

# **Machine Learning-enhanced feasibility studies on the production of new particles with preferential couplings to third generation fermions at the LHC**

Thesis submitted in partial fulfillment  
of the requirements for the degree of  
**Doctor in Sciences - Physics**

**Presented by:**  
Cristian Fernando Rodríguez Cruz

**Directed by:**  
Prof. Andrés Florez (advisor)  
Universidad de los Andes  
Prof. Joel Jones-Perez (co-advisor)  
Pontificia Universidad Católica del Perú

High Energy Physics Research Group  
Phenomenology of Particle Physics

**Universidad de los Andes**  
Faculty of Sciences  
Department of Physics  
Bogotá D.C., Colombia  
September 22, 2025. Pre-alpha version



### **Dedication**

to be added



## Acknowledgements

to be added



## Abstract

The Standard Model (SM) of particle physics is the most successful framework for describing the subatomic world. It is continuously tested in experiments worldwide, with the Large Hadron Collider (LHC) being the flagship project in this endeavor. One of the primary goals of the LHC is to precisely measure SM parameters and search for deviations that could signal new physics.

In recent years, reported anomalies, such as those in B-meson decays from LHCb, BaBar, and Belle experiments, along with the potential discrepancy in the muon's magnetic moment ( $g - 2$ ) from Fermilab, suggest a violation of lepton flavor universality (LFU). These observations provide a compelling window into physics beyond the SM. Among the proposed SM extensions to explain LFU violation, many introduce new particles with preferential couplings to third and second-generation fermions. Popular candidates include heavy states such as  $Z'$  bosons,  $\phi'$  scalars, and leptoquarks (LQs), among others.

This work presents two phenomenological studies proposing different strategies to probe new models, such as the 4321 [1],  $U(1)_{T_R^3}$  [2], that extend the SM particle content to explain clues on LFU violation. The studies use benchmark scenarios where the new particle fields have preferential couplings to second and third generation SM-fermions, depending on the model. The hypothetical signal and background samples are generated using Monte Carlo simulations, emulating the current running conditions of the LHC and the performance of the CMS detector. The expected sensitivity for the different signal models under study is obtained by performing a detailed analysis of the available (non-excluded) experimental phase-space, boosted by machine learning (ML) techniques to optimize the discovery potential for these exotic states.



---

## CONTENTS

Contents	ix
INTRODUCTION	1
1 STANDARD MODEL OF PARTICLE PHYSICS	5
1.1 Fields . . . . .	5
1.1.1 Interactions and Symmetries . . . . .	7
1.2 Standard Model . . . . .	11
1.2.1 Particle Content and Gauge Group . . . . .	13
1.2.2 Gauge Bosons . . . . .	15
1.2.3 Matter Fields . . . . .	16
1.2.4 Electroweak Symmetry Breaking . . . . .	17
1.3 Deficiencies of Standard Model and New Physics . . . . .	19
1.3.1 Theoretical problems . . . . .	19
1.3.2 Cosmological problems . . . . .	20
1.3.3 Phenomenological problems . . . . .	21
2 PHENOMENOLOGICAL FRAMEWORK	23
2.1 Detectors and Subsystems . . . . .	24
2.1.1 Collision Parameters . . . . .	29
2.2 Jets Reconstruction . . . . .	30
2.2.1 Jet algorithms . . . . .	31
2.2.2 $\tau$ Tagging at Multipurpose Detectors . . . . .	33
2.2.3 B Tagging at Multipurpose Detectors . . . . .	34
2.3 The CMS Detector . . . . .	35
2.4 Phenomenological Pipeline . . . . .	37
2.5 Measurement of the Power of an Analysis . . . . .	40
2.6 ML enhanced signal-background discrimination . . . . .	44
2.6.1 XGBoost: Optimized Gradient Boosting . . . . .	46
2.6.2 Standard ML Analysis Workflow . . . . .	49
3 $U(1)_{T_R^3}$ GAUGE EXTENSION OF THE STANDARD MODEL	51
3.1 Current Exclusion Limits on Vector-Like Quarks . . . . .	53
3.2 The Minimal $U(1)_{T_R^3}$ Model . . . . .	54
3.2.1 The Universal Seesaw Mechanism . . . . .	55
3.2.2 Minimal UV-complete Theory . . . . .	56
3.3 Samples and Simulation . . . . .	57
3.4 Data Analysis Using Machine Learning . . . . .	62
3.5 Results . . . . .	67
3.6 Discussion . . . . .	72
4 ON VECTORIAL LEPTOQUARKS SENSITIVITY AT THE LHC	75
4.1 A Simplified Model for the $U_1$ Leptoquark . . . . .	78
4.2 Search strategy and simulation . . . . .	82
4.3 Results . . . . .	89
4.4 Discussion and conclusions . . . . .	95

<b>A SUPERVISED LEARNING</b>	<b>99</b>
<b>A.1 Mathematical Framework of Supervised Learning</b>	100
<b>A.1.1 The Learning Problem</b>	100
<b>A.1.2 Hypothesis Classes and Model Complexity</b>	100
<b>A.1.3 Regularization and Occam's Razor</b>	101
<b>A.2 The Classification Problem</b>	101
<b>A.2.1 Binary Classification</b>	102
<b>A.2.2 Support Vector Machines</b>	102
<b>A.2.3 Multi-class Classification</b>	103
<b>A.3 The Regression Problem</b>	103
<b>A.3.1 Linear Regression</b>	103
<b>A.3.2 Polynomial Regression and Feature Engineering</b>	104
<b>A.3.3 Robust Regression</b>	104
<b>A.3.4 Performance Evaluation</b>	105
<b>A.4 Boosted Decision Trees</b>	105
<b>A.4.1 Decision Trees: The Foundation</b>	106
<b>A.4.2 The Boosting Paradigm</b>	106
<b>A.4.3 Gradient Boosting Framework</b>	107
<b>A.4.4 XGBoost: Extreme Gradient Boosting</b>	107
<b>A.4.5 Feature Importance and Interpretability</b>	109
<b>A.5 Neural Networks</b>	109
<b>A.5.1 The Perceptron: Building Block of Neural Networks</b>	109
<b>A.5.2 Multi-Layer Perceptrons</b>	110
<b>A.5.3 Backpropagation Algorithm</b>	110
<b>A.5.4 Optimization Algorithms</b>	110
<b>A.5.5 Regularization Techniques</b>	111
<b>A.5.6 Neural Networks in Particle Physics</b>	112
<b>A.5.7 Network Architectures for Particle Physics</b>	112
<b>A.6 Comparative Analysis: BDTs vs Neural Networks</b>	113
<b>A.6.1 Performance Comparison</b>	113
<b>A.6.2 Interpretability</b>	113
<b>A.6.3 Robustness and Generalization</b>	113
<b>A.6.4 Computational Considerations</b>	113
<b>B ON THE UNIVERSAL SEESAW MECHANISM</b>	<b>115</b>
<b>B.1 Diagonalization and Field Redefinitions</b>	115
<b>B.2 Symmetry Breaking and the Mass Matrix</b>	116
<b>B.3 Bi-Unitary Transformation and Mass Eigenvalues</b>	117
<b>B.4 Yukawa Coupling Enhancement and Perturbativity</b>	117
<b>B.5 Exact Diagonalization and Mixing Angles</b>	118
<b>C THE 4321 MODEL</b>	<b>121</b>
<b>BIBLIOGRAPHY</b>	<b>131</b>

---

## INTRODUCTION

---

The pursuit of a fundamental description of nature’s building blocks and their interactions is a central endeavor of modern physics. This quest has led to the development of the Standard Model (SM) of particle physics, a quantum field theory that encapsulates our current understanding of the subatomic world. With breathtaking precision, the SM describes the electromagnetic, weak, and strong nuclear forces and classifies all known elementary particles. Its triumphs are undeniable, crowned by the landmark discovery of the Higgs boson at the Large Hadron Collider (LHC) in 2012, which confirmed the mechanism for generating mass and represented the final piece of the SM puzzle.

Yet, for all its success, the Standard Model is universally acknowledged to be an incomplete theory. It offers no candidate for dark matter, cannot account for the asymmetry between matter and antimatter in the universe, does not incorporate gravity, and leaves the mass of the Higgs boson itself unnaturally unstable under quantum corrections—a problem known as the hierarchy problem. These profound theoretical shortcomings provide a clear motivation for physics beyond the Standard Model (BSM). However, the most compelling guide for this search has always come from experimental data itself.

The primary mission of the LHC is not only to consolidate the SM but to probe its boundaries and discover new physics. While no direct evidence of new particles has been found so far, a series of subtle but persistent discrepancies—termed “anomalies”—have emerged from experiments worldwide, suggesting a potential crack in the SM’s foundation.

A particularly intriguing set of these anomalies points towards a violation of Lepton Flavor Universality (LFU). The SM predicts that the electroweak force couples with identical strength to the three charged leptons (electrons, muons, and taus), a fundamental principle known as LFU. The most significant and long-standing hints of LFU violation come from measurements of semileptonic B-meson decays. The ratios  $R(D^{(*)}) = \mathcal{B}(B \rightarrow D^{(*)}\tau\nu_\tau)/\mathcal{B}(B \rightarrow D^{(*)}\ell\nu_\ell)$ , where  $\ell$  is a muon or electron, have been measured by the BaBar, Belle, and LHCb collaborations to consistently exceed the SM predictions by a combined significance of approximately  $3\sigma$ - $4\sigma$ . This deviation suggests that B mesons are more likely to decay to a final state containing a tau lepton than the SM allows, providing a compelling hint of new physics that couples preferentially to the third generation. Furthermore, the longstanding discrepancy in the muon’s anomalous magnetic moment ( $g - 2$ ), recently confirmed with increased precision by the Fermilab experiment, adds another layer of

intrigue, as it also hints at new physics potentially coupled preferentially to the second generation.

While each anomaly individually requires careful scrutiny, their collective persistence has generated significant excitement, as they seem to point towards new physics that breaks lepton flavor universality, potentially involving enhanced couplings to heavier fermions.

The pattern of these LFU-violating anomalies has inspired a vast landscape of theoretical models extending the SM. A common thread among the most promising explanations is the introduction of new heavy particles that mediate interactions with non-universal couplings to the different generations of fermions. This generational hierarchy is crucial to evade tight constraints from precision measurements on electrons (first generation) while affecting processes involving muons and taus.

In this thesis, we contextualize and present two of our phenomenological studies that propose different strategies to probe new physics models, such as the 4321 [1] and  $U(1)_{T_R^3}$  [2] models, which extend the SM particle content to explain the observed LFU violation. These models introduce new particles with preferential couplings to second and third-generation fermions, making them prime candidates for explaining the experimental anomalies.

The experimental challenge lies in probing these models at the LHC. The proposed new particles are often heavy, leading to low production rates, and their decay signatures are complex and overwhelmed by enormous Standard Model backgrounds. Given the immense number of theoretical possibilities and the finite resources available to experimental collaborations, it is impossible to pursue every potential signature with equal vigor. This is where **phenomenological feasibility studies** become critical. They provide a vital bridge between theory and experiment by performing a detailed *a priori* assessment of the discovery potential for a given signal model. By using Monte Carlo simulations to emulate the detector response and analysis chain, these studies can identify the most promising signatures, optimize event selection criteria, and estimate the sensitivity achievable with the available data. This process is essential for prioritizing the experimental program, justifying the dedication of significant computing and human resources to a particular search, and ultimately guiding the LHC experiments towards the most well-motivated and detectable signals of new physics.

This thesis contributes to this effort by presenting two dedicated phenomenological studies that propose and develop novel strategies to probe the 4321 and  $U(1)_{T_R^3}$  models at the LHC. The work is situated at the intersection of theoretical model-building and experimental high-energy physics, with the explicit goal of assessing the feasibility of these searches.

The core methodology of this research involves:

1. Defining **benchmark scenarios** within each model, selecting specific mass points and coupling structures that explain the LFU anomalies while remaining experimentally viable.

2. Using **Monte Carlo simulation** to accurately generate the hypothetical signal processes alongside the dominant SM background processes, emulating the run conditions of the LHC and the performance of the CMS detector.
3. Performing a detailed analysis of the available experimental phase-space. Given the high-dimensionality of the final states (e.g., involving multiple jets, leptons, and missing energy) and the complex, overlapping kinematical distributions of signal and background, traditional “cut-and-count” analyses are often sub-optimal. To overcome this, we employ advanced **Machine Learning (ML) techniques**, specifically supervised learning algorithms such as Boosted Decision Trees (BDTs) or Deep Neural Networks (DNNs). These algorithms are trained to learn the complex, non-linear correlations between many kinematic variables (e.g., invariant masses, angular separations, transverse momenta) to construct powerful discriminators that optimally separate the rare signal events from the large and diverse SM backgrounds. This ML-enhanced approach significantly boosts the analysis sensitivity, allowing for the detection of weaker signals or the setting of more stringent limits than would otherwise be possible.
4. Deriving the **expected sensitivity** for each model, establishing the exclusion limits or discovery potential that the LHC experiments could achieve with the current dataset. This final step is the ultimate quantitative measure of the search’s feasibility.

The structure of this thesis is as follows. We begin by establishing the theoretical foundation with a review of the Standard Model in Chapter 2. Chapter 3 then details the experimental context, describing the LHC and the CMS detector, and introduces the general analysis techniques employed, including a discussion on the application of Machine Learning in high-energy physics. The original phenomenological work of this thesis is presented in the subsequent chapters: Chapter 4 details a search for new physics in the process  $pp \rightarrow t\bar{t}\mu^+\mu^-$ , while Chapter 5 presents a search for vector leptoquarks in the process  $pp \rightarrow \tau^+\tau^- + b\text{-jets}$ . Finally, Chapter 6 concludes by summarizing our findings and discussing their implications for the field, along with an outlook on future prospects.



