta(1-z)\ , \tag{A.4}$$

where, for $SU(3)$ $C_F = \frac{4}{3}$, $C_A = 3$, $T_R = \frac{1}{2}$, and $b_0 = \frac{11}{6}C_A - \frac{2}{3}n_f T_R$. The splitting function of gluon into quark and anti-quark will not have any divergences.

Appendix B

Reference Formulae

The relation between beta and gamma function is as follow

$$B(\alpha,\beta) = \int_0^1 dx\ x^{\alpha-1}(1-x)^{\beta-1} = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \tag{B.1}$$

$$\begin{aligned} \Gamma(1-\epsilon) &= -\epsilon\ \Gamma(-\epsilon) \\ \text{or, } \Gamma(-\epsilon) &= -\frac{1}{\epsilon}\Gamma(1-\epsilon) \end{aligned} \tag{B.2}$$

$$\begin{aligned} \Gamma(1+\epsilon) &= \epsilon\ \Gamma(\epsilon) = \epsilon\ (-1+\epsilon)\ \Gamma(-1+\epsilon) \\ \text{or, } \Gamma(-1+\epsilon) &= -\frac{\Gamma(1+\epsilon)}{\epsilon\ (1-\epsilon)} \end{aligned} \tag{B.3}$$

$$\begin{aligned} &\Gamma(2-\epsilon) = (1-\epsilon)\Gamma(1-\epsilon) \\ &\Gamma(3-2\epsilon) = (2-2\epsilon)\Gamma(2-2\epsilon) = (2-2\epsilon)(1-2\epsilon)\Gamma(1-2\epsilon) \end{aligned} \tag{B.4}$$

$$A^{\epsilon} = e^{\epsilon\ lnA} = 1 + \epsilon\ lnA + \frac{(\epsilon\ lnA)^2}{2!} + \cdots \tag{B.5}$$

$$\mathrm{Re}(-1)^{-\epsilon} = \mathrm{Re}[e^{-i\epsilon\pi}] = \mathrm{Re}[1 - i\epsilon\pi - \frac{1}{2}\pi^2\epsilon^2 + \cdots] = 1 - \frac{1}{2}\pi^2\epsilon^2 + \mathcal{O}(\epsilon^4) \tag{B.6}$$

B.1 Scalar integrals over Feynman parameters

Using the scalar integral [23] we have

A.

$$\begin{aligned} I_1 = &\ 2\int_0^1 dx \int_0^{1-x} dy \int \frac{d^dl}{(2\pi)^d}\frac{1}{(l^2-\Delta)^3} \\ = &\ 2\int_0^1 dx \int_0^{1-x} dy \frac{-i}{(4\pi)^{d/2}}\ \frac{\Gamma(3-\frac{d}{2})}{2!}(\frac{1}{\Delta})^{3-d/2} \end{aligned}$$

For $\Delta = -xy\hat{s}_{12}$, we have the following

$$\begin{aligned} I_1 &= 2\frac{-i}{(4\pi)^{d/2}}\frac{\Gamma(3-\frac{d}{2})}{2}(-\frac{1}{\hat{s}_{12}})^{3-d/2}\int_0^1 dx\ x^{d/2-3}\int_0^{1-x} dy\ y^{d/2-3} \\ &= \frac{-i\ \Gamma(3-\frac{d}{2})}{(4\pi)^{d/2}}(-\frac{1}{\hat{s}_{12}})^{3-d/2}\int_0^1 dx\ x^{d/2-3}\ \frac{(1-x)^{d/2-2}}{d/2-2} \end{aligned}$$

Using Equation B.1, and putting $d = 4-2\epsilon$, we have

$$\begin{aligned} I_1 &= \frac{-i\ \Gamma(1+\epsilon)}{(4\pi)^{2-\epsilon}}(-\frac{1}{\hat{s}_{12}})^{1+\epsilon}\ \frac{B(-\epsilon,\ 1-\epsilon)}{-\epsilon} \\ &= \frac{i}{(4\pi)^2}(-\frac{1}{\hat{s}_{12}})(-\frac{\hat{s}_{12}}{4\pi})^{-\epsilon}\ \frac{\Gamma(1+\epsilon)}{\epsilon}\frac{\Gamma(-\epsilon)\Gamma(1-\epsilon)}{\Gamma(1-2\epsilon)} \end{aligned}$$

Using Equation B.2, we have

$$\begin{aligned} I_1 &= -\frac{i}{(4\pi)^2}(-\frac{1}{\hat{s}_{12}})(-\frac{\hat{s}_{12}}{4\pi})^{-\epsilon}\ \frac{\Gamma(1+\epsilon)}{\epsilon}\frac{\Gamma(1-\epsilon)\Gamma(1-\epsilon)}{\epsilon\ \Gamma(1-2\epsilon)} \\ &= \frac{i}{(4\pi)^2}(\frac{1}{\hat{s}_{12}})(-\frac{\hat{s}_{12}}{4\pi})^{-\epsilon}\ \frac{1}{\epsilon^2}\frac{\Gamma(1+\epsilon)[\Gamma(1-\epsilon)]^2}{\Gamma(1-2\epsilon)} \end{aligned}$$

$$\boxed{I_1 = \frac{i}{(4\pi)^2}(\frac{1}{\hat{s}_{12}})(-\hat{s}_{12})^{-\epsilon}\ \frac{C_\Gamma}{\epsilon^2}} \tag{B.7}$$

$$\text{where, } C_\Gamma = (4\pi)^\epsilon\frac{\Gamma(1+\epsilon)[\Gamma(1-\epsilon)]^2}{\Gamma(1-2\epsilon)} \tag{B.8}$$

Triangle scalar integral over Feynman parameters gives double pole (soft-collinear).

B.

$$\begin{aligned} I_2 &= \ 2\int_0^1 dx\int_0^{1-x} dy\ \int\frac{d^dl}{(2\pi)^d}\frac{\Delta}{(l^2-\Delta)^3} \\ &= 2\int_0^1 dx\int_0^{1-x} dy\frac{-i}{(4\pi)^{d/2}}\ \frac{\Gamma(3-\frac{d}{2})}{2!}(\frac{1}{\Delta})^{2-d/2} \\ &= \frac{-i}{(4\pi)^{d/2}}\Gamma(3-\frac{d}{2})(-\frac{1}{\hat{s}_{12}})^{2-d/2}\int_0^1 dx\ x^{d/2-2}\int_0^{1-x} dy\ y^{d/2-2} \\ &= \frac{-i\ \Gamma(3-\frac{d}{2})}{(4\pi)^{d/2}}(-\hat{s}_{12})^{d/2-2}\int_0^1 dx\ x^{d/2-2}\ \frac{(1-x)^{d/2-1}}{d/2-1} \end{aligned}$$

Using Equation B.1, and putting $d = 4-2\epsilon$, we have

$$\begin{aligned} I_2 &= \frac{-i\ \Gamma(1+\epsilon)}{(4\pi)^{2-\epsilon}}(-\hat{s}_{12})^{-\epsilon}\ \frac{B(1-\epsilon,\ 2-\epsilon)}{1-\epsilon} \\ &= \frac{-i}{(4\pi)^2}(-\frac{\hat{s}_{12}}{4\pi})^{-\epsilon}\ \frac{\Gamma(1+\epsilon)}{(1-\epsilon)}\ \frac{\Gamma(1-\epsilon)\Gamma(2-\epsilon)}{\Gamma(3-2\epsilon)} \end{aligned}$$

Using Equation B.4, we have

$$I_2 = -\frac{i}{(4\pi)^2}(-\frac{\hat{s}_{12}}{4\pi})^{-\epsilon}\ \frac{\Gamma(1+\epsilon)}{(1-\epsilon)}\frac{\Gamma(1-\epsilon)(1-\epsilon)\Gamma(1-\epsilon)}{(2-2\epsilon)(1-2\epsilon)\ \Gamma(1-2\epsilon)}$$

$$\boxed{I_2 = -\frac{i}{(4\pi)^2}(-\hat{s}_{12})^{-\epsilon}\ \frac{C_\Gamma}{2(1-\epsilon)(1-2\epsilon)}} \tag{B.9}$$

I_2 is finite at $\epsilon \to 0$.

C.

$$\begin{aligned}
I_3 =&\ 2\int_0^1 dx \int_0^{1-x} dy\ \int \frac{d^dl}{(2\pi)^d}\frac{l^2}{(l^2-\Delta)^3} \\
=&\ 2\int_0^1 dx \int_0^{1-x} dy \frac{i}{(4\pi)^{d/2}}\ \frac{d}{2}\ \frac{\Gamma(2-\frac{d}{2})}{2!}(\frac{1}{\Delta})^{2-d/2} \\
=&\ \frac{i}{(4\pi)^{d/2}}\ \frac{d}{2}\ \Gamma(2-\frac{d}{2})(-\hat{s}_{12})^{d/2-2}\ \int_0^1 dx\ x^{d/2-2}\int_0^{1-x} dy\ y^{d/2-2} \\
=&\ \frac{i}{(4\pi)^{d/2}}\ \frac{d}{2}\ \Gamma(2-\frac{d}{2})(-\hat{s}_{12})^{d/2-2}\ \int_0^1 dx\ x^{d/2-2}\frac{(1-x)^{d/2-1}}{d/2-1} \\
=&\ \frac{i}{(4\pi)^2}\ \frac{d}{2}(-\frac{\hat{s}_{12}}{4\pi})^{-\epsilon}\ \Gamma(\epsilon)\ \frac{\Gamma(1-\epsilon)\Gamma(2-\epsilon)}{(1-\epsilon)\ \Gamma(3-2\epsilon)}
\end{aligned}$$

$$I_3 = \frac{i}{(4\pi)^2}\ \frac{d}{2}(-\frac{\hat{s}_{12}}{4\pi})^{-\epsilon}\ \frac{\Gamma(1+\epsilon)}{\epsilon}\ \frac{\Gamma(1-\epsilon)(1-\epsilon)\ \Gamma(1-\epsilon)}{(1-\epsilon)\ (2-2\epsilon)(1-2\epsilon)\Gamma(1-2\epsilon)}$$

$$\boxed{I_3 = \frac{i}{(4\pi)^2}(-\hat{s}_{12})^{-\epsilon}\ \frac{d}{4}\ \frac{1}{\epsilon}\ \frac{C_\Gamma}{(1-\epsilon)(1-2\epsilon)}} \tag{B.10}$$

I_3 has a single pole.

D. Bubble integral over the Feynman parameter

$$I_4 = \int_0^1 dx\ \int \frac{d^dl}{(2\pi)^d}\frac{1}{(l^2-\Delta)^2}, \qquad (\Delta = \ -x(1-x)\hat{s}_{12})$$

$$\begin{aligned}
I_4 =&\ \int_0^1 dx\ \frac{i}{(4\pi)^{d/2}}\ \frac{\Gamma(2-\frac{d}{2})}{1!}(\frac{1}{\Delta})^{2-d/2} \\
=&\ \frac{i}{(4\pi)^{d/2}}\ \Gamma(2-\frac{d}{2})(-\hat{s}_{12})^{d/2-2}\ \int_0^1 dx\ x^{d/2-2}(1-x)^{d/2-2} \\
=&\ \frac{i}{(4\pi)^{d/2}}\ \Gamma(2-\frac{d}{2})(-\hat{s}_{12})^{d/2-2}\ B(d/2-1,\ d/2-1) \\
=&\ \frac{i}{(4\pi)^2}\ (-\frac{\hat{s}_{12}}{4\pi})^{-\epsilon}\ \Gamma(\epsilon)\ \frac{\Gamma(1-\epsilon)\Gamma(1-\epsilon)}{\Gamma(2-2\epsilon)}
\end{aligned}$$

$$\boxed{I_4 = \frac{i}{(4\pi)^2}(-\hat{s}_{12})^{-\epsilon}\ \frac{1}{\epsilon}\ \frac{C_\Gamma}{(1-2\epsilon)}} \tag{B.11}$$

I_4 has a single pole.

E.

$$I_5 = \int_0^1 dx \int \frac{d^d l}{(2\pi)^d} \frac{\Delta}{(l^2-\Delta)^2}, \qquad (\Delta = \ - x(1-x)\hat{s}_{12})$$
$$= \int_0^1 dx \; \frac{i}{(4\pi)^{d/2}} \; \frac{\Gamma(2-\frac{d}{2})}{1!} (\frac{1}{\Delta})^{1-d/2}$$

$$I_5 = \frac{i}{(4\pi)^{d/2}} \; \Gamma(2-\frac{d}{2})(-\hat{s}_{12})^{d/2-1} \; B(d/2, \; d/2)$$
$$= \frac{i}{(4\pi)^2} \; (-\frac{\hat{s}_{12}}{4\pi})^{-\epsilon} \; (-\hat{s}_{12}) \; \Gamma(\epsilon) \; \frac{\Gamma(2-\epsilon)\Gamma(2-\epsilon)}{\Gamma(4-2\epsilon)}$$
$$= \frac{i}{(4\pi)^2} \; (-\frac{\hat{s}_{12}}{4\pi})^{-\epsilon} \; (-\hat{s}_{12}) \; \frac{\Gamma(1+\epsilon)}{\epsilon} \; \frac{(1-\epsilon)^2 \; \Gamma(1-\epsilon)\Gamma(1-\epsilon)}{(3-2\epsilon)(2-2\epsilon)(1-2\epsilon)\Gamma(1-2\epsilon)}$$

$$\boxed{I_5 = -\frac{i}{(4\pi)^2} \; \frac{\hat{s}_{12}}{6}(-\hat{s}_{12})^{-\epsilon} \; \frac{1}{\epsilon} \; \frac{(1-\epsilon) \; C_\Gamma}{(1-\frac{2\epsilon}{3})(1-2\epsilon)}} \tag{B.12}$$

F.

$$I_6 = \int_0^1 dx \int \frac{d^d l}{(2\pi)^d} \frac{l^2}{(l^2-\Delta)^2}, \qquad (\Delta = \ - x(1-x)\hat{s}_{12})$$
$$= \int_0^1 dx \; \frac{-i}{(4\pi)^{d/2}} \; \frac{d}{2} \; \frac{\Gamma(1-\frac{d}{2})}{1!} (\frac{1}{\Delta})^{1-d/2}$$

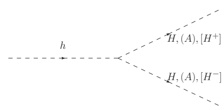
$$I_6 = -\frac{i}{(4\pi)^{d/2}} \; \frac{d}{2} \; \Gamma(1-\frac{d}{2})(-\hat{s}_{12})^{d/2-1} \; B(d/2, \; d/2)$$
$$= -\frac{i}{(4\pi)^2} \; \frac{d}{2} \; (-\frac{\hat{s}_{12}}{4\pi})^{-\epsilon} \; (-\hat{s}_{12}) \; \Gamma(-1+\epsilon) \; \frac{\Gamma(2-\epsilon)\Gamma(2-\epsilon)}{\Gamma(4-2\epsilon)}$$

Using Equation B.3, we have

$$\boxed{I_6 = -\frac{i}{(4\pi)^2} \; \frac{d}{2} \; \frac{\hat{s}_{12}}{6}(-\hat{s}_{12})^{-\epsilon} \; \frac{1}{\epsilon} \; \frac{C_\Gamma}{(1-\frac{2\epsilon}{3})(1-2\epsilon)}} \tag{B.13}$$

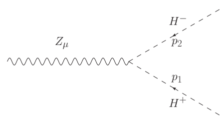
Appendix C

Feynman Rules of IDM



$$= -i\Lambda v = \begin{cases} -i\lambda_c v \ \text{ for } AA \\ -i\lambda_L v \ \text{ for } HH \\ -i\lambda_3 v \ \text{ for } H^+H^- \end{cases} \tag{C.1}$$

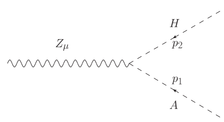
where $\lambda_{L/c} = (\lambda_3 + \lambda_4 \pm \lambda_5)$.



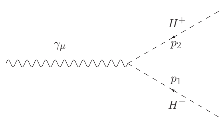
$$= i\, e\, (p_1 - p_2)_\mu \frac{\cos(2\theta_W)}{2\, C_W\, S_W} \tag{C.2}$$

where θ_W is the Weinberg angle. S_W and C_W correspond to the Sine and Cosine of the Weinberg angle respectively:

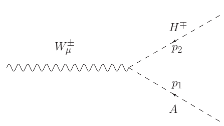
$$S_W = \sqrt{1 - (M_W/M_Z)^2} \ \text{ and } \ C_W = \sqrt{1 - S_W^2}\,. \tag{C.3}$$



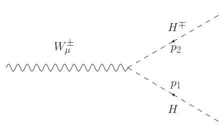
$$= e\, \frac{(p_1 - p_2)_\mu}{2\, C_W\, S_W}$$



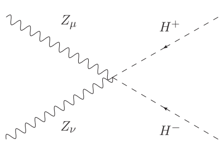
$$= i\, e\, (p_2 - p_1)_\mu \tag{C.4}$$



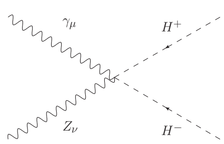
$$= -e\, \frac{(p_1 - p_2)_\mu}{2\, S_W}$$



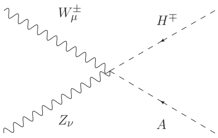
$$= \pm\, i\, e\, \frac{(p_1 - p_2)_\mu}{2\, S_W} \tag{C.5}$$



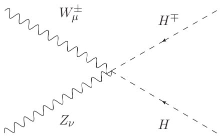
$$= i\, e^2\, \frac{\cos^2(2\theta_W)\, g_{\mu\nu}}{2\, C_W^2\, S_W^2}$$



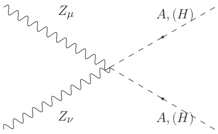
$$= i\, e^2\, \frac{\cos(2\theta_W)\, g_{\mu\nu}}{C_W\, S_W} \tag{C.6}$$



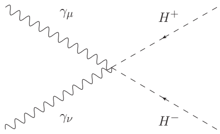
$$= \pm\, e^2 \,\frac{g_{\mu\nu}}{2\, C_W}$$



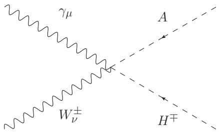
$$= -i\, e^2 \,\frac{g_{\mu\nu}}{2\, C_W} \tag{C.7}$$



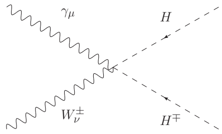
$$= i\, e^2 \,\frac{g_{\mu\nu}}{2\, C_W^2 S_W^2}$$



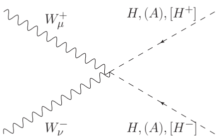
$$= i\; 2e^2 \; g_{\mu\nu} \tag{C.8}$$



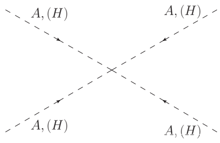
$$= \mp\, e^2 \,\frac{g_{\mu\nu}}{2\, S_W}$$



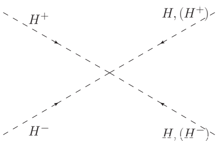
$$= i\, e^2 \,\frac{g_{\mu\nu}}{2\, S_W} \tag{C.9}$$



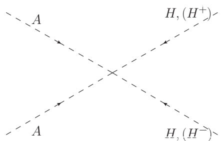
$$= i\, e^2 \,\frac{g_{\mu\nu}}{2\, S_W^2}$$



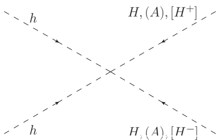
$$= -i\; 6\lambda_2 \tag{C.10}$$



$$= -i\; 2\lambda_2 \; (-i\; 4\lambda_2)$$



$$= -i\; 2\lambda_2 \tag{C.11}$$



$$= -i\; \lambda_L, (-i\; \lambda_c), [-i\; \lambda_3] \tag{C.12}$$

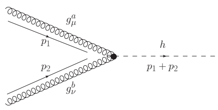
The effective Lagrangian accounts for Higgs interactions with gluons in the heavy top mass limit can be written as,

$$\mathcal{L}_{HEFT} = -\frac{1}{4} C_{eff}\; h\; G^a_{\mu\nu} G^{a\mu\nu} \; . \tag{C.13}$$

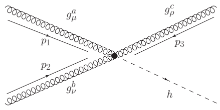
Here, $G^a_{\mu\nu}$ represents QCD field strength tensor and

$$C_{eff} = \frac{\alpha_s}{3\pi v}\; (1 + \frac{11}{4} \frac{\alpha_s}{\pi}) = C_0\; (1 + \frac{11}{4} \frac{\alpha_s}{\pi}) \; . \tag{C.14}$$

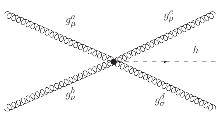
The Feynman rules for the different vertices of this effective Lagrangian are listed below.



$$= -iC_{eff}\delta^{ab}\,(p_1^\nu p_2^\mu - g^{\mu\nu}\;p_1.p_2) \tag{C.15}$$



$$\begin{aligned}= &- g_s\, C_{eff} f^{abc}[(p_1-p_2)^\rho\, g^{\mu\nu} \\ &+ (p_2-p_3)^\mu\, g^{\nu\rho} + (p_3-p_1)^\nu\, g^{\rho\mu}]\end{aligned} \tag{C.16}$$



$$\begin{aligned}= ig_s^2\, C_{eff}[&f^{abe} f^{cde}(g^{\mu\sigma} g^{\nu\rho} - g^{\mu\rho} g^{\nu\sigma}) \\ &+ f^{ace} f^{bde}(g^{\mu\sigma} g^{\nu\rho} - g^{\mu\nu} g^{\rho\sigma}) \\ &+ f^{ade} f^{bce}(g^{\mu\rho} g^{\nu\sigma} - g^{\mu\nu} g^{\rho\sigma})]\end{aligned} \tag{C.17}$$

Appendix D

Feynman Diagrams of top-philic hybrid KSVZ

- Annihilation channels of scalar dark matter S_1 are shown in Figure D.1.
- Co-annihilation channels of scalar dark matter S_1 are shown in Figure D.2.
- Annihilation channels of vector-like quark Ψ are shown in Figure D.3.
- Spin independent elastic scattering between dark matter (S_1) and nucleon channels are shown in Figure D.4.
- Diagrams contributing to the $D^0 - \bar{D}^0$ mixing are shown in Figure D.5.

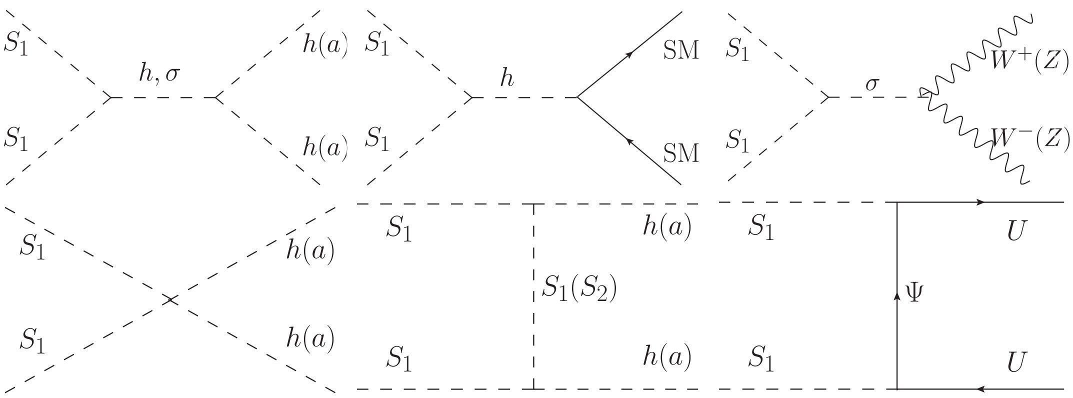


Figure D.1: Annihilation channels of scalar dark matter S_1. U denotes the SM up-type quark ($U \equiv u, c, t, \bar{u}, \bar{c}, \bar{t}$)