# 1

---

## STANDARD MODEL OF PARTICLE PHYSICS

---

The Standard Model (SM) of particle physics is a quantum field theory (QFT) that describes matter as fermionic particles and their fundamental interactions through the gauge principle. In this framework, force-carrying particles—vector bosons with integer spin, arising from the adjoint representation of symmetry groups (*gauge groups*)—mediate the interactions between matter particles [3, 4]. This elegant formulation unifies three of the four fundamental forces in nature.

In this chapter, we contextualize the SM by introducing the basic concepts of quantum field theory, including the notion of fields and symmetries. We then present the particle content of the SM, its gauge group, and the Lagrangian density that describes its dynamics. The Higgs mechanism and its role in providing mass to the weak gauge bosons and fermions are also discussed. Finally, we address the main deficiencies of the Standard Model and review the experimental evidence that motivates the search for physics beyond the SM.

### 1.1 FIELDS

---

Relativistic quantum fields are degrees of freedom in QFT. Formally, they are *operator-valued functions of the spacetime that transform under a representation of the Lorentz group within an invariant subspace* [5, 6]. The different representations of the Lorentz group are mainly characterized by their spin, and their fields obey a different equation of motion (see table 1.1).

In classical field theory, a variational principle is established which generates the equations that govern the dynamics of the different fields in a theory, *the equations of motion*. Hamilton's principle, or principle of minimal action, indicates that all possible physical configurations for a set of fields  $\varphi^I$ , with  $I = 1, 2, 3, \dots, n$ , are those where the integral of the action  $S$  is a minimal [7, 8]:

$$S = \int \mathcal{L}(\varphi^I, \partial_\mu \varphi^I) d^4x, \quad (1.1)$$

here,  $d^4x = dx^0 dx^1 dx^2 dx^3$  and  $x \equiv (ct, x^1, x^2, x^3) \equiv (x^0, x^1, x^2, x^3) \in \mathcal{M}^4$  are the space-time coordinates in the Minkowskian spacetime  $\mathcal{M}^4$ , and the function  $\mathcal{L}(\varphi^I, \partial_\mu \varphi^I)$  is called *the Lagrangian density* of a theory [3, 7]. The problem in classical field dynamics is to find the functions  $\varphi^I(x)$  in

a space-time  $\mathcal{M}^4$ , fixing their boundary conditions. The solution to this classical problem is given by the Euler-Lagrange equations:

$$\frac{\partial \mathcal{L}}{\partial \varphi^I} - \frac{\partial}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi^I)} = 0, \quad (1.2)$$

and they are used to obtain the equations of motion of the set of fields  $\varphi^I$  [8].

While in classical field theory the Euler-Lagrange equations directly determine the dynamics, in quantum field theory the picture changes: if we adopt the path-integral formulation [9, 10], the idea of an equation of motion vanishes and we move on to searching for correlations between free particle states. However, the notion of action remains the cornerstone in the description of these observables.

Explicitly, the correlation functions are calculated through the Lehmann-Symanzik-Zimmermann (LSZ) reduction formula, which connects these correlators with physical scattering amplitudes. These are computed from the path integral [11, 12]

$$\begin{aligned} Z[J] &= \langle \text{out}, 0 | 0, \text{in} \rangle \\ &= \mathcal{N} \int \mathcal{D}(\varphi, \bar{\varphi}) e^{iS[\varphi]} e^{i \int J_1 \varphi^I d^4x} \\ &= \mathcal{N} \int \mathcal{D}(\varphi, \bar{\varphi}) e^{i \int d^4x \mathcal{L}} e^{i \int J_1 \varphi^I d^4x}, \end{aligned} \quad (1.3)$$

taken over the space of fields  $\varphi$  with an appropriate measure  $\mathcal{D}(\varphi, \bar{\varphi})$  and normalized by  $\mathcal{N}$ . The quantity  $Z$  is known as the partition function of the theory and gives the transition amplitude from the initial vacuum  $|0, \text{in}\rangle$  to the final vacuum  $|0, \text{out}\rangle$  in the presence of a source  $J(x)$  producing particles [13].

Name	Field	Spin	Free-Lagrangian
Klein-Gordón	$\phi$	0	$\mathcal{L} = (\partial^\mu \bar{\phi} \partial_\mu \phi - m^2 \bar{\phi} \phi)$
Dirac	$\chi$	1/2	$\mathcal{L} = \bar{\chi} (i \gamma^\mu \partial_\mu - m) \chi$
Proca (Massive Vector)	$A^\mu$	1	$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} + \frac{1}{2} m^2 A^\mu A_\mu$

**Table 1.1:** Some relevant representations of the Lorentz group in 4-dimensional space-time. In this notation  $\eta_{\mu\nu} = \text{diag}(1, -1, -1, -1)$ ,  $\gamma^\mu$  are the Dirac matrices,  $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$  is the abelian field strength tensor. All equations are written in natural units with  $c = \hbar = 1$ . Fields are shown in their standard representations.

Therefore, the dynamics, at both the classical and quantum levels, are entirely determined by the Lagrangian density. For free fields (i.e., non-interacting), the Lagrangian is quadratic in the fields and the path integral can be evaluated exactly. Table 1.1 records the Lagrangian density for these free fields. However, to describe physics, we must include interactions, which render the path integral impossible to compute exactly.

The framework of *perturbation theory* addresses this by expanding the interaction part of the Lagrangian as a power series. This expansion is organized using *Feynman diagrams*, which provides a pictorial representation of each term, and a set of *Feynman rules*, which provides a precise dictionary to translate these diagrams into mathematical expressions for scattering amplitudes [10, 12]. The importance of these rules cannot be overstated, as they are the practical computational tools of perturbative QFT.

In this paradigm, our task is to propose a Lagrangian density for a set of fields that correctly models the propagation and interactions of fundamental particles. The free part defines the particle content and propagators, while the interaction part defines the vertices and possible scattering processes.

### 1.1.1 INTERACTIONS AND SYMMETRIES

The form of the Lagrangian density is not arbitrary: it is shaped by a small set of physical and mathematical principles. These principles act as “rules” that guide the construction of consistent theories, ensuring both their internal consistency and their predictive power. In particular, if we want a relativistic and renormalizable theory, the Lagrangian must satisfy several conditions that strongly restrict the kind of terms that can appear.

The need for these restrictions is evident from the path integral formulation itself. If we split the Lagrangian into a free part and an interaction part in the form

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_{\text{int}}. \quad (1.4)$$

The generating functional  $Z[J]$  can then be expressed as an operator acting on the free functional  $Z_0[J]$ :

$$Z[J] = \mathcal{N} \exp \left[ i \int d^4x \mathcal{L}_{\text{int}} \left( -i \frac{\delta}{\delta J(x)} \right) \right] Z_0[J]. \quad (1.5)$$

The exponential operator generates an infinite perturbation series. The  $n$ -point correlation function is found by taking functional derivatives of  $Z[J]$  with respect to the sources  $J(x_i)$  and setting  $J = 0$ .

Each term in this series is represented by a **Feynman diagram**:

- **External Lines:** Represent incoming and outgoing physical particles.
- **Internal Lines:** Represent virtual particles propagating between interactions, corresponding to the free-field propagators derived from  $\mathcal{L}_0$ .
- **Vertices:** Represent interactions, derived from the terms in  $\mathcal{L}_{\text{int}}$ . Each vertex has an associated coupling constant and enforces momentum conservation.

For this series to be a predictive and well-defined computational tool, the individual terms must yield finite results. This requirement of *renormalizability* is a powerful constraint on  $\mathcal{L}_{\text{int}}$ . Furthermore, the structure of both

$\mathcal{L}_0$  and  $\mathcal{L}_{\text{int}}$  is profoundly constrained by the requirement that the theory possesses certain *symmetries*.

To begin with, relativistic invariance demands that the equations of motion look the same in all inertial frames. This requirement is implemented by asking the action to be invariant under Poincaré transformations [14]. Equivalently, the Lagrangian density must transform as a Lorentz scalar and may change under translations at most by a total derivative [8].

Another basic condition is Hermiticity: the Lagrangian density must be Hermitian so that observables are real and the time evolution of the theory is unitary [12, 14]. In addition, dimensional analysis places further restrictions. In natural units,  $\mathcal{L}$  carries dimensions of [mass]<sup>4</sup> (an energy density). This means that the interaction terms that we can add must be such that the overall operator has the correct dimension, which already rules out many possibilities.

In quantum field theory, loop corrections to scattering amplitudes typically produce divergences. A theory is called renormalizable if all these divergences can be absorbed into a redefinition of a *finite set* of physical parameters (such as masses and couplings). In practice, this requirement translates into a restriction on the operators that may appear in the Lagrangian: only terms of mass dimension  $\leq 4$  lead to renormalizable interactions. Higher-dimensional operators are still allowed, but they correspond to *effective* interactions that are suppressed at low energies and signal the presence of new physics at higher scales [10, 12].

A classical symmetry of the Lagrangian may not always survive the process of quantization. If it fails to do so, it is said to be anomalous. Chiral anomalies, specifically, arise from the regularization of fermion loops in triangle diagrams and can break the gauge symmetry at the quantum level. Since the gauge symmetry is the very principle that dictates the form of interactions and removes unphysical states, its violation would destroy the renormalizability and unitarity of the theory. Therefore, the particle content must be carefully chosen so that these potential anomalies cancel amongst the fermions, a non-trivial condition famously satisfied by the quarks and leptons of the Standard Model.

Furthermore, the stability of the vacuum is a prerequisite for a physically meaningful theory. This is ensured by demanding that the scalar potential, which governs the self-interactions of the Higgs field, is bounded from below. If the potential were unbounded, it would imply that the system could lower its energy indefinitely by rolling down the potential to field values of ever-greater magnitude, meaning no stable ground state could exist. For a renormalizable potential, this stability condition typically translates into the requirement that the quartic coupling constant  $\lambda > 0$ . However, this condition must hold not just at tree-level but also at the quantum level, as running couplings can change sign at different energy scales, potentially leading to metastability or instability of the vacuum.

Summarizing, the main constraints that a relativistic and renormalizable Lagrangian density must satisfy are:

- **Poincaré (global) invariance:** the action must be invariant under Lorentz transformations and translations; the Lagrangian density is a Lorentz scalar and may change by at most a total derivative [8, 14].
- **Hermiticity:**  $\mathcal{L}$  must be Hermitian to ensure real observables and unitary evolution [12, 14].
- **Renormalizability and Operator Dimension:** The theory must be perturbatively renormalizable, meaning all ultraviolet divergences can be absorbed into a finite number of parameters. This requirement, determined via power-counting arguments, restricts interaction operators to have a **mass dimension**  $\leq 4$ . In natural units, where  $\mathcal{L}$  has dimension [mass]<sup>4</sup>, this allows only Yukawa couplings (dim 4), scalar  $\phi^4$  interactions (dim 4), and gauge interactions (dim 4), while forbidding non-renormalizable operators like  $\phi^6$  (dim 6) [10, 12].
- **Absence of chiral anomalies:** gauge symmetries must be free of chiral (gauge) anomalies to ensure the consistency and unitarity of the quantum theory [10, 12, 15]. In the Standard Model, the particle content is such that all gauge anomalies cancel exactly.
- **Stability of the potential:** the scalar potential must be bounded from below to guarantee the existence of a stable vacuum state. This typically requires that the quartic couplings in the potential are positive at the relevant energy scales.

These constraints drastically reduce the number of possible terms in the Lagrangian. As a result, the renormalizable interaction structures that typically arise are limited to: Yukawa couplings between fermions and scalars, scalar self-interactions (up to quartic order), and gauge interactions between matter fields and vector bosons.

Despite these powerful constraints, a vast number of possible interaction terms between the allowed fields remain. To further restrict the form of the Lagrangian and to describe fundamental forces, the concept of *symmetry*—specifically *gauge symmetry*—has proven to be our most powerful guiding principle.

The procedure is systematic: first, the spin 0 and spin 1/2 fields are organized into representations of a unitary (gauge) group  $G$  such that the Lagrangian density is a global scalar (invariant) under  $G$ . This global symmetry is then “promoted” to a *local symmetry* (where the group parameters can vary in spacetime) by replacing the ordinary derivatives  $\partial_\mu$  with *covariant derivatives*  $\mathcal{D}_\mu$  that incorporate new *gauge fields*  $B_\mu^A$  [4, 16–19]. This “promotion” is described in more detail below.

Given a Lagrangian density  $\mathcal{L}(\varphi_i, \partial_\mu \varphi_i)$ , a given field  $\varphi$  is said to be *globally symmetric* under unitary transformations if the action is invariant under the variations of the fields  $\varphi^I$  which are given, at the infinitesimal level, by:

$$\delta_G \varphi^I = i\theta^A (T_A)^I_J \varphi^J, \quad (1.6)$$

*In QFT, Poincaré invariance is assumed to be global.*

*Promoting it to a local symmetry leads to gravity, with spin-2 fields (the graviton) as mediators.*

*Perturbatively, such a theory is not renormalizable, so it lacks predictivity at high energies, although it can still be understood as an effective field theory.*

where  $\theta^A$  are constant parameters of the transformation and the  $T_A$  are the generators of the group  $G$  in the appropriate representation. The corresponding finite unitary transformation is

$$U_G \equiv U(\theta) = \exp(i\theta^A T_A). \quad (1.7)$$

The generators  $T_A$  satisfy the Lie algebra of  $G$ :

$$[T_A, T_B] = i f_{AB}^C T_C, \quad (1.8)$$

where  $f_{AB}^C$  are the structure constants of  $G$ .