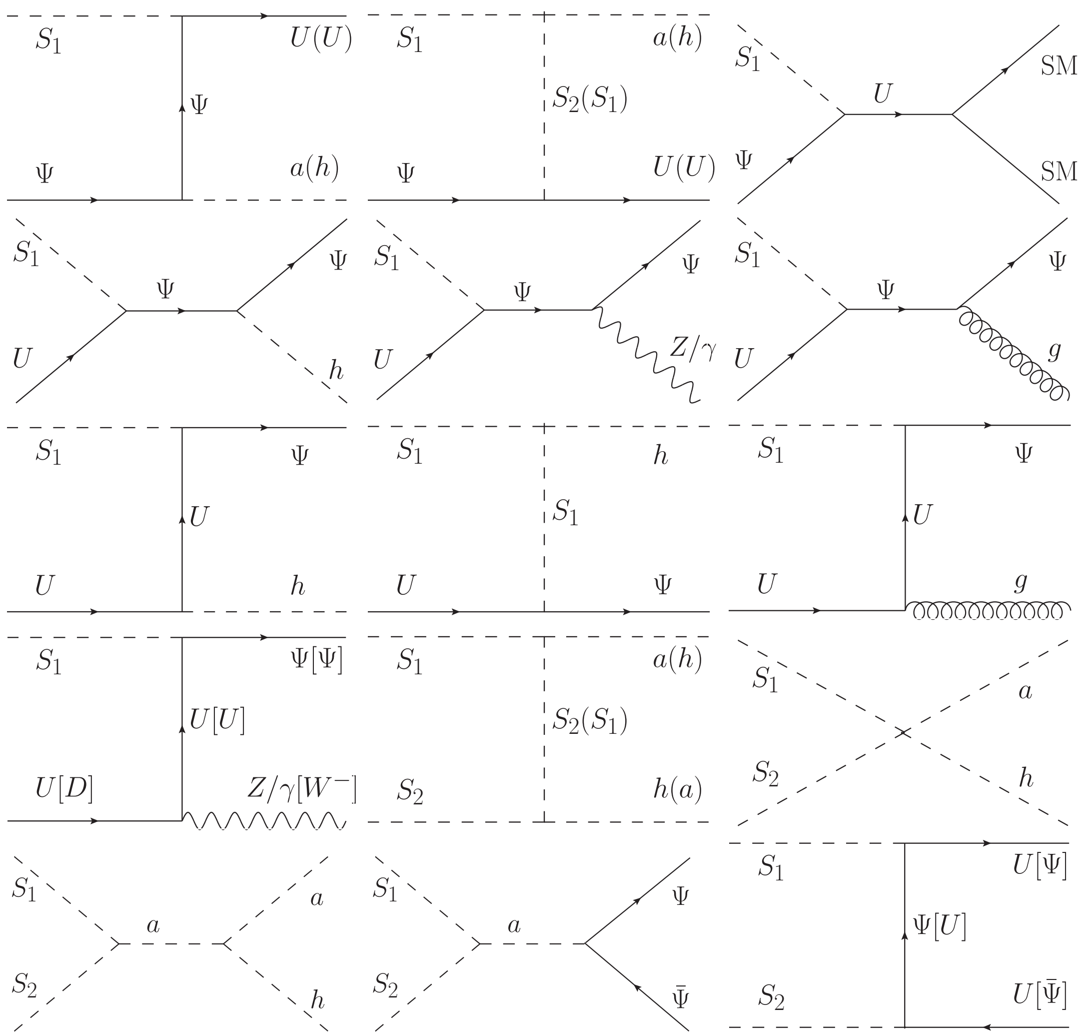


Figure D.2: Co-annihilation channels of scalar dark matter S_1. U and D denote the SM up-type and down-type quark, respectively; $U \equiv u, c, t, \bar{u}, \bar{c}, \bar{t}$, $D \equiv d, s, b, \bar{d}, \bar{s}, \bar{b}$

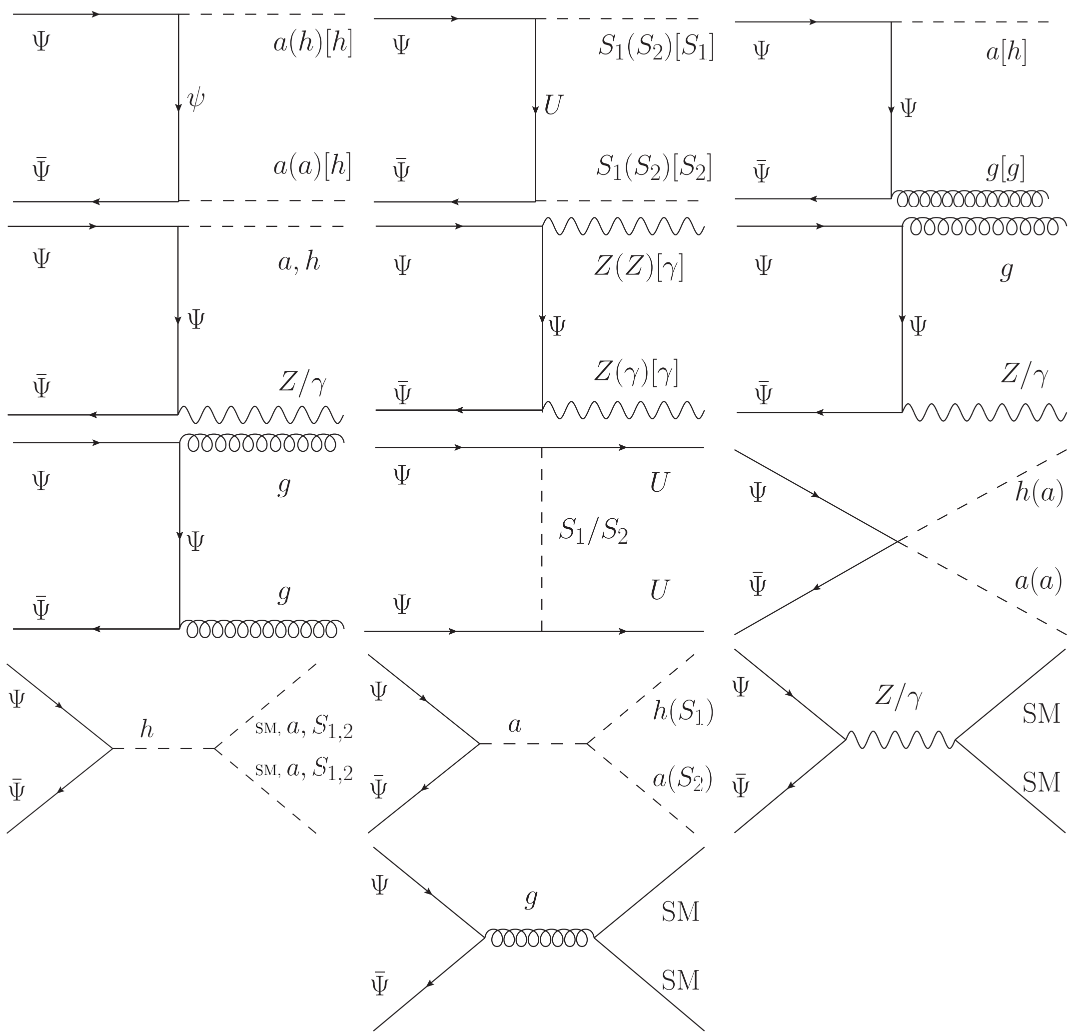


Figure D.3: Annihilation channels of vectorlike quark Ψ. U denotes the SM up-type quark ($U \equiv u, c, t, \bar{u}, \bar{c}, \bar{t}$)

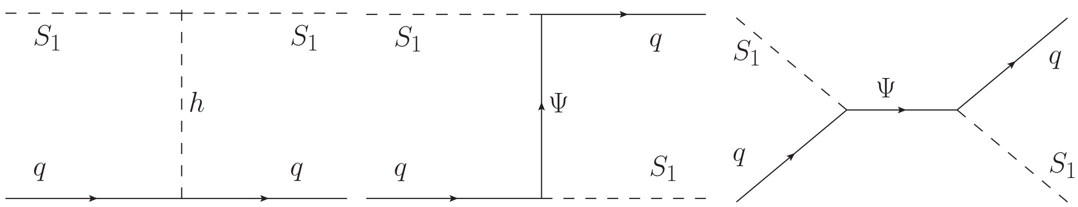


Figure D.4: Spin independent elastic scattering between dark matter (S_1) and nucleon

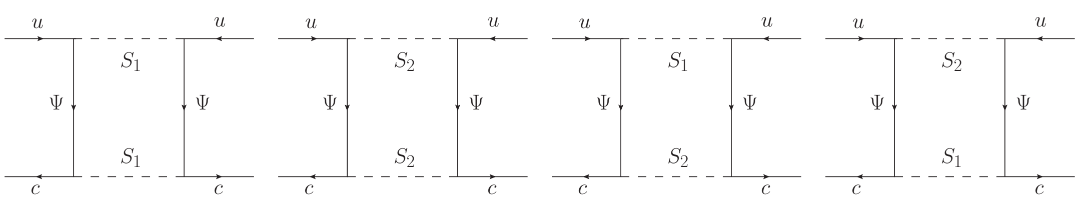


Figure D.5: Diagrams contributing to the $D^0 - \bar{D}^0$ mixing.

Appendix E

Direct detection channels of the extended KSVZ model

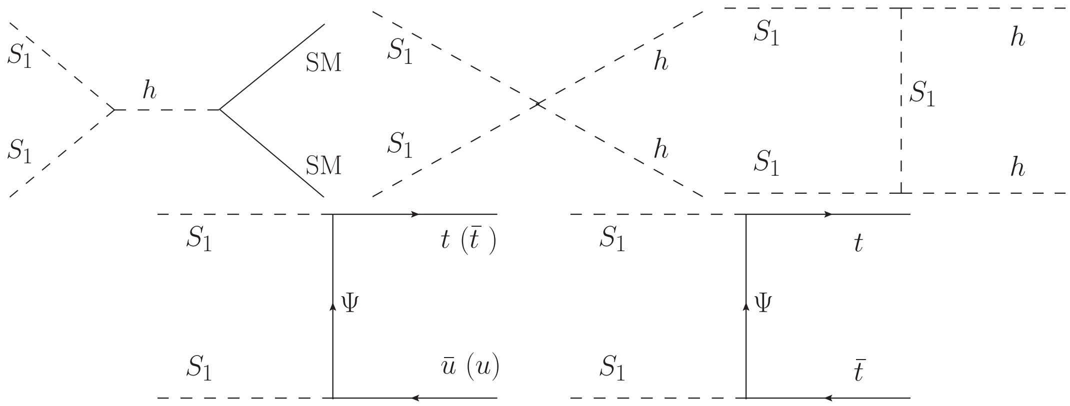


Figure E.1: Feynman diagrams of scalar dark matter annihilation into Standard Model particles are presented.

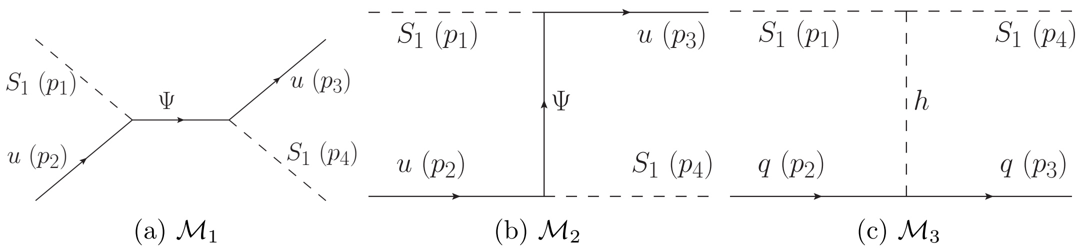


Figure E.2: Scattering diagrams between scalar dark matter and the nucleon.

Direct detection channels: Three different channels (Figure E.2) are possible at the tree level for the scattering process $S_1(p_1)\ u(p_2) \to S_1(p_4)\ u(p_3)$, VLQ-mediated s-channel, VLQ-mediated t-channel, and Higgs-mediated t-channel diagrams. The total cross section comprises the amplitude square of the individual

channels and the interference between different diagrams. The interference with different diagrams and the amplitude square of the individual diagrams are provided below.
Amplitude square of VLQ-mediated s-channel diagram:

$$\mathcal{M}_1^\dagger \mathcal{M}_1 = \frac{f^4}{4} \frac{\mathcal{N} + m_u^2 \ (p_1.p_2 + p_1.p_3)}{[(p_1+p_2)^2 - M_\Psi^2 \ ]^2} \tag{E.1}$$

Amplitude square of VLQ-mediated t-channel diagram:

$$\mathcal{M}_2^\dagger \mathcal{M}_2 = \frac{f^4}{4} \frac{\mathcal{N} - m_u^2 \ (p_1.p_2 + p_1.p_3)}{[(p_3-p_1)^2 - M_\Psi^2 \ ]^2} \tag{E.2}$$

Interference between VLQ-mediated s and t-channel diagrams:

$$2\mathcal{M}_1^\dagger \mathcal{M}_2 = -2 \times \frac{f^4}{4} \frac{\mathcal{N} + m_u^2 \ (-p_1.p_2 + p_1.p_3)}{[(p_1+p_2)^2 - M_\Psi^2 \ ][(p_3-p_1)^2 - M_\Psi^2 \ ]} \tag{E.3}$$

Where $p_2^2 = p_3^2 = m_u^2$, $p_1^2 = p_4^2 = M_{S_1}^2$, and $\mathcal{N}$ is given below.

$$\mathcal{N} = 2(p_1.p_3)(p_1.p_2) + M_{S_1}^2 \ (p_1.p_3 - p_1.p_2 - m_u^2) + m_u^4 \tag{E.4}$$

Amplitude square of Higgs-mediated t-channel diagram:

$$\mathcal{M}_3^\dagger \mathcal{M}_3 = 2m_q^2 \lambda_{SH}^2 \cos^2\theta \ \frac{p_1.p_3 - p_1.p_2 - 2m_u^2}{[(p_4-p_1)^2 - M_h^2 \ ]^2} \tag{E.5}$$

Interference between VLQ-mediated s-channel and Higgs-mediated t-channel diagrams:

$$2\mathcal{M}_1^\dagger \mathcal{M}_3 = 2m_u^2 \lambda_{SH} \cos\theta f^2 \ \frac{p_1.p_2 + m_u^2}{[(p_1+p_2)^2 - M_\Psi^2 \ ][(p_4-p_1)^2 - M_h^2 \ ]} \tag{E.6}$$

Interference between VLQ-mediated t-channel and Higgs-mediated t-channel diagrams:

$$2\mathcal{M}_2^\dagger \mathcal{M}_3 = - \ 2m_u^2 \lambda_{SH} \cos\theta f^2 \ \frac{p_1.p_3 - m_u^2}{[(p_3-p_1)^2 - M_\Psi^2 \ ][(p_4-p_1)^2 - M_h^2 \ ]} \tag{E.7}$$

Appendix F

Reference Formulae of top Polarization

F.0.1 Distribution of daughter of top quark

Here we will find the differential distribution of decay width of right handed top in its rest frame. The distribution for left handed particle can be obtained similarly.

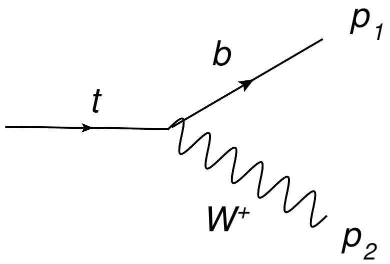


Figure F.1: The Feynman diagram for top decay.

The matrix element can be written as

$$M = \bar{u}_b(p_1)\frac{ig}{\sqrt{2}}\gamma_\mu P_L u_t(m_t)\epsilon^{\mu *}_{W^+}(p_2) \tag{F.1}$$

So, $|M|^2$ can be written as

$$\begin{aligned} |M|^2 &= \frac{g^2}{2}\bar{u}_b(p_1)\gamma_\mu P_L u_t(m_t)\bar{u}_t(m_t)P_R\gamma_\nu u_b(p_1)\epsilon^{\mu *}_{W^+}(p_2)\epsilon^{\nu}_{W^+}(p_2) \\ \sum_{final\ spins}|M|^2 &= -\frac{g^2}{2}Tr[\gamma_\mu P_L u_t(m_t)\bar{u}_t(m_t)P_R\gamma_\nu(\not{p}_1 + m_b)](g^{\mu\nu} - \frac{p^\mu p^\nu}{M_W^2}) \end{aligned} \tag{F.2}$$

In the Weyl basis, $\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix}$ and $\gamma^5 = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix}$, where $\sigma^\mu = (I, \boldsymbol{\sigma})$

and $\bar{\sigma}^\mu = (I, -\boldsymbol{\sigma})$. The spinor in the rest frame is given by

$$u_t^s(m_t) = \sqrt{m_t}\begin{pmatrix}\xi^s \\ \xi^s\end{pmatrix}$$

Using above expressions, Equation F.2 can be written as

$$\begin{aligned}\sum_{final\ \ spins}|M|^2 &= -\frac{g^2 m_t}{2}Tr[\begin{pmatrix}0 & 0\\ \bar{\sigma}_\mu & 0\end{pmatrix}\begin{pmatrix}\xi^s\xi^{s\dagger} & \xi^s\xi^{s\dagger}\\ \xi^s\xi^{s\dagger} & \xi^s\xi^{s\dagger}\end{pmatrix}\begin{pmatrix}0 & 0\\ \bar{\sigma}_\nu & 0\end{pmatrix}\not{p}_1](g^{\mu\nu} - \frac{p^\mu p^\nu}{M_W^2})\\ &= -\frac{g^2 m_t}{2}Tr[\bar{\sigma}^\mu\xi^s\xi^{s\dagger}\bar{\sigma}_\mu\sigma\cdot p_1 - \frac{\bar{\sigma}\cdot p_2\xi^s\xi^{s\dagger}\bar{\sigma}\cdot p_2\sigma\cdot p_1}{M_W^2}]\end{aligned} \tag{F.3}$$

For the spin up top $\xi^s\xi^{s\dagger}$ can be written as $\frac{I+\sigma^3}{2}$. So Equation F.3 can be written as

$$\sum_{final\ \ spins}|M|^2 = -\frac{g^2 m_t}{4}Tr[\bar{\sigma}^\mu(I+\sigma^3)\bar{\sigma}_\mu\sigma\cdot p_1 - \frac{\bar{\sigma}\cdot p_2(I+\sigma^3)\bar{\sigma}\cdot p_2\sigma\cdot p_1}{M_W^2}] \tag{F.4}$$

Using $\sigma^i\sigma^j + \sigma^j\sigma^i = 2\delta^{ij}I$, the following can be proven:

$$\bar{\sigma}^\mu\bar{\sigma}_\mu = -2I \tag{F.5}$$

$$\bar{\sigma}^\mu\sigma^3\bar{\sigma}_\mu = 2\sigma^3 \tag{F.6}$$

$$\bar{\sigma}\cdot p_2\bar{\sigma}\cdot p_2 = (p_2^0)^2 I + \vec{p_2}^2 I + 2\sigma^i p_2^i p_2^0 \tag{F.7}$$

$$\bar{\sigma}\cdot p_2\sigma^3\bar{\sigma}\cdot p_2 = (p_2^0)^2\sigma^3 + 2p_2^3 p_2^0 I - \sigma^3\vec{p_2}^2 + 2\sigma^i p_2^3 p_2^i \tag{F.8}$$

Again the following relations can be proven easily

$$Tr(\sigma^i\sigma^j) = 2\delta^{ij} \tag{F.9}$$

$$Tr(\sigma^i) = 0 \tag{F.10}$$

$$Tr(\sigma\cdot a) = 2a^0 \tag{F.11}$$

The different parts of Equation F.4 can be obtained using Equation F.5-Equation F.11. After using them, we have

$$Tr[\bar{\sigma}^\mu\bar{\sigma}_\mu\sigma\cdot p_1] = -4p_1^0 \tag{F.12}$$

$$Tr[\bar{\sigma}^\mu\sigma^3\bar{\sigma}_\mu\sigma\cdot p_1] = -4p_1^3 \tag{F.13}$$

$$Tr[\bar{\sigma}\cdot p_2\bar{\sigma}\cdot p_2\sigma\cdot p_1] = 2p_1^0(p_2^0)^2 + 2p_1^0\vec{p_2}^2 - 4\vec{p_1}\cdot\vec{p_2}p_2^0 \tag{F.14}$$

$$Tr[\bar{\sigma}\cdot p_2\sigma^3\bar{\sigma}\cdot p_2\sigma\cdot p_1] = -2(p_2^0)^2 p_1^3 + 4p_2^3 p_2^0 p_1^0 + 2p_1^3\vec{p_2}^2 - 4p_2^3\vec{p_1}\cdot\vec{p_2} \tag{F.15}$$

Using energy conservation in the top rest frame,

$$
\begin{aligned}
m_t &= |\vec{p}_1| + \sqrt{\vec{p}_1^{\,2} + m_W^2} \\
|\vec{p}_1| &= \frac{m_t^2 - m_W^2}{2m_t}
\end{aligned}
\tag{F.16}
$$

Equation F.14 can be written as

$$
\begin{aligned}
&2p_1^0(p_2^0)^2 + 2p_1^0\vec{p}_2^{\,2} - 4\vec{p}_1\cdot\vec{p}_2 p_2^0 \\
=&2p_1^0(m_W^2 + 2\vec{p}_1^{\,2}) + 4\vec{p}_1\cdot\vec{p}_1(m_t - p_1^0) \\
=&2|\vec{p}_1|(m_W^2 + 2|\vec{p}_1|m_t) = 2|\vec{p}_1|m_t^2
\end{aligned}
\tag{F.17}
$$

Equation F.15 can be written as

$$
\begin{aligned}
&-2(p_2^0)^2p_1^3 + 4p_2^3p_2^0p_1^0 + 2p_1^3\vec{p}_2^{\,2} - 4p_2^3\vec{p}_1\cdot\vec{p}_2 \\
=&-2p_1^3((p_2^0)^2 - \vec{p}_2^{\,2}) + 4p_2^3(p_2^0p_1^0 - \vec{p}_1\cdot\vec{p}_2) \\
=&-2p_1^3(m_W^2) - 2p_1^3(m_t^2 - m_W^2) = -2p_1^3m_t^2
\end{aligned}
\tag{F.18}
$$

Using the above formulae in Equation F.4, we have

$$
\begin{aligned}
\sum_{final\ spins}|M|^2 &= -\frac{g^2m_t}{4}[-4p_1^0 - 4p_1^3 - \frac{2|\vec{p}_1|m_t^2 - 2p_1^3m_t^2}{M_W^2}] \\
&= g^2m_t[(1+\frac{m_t^2}{2M_W^2}) + (1-\frac{m_t^2}{2M_W^2})\cos\theta_b] \\
&= g^2m_t(1+\frac{m_t^2}{2M_W^2})[1+\frac{2M_W^2 - m_t^2}{2M_W^2+m_t^2}\cos\theta_b] \\
&= g^2m_t(1+\frac{m_t^2}{2M_W^2})[1+k_b\ \cos\theta_b],
\end{aligned}
$$

where $k_b = \frac{2M_W^2 - m_t^2}{2M_W^2+m_t^2} = -0.4$, spin analyzing power of b-quark.
The differential distribution of decay width for right handed top quark is given

by

$$\begin{aligned}
\frac{d\Gamma}{d\cos\theta_b} &= \frac{1}{2m_t^2}\frac{|\vec{p}_1|}{8\pi}\sum_{final\ spins}|M|^2 \\
&= \frac{1}{2m_t^2}\frac{1}{8\pi}\frac{m_t^2-m_W^2}{2m_t}g^2 m_t(1+\frac{m_t^2}{2M_W^2})[1+k_b\ \cos\theta_b] \\
&= \frac{g^2}{32\pi}\frac{(m_t^2-m_W^2)(2M_W^2+m_t^2)}{m_t^2 m_W^2}[1+k_b\ \cos\theta_b]
\end{aligned}$$

For the spin down top, $\xi^s\xi^{s\dagger}$ can be written as $\frac{I-\sigma^3}{2}$. So it is easy follow that in the above expression there will be a minus sign in front of k_b for this case.