To promote the global symmetry to a local one ( $\theta^A \rightarrow \theta^A(x)$ ), the ordinary derivative  $\partial_\mu$  is replaced by a *covariant derivative*  $D_\mu$ . This new derivative is designed to transform covariantly under the gauge group, meaning  $D_\mu \varphi \rightarrow U(x)(D_\mu \varphi)$ , so that the kinetic terms  $\mathcal{L}_{\text{kin}} \sim (D_\mu \varphi)^\dagger (D^\mu \varphi)$  remain invariant. This is achieved by introducing a gauge field  $B_\mu^A$  for each generator  $T_A$  and defining:

$$D_\mu = \partial_\mu - ig B_\mu^A T_A, \quad (1.9)$$

where  $g$  is the gauge coupling constant. The transformation law for the gauge fields that ensures the covariant transformation of  $D_\mu$  is:

$$\delta B_\mu^A = \partial_\mu \theta^A + g f_{BC}^A \theta^B B_\mu^C. \quad (1.10)$$

The introduction of the gauge fields  $B_\mu^A$  necessitates the addition of a kinetic term for them to the Lagrangian. This is constructed from the *field strength tensor*  $F_{\mu\nu}^A$ , defined as the curvature of the covariant derivative:

$$F_{\mu\nu}^A T_A = -\frac{i}{g} [D_\mu, D_\nu] = \partial_\mu B_\nu^A - \partial_\nu B_\mu^A + g f_{BC}^A B_\mu^B B_\nu^C. \quad (1.11)$$

The gauge-invariant kinetic Lagrangian is then:

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4} \delta_{AB} F_{\mu\nu}^A F^{\mu\nu B}. \quad (1.12)$$

Often, the rescaling  $B_\mu^A \rightarrow g B_\mu^A$  is performed, which moves the coupling constant  $g$  from the kinetic term to the covariant derivative, resulting in the more conventional form  $D_\mu = \partial_\mu - ig B_\mu^A T_A$  and  $\mathcal{L}_{\text{gauge}} = -\frac{1}{4g^2} \delta_{AB} F_{\mu\nu}^A F^{\mu\nu B}$ .

A general, archetypal Lagrangian embodying these structures can be written as:

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu}^A F^{A\mu\nu} + i \bar{\psi}^i \gamma^\mu D_\mu \psi^i + \left( \bar{\psi}_L^j \Gamma_k^j \Phi \psi_R^k + \text{h.c.} \right) + |D_\mu \Phi|^2 - V(\Phi) \quad (1.13)$$

The terms correspond to: the kinetic term for gauge fields ( $F_{\mu\nu}^A$ ); the kinetic term for fermions  $\psi^i$  ( $D_\mu$  is the gauge covariant derivative); Yukawa interactions between left- and right-handed fermions and scalars ( $\Gamma_k^j$  is a Yukawa coupling matrix and  $\Phi$  is a scalar field); the kinetic term for scalars; and the scalar potential  $V(\Phi)$ , which for a renormalizable, stable theory is  $V(\Phi) = \mu^2 |\Phi|^2 + \lambda |\Phi|^4$  with  $\lambda > 0$ .

Note the absence of explicit mass terms for the gauge fields ( $\sim M^2 B_\mu B^\mu$ ) and fermions ( $\sim m \bar{\psi} \psi$ ). These are forbidden by gauge invariance for non-abelian fields and for chiral fermions. Mass terms can be generated via spontaneous symmetry breaking, as discussed below.

### Example

As a way of illustration let us consider a renormalizable theory with a real scalar  $\phi$  and a Dirac spinor  $\psi$ , and suppose that this theory is globally invariant under  $U(1)$  phase transformations, i.e. the fields  $\varphi \in \{\phi, \psi\}$  transform as  $\varphi \mapsto e^{i\theta}\hat{Q}\varphi$  such that  $\hat{Q}\psi = q\psi$  and  $\hat{Q}\phi = 0$ . The free Lagrangian density is

$$\mathcal{L}_{\text{free}} = \frac{1}{2}\partial^\mu\phi\partial_\mu\phi - \frac{1}{2}m^2\phi^2 + \bar{\psi}(i\gamma_\mu\partial^\mu - m)\psi. \quad (1.14)$$

To add globally symmetric interaction terms, we must consider operators of mass dimension  $\leq 4$ . The most general renormalizable Lagrangian, invariant under the global  $U(1)$  symmetry, is

$$\begin{aligned} \mathcal{L}_{\text{global}} &= \frac{1}{2}\partial^\mu\phi\partial_\mu\phi - V(\phi) + \bar{\psi}(i\gamma_\mu\partial^\mu - m)\psi + k_1\phi\bar{\psi}\psi, \\ V(\phi) &= \frac{1}{2}m^2\phi^2 + \frac{\alpha}{3!}\phi^3 + \frac{\lambda}{4!}\phi^4. \end{aligned} \quad (1.15)$$

The cubic and quartic terms in  $V(\phi)$  are allowed as  $\phi$  is neutral. The Yukawa coupling  $k_1\phi\bar{\psi}\psi$  is also gauge-invariant since the charges of  $\bar{\psi}$ ,  $\phi$ , and  $\psi$  sum to zero ( $-q + 0 + q = 0$ ).

Promoting the global symmetry to a local one ( $\theta \rightarrow \theta(x)$ ) requires introducing a gauge field  $A_\mu$  and replacing ordinary derivatives with covariant derivatives:

$$\mathcal{D}_\mu\varphi = (\partial_\mu - igA_\mu\hat{Q})\varphi \implies \begin{cases} \mathcal{D}_\mu\phi = \partial_\mu\phi, & (\text{since } \hat{Q}\phi = 0) \\ \mathcal{D}_\mu\psi = (\partial_\mu - igqA_\mu)\psi. \end{cases} \quad (1.16)$$

The field strength tensor for the abelian  $U(1)$  field is defined as  $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ . The locally invariant Lagrangian is then:

$$\mathcal{L}_{\text{local}} = \frac{1}{2}\mathcal{D}^\mu\phi\mathcal{D}_\mu\phi - V(\phi) + \bar{\psi}i\gamma_\mu\mathcal{D}^\mu\psi - m\bar{\psi}\psi + k_1\phi\bar{\psi}\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \quad (1.17)$$

With these ingredients and principles, we are now equipped to understand the structure of the Standard Model Lagrangian, which will be discussed in the next section.

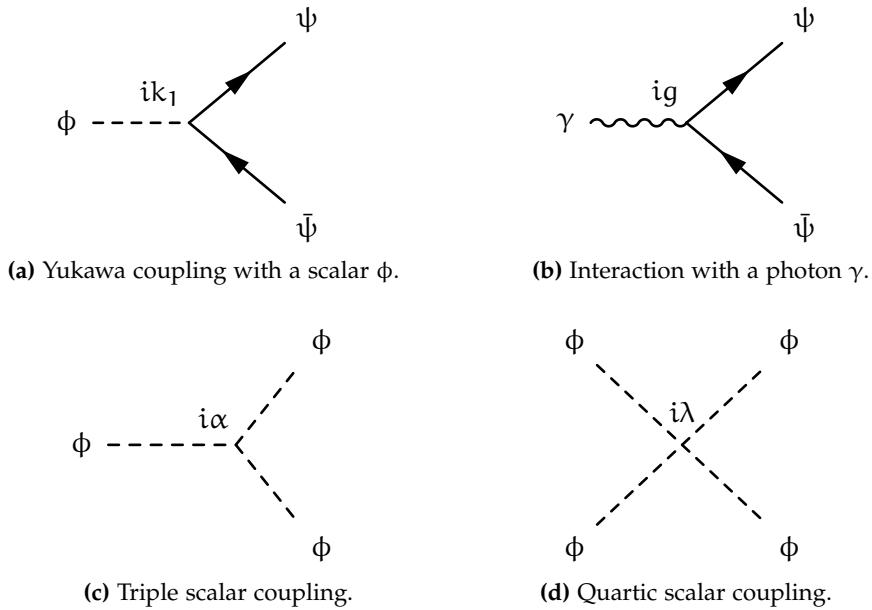
## 1.2 STANDARD MODEL

---

Fragment extracted and adapted from [20]

To contextualize the SM let me place us in 1965. Tomonaga, Feynman, and Schwinger have just won the Nobel prize for their independent contributions on the development of the Quantum Electrodynamics theory [21]. They calculated the magnetic moment of the electron and other observables using quantum field theory and renormalization to separate out the infinities of the theory from a finite contribution [22] showing that renormalized

*Note that for this vector-like  $U(1)$  theory, the explicit fermion mass term  $m\bar{\psi}\psi$  is gauge-invariant. This will not be the case for chiral gauge theories like the Standard Model.*



**Figure 1.1:** Feynman diagrams for Yukawa coupling, gauge boson coupling and quartic scalar coupling.

gauge theories agree with experiment up to very high precision (to more than 13 significant digits)[23].

Unfortunately, in 1965, the models explaining radioactive decay and the strong interaction were not renormalizable. The leading theory was called *the chiral V – A universal model of weak decays* featuring four-fermion interactions in the combination of vector minus axial currents. The V – A model could not be mathematically broken down into a finite and an infinite component. Although gauge theory and renormalization explained the interaction of electrons with photons, gauge theory was not able to address the strong and weak forces. These forces were known to be short-range forces. To make a force have a short range in QFT, the mediating boson needed a mass. The Yukawa theory of scalar fields included such a term as an early model for the strong force with short range. The force law then fell off as  $\exp(-rm)/r^2$  with both the classic inverse square law multiplied by an exponential dampening with distance parameterized by the mass  $m$ . To give a gauge boson  $A_\mu$  a short range, the Lagrangian would need a mass term such as  $m_A^2 A_\mu A^\mu$ . This term violates gauge symmetry because when  $A \mapsto A_\mu + \epsilon_\mu$  we see that  $A_\mu A^\mu \neq A'_\mu A'^\mu$ . Naively, one would think that gauge symmetry blocks all gauge bosons from having mass; and therefore, all gauge theories (Abelian and the non-Abelian ones) would obey force laws that scale as  $1/r^2$ . This would mean that all gauge theories would represent long-range forces similar to gravity and electromagnetism (each

of which is mediated by a massless boson)<sup>1</sup>. There are two known solutions to this quandary:

1. The Higgs mechanism which gives renormalizable gauge bosons mass without violating gauge symmetry.
2. A spontaneously created mass gap phenomena associated with non-Abelian gauge theories, which is not fully understood yet, and seems to be related to the confinement of individual quarks.

The SM chooses (1) the Higgs mechanism for the weak force, and (2) for QCD.

### 1.2.1 PARTICLE CONTENT AND GAUGE GROUP

First, let us talk about the chiral nature of particles: Massive half-spin particles are described at the fundamental level by a Dirac spinorial field, see table 1.1. However, Dirac spinors do not transform under an irreducible representation of the Lorentz group. Spinors can be decomposed into two components that do transform under irreducible representations of the Lorentz group: two *Weyl spinors*. The left and right chiral projectors,  $P_L$  and  $P_R$ , take a Dirac spinor and project it onto each of these invariant subspaces. For a massless Dirac spinor, the left and right components are dynamically decoupled, *i.e.* which are independent fields obeying independent Lagrangian densities; for example, the left component of a massless spinor has the Lagrangian  $\mathcal{L} = -i\bar{\psi}\not{D}P_L\psi$  (For more details see Appendix A at [6]).

The discovery of parity asymmetry in radioactive decays [24] indicates that the chiral description of weak interactions couples differently to the left and right chiral components of half-spin particles. Indeed, the chirality of the fermionic spectrum is possibly one of the deepest properties of the Standard Model. Describing particles in terms of Dirac spinors, it means that left- and right-chirality components actually have different EW quantum numbers. This is compatible with a gauge symmetry only if half-spin particles are considered to be massless, at least without a Dirac mass  $m\bar{f}_R f_L + h.c.$  Nevertheless, half-integer spin fundamental particles, such as the electron, have a well-measured mass. Therefore, the reconciliation of chiral asymmetry and mass lies in the Higgs mechanism, where the masses of the particles result from an effective Yukawa coupling with a scalar, the Higgs boson.

---

<sup>1</sup> In 1954 when Yang was first giving a presentation on non-Abelian gauge theories, Pauli interrupted the talk. Pauli wanted to know what the mass of the non-Abelian gauge boson was. Pauli was so insistent that Yang eventually sat down. Pauli realized that a mass term violated gauge symmetry; the mass terms were needed for short-range forces; non-Abelian gauge theories seemed like they should have long-range forces; and therefore, they probably do not explain strong or weak forces. In short, people no less than Pauli felt gauge symmetry's properties made them unlikely candidates for the a short-range force needed to explain the strong and weak forces [20]

With this in mind, the SM has a content of matter fields from three generations (or families) of quarks  $q$  and leptons  $\ell$ , described as Weyl 2-component spinors, with the structure

$$q_L = \begin{pmatrix} u_L^i \\ d_L^i \end{pmatrix}, u_R^i, d_R^i, \quad \ell_L = \begin{pmatrix} v_L^i \\ e_L^i \end{pmatrix}, e_R^i; \quad i = 1, 2, 3. \quad (1.18)$$

All these particles transform under a group  $U(1)$  with different associated (hyper)charges. The doublets formed by the left components of the fields transform under the representation of two components of a  $SU(2)$  group. The right components do not transform under  $SU(2)$ , therefore they are singlets. In addition, each quark in  $q_L$  transforms as color triplets under  $SU(3)$ , while  $u_R, d_R$  transforms as conjugate triplets. Leptons, on the other hand, turn out to be colored singlets. Gauge quantum numbers of the Standard Model fermions are shown in table 1.2.