F.0.2 Relation between $\cos\ \theta_{\mathrm{b}}'$ and z

In the following, quantities in the lab frame will be denoted by unprimed symbols, whereas in the top rest frame they will be denoted by primed symbols[a]. So in the rest frame of the top quark the angle of bottom quark's direction of motion with the boost direction of top quark is given by

$$\cos\ \theta_{\mathrm{b}}' = \frac{p_b^{z'}}{|\vec{p_b}'|}, \tag{F.19}$$

where z and z' axes are along the direction of motion of the top quark in the lab frame.

Using Lorentz transformation between two frames with $\beta = \frac{|\vec{p}_t|}{E_t}$ and $\gamma = \frac{1}{\sqrt{1-\beta^2}} = \frac{E_t}{m_t}$

$$p_b^{z'} = -\gamma\beta E_b + \gamma p_b^z \tag{F.20}$$

Using energy conservation in the lab frame,

$$\begin{aligned}
E_t &= E_b + \sqrt{(\vec{p}_t-\vec{p}_b)^2+m_W^2} \\
p_b^z &= -\frac{m_t^2+m_b^2-m_W^2-2E_tE_b}{2|\vec{p}_t|} \\
&= \frac{E_b}{\beta} - \frac{m_t^2+m_b^2-m_W^2}{2\beta\gamma m_t}
\end{aligned} \tag{F.21}$$

Using Equation F.21 in Equation F.20, we have

[a]Note in the main text, we did not use any prime for the angle in the rest frame. So the $\cos\ \theta_{\mathrm{b}}'$ here is same as $\cos\ \theta_{\mathrm{b}}$ in the main text.

$$
\begin{aligned}
p_b^{z'} &= -\gamma\beta E_b + \gamma\frac{E_b}{\beta} - \frac{m_t^2 + m_b^2 - m_W^2}{2\beta m_t} \\
&= \frac{zE_t}{\gamma\beta} - \frac{m_t^2 + m_b^2 - m_W^2}{2\beta m_t} \\
&= \frac{zm_t}{\beta} - \frac{m_t^2 + m_b^2 - m_W^2}{2\beta m_t}
\end{aligned}
$$

Assuming $m_b = 0$,

$$
p_b^{z'} = \frac{1}{\beta}(zm_t - \frac{m_t^2 - m_W^2}{2m_t}) \tag{F.22}
$$

Using energy conservation in the top rest frame,

$$
\begin{aligned}
m_t &= E_b' + \sqrt{\vec{p_b}'^2 + m_W^2} \\
|\vec{p_b}'| &= \frac{m_t^2 - m_W^2}{2m_t}
\end{aligned} \tag{F.23}
$$

Using Equation F.22 and Equation F.23, in Equation F.19, we have

$$
\cos\theta_{\mathrm{b}}' = \frac{1}{\beta}(\frac{2m_t^2}{m_t^2 - m_W^2}z - 1) \tag{F.24}
$$

Bibliography

[1] S. L. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl. Phys.* **22** (1961) 579–588.

[2] N. Cabibbo, *Unitary symmetry and leptonic decays*, *Phys. Rev. Lett.* **10** (Jun, 1963) 531–533.

[3] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, *Phys. Rev. Lett.* **13** (Aug, 1964) 321–323.

[4] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Global conservation laws and massless particles*, *Phys. Rev. Lett.* **13** (Nov, 1964) 585–587.

[5] A. Salam and J. C. Ward, *Gauge theory of elementary interactions*, *Phys. Rev.* **136** (Nov, 1964) B763–B768.

[6] S. Weinberg, *A model of leptons*, *Phys. Rev. Lett.* **19** (Nov, 1967) 1264–1266.

[7] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652–657.

[8] D. J. Gross and F. Wilczek, *Asymptotically free gauge theories. i*, *Phys. Rev. D* **8** (Nov, 1973) 3633–3652.

[9] D. J. Gross and F. Wilczek, *Asymptotically free gauge theories. ii*, *Phys. Rev. D* **9** (Feb, 1974) 980–993.

[10] G. 't Hooft, *Renormalizable Lagrangians for Massive Yang-Mills Fields*, *Nucl. Phys. B* **35** (1971) 167–188.

[11] P. W. Anderson, *Plasmons, gauge invariance, and mass*, *Phys. Rev.* **130** (Apr, 1963) 439–442.

[12] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, *Phys. Rev. Lett.* **13** (Oct, 1964) 508–509.

[13] F. Herzog, B. Ruijl, T. Ueda, J. A. M. Vermaseren and A. Vogt, *The five-loop beta function of Yang-Mills theory with fermions*, *JHEP* **02** (2017) 090, [1701.01404].

[14] B. Odom, D. Hanneke, B. D'Urso and G. Gabrielse, *New measurement of the electron magnetic moment using a one-electron quantum cyclotron*, *Phys. Rev. Lett.* **97** (Jul, 2006) 030801.

[15] D. J. Gross and R. Jackiw, *Effect of anomalies on quasi-renormalizable theories*, *Phys. Rev. D* **6** (Jul, 1972) 477–493.

[16] C. Bouchiat, J. Iliopoulos and P. Meyer, *An Anomaly Free Version of Weinberg's Model*, *Phys. Lett. B* **38** (1972) 519–523.

[17] M. L. Perl, G. S. Abrams, A. M. Boyarski, M. Breidenbach, D. D. Briggs, F. Bulos et al., *Evidence for anomalous lepton production in* $e^+ - e^-$ *annihilation*, *Phys. Rev. Lett.* **35** (Dec, 1975) 1489–1492.

[18] P. Langacker and M. Luo, *Implications of precision electroweak experiments for* m_t, ρ_0, $sin^2\theta_W$, *and grand unification*, *Phys. Rev. D* **44** (Aug, 1991) 817–822.

[19] 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. B* **716** (2012) 1–29, [1207.7214].

[20] CMS collaboration, S. Chatrchyan et al., *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30–61, [1207.7235].

[21] Planck collaboration, N. Aghanim et al., *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6, [1807.06209].

[22] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, *Pisma Zh. Eksp. Teor. Fiz.* **5** (1967) 32–35.

[23] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995.

[24] F. Zwicky, *Die Rotverschiebung von extragalaktischen Nebeln*, *Helv. Phys. Acta* **6** (1933) 110–127.

[25] Planck collaboration, N. Aghanim et al., *Planck 2018 results. VI. Cosmological parameters*, 1807.06209.

[26] 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].

[27] G. Bélanger, F. Boudjema, A. Goudelis, A. Pukhov and B. Zaldivar, *micrOMEGAs5.0 : Freeze-in*, *Comput. Phys. Commun.* **231** (2018) 173–186, [1801.03509].

[28] J. D. Lewin and P. F. Smith, *Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil*, *Astropart. Phys.* **6** (1996) 87–112.

[29] D. G. Cerdeno and A. M. Green, *Direct detection of WIMPs*, 1002.1912.

[30] T. Lin, *Dark matter models and direct detection*, *PoS* **333** (2019) 009, [1904.07915].

[31] J. I. Read, *The Local Dark Matter Density*, *J. Phys. G* **41** (2014) 063101, [1404.1938].

[32] F. J. Kerr and D. Lynden-Bell, *Review of galactic constants*, *Mon. Not. Roy. Astron. Soc.* **221** (1986) 1023.

[33] M. C. Smith et al., *The RAVE Survey: Constraining the Local Galactic Escape Speed*, *Mon. Not. Roy. Astron. Soc.* **379** (2007) 755–772, [astro-ph/0611671].

[34] M. I. Gresham and K. M. Zurek, *Effect of nuclear response functions in dark matter direct detection*, *Phys. Rev. D* **89** (2014) 123521, [1401.3739].

[35] R. H. Helm, *Inelastic and elastic scattering of 187-mev electrons from selected even-even nuclei*, *Phys. Rev.* **104** (Dec, 1956) 1466–1475.

[36] J. R. Ellis, K. A. Olive and C. Savage, *Hadronic Uncertainties in the Elastic Scattering of Supersymmetric Dark Matter*, *Phys. Rev. D* **77** (2008) 065026, [0801.3656].

[37] J. Ellis, N. Nagata and K. A. Olive, *Uncertainties in WIMP Dark Matter Scattering Revisited*, *Eur. Phys. J. C* **78** (2018) 569, [1805.09795].

[38] M. Schumann, *Direct Detection of WIMP Dark Matter: Concepts and Status*, *J. Phys. G* **46** (2019) 103003, [1903.03026].

[39] J. Silk and M. Srednicki, *Cosmic-ray antiprotons as a probe of a photino-dominated universe*, *Phys. Rev. Lett.* **53** (Aug, 1984) 624–627.

[40] F. W. Stecker and A. J. Tylka, *The Cosmic Ray Anti-proton Spectrum From Dark Matter Annihilation and Its Astrophysical Implications: A New Look*, *Astrophys. J. Lett.* **336** (1989) L51–L54.

[41] H. Baer and S. Profumo, *Low energy antideuterons: shedding light on dark matter*, *JCAP* **12** (2005) 008, [astro-ph/0510722].

[42] F. Donato, N. Fornengo and D. Maurin, *Antideuteron fluxes from dark matter annihilation in diffusion models*, *Phys. Rev. D* **78** (Aug, 2008) 043506.

[43] J. E. Gunn, B. W. Lee, I. Lerche, D. N. Schramm and G. Steigman, *Some Astrophysical Consequences of the Existence of a Heavy Stable Neutral Lepton*, *Astrophys. J.* **223** (1978) 1015–1031.

[44] Y. B. Zeldovich, A. A. Klypin, M. Y. Khlopov and V. M. Chechetkin, *Astrophysical constraints on the mass of heavy stable neutral leptons*, *Sov. J. Nucl. Phys.* **31** (1980) 664–669.

[45] M. Cirelli, N. Fornengo, T. Montaruli, I. A. Sokalski, A. Strumia and F. Vissani, *Spectra of neutrinos from dark matter annihilations*, *Nucl. Phys. B* **727** (2005) 99–138, [hep-ph/0506298].

[46] K. Kohri, A. Mazumdar, N. Sahu and P. Stephens, *Probing Unified Origin of Dark Matter and Baryon Asymmetry at PAMELA/Fermi*, *Phys. Rev.* **D80** (2009) 061302, [0907.0622].

[47] Fermi-LAT, MAGIC collaboration, M. L. Ahnen et al., *Limits to dark matter annihilation cross-section from a combined analysis of MAGIC and Fermi-LAT observations of dwarf satellite galaxies*, *JCAP* **1602** (2016) 039, [1601.06590].

[48] B. Eiteneuer, A. Goudelis and J. Heisig, *The inert doublet model in the light of Fermi-LAT gamma-ray data – a global fit analysis*, 1705.01458.

[49] ATLAS collaboration, *Combination of searches for invisible decays of the Higgs boson using 139 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment*, 2301.10731.

[50] C. Abel, S. Afach, N. J. Ayres, C. A. Baker, G. Ban, G. Bison et al., *Measurement of the permanent electric dipole moment of the neutron*, *Phys. Rev. Lett.* **124** (Feb, 2020) 081803.

[51] A. Hook, *TASI Lectures on the Strong CP Problem and Axions*, *PoS* **TASI2018** (2019) 004, [1812.02669].

[52] H.-Y. Cheng, *The Strong CP Problem Revisited*, *Phys. Rept.* **158** (1988) 1.

[53] V. Baluni, *CP Violating Effects in QCD*, *Phys. Rev. D* **19** (1979) 2227–2230.

[54] M. Dine, *TASI lectures on the strong CP problem*, in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 2000): Flavor Physics for the Millennium*, pp. 349–369, 6, 2000. `hep-ph/0011376`.

[55] J. R. Ellis and M. K. Gaillard, *Strong and Weak CP Violation*, *Nucl. Phys. B* **150** (1979) 141–162.

[56] J. E. Kim, *Weak Interaction Singlet and Strong CP Invariance*, *Phys. Rev. Lett.* **43** (1979) 103.

[57] M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, *Can Confinement Ensure Natural CP Invariance of Strong Interactions?*, *Nucl. Phys. B* **166** (1980) 493–506.

[58] G. Raffelt and D. Seckel, *Bounds on Exotic Particle Interactions from SN 1987a*, *Phys. Rev. Lett.* **60** (1988) 1793.

[59] P. Langacker, *Grand Unified Theories and Proton Decay*, *Phys. Rept.* **72** (1981) 185.