Field	$SU(3)_C$	$SU(2)_L$	$U(1)_Y$	$U(1)_{EM}$
$q_L^i = (u^i, d^i)_L$	3	2	+1/3	(2/3, -1/3)
$u_R^i$	3	1	+4/3	+2/3
$d_R^i$	3	1	-2/3	-1/3
$\ell_L^i = (v^i, e^i)_L$	1	2	-1	(0, -1)
$e_R^i$	1	1	-2	-1
$H = (H^+, H^0)$	1	2	+1	(+1, 0)

**Table 1.2:** Gauge quantum numbers of Standard Model quarks, leptons and the Higgs scalar.

Then, we consider the Standard Model as a quantum field theory based on a gauge group

$$G_{SM} = SU(3)_C \times SU(2)_L \times U(1)_Y, \quad (1.19)$$

with  $SU(3)_C$  describing strong interactions via Quantum Chromodynamics (QCD), and  $SU(2)_L \times U(1)_Y$  describing electroweak (EW) interactions. Gauge vector bosons that result from taking this group locally are eight gluons ( $G^a$ ) from each  $t^a$  color-generator of  $SU(3)_C$ , and a linear combination of the three ( $W^\pm, Z$ ) weak bosons and the ( $\gamma$ ) electromagnetic photon from the three  $T^i$  isospin-generators of  $SU(2)_L$  and  $Y$  hyper-charge-generator of  $U(1)_Y$ .

Electroweak symmetry is spontaneously broken into electromagnetic symmetry  $U(1)_{EM}$  via the Higgs mechanism and the Higgs boson  $H$ . The hypercharges  $Y$  of the Standard Model fermions in table 1.2 are related to their usual electric charges by the Gell-Mann–Nishijima relation [25]

$$Q_{EM} = \frac{1}{2}Y + T_3, \quad (1.20)$$

where  $T_3 \doteq \text{diag}(\frac{1}{2}, -\frac{1}{2})$  is an  $SU(2)_L$  generator. Thus, they reproduce electric charge quantization, e.g. the equality in magnitude of the proton and electron charges. Although these hypercharge assignments look rather ad hoc, their values are dictated by the quantum consistency of the theory <sup>2</sup>.

### 1.2.2 GAUGE BOSONS

The Lie algebra of the gauge group  $SU(3) \times SU(2) \times U(1)$  is

$$\begin{aligned}[t^a, t^b] &= if^{abc}t_c, \\ [T^i, T^j] &= i\epsilon^{ijk}T_k, \\ [T^i, Y] &= [t^a, T^j] = [t^a, Y] = 0,\end{aligned}\tag{1.21}$$

where  $f^{abc}$  and  $\epsilon^{ijk}$  are the structure constants of  $SU(3)$  and  $SU(2)$ . And therefore, the gauge fields  $G_\mu$ ,  $W_\mu$ , and  $B_\mu$  must transform in the adjoint representation:

$$\begin{aligned}\delta B_\mu &= \partial_\mu \theta, \\ \delta W_\mu^i &= \partial_\mu \theta^i - g\epsilon^{ijk}\theta^j W_\mu^k, \\ \delta G_\mu^a &= \partial_\mu \epsilon^a - g_s f^{abc}\epsilon^b G_\mu^c.\end{aligned}\tag{1.22}$$

Then, the curvature strength tensors are

$$\begin{aligned}G_{\mu\nu}^a &= \partial_\mu G_\nu^a - \partial_\nu G_\mu^a + g_s f^{abc} G_\mu^b G_\nu^c \\ W_{\mu\nu}^i &= \partial_\mu W_\nu^i - \partial_\nu W_\mu^i + g\epsilon^{ijk} W_\mu^j W_\nu^k \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu\end{aligned}\tag{1.23}$$

and the “kinetic” term for gauge fields in the Lagrangian is

$$\mathcal{L}_{\text{Gauge}} = -\frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu} - \frac{1}{4} W_{\mu\nu}^i W_i^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}.\tag{1.24}$$

while these kinetic terms induce vertices between gauge bosons and in turn do not take into account the masses for such vector bosons, the Higgs mechanism produces the masses for them and gives us the linear combination to the physical bosons  $W^\pm, Z, \gamma$ :

$$\left\{ \begin{array}{l} W_\mu^+ = \frac{1}{\sqrt{2}} (W_\mu^1 - iW_\mu^2) \\ W_\mu^- = \frac{1}{\sqrt{2}} (W_\mu^1 + iW_\mu^2) \\ Z_\mu = c_w W_\mu^3 - s_w B_\mu \\ A_\mu = s_w W_\mu^3 + c_w B_\mu \end{array} \right. \quad \text{where} \quad \left\{ \begin{array}{l} s_w = \sin \theta_w = \frac{g}{\sqrt{g^2 + g'^2}} \\ c_w = \cos \theta_w = \frac{g'}{\sqrt{g^2 + g'^2}} \end{array} \right. \tag{1.25}$$

where to avoid confusion with Dirac matrices, we denote as  $A_\mu$  the electromagnetic potential.

---

<sup>2</sup> It is indeed easy to check that these are (modulo an irrelevant overall normalization) the only (family independent) assignments canceling all potential triangle gauge anomalies.

### 1.2.3 MATTER FIELDS

We refer to the fermionic fields of the SM as the matter fields. We distinguish fermions in these two categories: leptons, fermions that do not have strong interaction, and quarks that interact both strongly and electroweakly. In table 1.3, we can see that there are six leptons, three charged and three neutral: each charged lepton has an associated neutrino forming between them doublets of  $SU(2)_L$  and similarly for quarks.

According to the SM, there are three generations of fermions. Each generation contains a doublet of leptons and a doublet of quarks. Among generations, particles differ by their flavour quantum number and mass, but their strong and electrical interactions are identical. Moreover, the flavour quantum number is a quantity conserved by all interactions except for the weak interaction. Each generation is more massive than the previous one. The second and third generations are unstable and they disintegrate into the first generation. This is why ordinary matter is composed of the first generation. All three generations are produced in nuclear reactors, colliders, and cosmic rays.

Fermion categories		Elementary particle generation		
Type	Subtype	First	Second	Third
Quarks ( $q$ )	up-type	(u) up	(c) charm	(t) top
	down-type	(d) down	(s) strange	(b) bottom
Leptons ( $\ell$ )	charged	(e) electron	( $\mu$ ) muon	( $\tau$ ) tauon
	neutrino	( $\nu_e$ )	( $\nu_\mu$ )	( $\nu_\tau$ )

**Table 1.3:** Three generations of fermions according to the Standard Model of particle physics. Each generation containing two types of leptons and two types of quarks.

Under all the constraints on local gauge invariance and renormalizability of the theory, the fermionic Lagrangian for SM is given by

$$\mathcal{L}_{\text{Fer}} = i\bar{\ell}_L^j \mathcal{D}\ell_L^j + i\bar{e}_R^j \mathcal{D}e_R^j + i\bar{q}_L^j \mathcal{D}q_L^j + i\bar{u}_R^j \mathcal{D}u_R^j + i\bar{d}_R^j \mathcal{D}d_R^j \quad (1.26)$$

where  $\mathcal{D} \equiv \gamma^\mu \mathcal{D}_\mu$  with covariant derivative

$$\mathcal{D}_\mu = \partial_\mu - ig_s t_a G_\mu^a - ig T_i W_\mu^i - ig' \frac{Y}{2} B_\mu, \quad (1.27)$$

and gauge fields  $G^a$ ,  $W^i$ , and  $B$  acting on each kind of fermion via

$$\begin{aligned}\mathcal{D}_\mu \ell_L^i &= \left( \partial_\mu - ig T_j W_\mu^j + i \frac{g'}{2} B_\mu \right) \ell_L^i \\ \mathcal{D}_\mu e_R^i &= \left( \partial_\mu - ig' B_\mu \right) e_R^i \\ \mathcal{D}_\mu q_L^i &= \left( \partial_\mu - ig_s t_a G_\mu^a - ig T_j W_\mu^j - i \frac{g'}{6} B_\mu \right) q_L^i \\ \mathcal{D}_\mu u_R^i &= \left( \partial_\mu - ig_s t_a G_\mu^a - i \frac{2g'}{3} B_\mu \right) u_R^i \\ \mathcal{D}_\mu d_R^i &= \left( \partial_\mu - ig_s t_a G_\mu^a + i \frac{g'}{3} B_\mu \right) d_R^i\end{aligned}\quad (1.28)$$

which couples the fermions to the gauge bosons.

#### 1.2.4 ELETROWEAK SYMMETRY BREAKING

In the SM, the electroweak symmetry  $SU(2)_L \times U(1)_Y$  is spontaneously broken down to the electromagnetic  $U(1)_{EM}$  symmetry by a complex scalar Higgs field transforming as a  $SU(2)_L$  doublet  $H = (H^+, H^0)$  and with hypercharge +1. Its dynamics are parametrized in terms of a potential, devised to trigger a non-vanishing Higgs vacuum expectation value (vev)  $v$ :

$$V = -\mu^2 |H|^2 + \lambda |H|^4 \Rightarrow v^2 \equiv \langle |H| \rangle^2 = \mu^2 / 2\lambda. \quad (1.29)$$

The vev defines the electrically neutral direction and is set to  $\langle H^0 \rangle \approx 170\text{GeV}$  in order to generate the vector boson masses. Simultaneously, it produces masses for quarks and leptons through the Yukawa couplings:

$$\mathcal{L}_{Yuk} = y_u^{ij} \bar{q}_L^i u_R^j \tilde{H} + y_d^{ij} \bar{q}_L^i d_R^j H + y_\ell^{ij} \bar{\ell}_L^i e_R^j H + h.c. \quad (1.30)$$

where  $y_{u,d,l}$  are  $3 \times 3$  complex coupling matrices. These interactions are actually the most general consistent with gauge invariance and renormalizability, and accidentally are invariant under the global symmetries related to the baryon number  $B$  and the three family lepton numbers  $L_i$ <sup>3</sup>. When  $H$  acquires a vacuum expectation value,  $\langle H \rangle = (0, v/\sqrt{2})$ ,  $\mathcal{L}_{Yuk}$  yields mass terms for the quarks and leptons. For quarks, the physical states are obtained by diagonalizing  $y_{u,d}$  by four unitary matrices,  $V_{L,R}^{u,d}$ , as  $M_{diag}^f = V_L^f Y^f V_R^{f\dagger} (v/\sqrt{2})$ ,  $f = u, d$ . As a result, the charged-current  $W^\pm$  interactions couple to the physical  $u_{Lj}$  and  $d_{Lk}$  quarks with couplings given by

---

<sup>3</sup> Regarding the Standard Model as an effective theory, non-renormalizable operators violating these symmetries may, however, be present.

$$\mathcal{L}_{\text{Fer}} \supset \frac{-g}{\sqrt{2}} (\bar{u}_L, \bar{c}_L, \bar{t}_L) \gamma^\mu W_\mu^+ V_{\text{CKM}} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} + \text{h.c.},$$

$$V_{\text{CKM}} \equiv V_L^u V_L^{d\dagger} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}.$$
(1.31)

However, in both flavour-changing charged and neutral currents, the weak interaction at play deals with lepton flavours in a universal manner. This property is known as *Lepton Flavour Universality*; whereas quarks are treated on a different footing due to the CKM matrix. This universality of lepton couplings is assumed when determining the CKM parameters, in particular to combine results from semileptonic and leptonic decays that involve  $e$ ,  $\mu$ , and/or  $\tau$  leptons.

The Lagrangian of the scalar sector is simply

$$\mathcal{L}_H = \mathcal{D}_\mu H^\dagger \mathcal{D}^\mu H - V(H^\dagger, H)$$
(1.32)

where  $\mathcal{D}_\mu H = (\partial_\mu + igT_a W_\mu^a + ig' \frac{Y}{2} B_\mu) H$ , then

$$\begin{aligned} \mathcal{L}_{\langle H \rangle} &= -\frac{1}{8} \begin{pmatrix} 0 & v \end{pmatrix} \begin{pmatrix} gW_\mu^3 - g'B_\mu & g(W_\mu^1 - iW_\mu^2) \\ g(W_\mu^1 + iW_\mu^2) & -gW_\mu^3 - g'B_\mu \end{pmatrix}^2 \begin{pmatrix} 0 \\ v \end{pmatrix} \\ &= -\frac{1}{8} v^2 V_\mu^T \begin{pmatrix} g^2 & 0 & 0 & 0 \\ 0 & g^2 & 0 & 0 \\ 0 & 0 & g^2 & -g'g \\ 0 & 0 & -g'g & g'^2 \end{pmatrix} V^\mu \end{aligned}$$
(1.33)

where  $V_\mu^T = (W_\mu^1, W_\mu^2, W_\mu^3, B_\mu)$ . Diagonalizing this mass matrix, we have that the mass eigenvalues are  $0, -\frac{1}{8}v^2g^2, -\frac{1}{8}v^2g^2$ , and  $-\frac{1}{8}v^2(g^2 + g'^2)$ . The massless boson is the photon, the most massive is the Z boson, and the two intermediate vectors correspond to the bosons  $W^+$  and  $W^-$ , that transform under a representation of the unbroken generator  $Q_{EM}$ .

Having said that, so far, it is enough to understand how the standard model of particle physics as a relativistic field theory describes the interactions of fundamental matter articles via the fundamental forces, mediated by the force carrying particles, the vector bosons. The Higgs boson, also a fundamental Standard Model particle, plays a central role in the mechanism that determines the masses of the photon and weak bosons, as well as the rest of the standard model particles.

Since then, the standard model has faced several experimental tests and has had unprecedented success in explaining the measurements made so far; it has also been a powerful predictive theory. The Standard model has proven successfully at describing many features of nature that we measure in our experiments. The most famous example is the agreement

of the Standard Model prediction and the experimental measurement of the electron magnetic dipole moment to within twelve significant figures of accuracy [26]. The 2012 discovery of the Higgs boson was the culmination of almost fifty years of searching for the particle first predicted to exist in 1965 and first incorporated into the Standard Model in 1967 with Glashow, Weinberg, and Salam's unification of the electromagnetic and weak forces [27, 28]. With the 2012 Higgs discovery, the full predicted particle spectrum of the Standard Model was finally observed.

### 1.3 DEFICIENCIES OF STANDARD MODEL AND NEW PHYSICS

---

#### Pending to be updated

While these and other successes of the Standard Model are an achievement for the field of particle physics, it is well known that this cannot be the ultimate theory of fundamental particles and interactions. Even though the Standard Model is currently the best description there is of the subatomic world, it does not explain the complete picture; there are also important questions that it does not answer and it is also surrounded by different irregularities. Some of them are completely incompatible with the current Standard Model, and strongly suggest that the Standard Model requires a consistent extension to solve experimental and theoretical problems that we will label as the cosmological problems, phenomenological problems, and theoretical problems. Below we will list very briefly the main representatives of these categories.

##### 1.3.1 THEORETICAL PROBLEMS

**HIERARCHY PROBLEM** Is the problem concerning the large discrepancy between aspects of the weak force and gravity. Both of these forces involve constants of nature, the Fermi constant for the weak force and the Newtonian constant of gravitation for gravity. If the Standard Model is used to calculate the quantum corrections to Fermi's constant, it appears that Fermi's constant is surprisingly large and is expected to be closer to Newton's constant unless there is a delicate cancellation between the bare value of Fermi's constant and the quantum corrections to it.

In the Standard Model context, the Higgs boson is much lighter than the energy scale on which the standard model is considered valid (ideally the Plank mass), and the quantum corrections to the Higgs mass are on the order of this energy scale; it would inevitably make the Higgs and fermions masses huge, comparable to the scale at which new physics appears, unless there is an incredible fine-tuning cancellation between the quadratic radiative corrections and the bare mass. This level of fine-tuning is deemed unnatural.

**STRONG CP PROBLEM** QCD Lagrangian supports a term associated with the strength tensor dual for gluons that break CP symmetry in the strong interaction sector. Experimentally, however, no such violation has been found, implying that the coefficient of this term is fine tuned to zero.

**QUANTUM TRIVIALITY** Suggests that it may not be possible to create a consistent quantum field theory involving elementary scalar Higgs particles because for high momentum particles the renormalization presents inconsistencies unless the renormalization of the charges becomes null, and therefore not interacting, *i.e.* trivial. Nevertheless, because the Higgs boson plays a central role in the Standard Model of particle physics, the question of triviality in Higgs models is of great importance.

**NUMBER OF PARAMETERS AND UNEXPLAINED RELATIONS** In total, the standard model has too many free parameters (19 in total) that are obtained experimentally, and there are indications that several of them may be correlated, however the origin of these correlations is beyond the standard model.

For example, Yoshio Koide's empirical formula [29]

$$\frac{m_e + m_\mu + m_\tau}{(\sqrt{m_e} + \sqrt{m_\mu} + \sqrt{m_\tau})^2} = 0.666661(7) \approx \frac{2}{3}$$

seems to indicate that there is a way to predict the masses of leptons.

### 1.3.2 COSMOLOGICAL PROBLEMS

**GRAVITY** Although the Standard Model describes the three important fundamental forces at the subatomic scale, it does not include gravity. However, at larger scales, gravity becomes present and is described by Einstein's theory of general relativity, in which gravity rather than a force is a property that measures the deformation of spacetime then, the most of the conventional machinery of perturbative QFT is profoundly incompatible with the general relativistic framework [30], and a theory of quantum gravity with which we are enabled to perform calculations has yet to be discovered.

**DARK MATTER** Within the framework of Einstein's general relativity, the cosmological standard model ( $\Lambda$ CDM) is, like the standard model of particle physics, one of the most successful theories of the 20th century.  $\Lambda$ CDM it is based on a very specific density of matter that can be explained with ordinary matter from the standard model of particles, baryonic matter; according to  $\Lambda$ CDM, in addition to baryonic matter, there is a kind of matter five times more abundant than baryonic matter, which does not interact electrically (therefore it is dark) and non-relativistic (therefore it is cold), known as cold

dark matter (CDM). Yet, the Standard Model does not supply any fundamental particles that are good dark matter candidates.

**DARK ENERGY** Moreover, according to Lambda CDM only 31% of the energy that makes up the universe is matter, the remaining 69% of the universe's energy should consist of the so-called dark energy, a constant energy density for the vacuum ( $\Lambda$ ). If we try to explain dark energy in terms of vacuum energy only from the standard model lead to a mismatch of 120 orders of magnitude [31], sometimes called *The Worst Theoretical Prediction in the History of Physics* [32]; a bit sensationalist title to indicate the fact that we do not fully understand the composition of the particle spectrum of the universe.

**MATTER-ANTIMATTER ASYMMETRY** In the observable universe there is more matter than antimatter. In 1967, Andrei Sakharov proposed a set of three necessary conditions that a baryon-generating interaction must satisfy to produce matter and antimatter at different rates [33]. While the standard model can satisfy these three conditions [34, 35], it satisfies them at three different energy scales and therefore presents difficulties in the capability to explain the matter-antimatter asymmetry [20].

### 1.3.3 PHENOMENOLOGICAL PROBLEMS

**NEUTRINO MASSES** In the standard model, the right chiral component of neutrinos is not part of the composition of fermionic fields because if they were present they would not interact and consequently neutrinos have no mass. However, the precision measurement [36] of the mixing matrix for neutrino oscillations has shown that neutrinos change flavour in free flight and in turn that the three neutrino flavours cannot have identical mass, meaning that all three cannot have zero mass. There is no single way to extend the standard model to include masses to neutrinos and even more to explain their value so close to zero and results in the open problem confirmed at the phenomenological level present in the standard model.

**ANOMALOUS B-MESONS DECAY** A B-meson is a bound state made up of an quark-antiquark pair where one of them comes from a b-quark. Various experimental results [37–41] have suggested a surplus over Standard Model predictions in its decays to D-mesons along with a  $\tau$ ,  $\nu_\tau$  doublet. While none of them have reached the statistical threshold of  $5\sigma$  to declare a break from the standard model, the Capdevila's meta-analysis of all available data reported a  $5.0\sigma$  deviation from SM [42].

**ANOMALOUS MAGNETIC DIPOLE MOMENT OF MUON** Unlike the extraordinary agreement between theory and experiment with the magnetic

dipole moment of the electron [26]; in the case of the muon, the measurement of Fermilab’s Muon g-2 experiment has presented an apparent discrepancy with an accuracy of  $4.2\sigma$  [43, 44] which strengthen evidence of new physics in the muon sector and apparently in the violation of lepton universality of the standard model.

**ANOMALOUS MASS OF THE W BOSON** Results from the CDF Collaboration, reported in April 2022, indicate that the mass of a W boson exceeds the mass predicted by the Standard Model with a significance of  $7\sigma$  [45]. However, this very highly accurate result, unlike the anomaly in B-meson Decay, is in tension with the results of Atlas, LHCb, LEP and Do II [46–49]. Certainly, a review of all the information we possess so far must be done to determine if this anomaly is a window into new physics beyond the standard model.

**CCA AND  $q\bar{q} \mapsto e^+e^-$**  It has been observed that certain nuclear beta decays happen less frequently than expected [50]. This tension, called the Cabibbo Angle anomaly (CAA), displays a significance around  $3\sigma$  [23], and can again be interpreted as a sign that electrons and muons behave more differently than predicted by the SM [51]. Furthermore, the CMS experiment at CERN observed more very high-energetic electrons in proton-proton collisions ( $q\bar{q} \rightarrow e^+e^-$ ) compared to muons than expected [52].

# 2

---

## PHENOMENOLOGICAL FRAMEWORK FOR LHC SEARCHES AND ANALYSIS

---

Since its formulation, the Standard Model (SM) has proven remarkably successful in describing the fundamental particles and interactions, and its parameters have been measured with increasing precision over several decades [23, 27]. However, as discussed in the previous chapter, various theoretical and experimental observations suggest that the SM is incomplete [53, 54]. As outlined previously, this is motivated by theoretical shortcomings such as the hierarchy problem [55, 56], the absence of a dark matter candidate [57], and non-zero neutrino masses [58] as well as by experimental anomalies [59–61]. These limitations motivate the search for new physics (NP) beyond the SM (BSM) [54, 62].

The search for BSM physics proceeds along two main axes: the construction of theoretical extensions to the SM [54, 56, 62], and the development of experimental methods to probe them [63–65]. A necessary condition for any viable BSM model is consistency with existing experimental data, which places strong constraints on its parameter space [23, 66, 67]. These constraints include lower limits on the masses of new particles from direct searches at high-energy colliders [66, 67], and upper bounds on couplings and mixing angles from precision measurements at both high and low energies [39, 48], which are sensitive to virtual corrections [68].

Phenomenology connects theoretical models to experimental observables by calculating cross sections, decay rates, and other signatures for given model parameters [63, 69, 70]. A critical function of this field is to assess the experimental feasibility of BSM scenarios—evaluating whether predicted signals would be observable above background processes given the capabilities of current and future detectors [63–65]. This involves estimating production rates [63, 69, 71], modeling detector acceptance and efficiency [65, 72–74], and developing discrimination variables to maximize signal-to-background ratios [75–78]. This feasibility assessment is essential for designing analysis strategies, particularly at the Large Hadron Collider (LHC), where signals of new physics must be discriminated from large Standard Model backgrounds [63, 65].

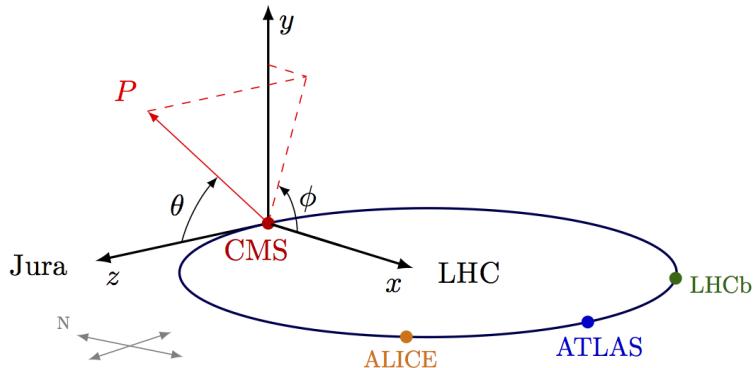
This feasibility assessment is essential for designing analysis strategies at the Large Hadron Collider (LHC), a proton-proton ( $\text{pp}$ ) collider operating since 2009 [72, 73, 79]. The LHC has provided data at center-of-mass energies from 7 TeV to 13.6 TeV [67, 74]. During Run I (2010–2013), operations at 7–8 TeV led to the discovery of the Higgs boson using a

dataset corresponding to an integrated luminosity of roughly  $30 \text{ fb}^{-1}$  [80, 81]. Run II (2015–2018) significantly expanded this dataset, collecting approximately  $140 \text{ fb}^{-1}$  at 13 TeV [67, 74]. Run III (2022–2025) is currently underway at 13.6 TeV and is more than doubling the available data, with a target of over  $300 \text{ fb}^{-1}$  [74]. Future operations will be dominated by the High-Luminosity LHC (HL-LHC), starting around 2029, which is designed to accumulate an unprecedented integrated luminosity of  $3000 \text{ fb}^{-1}$  [82, 83]. This vast increase in data volume enables searches for exceedingly rare processes but also requires discriminating potential signals of new physics from correspondingly large and complex SM backgrounds, making sophisticated phenomenological tools increasingly important [63, 65].