[60] H. Georgi and S. L. Glashow, *Unity of All Elementary Particle Forces*, *Phys. Rev. Lett.* **32** (1974) 438–441.

[61] J. C. Pati and A. Salam, *Lepton Number as the Fourth Color*, *Phys. Rev.* **D10** (1974) 275–289.

[62] J. C. Pati and A. Salam, *Unified Lepton-Hadron Symmetry and a Gauge Theory of the Basic Interactions*, *Phys. Rev. D* **8** (1973) 1240–1251.

[63] J. Wudka, *COMPOSITE LEPTOQUARKS*, *Phys. Lett. B* **167** (1986) 337–342.

[64] "The large hadron collider." `https://home.cern/science/accelerators/large-hadron-collider`.

[65] ATLAS collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.

[66] CMS collaboration, S. Chatrchyan et al., *The CMS Experiment at the CERN LHC*, *JINST* **3** (2008) S08004.

[67] ALICE collaboration, K. Aamodt et al., *The ALICE experiment at the CERN LHC*, *JINST* **3** (2008) S08002.

[68] LHCb collaboration, A. A. Alves, Jr. et al., *The LHCb Detector at the LHC*, *JINST* **3** (2008) S08005.

[69] 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].

[70] 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].

[71] CMS collaboration, S. Chatrchyan et al., *Identification of b-Quark Jets with the CMS Experiment*, *JINST* **8** (2013) P04013, [1211.4462].

[72] CMS collaboration, V. Khachatryan et al., *The CMS trigger system*, *JINST* **12** (2017) P01020, [1609.02366].

[73] ATLAS collaboration, M. Aaboud et al., *Performance of the ATLAS Trigger System in 2015*, *Eur. Phys. J. C* **77** (2017) 317, [1611.09661].

[74] ATLAS collaboration, G. Aad et al., *Performance of the ATLAS muon triggers in Run 2*, *JINST* **15** (2020) P09015, [2004.13447].

[75] D. Bertolini, P. Harris, M. Low and N. Tran, *Pileup Per Particle Identification*, *JHEP* **10** (2014) 059, [1407.6013].

[76] M. Cacciari, G. P. Salam and G. Soyez, *SoftKiller, a particle-level pileup removal method*, *Eur. Phys. J. C* **75** (2015) 59, [1407.0408].

[77] ATLAS collaboration, G. Aad et al., *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, *Eur. Phys. J. C* **76** (2016) 581, [1510.03823].

[78] CMS collaboration, V. Khachatryan et al., *Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV*, *JINST* **12** (2017) P02014, [1607.03663].

[79] C. Frye, A. J. Larkoski, M. D. Schwartz and K. Yan, *Precision physics with pile-up insensitive observables*, 1603.06375.

[80] CMS collaboration, A. M. Sirunyan et al., *Pileup mitigation at CMS in 13 TeV data*, *JINST* **15** (2020) P09018, [2003.00503].

[81] S. Dawson, *Radiative corrections to Higgs boson production*, *Nucl. Phys. B* **359** (1991) 283–300.

[82] A. Djouadi, M. Spira and P. M. Zerwas, *Production of Higgs bosons in proton colliders: QCD corrections*, *Phys. Lett. B* **264** (1991) 440–446.

[83] R. V. Harlander and W. B. Kilgore, *Next-to-next-to-leading order Higgs production at hadron colliders*, *Phys. Rev. Lett.* **88** (2002) 201801, [hep-ph/0201206].

[84] R. K. Ellis and S. Veseli, *Strong radiative corrections to W b anti-b production in p anti-p collisions*, *Phys. Rev. D* **60** (1999) 011501, [hep-ph/9810489].

[85] F. Febres Cordero, L. Reina and D. Wackeroth, *NLO QCD corrections to W boson production with a massive b-quark jet pair at the Tevatron p anti-p collider*, *Phys. Rev. D* **74** (2006) 034007, [hep-ph/0606102].

[86] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, *Sov. J. Nucl. Phys.* **15** (1972) 438–450.

[87] L. N. Lipatov, *The parton model and perturbation theory*, *Yad. Fiz.* **20** (1974) 181–198.

[88] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, *Sov. Phys. JETP* **46** (1977) 641–653.

[89] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl. Phys. B* **126** (1977) 298–318.

[90] G. C. Blazey et al., *Run II jet physics*, in *Physics at Run II: QCD and Weak Boson Physics Workshop: Final General Meeting*, pp. 47–77, 5, 2000. hep-ex/0005012.

[91] G. P. Salam and G. Soyez, *A Practical Seedless Infrared-Safe Cone jet algorithm*, *JHEP* **05** (2007) 086, [0704.0292].

[92] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, *Eur. Phys. J. C* **72** (2012) 1896, [1111.6097].

[93] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, *Longitudinally invariant K_t clustering algorithms for hadron hadron collisions*, Nucl. Phys. B **406** (1993) 187–224.

[94] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. D **48** (1993) 3160–3166, [hep-ph/9305266].

[95] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, *Better jet clustering algorithms*, JHEP **08** (1997) 001, [hep-ph/9707323].

[96] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross-sections in deep inelastic scattering*, in *Workshop on Monte Carlo Generators for HERA Physics (Plenary Starting Meeting)*, pp. 270–279, 4, 1998. hep-ph/9907280.

[97] G. P. Salam, *Towards Jetography*, Eur. Phys. J. C **67** (2010) 637–686, [0906.1833].

[98] M. Cacciari, G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, [0802.1189].

[99] L. G. Almeida, S. J. Lee, G. Perez, I. Sung and J. Virzi, *Top Jets at the LHC*, Phys. Rev. D **79** (2009) 074012, [0810.0934].

[100] S. Catani, L. Trentadue, G. Turnock and B. R. Webber, *Resummation of large logarithms in e+ e- event shape distributions*, Nucl. Phys. B **407** (1993) 3–42.

[101] M. Dasgupta and G. P. Salam, *Resummation of nonglobal QCD observables*, Phys. Lett. B **512** (2001) 323–330, [hep-ph/0104277].

[102] D. Krohn, J. Thaler and L.-T. Wang, *Jet Trimming*, JHEP **02** (2010) 084, [0912.1342].

[103] S. D. Ellis, C. K. Vermilion and J. R. Walsh, *Techniques for improved heavy particle searches with jet substructure*, Phys. Rev. D **80** (2009) 051501, [0903.5081].

[104] S. D. Ellis, C. K. Vermilion and J. R. Walsh, *Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches*, Phys. Rev. D **81** (2010) 094023, [0912.0033].

[105] J. M. Butterworth, A. R. Davison, M. Rubin and G. P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, Phys. Rev. Lett. **100** (2008) 242001, [0802.2470].

[106] J. Thaler and K. Van Tilburg, *Identifying Boosted Objects with N-subjettiness*, *JHEP* **03** (2011) 015, [1011.2268].

[107] J. Thaler and K. Van Tilburg, *Maximizing Boosted Top Identification by Minimizing N-subjettiness*, *JHEP* **02** (2012) 093, [1108.2701].

[108] I. W. Stewart, F. J. Tackmann and W. J. Waalewijn, *N-Jettiness: An Inclusive Event Shape to Veto Jets*, *Phys. Rev. Lett.* **105** (2010) 092002, [1004.2489].

[109] B. P. Roe, H.-J. Yang, J. Zhu, Y. Liu, I. Stancu and G. McGregor, *Boosted decision trees, an alternative to artificial neural networks*, *Nucl. Instrum. Meth. A* **543** (2005) 577–584, [physics/0408124].

[110] Y. Freund, *Boosting a weak learning algorithm by majority*, *Information and Computation* **121** (1995) 256–285.

[111] Y. Freund and R. E. Schapire, *Experiments with a new boosting algorithm*, in *International Conference on Machine Learning*, 1996.

[112] A. Hocker et al., *TMVA - Toolkit for Multivariate Data Analysis*, physics/0703039.

[113] R. Barbieri, L. J. Hall and V. S. Rychkov, *Improved naturalness with a heavy Higgs: An Alternative road to LHC physics*, *Phys. Rev. D* **74** (2006) 015007, [hep-ph/0603188].

[114] M. Cirelli, N. Fornengo and A. Strumia, *Minimal dark matter*, *Nucl. Phys. B* **753** (2006) 178–194, [hep-ph/0512090].

[115] A. Goudelis, B. Herrmann and O. Stål, *Dark matter in the Inert Doublet Model after the discovery of a Higgs-like boson at the LHC*, *JHEP* **09** (2013) 106, [1303.3010].

[116] N. Blinov, J. Kozaczuk, D. E. Morrissey and A. de la Puente, *Compressing the Inert Doublet Model*, *Phys. Rev. D* **93** (2016) 035020, [1510.08069].

[117] M. A. Díaz, B. Koch and S. Urrutia-Quiroga, *Constraints to Dark Matter from Inert Higgs Doublet Model*, *Adv. High Energy Phys.* **2016** (2016) 8278375, [1511.04429].

[118] A. Belyaev, G. Cacciapaglia, I. P. Ivanov, F. Rojas-Abatte and M. Thomas, *Anatomy of the Inert Two Higgs Doublet Model in the light of the LHC and non-LHC Dark Matter Searches*, *Phys. Rev. D* **97** (2018) 035011, [1612.00511].

[119] A. Bhardwaj, P. Konar, T. Mandal and S. Sadhukhan, *Probing the inert doublet model using jet substructure with a multivariate analysis*, *Phys. Rev. D* **100** (2019) 055040, [1905.04195].

[120] I. F. Ginzburg, K. A. Kanishev, M. Krawczyk and D. Sokolowska, *Evolution of Universe to the present inert phase*, *Phys. Rev. D* **82** (2010) 123533, [1009.4593].

[121] B. Świeżewska, *Yukawa independent constraints for two-Higgs-doublet models with a 125 GeV Higgs boson*, *Phys. Rev. D* **88** (2013) 055027, [1209.5725].

[122] A. Arhrib, R. Benbrik and N. Gaur, *$H \to \gamma\gamma$ in Inert Higgs Doublet Model*, *Phys. Rev. D* **85** (2012) 095021, [1201.2644].

[123] M. E. Peskin and T. Takeuchi, *Estimation of oblique electroweak corrections*, *Phys. Rev. D* **46** (1992) 381–409.

[124] J. Haller, A. Hoecker, R. Kogler, K. Mönig, T. Peiffer and J. Stelzer, *Update of the global electroweak fit and constraints on two-Higgs-doublet models*, *Eur. Phys. J. C* **78** (2018) 675, [1803.01853].

[125] E. Lundstrom, M. Gustafsson and J. Edsjo, *The Inert Doublet Model and LEP II Limits*, *Phys. Rev. D* **79** (2009) 035013, [0810.3924].

[126] G. Belanger, B. Dumont, A. Goudelis, B. Herrmann, S. Kraml and D. Sengupta, *Dilepton constraints in the Inert Doublet Model from Run 1 of the LHC*, *Phys. Rev. D* **91** (2015) 115011, [1503.07367].

[127] A. Pierce and J. Thaler, *Natural Dark Matter from an Unnatural Higgs Boson and New Colored Particles at the TeV Scale*, *JHEP* **08** (2007) 026, [hep-ph/0703056].

[128] ATLAS collaboration, G. Aad et al., *Search for charginos and neutralinos in final states with two boosted hadronically decaying bosons and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev. D* **104** (2021) 112010, [2108.07586].

[129] S. Banerjee, F. Boudjema, N. Chakrabarty and H. Sun, *Relic density of dark matter in the inert doublet model beyond leading order for the low mass region: 1. Renormalisation and constraints*, *Phys. Rev. D* **104** (2021) 075002, [2101.02165].

[130] Particle Data Group collaboration, M. Tanabashi et al., *Review of Particle Physics*, *Phys. Rev. D* **98** (2018) 030001.

[131] ATLAS collaboration, G. Aad et al., *Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment*, Eur. Phys. J. C **76** (2016) 6, [1507.04548].

[132] CMS collaboration, A. M. Sirunyan et al., *Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **11** (2018) 185, [1804.02716].

[133] ATLAS collaboration, *Combination of searches for invisible Higgs boson decays with the ATLAS experiment*, .