## 2.1 DETECTORS AND SUBSYSTEMS

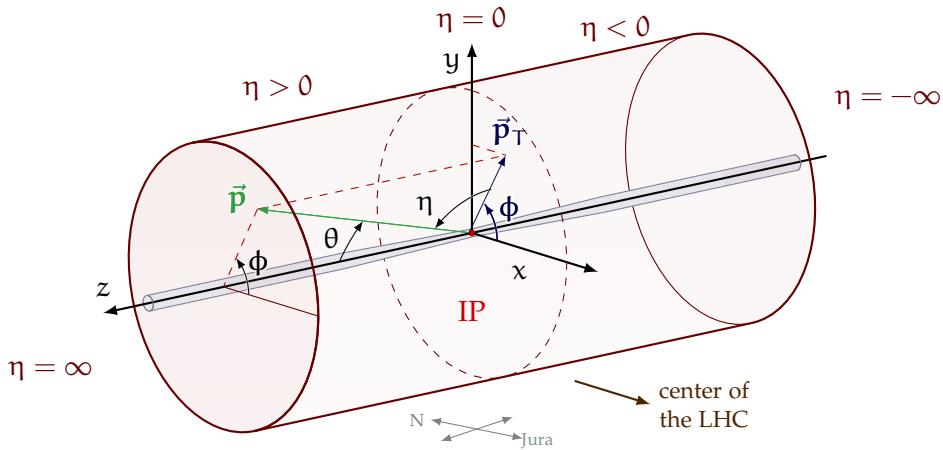
When two particle bunches from colliding beams cross each other, they generate individual interactions known as events [72, 73]. At the LHC, the beam intensity is so high that multiple interactions can take place in a single event; this phenomenon is referred to as pile-up [83, 84]. In other words, the probability that several proton-proton interactions occur within the same bunch crossing is non-negligible, leading to multiple overlapping events in a single detector readout [70, 85]. These collisions occur at four main interaction points, each hosting a large particle detector designed to record and analyze the outcomes [72, 73].

Among them, the Compact Muon Solenoid (CMS) and ATLAS are the largest and most comprehensive experiments [72, 73]. Both are multipurpose detectors with broad physics programs, capable of exploring a wide range of phenomena [72, 73]. They perform precision measurements within the electroweak sector of the Standard Model [23], probe the dynamics of quarks and gluons (including through heavy-ion collisions) [65], and conduct extensive searches for physics beyond the Standard Model using pp collision data [80, 81]. While CMS and ATLAS differ in their detector designs and reconstruction strategies, their physics goals are largely overlapping, and their results are complementary [72, 73].



**Figure 2.1:** Coordinate system employed by the CMS experiment (retrieved from [86]).

Throughout this work, phenomenological studies and comparisons are primarily developed in the context of CMS, although several results from ATLAS are also referenced, given the close alignment in sensitivity and scope [72, 73]. Measurements performed at CMS adopt a right-handed coordinate system with its origin at the nominal collision point [73]. The z-axis is defined along the beam direction, the x-axis points radially inward toward the center of the LHC ring, and the y-axis points vertically upward [73]. The azimuthal angle  $\phi$  is measured in the transverse (xy) plane from the x-axis, while the polar angle  $\theta$  is measured from the z-axis, as shown in Fig. 2.1 [73]. Moreover, for kinematic analysis at hadron colliders, the Cartesian coordinate system is often reparameterized into quantities that are more physically meaningful and experimentally convenient as shown in Fig. 2.2 [87]:



**Figure 2.2:** Detailed reparametrization of the coordinate system employed by the CMS experiment (retrieved from [86])

**PSEUDO-RAPIDITY** ( $\eta$ ) Instead of using the polar angle, CMS measurements involve the pseudo-rapidity [23, 73], defined by

$$\eta = -\ln \left( \tan \frac{\theta}{2} \right)$$

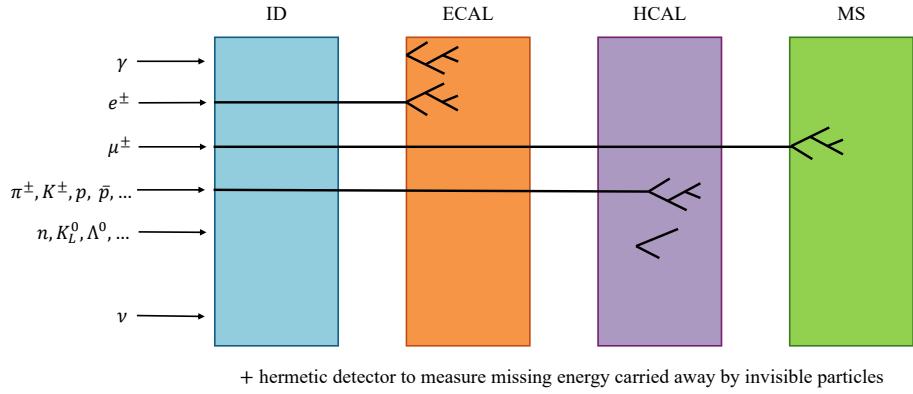
The main advantage of using the pseudo-rapidity is that distributions over it tend to be closer to a uniform distribution than those over the polar angle [87], see Fig. 2.2. Furthermore, the difference in pseudo-rapidity is invariant under Lorentz boosts along the beam direction [23].

**TRANSVERSE MOMENTUM** ( $p_T$ ) Refers to the component of momentum which is perpendicular to the beam line [87]. It is usually preferred over full momentum because momentum along the beamline may just be left over from the beam particles, while the transverse momentum is always associated with whatever physics happened at the vertex [87], see Fig. 2.2.

**AZIMUTHAL ANGLE** ( $\phi$ ) Measures the angle in the transverse plane relative to the  $x$ -axis, providing the directional component perpendicular to the beam line [73].

Together, the triplet  $(p_T, \phi, \eta)$  forms a natural coordinate system that fully describes a particle's three-momentum vector at a hadron collider [23, 87]. The full four-momentum  $(E, p_x, p_y, p_z)$  can be reconstructed from these quantities, typically supplemented by either the particle's mass hypothesis (for identified particles like electrons or muons) or the energy deposited in the calorimeters (for neutral objects like photons or jets) [70, 88, 89]. This  $(p_T, \phi, \eta)$  system serves as the fundamental framework for defining physical objects, calculating event variables, and performing analyses at the LHC, providing both experimental convenience and physical insight into the collision dynamics [70, 87].

A key challenge is isolating the primary hard interaction from the additional concurrent pile-up interactions [83, 84]. This is accomplished by reconstructing distinct interaction vertices along the beam direction and associating charged particles to their point of origin using the CMS tracking and vertexing algorithms [87, 90]. The ultimate aim of the reconstruction chain is to identify all stable particles produced in the collision and measure their four-momenta, thereby enabling the identification of the underlying fundamental process [87].

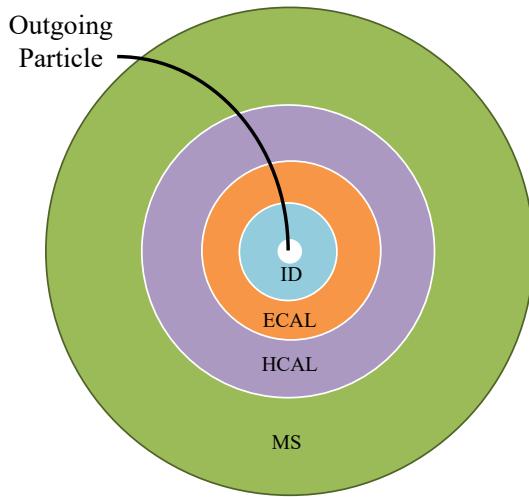


**Figure 2.3:** Illustration of high-energy particles being identified by consecutive types of subdetectors in a typical collider experiment. The curvature of the tracks in the magnetic field is not shown for simplicity. Representation of which particles and kinds of detectors are used in a multipurpose detector such as CMS or ATLAS.

However, the reconstruction is complicated by several factors [73, 87]. The initial state of the colliding protons is not fully known, as they are composite particles made up of quarks and gluons (collectively referred to as partons) [71, 91]. The fraction of the proton's momentum carried by each parton is described by parton distribution functions (PDFs), which are determined experimentally and recommended for LHC Run II/III use by PDF4LHC efforts [71, 92]. Consequently, the total momentum along the beam axis ( $z$ ) is not balanced on an event-by-event basis [91]. Furthermore,

not all particles are stable enough to reach the detector; some decay before being detected, and only their decay products are observed [23]. The design of a collider experiment, illustrated in Fig. 2.3, is optimized for the identification and energy measurement of the particles produced in high-energy collisions [65, 73].

Finally, some hypothetical particles, such as those comprising dark matter, along with known neutrinos, interact very weakly with matter and escape direct detection [23, 57]. Therefore, a hermetic detector design is crucial to infer their presence by accurately measuring the imbalance of energy and momentum in the transverse plane, referred to as missing transverse momentum [87, 93].



**Figure 2.4:** Schematic representation of a transverse section of a generic multipurpose detector. The inner detector (ID) is used to measure the trajectories of charged particles, the electromagnetic calorimeter (ECAL) measures the energy of photons and electrons, the hadronic calorimeter (HCAL) measures the energy of hadrons, and the muon system (MS) identifies and measures muons. The missing transverse momentum (MET) is inferred from the momentum imbalance in the transverse plane.

In this way, a typical collider experiment comprises several main detector subsystems that are used jointly to detect and measure the properties of particles produced in the collision [65, 72, 73, 87]. A *schematic representation* of such a generic multipurpose detector is shown in Fig. 2.4 [65, 73]. The detector features an "onion-like" design of several concentric layers, each optimized to identify different types of particles and measure their properties [73, 87].

The innermost subsystem, the inner detector (ID) or tracker, is immersed in a strong axial magnetic field (typically 1–4 T) [73, 90]. It is designed to reconstruct the trajectories of charged particles, which are bent by the magnetic field [87, 90]. The direction and curvature of these

trajectories, called **tracks**, yield the particle's momentum vector and electric charge [23, 90]. The most common long-lived charged particles from the SM are leptons (electrons  $e$  and muons  $\mu$ ) and hadrons (pions  $\pi$ , kaons  $K$ , and protons  $p$ ) [23]. In some detectors, the ID is complemented by a Cherenkov light detector (RICH) to measure particle velocity and aid particle identification [23, 94]. Combined with the momentum measurement, this velocity helps determine the particle mass, allowing for differentiation between pions, kaons, and protons [23, 94].

After the tracker, particles enter the electromagnetic calorimeter (ECAL), which is designed to fully absorb photons, electrons, and positrons [73, 88]. These particles deposit all their energy in the ECAL by initiating an electromagnetic shower via bremsstrahlung and  $e^+e^-$  pair production [88]. Electrons are identified as charged tracks that point to a compact, high-energy deposit in the ECAL [88].

The hadronic calorimeter (HCAL) surrounds the ECAL and is built to absorb hadrons and measure their energy through hadronic interactions [65, 73]. High-energy quarks and gluons hadronize into collimated sprays of hadrons known as **jets**; a jet's energy is measured by combining calorimeter deposits with track momenta, an approach formalized in particle-flow reconstruction [70, 85, 87].

Muons are unique as they can penetrate the calorimeters; a dedicated muon system outside the calorimeters identifies and measures muons, and muon tracks in the ID are matched to tracks in the muon chambers [73, 89]. Since the detector is nearly hermetic, momentum conservation in the transverse plane is a powerful tool: any significant imbalance (missing transverse momentum, MET) signals undetected neutral particles such as neutrinos or potential dark-matter candidates [87, 93].

Since the detector is nearly hermetic (covering almost the full solid angle), momentum conservation in the plane transverse to the beam line ( $x$ - $y$  plane) is a powerful tool. The vector sum of the momenta in the transverse plane ( $\vec{p}_T$ ) of all detected particles should be zero. Any significant imbalance indicates the presence of undetected, neutral particles that did not interact with the detector, such as neutrinos or hypothetical new particles. This imbalance is called missing transverse momentum (MET) and is formally defined as:

$$\vec{p}_T^{\text{miss}} \equiv - \sum_i \vec{p}_{T,i}$$

where the sum runs over all reconstructed particles (e.g., leptons, photons, jets) or calorimeter deposits in the event.

This detector design, optimized for identifying and measuring Standard Model particles, also makes it a powerful instrument for searching for new physics beyond the SM through unusual signatures or an excess of events with large MET.

### 2.1.1 COLLISION PARAMETERS

A aim of particle physics experiments is to quantify how frequently different processes occur and to characterize the properties of the particles involved. These quantities are expressed in terms of **cross-sections**, normalized to the **luminosity** delivered by the accelerator.

The cross-section ( $\sigma$ ) quantifies the probability for a specific process to occur. Formally, it represents the effective area of a target particle presented to an incoming beam particle for an interaction to happen. It has units of area, typically barn (b), where  $1 \text{ b} = 10^{-28} \text{ m}^2$ .

In the context of proton-proton (pp) collisions at the LHC, the concept is generalized. Since both colliding particles are composite, the cross-section for a specific process is calculated by considering the interactions between their constituent partons (quarks and gluons). The total cross-section for a process  $\text{pp} \rightarrow X$  is given by the convolution of the parton distribution functions (PDFs) and the partonic cross-section  $\hat{\sigma}_{ij \rightarrow X}$ :

$$\sigma(\text{pp} \rightarrow X) = \sum_{i,j} \int_0^1 dx_1 dx_2 f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) \hat{\sigma}_{ij \rightarrow X}(\hat{s}, \mu_F^2, \mu_R^2), \quad (2.1)$$

where:

- The sum runs over all possible parton types  $i, j$  (e.g.,  $u, d, g$ ) in the two protons.
- $f_i(x, \mu_F^2)$  is the PDF, representing the probability density to find a parton of type  $i$  carrying a fraction  $x$  of the proton's momentum at a factorization scale  $\mu_F$ .
- $\hat{s} = x_1 x_2 s$  is the square of the center-of-mass energy for the colliding partons, with  $s$  being the square of the pp center-of-mass energy (e.g., 13.6 TeV).
- $\mu_R$  is the renormalization scale.
- $\hat{\sigma}_{ij \rightarrow X}$  is the partonic cross-section for the hard scattering process  $ij \rightarrow X$ .

Then, in one hand, the cross-section  $\sigma$  is a theoretical quantity that encapsulates the fundamental physics of the interaction, independent of the accelerator's performance. On the other hand, the **luminosity** ( $\mathcal{L}$ ) is a property of the particle accelerator and beams. It measures the density of particles in the colliding beams and thus the rate at which interactions can occur. The instantaneous luminosity is defined by:

$$\mathcal{L} = \frac{fn_1 n_2}{4\pi\sigma_x\sigma_y}, \quad (2.2)$$

where  $f$  is the revolution frequency of the bunches,  $n_1$  and  $n_2$  are the numbers of particles in each bunch, and  $\sigma_x$  and  $\sigma_y$  are the transverse

dimensions of the beams at the interaction point. The integrated luminosity is the integral of the instantaneous luminosity over time:

$$L = \int \mathcal{L} dt. \quad (2.3)$$

The primary unit of integrated luminosity is the inverse barn ( $b^{-1}$ ), commonly  $fb^{-1}$ .

The expected number of events ( $N$ ) for a given process in a collider experiment connects the cross-section and integrated luminosity through the relation

$$N = \sigma \cdot L \cdot \epsilon. \quad (2.4)$$

Here,  $\epsilon$  is the product efficiency  $\times$  acceptance of the detector for the process in question [65, 87]. It estimates the fraction of produced events that are actually detected and reconstructed, accounting for detector geometry, resolution, and analysis selection criteria [87].

On one side, if a process is known, we can measure the number of events  $N$ , know the integrated luminosity  $L$  from accelerator measurements [74], and calculate the efficiency  $\epsilon$  from simulation (event generation, parton shower and detector simulation) [63, 65, 69]. Equation 2.4 is then solved for  $\sigma$  to provide a measurement of the production rate, with statistical interpretation based on likelihood methods [95]. On the other side, if we are searching for a new process not predicted by the Standard Model, we calculate the expected background  $N_{bkg}$  from known SM processes using Monte Carlo and data-driven techniques [63, 70]. We then observe a number of events  $N_{obs}$ . Any significant excess,  $N_{obs} - N_{bkg}$ , can be interpreted as a potential signal, and Equation 2.4 can be used to set an upper limit on the cross-section  $\sigma_{BSM}$  of the hypothetical new process using standard limit-setting procedures (CLs, profile-likelihood, treatments of nuisances) [96–98].

## 2.2 JETS RECONSTRUCTION

---

Quarks and gluons are never observed as free particles because of colour confinement [99, 100]. Nevertheless, perturbative QCD treats them as the relevant short-distance degrees of freedom: factorization theorems and asymptotic freedom justify computing hard-scattering matrix elements for incoming and outgoing partons even though QCD becomes non-perturbative at low scales [91]. The strong coupling  $\alpha_s$  grows large and effectively “blows up” around the confinement scale  $\Lambda_{QCD}$  [23]; consequently, something must happen to quarks and gluons before they reach the detector [69]. In practice, the gluon and all quarks except the top hadronize, producing cascades of baryons and mesons that themselves undergo further decays; hadronization is modelled e.g. with the Lund string or cluster models [69, 99, 100]. At the LHC, these hadrons typically carry energies comparable to the electroweak scale, and relativistic boosts tend to

collimate their decay products into narrow bunches [101]. Those collimated collections of hadrons are the jets we measure at hadron colliders and the objects we use to infer the partons produced in the hard interaction [85, 101].

Each high-energy parton produced in a collision, such as a quark from the process  $gg \rightarrow q\bar{q}$ , undergoes hadronization over a distance scale of  $\sim 10^{-15}$  m, producing a jet of hadrons [69, 99]. The energy composition of these jets is phenomenologically well established and is the basis of particle-flow reconstruction: on average roughly  $\sim 60\%$  of the jet energy is carried by charged particles (mostly  $\pi^\pm, K^\pm$ ),  $\sim 30\%$  by photons (from  $\pi^0 \rightarrow \gamma\gamma$ ) and  $\sim 10\%$  by neutral hadrons [87]. In high-energy jets, the particles can be too collimated to be resolved individually in coarse calorimeter segmentation; nevertheless, the jet four-momentum is reconstructed from clustered PF candidates or calorimeter deposits and then corrected using jet energy corrections derived from simulation and in-situ data [65, 70, 87].

Phenomenologically one usually assumes that each high-energy parton yields a jet and that the measured jet four-momentum can, to useful accuracy, be related to the original parton four-momentum [102, 103]. Jets are therefore defined operationally using recombination (clustering) algorithms such as Cambridge–Aachen [104] or the (anti-)k<sub>T</sub> family [85]. Experimentally this means grouping a large number of energy depositions (or particle-flow candidates) observed in the calorimeters and tracker into a much smaller set of jets or sub-jets [87]. Nothing in the raw detector data, however, indicates a priori how many jets there should be: the clustering procedure and the choice of a resolution scale fix the outcome [101]. In practice one must either specify the desired number of final jets or choose a resolution/stop criterion (for example a distance parameter R, a clustering distance cut, or a jet-mass/sub-jet-resolution threshold) that determines the smallest substructure to be considered a separate parton-like object [105].

Modern reconstruction at the LHC typically uses particle-flow (PF) candidates as input together with infrared- and collinear-safe clustering algorithms to define jet four-momenta [70, 87]. The anti-k<sub>T</sub> algorithm [85], implemented in FastJet [70], is widely used in ATLAS and CMS; it groups candidates by proximity in the rapidity–azimuth ( $y, \phi$ ) plane with a typical distance parameter  $R \sim 0.4\text{--}0.6$  and is relatively insensitive to soft radiation and pileup when combined with area-based subtraction techniques [106]. After clustering, jet energy corrections (JEC) derived from simulation and in-situ calibrations compensate for detector response, pileup, and underlying-event effects [107], while jet-substructure and tagging algorithms (mass-drop, N-subjettiness, SoftDrop, etc.) help infer the flavour and origin of the initiating parton [105, 108, 109].

### 2.2.1 JET ALGORITHMS

Recombination (or sequential clustering) algorithms formalise the intuitive idea that parton showering produces collinear and soft splittings [101,

[102, 104]: two nearby and kinematically compatible sub-jets are merged if they are more likely to have originated from a single parton [102, 104]. A practical implementation requires a measure of “distance” between objects [70, 102]; common choices combine an angular separation in the rapidity–azimuth plane,  $\Delta R_{ij}$ , with a transverse-momentum weighting [85, 102]. Typical distance measures are [85, 102, 104]

$$\begin{aligned} k_T : \quad y_{ij} &= \frac{\Delta R_{ij}}{R} \min(p_{T,i}, p_{T,j}), \quad y_{iB} = p_{T,i}, \\ C/A : \quad y_{ij} &= \frac{\Delta R_{ij}}{R}, \quad y_{iB} = 1, \\ \text{anti-}k_T : \quad y_{ij} &= \frac{\Delta R_{ij}}{R} \min(p_{T,i}^{-1}, p_{T,j}^{-1}), \quad y_{iB} = p_{T,i}^{-1}. \end{aligned} \quad (2.5)$$

The parameter  $R$  balances jet–jet and jet–beam criteria and sets the geometric size of jets; in LHC analyses, typical values are  $R \sim 0.4\text{--}0.7$  depending on the physics target [70].

Two operational modes are useful to distinguish. In an exclusive algorithm, one supplies a resolution scale  $y_{cut}$  and proceeds iteratively:

1. compute  $y^{\min} = \min_{i,j}\{y_{ij}, y_{iB}\}$ ;
2. if  $y^{\min} = y_{ij} < y_{cut}$  merge  $i$  and  $j$  and repeat;
3. if  $y^{\min} = y_{iB} < y_{cut}$  remove  $i$  as beam radiation and repeat;
4. stop when  $y^{\min} > y_{cut}$  and keep remaining sub-jets as jets.

An inclusive algorithm omits  $y_{cut}$  and instead declares a sub-jet a final-state jet when its jet–beam distance is the smallest quantity; iteration continues until no inputs remain. Inclusive algorithms therefore produce a variable number of jets, while exclusive algorithms deliver a scale-dependent fixed set.

A practical question is how to combine the kinematics of merged objects. The most common choice in modern experiments is the E-scheme: four-vectors are added, which preserves energy–momentum and yields a physical jet mass useful for substructure and boosted-object tagging. An alternative is to sum three-momenta and rescale the energy to enforce a massless jet; this can be appropriate when the analysis targets massless parton kinematics, but it discards potentially useful jet-mass information.

From a theoretical and experimental viewpoint, important properties are infrared and collinear safety: a jet algorithm should give stable results under the emission of soft particles or collinear splittings. The  $k_T$ , C/A and anti- $k_T$  families are constructed to satisfy these requirements. Their practical behavior differs:  $k_T$  naturally follows the physical shower history soft-first clustering, C/A is purely geometric useful for declustering and substructure studies, while anti- $k_T$  produces regular, cone-like jets that are robust and convenient experimentally.

Corrections for pileup and the underlying event are necessary at the LHC. These corrections depend on the jet area and are typically performed

by estimating an event-wide transverse-momentum density and subtracting the corresponding contribution proportional to the jet area. Finally, because inclusive algorithms can produce jets arbitrarily close to the beam, a minimum jet  $p_T$  threshold, commonly 20–100 GeV depending on the analysis, is imposed to ensure experimental observability and theoretical control.

### 2.2.2 $\tau$ TAGGING AT MULTIPURPOSE DETECTORS

The  $\tau$  lepton decays hadronically with a probability of  $\sim 65\%$ , producing a narrow “ $\tau$ -jet” that contains only a few charged and neutral hadrons [23, 110]. Hadronic decays are dominated by one- and three-prong topologies and often include neutral pions that promptly convert to photons, giving a sizable electromagnetic fraction in the calorimeters [23, 111]. When the  $\tau$  momentum is large compared to its mass the decay products are highly collimated [78, 110]: for  $p_T > 50$  GeV roughly 90% of the visible energy is contained within a cone of radius  $R = \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2} = 0.2$  [110]. These properties motivate the use of small signal cones and narrow isolation annuli in reconstruction [78, 110].

Identification exploits three complementary classes of observables [78, 87, 110, 111]:

- Calorimetric isolation and shower-shape variables [110, 111]: hadronic  $\tau$  decays deposit localized energy in ECAL+HCAL [110]. Experiments use isolation sums and shape ratios to quantify peripheral activity [110, 111]. Example variables are

$$\Delta E_T^{12} = \frac{\sum_{0.1 < \Delta R < 0.2} E_{T,j}}{\sum_{\Delta R < 0.4} E_{T,i}}, \quad P_{ISOL} = \sum_{\Delta R < 0.40} E_T - \sum_{\Delta R < 0.13} E_T, \quad (2.6)$$

which suppress QCD jets that populate the isolation ring [110].

- Charged-track isolation and prong topology [87, 110]: the few, collimated charged tracks of a  $\tau$  allow powerful selections. A common procedure defines a matching cone of radius  $R_m$  around the calorimeter jet axis to select candidate tracks above a  $p_T^{\min}$  threshold; the leading track ( $tr_1$ ) defines a narrow signal cone  $R_S$  (1- or 3-prong hypotheses) and a larger isolation cone  $R_I$  is scanned for additional tracks [87, 110].
- Lifetime and vertexing observables [23, 90]: the finite  $\tau$  lifetime ( $c\tau \approx 87$   $\mu\text{m}$ ) produces displaced tracks and, for multi-prong decays, a reconstructible secondary vertex; impact-parameter significances and secondary-vertex properties are exploited to separate genuine  $\tau_h$  from prompt jets or leptons [23, 90].

Additional discriminants include the invariant mass of the visible decay products computed from tracks and calorimeter clusters, electromagnetic energy fractions (sensitive to  $\pi^0 \rightarrow \gamma\gamma$ ), and dedicated shower-strip

grouping for nearby photons. For example, invariant-mass reconstruction commonly uses a jet cone  $\Delta R_{\text{jet}} \lesssim 0.4$  while excluding calorimeter clusters matched to tracks by a minimum separation  $\Delta R_{\text{track}} \gtrsim 0.08$  to reduce double counting.

Reconstruction algorithms combine these inputs. CMS’s Hadron-Plus-Strips (HPS) and modern DeepTau methods explicitly build decay-mode hypotheses and use strip-clustering of photons plus multivariate or deep-learning discriminators to reject jets, electrons, and muons [78, 112]. ATLAS employs analogous calorimeter+track based MVAs and BDTs [113]. Typical working points trade efficiency versus background: medium points often give  $\tau_h$  efficiencies of order 50–70% with light-jet misidentification rates in the per-mille to percent range, depending on kinematics and pileup.

Practical implementations tune cone sizes, isolation thresholds, and MVA inputs to the kinematic region and analysis goals; the choice of working point is driven by the signal-to-background optimization for the search or measurement at hand.

### 2.2.3 B TAGGING AT MULTIPURPOSE DETECTORS

Jets originating from bottom quarks (b-jets) exhibit several distinctive properties that enable their identification. The relatively long lifetime of b hadrons (order 1.5 ps) produces displaced charged tracks and often reconstructible secondary vertices a few millimetres from the primary interaction point. The large b-hadron mass yields decay products with sizable transverse momentum relative to the jet axis, and semileptonic branching fractions produce soft electrons or muons inside the jet. These features form the basis for b-tagging [114].

Practical algorithms exploit individual signatures or combine them:

- **Track-counting:** counts tracks with large impact-parameter significance to identify a b-like topology [114].
- **Jet-probability:** evaluates the compatibility of the jet’s track impact-parameter distribution with the primary vertex hypothesis [114].
- **Secondary-vertex:** explicitly reconstructs displaced vertices and uses their kinematic properties (decay length significance, vertex mass) [114].
- **Soft-lepton taggers:** identify low- $p_T$  leptons inside jets from semileptonic b decays [114].