[134] WMAP collaboration, G. Hinshaw et al., *Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results*, Astrophys. J. Suppl. **208** (2013) 19, [1212.5226].

[135] Planck collaboration, P. A. R. Ade et al., *Planck 2013 results. XVI. Cosmological parameters*, Astron. Astrophys. **571** (2014) A16, [1303.5076].

[136] Planck collaboration, P. A. R. Ade et al., *Planck 2015 results. XIII. Cosmological parameters*, Astron. Astrophys. **594** (2016) A13, [1502.01589].

[137] Y. Mambrini, *Higgs searches and singlet scalar dark matter: Combined constraints from XENON 100 and the LHC*, Phys. Rev. D **84** (2011) 115017, [1108.0671].

[138] J. Giedt, A. W. Thomas and R. D. Young, *Dark matter, the CMSSM and lattice QCD*, Phys. Rev. Lett. **103** (2009) 201802, [0907.4177].

[139] LUX collaboration, D. S. Akerib et al., *Results from a search for dark matter in the complete LUX exposure*, Phys. Rev. Lett. **118** (2017) 021303, [1608.07648].

[140] XENON100 collaboration, E. Aprile et al., *Dark Matter Results from 225 Live Days of XENON100 Data*, Phys. Rev. Lett. **109** (2012) 181301, [1207.5988].

[141] CMS collaboration, V. Khachatryan et al., *Search for disappearing tracks in proton-proton collisions at $\sqrt{s} = 8$ TeV*, JHEP **01** (2015) 096, [1411.6006].

[142] C. Degrande, *Automatic evaluation of UV and R2 terms for beyond the Standard Model Lagrangians: a proof-of-principle*, *Comput. Phys. Commun.* **197** (2015) 239–262, [1406.3030].

[143] S. Frixione, Z. Kunszt and A. Signer, *Three jet cross-sections to next-to-leading order*, *Nucl. Phys. B* **467** (1996) 399–442, [hep-ph/9512328].

[144] G. Ossola, C. G. Papadopoulos and R. Pittau, *Reducing full one-loop amplitudes to scalar integrals at the integrand level*, *Nucl. Phys. B* **763** (2007) 147–169, [hep-ph/0609007].

[145] P. Artoisenet, R. Frederix, O. Mattelaer and R. Rietkerk, *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, *JHEP* **03** (2013) 015, [1212.3460].

[146] T. Sjostrand, L. Lonnblad and S. Mrenna, *PYTHIA 6.2: Physics and manual*, hep-ph/0108264.

[147] 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].

[148] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **06** (2002) 029, [hep-ph/0204244].

[149] R. D. Ball et al., *Parton distributions with LHC data*, *Nucl. Phys. B* **867** (2013) 244–289, [1207.1303].

[150] M. L. Mangano, M. Moretti, F. Piccinini and M. Treccani, *Matching matrix elements and shower evolution for top-quark production in hadronic collisions*, *JHEP* **01** (2007) 013, [hep-ph/0611129].

[151] S. Hoeche, F. Krauss, N. Lavesson, L. Lonnblad, M. Mangano, A. Schalicke et al., *Matching parton showers and matrix elements*, in *HERA and the LHC: A Workshop on the Implications of HERA for LHC Physics: CERN - DESY Workshop 2004/2005 (Midterm Meeting, CERN, 11-13 October 2004; Final Meeting, DESY, 17-21 January 2005)*, pp. 288–289, 2005. hep-ph/0602031. DOI.

[152] P. Konar, K. Kong and K. T. Matchev, $\sqrt{\hat{s}}_{min}$ *: A Global inclusive variable for determining the mass scale of new physics in events with missing energy at hadron colliders*, *JHEP* **03** (2009) 085, [0812.1042].

[153] P. Konar, K. Kong, K. T. Matchev and M. Park, *RECO level $\sqrt{s}_{min}$ and subsystem $\sqrt{s}_{min}$: Improved global inclusive variables for measuring the new physics mass scale in $\not{E}_T$ events at hadron colliders*, *JHEP* **06** (2011) 041, [1006.0653].

[154] A. J. Barr, T. J. Khoo, P. Konar, K. Kong, C. G. Lester, K. T. Matchev et al., *Guide to transverse projections and mass-constraining variables*, *Phys. Rev. D* **84** (2011) 095031, [1105.2977].

[155] R. D. Peccei and H. R. Quinn, *CP Conservation in the Presence of Instantons*, *Phys. Rev. Lett.* **38** (1977) 1440–1443.

[156] R. D. Peccei and H. R. Quinn, *Constraints Imposed by CP Conservation in the Presence of Instantons*, *Phys. Rev. D* **16** (1977) 1791–1797.

[157] J. Preskill, M. B. Wise and F. Wilczek, *Cosmology of the Invisible Axion*, *Phys. Lett. B* **120** (1983) 127–132.

[158] L. F. Abbott and P. Sikivie, *A Cosmological Bound on the Invisible Axion*, *Phys. Lett. B* **120** (1983) 133–136.

[159] M. Dine and W. Fischler, *The Not So Harmless Axion*, *Phys. Lett. B* **120** (1983) 137–141.

[160] D. J. E. Marsh, *Axion Cosmology*, *Phys. Rept.* **643** (2016) 1–79, [1510.07633].

[161] F. Wilczek, *Problem of Strong P and T Invariance in the Presence of Instantons*, *Phys. Rev. Lett.* **40** (1978) 279–282.

[162] S. Weinberg, *A New Light Boson?*, *Phys. Rev. Lett.* **40** (1978) 223–226.

[163] M. Dine, W. Fischler and M. Srednicki, *A Simple Solution to the Strong CP Problem with a Harmless Axion*, *Phys. Lett. B* **104** (1981) 199–202.

[164] A. R. Zhitnitsky, *On Possible Suppression of the Axion Hadron Interactions. (In Russian)*, *Sov. J. Nucl. Phys.* **31** (1980) 260.

[165] LUX collaboration, D. S. Akerib et al., *Results from a search for dark matter in the complete LUX exposure*, *Phys. Rev. Lett.* **118** (2017) 021303, [1608.07648].

[166] PandaX-II collaboration, A. Tan et al., *Dark Matter Results from First 98.7 Days of Data from the PandaX-II Experiment*, *Phys. Rev. Lett.* **117** (2016) 121303, [1607.07400].

[167] PandaX-II collaboration, X. Cui et al., *Dark Matter Results From 54-Ton-Day Exposure of PandaX-II Experiment*, *Phys. Rev. Lett.* **119** (2017) 181302, [1708.06917].

[168] XENON collaboration, E. Aprile et al., *Dark Matter Search Results from a One Ton-Year Exposure of XENON1T*, *Phys. Rev. Lett.* **121** (2018) 111302, [1805.12562].

[169] M. Garny, A. Ibarra, S. Rydbeck and S. Vogl, *Majorana Dark Matter with a Coloured Mediator: Collider vs Direct and Indirect Searches*, *JHEP* **06** (2014) 169, [1403.4634].

[170] O. Gedalia, Y. Grossman, Y. Nir and G. Perez, *Lessons from Recent Measurements of D0 - anti-D0 Mixing*, *Phys. Rev. D* **80** (2009) 055024, [0906.1879].

[171] T. Robens and T. Stefaniak, *LHC Benchmark Scenarios for the Real Higgs Singlet Extension of the Standard Model*, *Eur. Phys. J. C* **76** (2016) 268, [1601.07880].

[172] ATLAS Collaboration collaboration, *Combination of searches for invisible Higgs boson decays with the ATLAS experiment*, tech. rep., CERN, Geneva, Oct, 2020.

[173] F. Giacchino, A. Ibarra, L. Lopez Honorez, M. H. G. Tytgat and S. Wild, *Signatures from Scalar Dark Matter with a Vector-like Quark Mediator*, *JCAP* **1602** (2016) 002, [1511.04452].

[174] OPAL collaboration, G. Abbiendi et al., *Search for scalar top and scalar bottom quarks at LEP*, *Phys. Lett. B* **545** (2002) 272–284, [hep-ex/0209026].

[175] ATLAS collaboration, M. Marjanovic, *Search for squarks and gluinos with the ATLAS detector in final states with jets and missing transverse momentum using 20.3 fb^{-1} of $\sqrt{s} = 8$ TeV proton-proton collision data*, in *2nd Large Hadron Collider Physics Conference*, 8, 2014. 1408.5857.

[176] ATLAS collaboration, M. Aaboud et al., *Search for top-squark pair production in final states with one lepton, jets, and missing transverse momentum using 36 fb^{1} of $\sqrt{s} = 13$ TeV pp collision data with the ATLAS detector*, *JHEP* **06** (2018) 108, [1711.11520].

[177] ATLAS collaboration, M. Aaboud et al., *Search for a scalar partner of the top quark in the jets plus missing transverse momentum final state at*

$\sqrt{s}$=*13 TeV with the ATLAS detector*, *JHEP* **12** (2017) 085, [1709.04183].

[178] ATLAS collaboration, M. Aaboud et al., *Search for supersymmetry in events with b-tagged jets and missing transverse momentum in pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, *JHEP* **11** (2017) 195, [1708.09266].

[179] ATLAS collaboration, M. Aaboud et al., *Search for direct top squark pair production in final states with two leptons in* $\sqrt{s} = 13$ *TeV pp collisions with the ATLAS detector*, *Eur. Phys. J. C* **77** (2017) 898, [1708.03247].

[180] ATLAS collaboration, M. Aaboud et al., *Search for supersymmetry in final states with two same-sign or three leptons and jets using 36 fb*$^{-1}$ *of* $\sqrt{s} = 13$ *TeV pp collision data with the ATLAS detector*, *JHEP* **09** (2017) 084, [1706.03731].

[181] CMS collaboration, A. M. Sirunyan et al., *Search for new phenomena with the* M_{T2} *variable in the all-hadronic final state produced in proton–proton collisions at* $\sqrt{s} = 13$ *TeV*, *Eur. Phys. J. C* **77** (2017) 710, [1705.04650].

[182] CMS collaboration, A. M. Sirunyan et al., *Search for direct production of supersymmetric partners of the top quark in the all-jets final state in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, *JHEP* **10** (2017) 005, [1707.03316].

[183] CMS collaboration, A. M. Sirunyan et al., *Search for top squark pair production in pp collisions at* $\sqrt{s} = 13$ *TeV using single lepton events*, *JHEP* **10** (2017) 019, [1706.04402].

[184] CMS collaboration, A. M. Sirunyan et al., *Search for top squarks and dark matter particles in opposite-charge dilepton final states at* $\sqrt{s}$ = *13 TeV*, *Phys. Rev. D* **97** (2018) 032009, [1711.00752].

[185] CMS collaboration, A. M. Sirunyan et al., *Search for dark matter produced in association with heavy-flavor quark pairs in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, *Eur. Phys. J. C* **77** (2017) 845, [1706.02581].

[186] S. Colucci, B. Fuks, F. Giacchino, L. Lopez Honorez, M. H. G. Tytgat and J. Vandecasteele, *Top-philic Vector-Like Portal to Scalar Dark Matter*, *Phys. Rev.* **D98** (2018) 035002, [1804.05068].

[187] B. Dasgupta, E. Ma and K. Tsumura, *Weakly interacting massive particle dark matter and radiative neutrino mass from Peccei-Quinn symmetry*, *Phys. Rev. D* **89** (2014) 041702, [1308.4138].

[188] S. Chatterjee, A. Das, T. Samui and M. Sen, *Mixed WIMP-axion dark matter*, *Phys. Rev. D* **100** (2019) 115050, [1810.09471].

[189] K. J. Bae, J.-H. Huh and J. E. Kim, *Update of axion CDM energy*, *JCAP* **09** (2008) 005, [0806.0497].

[190] J. McDonald, *Gauge singlet scalars as cold dark matter*, *Phys. Rev.* **D50** (1994) 3637–3649, [hep-ph/0702143].

[191] ADMX collaboration, S. J. Asztalos et al., *A SQUID-based microwave cavity search for dark-matter axions*, *Phys. Rev. Lett.* **104** (2010) 041301, [0910.5914].

[192] D. Budker, P. W. Graham, M. Ledbetter, S. Rajendran and A. Sushkov, *Proposal for a Cosmic Axion Spin Precession Experiment (CASPEr)*, *Phys. Rev. X* **4** (2014) 021030, [1306.6089].

[193] Y. Kahn, B. R. Safdi and J. Thaler, *Broadband and Resonant Approaches to Axion Dark Matter Detection*, *Phys. Rev. Lett.* **117** (2016) 141801, [1602.01086].

[194] MADMAX Working Group collaboration, A. Caldwell, G. Dvali, B. Majorovits, A. Millar, G. Raffelt, J. Redondo et al., *Dielectric Haloscopes: A New Way to Detect Axion Dark Matter*, *Phys. Rev. Lett.* **118** (2017) 091801, [1611.05865].

[195] S. Profumo, L. Giani and O. F. Piattella, *An Introduction to Particle Dark Matter*, *Universe* **5** (2019) 213, [1910.05610].

[196] F. F. Deppisch, W. Liu and M. Mitra, *Long-lived Heavy Neutrinos from Higgs Decays*, *JHEP* **08** (2018) 181, [1804.04075].

[197] S. Baek, P. Ko and P. Wu, *Top-philic Scalar Dark Matter with a Vector-like Fermionic Top Partner*, *JHEP* **10** (2016) 117, [1606.00072].

[198] A. Das, P. Konar and A. Thalapillil, *Jet substructure shedding light on heavy Majorana neutrinos at the LHC*, *JHEP* **02** (2018) 083, [1709.09712].

[199] A. Bhardwaj, A. Das, P. Konar and A. Thalapillil, *Looking for Minimal Inverse Seesaw scenarios at the LHC with Jet Substructure Techniques*, *J. Phys. G* **47** (2020) 075002, [1801.00797].

[200] A. Bhardwaj, J. Dutta, P. Konar, B. Mukhopadhyaya and S. K. Rai, *Boosted jet techniques for a supersymmetric scenario with gravitino LSP*, *JHEP* **10** (2020) 083, [2007.00351].

[201] T. Plehn, M. Spannowsky, M. Takeuchi and D. Zerwas, *Stop Reconstruction with Tagged Tops*, *JHEP* **10** (2010) 078, [1006.2833].

[202] C. Muselli, M. Bonvini, S. Forte, S. Marzani and G. Ridolfi, *Top Quark Pair Production beyond NNLO*, *JHEP* **08** (2015) 076, [1505.02006].

[203] N. Kidonakis, *Theoretical results for electroweak-boson and single-top production*, *PoS* **DIS2015** (2015) 170, [1506.04072].

[204] S. Catani, L. Cieri, G. Ferrera, D. de Florian and M. Grazzini, *Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO*, *Phys. Rev. Lett.* **103** (2009) 082001, [0903.2120].

[205] G. Balossini, G. Montagna, C. M. Carloni Calame, M. Moretti, O. Nicrosini, F. Piccinini et al., *Combination of electroweak and QCD corrections to single W production at the Fermilab Tevatron and the CERN LHC*, *JHEP* **01** (2010) 013, [0907.0276].

[206] J. M. Campbell, R. K. Ellis and C. Williams, *Vector boson pair production at the LHC*, *JHEP* **07** (2011) 018, [1105.0020].

[207] C. G. Lester and D. J. Summers, *Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders*, *Phys. Lett. B* **463** (1999) 99–103, [hep-ph/9906349].

[208] P. Konar, K. Kong, K. T. Matchev and M. Park, *Superpartner Mass Measurement Technique using 1D Orthogonal Decompositions of the Cambridge Transverse Mass Variable M_{T2}*, *Phys. Rev. Lett.* **105** (2010) 051802, [0910.3679].

[209] S. Frixione, *A General approach to jet cross-sections in QCD*, *Nucl. Phys. B* **507** (1997) 295–314, [hep-ph/9706545].

[210] A. Ghosh, P. Konar and R. Roshan, *Top-philic dark matter in a hybrid KSVZ axion framework*, *JHEP* **12** (2022) 167, [2207.00487].

[211] W. Buchmuller, R. Ruckl and D. Wyler, *Leptoquarks in Lepton - Quark Collisions*, Phys. Lett. B **191** (1987) 442–448.

[212] BaBar collaboration, J. P. Lees et al., *Measurement of an Excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ Decays and Implications for Charged Higgs Bosons*, Phys. Rev. D **88** (2013) 072012, [1303.0571].

[213] BELLE collaboration, S. Choudhury et al., *Test of lepton flavor universality and search for lepton flavor violation in $B \to K\ell\ell$ decays*, JHEP **03** (2021) 105, [1908.01848].

[214] Belle collaboration, A. Abdesselam et al., *Measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with a semileptonic tagging method*, 1904.08794.

[215] Belle collaboration, A. Abdesselam et al., *Test of Lepton-Flavor Universality in $B \to K^*\ell^+\ell^-$ Decays at Belle*, Phys. Rev. Lett. **126** (2021) 161801, [1904.02440].

[216] LHCb collaboration, R. Aaij et al., *Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu)$*, Phys. Rev. Lett. **115** (2015) 111803, [1506.08614].

[217] LHCb collaboration, *Measurement of the ratios of branching fractions $\mathcal{R}(D^*)$ and $\mathcal{R}(D^0)$*, 2302.02886.

[218] B. Gripaios, A. Papaefstathiou, K. Sakurai and B. Webber, *Searching for third-generation composite leptoquarks at the LHC*, JHEP **01** (2011) 156, [1010.3962].

[219] K. Chandak, T. Mandal and S. Mitra, *Hunting for scalar leptoquarks with boosted tops and light leptons*, Phys. Rev. D **100** (2019) 075019, [1907.11194].

[220] A. Bhaskar, T. Mandal, S. Mitra and M. Sharma, *Improving third-generation leptoquark searches with combined signals and boosted top quarks*, Phys. Rev. D **104** (2021) 075037, [2106.07605].

[221] G. Belanger et al., *Leptoquark manoeuvres in the dark: a simultaneous solution of the dark matter problem and the $R_{D^{(*)}}$ anomalies*, JHEP **02** (2022) 042, [2111.08027].

[222] ATLAS collaboration, M. Aaboud et al., *Searches for third-generation scalar leptoquarks in $\sqrt{s}$ = 13 TeV pp collisions with the ATLAS detector*, JHEP **06** (2019) 144, [1902.08103].

[223] ATLAS collaboration, *Search for pair production of third-generation leptoquarks decaying into a bottom quark and a τ-lepton with the ATLAS detector*, 2303.01294.

[224] CMS collaboration, A. M. Sirunyan et al., *Search for singly and pair-produced leptoquarks coupling to third-generation fermions in proton-proton collisions at s=13 TeV*, *Phys. Lett. B* **819** (2021) 136446, [2012.04178].

[225] CMS collaboration, *The search for a third-generation leptoquark coupling to a τ lepton and a b quark through single, pair and nonresonant production at* $\sqrt{s} = 13$ TeV, .

[226] ATLAS collaboration, G. Aad et al., *Search for a scalar partner of the top quark in the all-hadronic $t\bar{t}$ plus missing transverse momentum final state at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **80** (2020) 737, [2004.14060].

[227] P. Bandyopadhyay, S. Dutta, M. Jakkapu and A. Karan, *Distinguishing Leptoquarks at the LHC/FCC*, *Nucl. Phys. B* **971** (2021) 115524, [2007.12997].

[228] V. D. Barger, J. Ohnemus and R. J. N. Phillips, *Spin Correlation Effects in the Hadroproduction and Decay of Very Heavy Top Quark Pairs*, *Int. J. Mod. Phys. A* **4** (1989) 617.

[229] K. Hagiwara and D. Zeppenfeld, *Helicity Amplitudes for Heavy Lepton Production in e+ e- Annihilation*, *Nucl. Phys. B* **274** (1986) 1–32.

[230] G. L. Kane, G. A. Ladinsky and C. P. Yuan, *Using the Top Quark for Testing Standard Model Polarization and CP Predictions*, *Phys. Rev. D* **45** (1992) 124–141.

[231] R. H. Dalitz and G. R. Goldstein, *The Decay and polarization properties of the top quark*, *Phys. Rev. D* **45** (1992) 1531–1543.

[232] Y. Sumino and S. Tsuno, *New method for reconstructing effective top quark spin*, *Phys. Lett. B* **633** (2006) 715–720, [hep-ph/0512205].

[233] R. M. Godbole, S. D. Rindani and R. K. Singh, *Lepton distribution as a probe of new physics in production and decay of the t quark and its polarization*, *JHEP* **12** (2006) 021, [hep-ph/0605100].

[234] W. Bernreuther, *Top quark physics at the LHC*, *J. Phys. G* **35** (2008) 083001, [0805.1333].

[235] J. Shelton, *Polarized tops from new physics: signals and observables*, *Phys. Rev. D* **79** (2009) 014032, [0811.0569].

[236] R. M. Godbole, S. D. Rindani, K. Rao and R. K. Singh, *Top polarization as a probe of new physics*, *AIP Conf. Proc.* **1200** (2010) 682–685, [0911.3622].

[237] R. M. Godbole, K. Rao, S. D. Rindani and R. K. Singh, *On measurement of top polarization as a probe of $t\bar{t}$ production mechanisms at the LHC*, *JHEP* **11** (2010) 144, [1010.1458].

[238] B. Bhattacherjee, S. K. Mandal and M. Nojiri, *Top Polarization and Stop Mixing from Boosted Jet Substructure*, *JHEP* **03** (2013) 105, [1211.7261].

[239] W. Bernreuther, P. González and C. Mellein, *Decays of polarized top quarks to lepton, neutrino and jets at NLO QCD*, *Eur. Phys. J. C* **74** (2014) 2815, [1401.5930].

[240] A. Prasath V, R. M. Godbole and S. D. Rindani, *Longitudinal top polarisation measurement and anomalous Wtb coupling*, *Eur. Phys. J. C* **75** (2015) 402, [1405.1264].

[241] S. Bhattacharya, M. Guchait and A. H. Vijay, *Boosted top quark tagging and polarization measurement using machine learning*, *Phys. Rev. D* **105** (2022) 042005, [2010.11778].

[242] D. Krohn, J. Shelton and L.-T. Wang, *Measuring the Polarization of Boosted Hadronic Tops*, *JHEP* **07** (2010) 041, [0909.3855].

[243] A. Papaefstathiou and K. Sakurai, *Determining the Helicity Structure of Third Generation Resonances*, *JHEP* **06** (2012) 069, [1112.3956].

[244] I. Doršner, S. Fajfer, A. Greljo, J. F. Kamenik and N. Košnik, *Physics of leptoquarks in precision experiments and at particle colliders*, *Phys. Rept.* **641** (2016) 1–68, [1603.04993].

[245] T. Mandal, S. Mitra and S. Seth, *Pair Production of Scalar Leptoquarks at the LHC to NLO Parton Shower Accuracy*, *Phys. Rev. D* **93** (2016) 035018, [1506.07369].

[246] D. Binosi and L. Theussl, *JaxoDraw: A Graphical user interface for drawing Feynman diagrams*, *Comput. Phys. Commun.* **161** (2004) 76–86, [hep-ph/0309015].

[247] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer and T. Reiter, *UFO - The Universal FeynRules Output*, Comput. Phys. Commun. **183** (2012) 1201–1214, [1108.2040].

[248] C. Bierlich et al., *A comprehensive guide to the physics and usage of PYTHIA 8.3*, 2203.11601.

[249] A. Czarnecki, J. G. Korner and J. H. Piclum, *Helicity fractions of W bosons from top quark decays at NNLO in QCD*, Phys. Rev. D **81** (2010) 111503, [1005.2625].

[250] ATLAS collaboration, G. Aad et al., *Measurement of the W boson polarization in top quark decays with the ATLAS detector*, JHEP **06** (2012) 088, [1205.2484].

[251] R. Allahverdi, M. Dalchenko, B. Dutta, A. Flórez, Y. Gao, T. Kamon et al., *Distinguishing Standard Model Extensions using Monotop Chirality at the LHC*, JHEP **12** (2016) 046, [1507.02271].

[252] A. D'Alise et al., *Standard model anomalies: lepton flavour non-universality, $g-2$ and W-mass*, JHEP **08** (2022) 125, [2204.03686].