Modern taggers combine many observables in multivariate or deep learning classifiers to maximize discrimination power. Contemporary approaches exploit rich, low level inputs (track by track and PF candidate information, vertex features and kinematics) and advanced network architectures (DeepCSV/DeepJet, RNN/sequence, graph/set networks) [77, 113, 114]. These developments yield measurable performance gains: modern deep classifiers typically improve b efficiency at fixed mistag rate relative

to classical taggers, and allow continuous discriminants with tunable operating points. Calibration with data-driven scale factors (from  $t\bar{t}$ , multijet or dilepton control samples) and propagation of associated systematic uncertainties remain essential for physics results [114].

- Deep feed-forward networks (e.g. DeepCSV/DeepJet) ingest a large set of high-level and per-track inputs to produce powerful binary or multi-class discriminants that separate  $b$ ,  $c$  and light-flavour jets.
- Sequence models and recurrent networks (RNN-based taggers) process an arbitrary ordered list of track-level variables, improving sensitivity by directly exploiting per-track correlations and order-dependent information (impact-parameter sequences, track kinematics).
- Graph- and set-based architectures and combined particle+vertex networks (sometimes referred to as “DeepFlavour”-style models) aggregate heterogeneous inputs and return per-flavour probabilities, enabling natural multi-classification and calibrated operating points.

These developments yield measurable performance gains: modern deep classifiers typically improve  $b$  efficiency at fixed mistag rate (or reduce mistag rates at fixed efficiency) relative to classical taggers. The continuous output of such networks permits analyses to choose operating points (loose/medium/tight) corresponding to desired efficiencies or mistag targets. Calibration remains essential: data-driven scale factors derived from control samples (e.g.  $t\bar{t}$ , multijet, dilepton) are applied to correct simulation, and systematic uncertainties from the calibration, flavour composition, and kinematic extrapolation are propagated to physics results.

Examples in use are CMS DeepCSV / DeepJet and ATLAS MV2 / DL1 [78, 113], which illustrate the transition from expert-designed high-level variables to large-scale machine learning leveraging low-level detector information. Typical medium working points yield  $b$ -tag efficiencies of order 60–80% with light-jet misidentification rates at or below the percent level; the precise choice of working point is tuned per analysis to optimise sensitivity while accounting for calibration and systematic uncertainties.

## 2.3 THE CMS DETECTOR

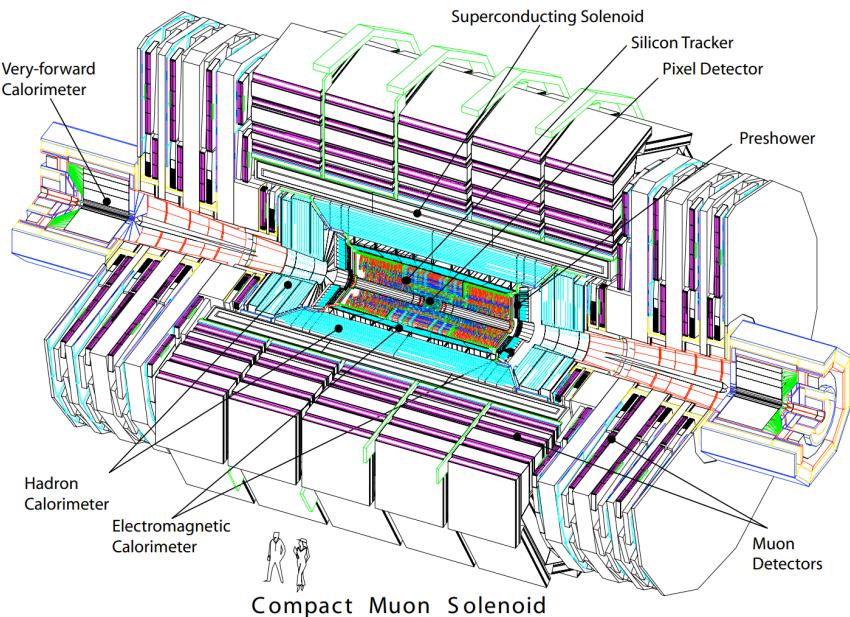
---

CMS is a general-purpose detector at the LHC [115]. With a length of 21.6 m, a diameter of 14.6 m, and a weight of 14,000 tonnes, its cylindrical geometry is divided into a central barrel section and two endcaps. This design provides hermetic coverage to accurately measure momentum and energy balance, which is crucial for identifying non-interacting particles like neutrinos through missing transverse energy.

The detector is constructed from concentric layers of sub-detectors, as illustrated in Figure 2.5. The innermost component is the silicon tracker,

comprising a pixel detector and silicon strip tracker. It reconstructs the trajectories of charged particles and measures their transverse momenta ( $p_T$ ) with a resolution of  $\approx 0.7\%$  for 10 GeV particles within a pseudorapidity range of  $|\eta| < 2.5$ .

Surrounding the tracker is the calorimetric system. The electromagnetic calorimeter (ECAL) is made of lead-tungstate crystals. It is designed to measure electrons and photons with a high resolution of  $\approx 0.6\%$  for 50 GeV electrons. The hadronic calorimeter (HCAL), located outside the ECAL, is a brass-scintillator sampling calorimeter that measures hadrons (e.g., charged pions, kaons, protons) with an energy resolution of  $\approx 18\%$  for 50 GeV pions. Together, the ECAL and HCAL cover  $|\eta| < 3$ . The coverage is extended to  $|\eta| < 5$  with steel and quartz-fiber hadron calorimeters in the forward regions.



**Figure 2.5:** Layout of the CMS experiment at the CERN LHC. (retrieved from [115]).

A key feature of CMS is its large superconducting solenoid, which encloses the tracker and calorimeters. The solenoid is constructed from a niobium-titanium alloy and cooled to 4.2 K with liquid helium. It generates a uniform magnetic field of 3.8 T throughout the tracking volume, enabling precise momentum measurement from the curvature of charged particle tracks.

The outermost system is dedicated to muon identification and measurement. Gas-ionization detectors are embedded in the steel flux-return yoke that surrounds the solenoid. This system provides triggering and tracking capabilities for muons up to  $|\eta| < 2.4$ . The combination of the inner tracker and the muon system allows for a robust identification and momentum measurement of muons across a wide kinematic range.

The geometrical segmentation of the barrel and endcaps defines the detector's acceptance in terms of pseudorapidity. The central barrel provides

optimal coverage for  $|\eta| \lesssim 1.5$ , while the endcaps extend the acceptance to  $|\eta| \lesssim 2.5$  for the tracker and calorimeters, and to  $|\eta| \lesssim 2.4$  for the muon system.

This segmentation impacts the detection efficiency. The silicon trackers are highly efficient in the barrel, where particles cross the layers perpendicularly. In the endcaps, the reduced hit multiplicity from shallow-angle traversals leads to a slight decrease in tracking efficiency and resolution. The calorimeters are also optimized to maintain performance across  $\eta$ , though the material budget and granularity vary.

Muon reconstruction performance exhibits regional differences. In the barrel, drift tubes (DTs) provide high spatial resolution, while in the endcaps, cathode strip chambers (CSCs) and resistive plate chambers (RPCs) are used to handle higher background rates and non-uniform magnetic fields. The assumed identification efficiency for muons (electrons) is 95% (85%), with a mis-identification rate of 0.3% (0.6%) [82, 88, 89].

For the identification of heavy-flavor jets, we adopt the DeepCSV algorithm [114]. We use its “medium” working point, which provides a b-tagging efficiency of 70% with a light-flavor jet misidentification rate of approximately 1% across the entire  $p_T$  spectrum. The “loose” (85% efficiency, 10% mis-id) and “tight” (45% efficiency, 0.1% mis-id) working points were also explored during the analysis optimization.

For hadronically decaying  $\tau$  leptons ( $\tau_h$ ), we use the DeepTau algorithm [78], which employs a deep neural network combining isolation and lifetime information to identify  $\tau_h$  decay modes. The “medium” working point is chosen for this analysis, providing a  $\tau_h$  identification efficiency of 70% and a misidentification rate of 0.5% for jets originating from light quarks and gluons. This working point was selected through an optimization process that maximized the discovery reach of the analysis.

## 2.4 PHENOMENOLOGICAL PIPELINE

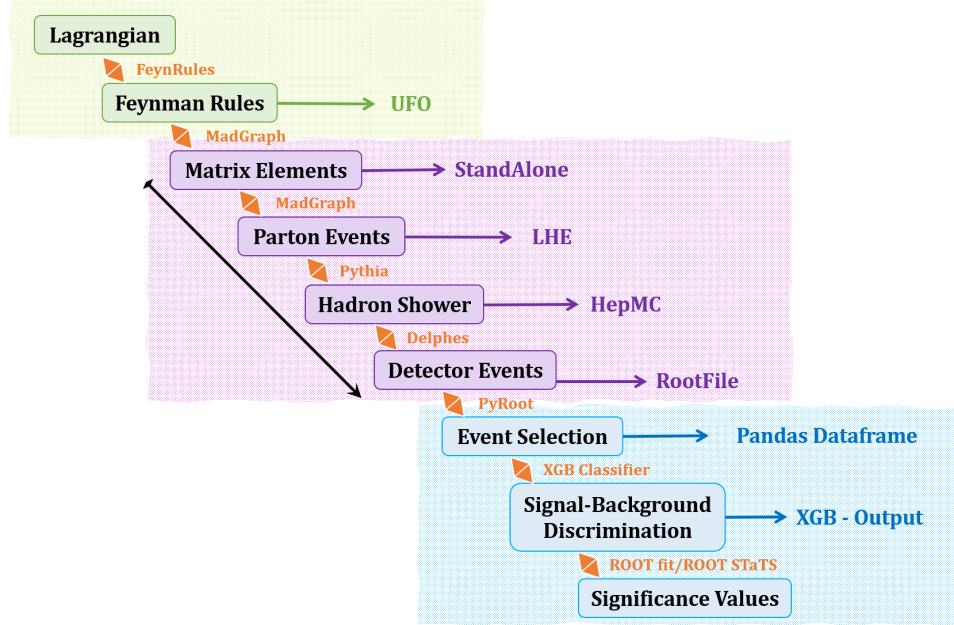
---

The estimation of signal and background event yields is performed through a comprehensive Monte Carlo (MC) simulation pipeline [63, 64, 116, 117]. This approach, a cornerstone of high-energy physics research, enables robust studies of Beyond the Standard Model (BSM) scenarios by emulating the entire data collection and processing chain of a collider experiment [65, 69]. The key advantages of this methodology include [63, 70]:

- The ability to perform automated calculations of theoretical quantities such as cross-sections and decay widths for complex processes.
- Conducting feasibility studies and optimizing analysis strategies prior to data acquisition.
- Estimating the efficiency of complex event selection criteria and the geometric acceptance of the detector.

- Predicting the rates and kinematical distributions of both irreducible and reducible background processes.
- Comparing and distinguishing between different theoretical hypotheses for a potential discovered signal.

The simulation workflow is modular, reflecting the logical progression from a theoretical Lagrangian to simulated detector-level observables [63, 64, 116, 117]. A schematic view of this pipeline is presented in Figure 2.6 [63, 65].



**Figure 2.6:** Schematic overview of the phenomenological MC pipeline: model definition (FeynRules) -> UFO export -> matrix-element generation (MadGraph) -> parton shower and hadronization (PYTHIA) -> fast detector simulation (DELPHES) -> analysis ntuples (ROOT).

The process begins with the implementation of the theoretical model in FeynRules (v2.3.43) [64, 116]. The Lagrangian of the new physics scenario, including all particle definitions, parameters, and interactions, is translated into a set of Feynman rules and exported in the Universal FeynRules Output (UFO) format [117], interoperable with modern matrix-element generators [63].

This UFO module, accompanied by a parameter card defining numerical values for masses and couplings, serves as input to MadGraph5\_aMC@NLO (v3.5.7) [63, 118]. Within MadGraph the hard process and corresponding matrix elements are generated and stored in Les Houches event (LHE) files; PDF choices (here NNPDF3.0 NLO [71]) and matching/merging settings (MLM/CKKW-type) are configured to control radiation and multi-jet overlap [119, 120]. Parton-level LHE events are passed to PYTHIA 8 for showering, hadronization and decays [69], and the generator output is exchanged in HepMC format for downstream processing [121].

To accurately model processes featuring significant interference effects between the new physics signal (e.g., a  $Z'$  boson) and the Standard Model backgrounds, the full squared amplitude (often referred to as the Signal-Discriminated Events or SDE strategy) is employed for the phase-space integration. The MadEvent submodule then generates unweighted parton-level events, which are stored in the Les Houches Event (LHE) format, containing the four-momenta of all final-state particles. The generation is optimized through careful configuration of the `run_card`, setting appropriate kinematic cuts on final-state partons to avoid wasting computational resources on events that would subsequently be rejected by the detector simulation.

Given the presence of additional jet radiation, the MLM matching scheme [119] is applied to mitigate the double-counting of jet emission between the matrix element calculation and the subsequent parton shower. This ensures a smooth transition between the hard process and softer radiative effects.

The parton-level LHE events are then passed to PYTHIA (v8.2.44) [69] for the modeling of QCD and QED radiation (parton showering), hadronization, and particle decays. This step translates the colored partons into stable, color-singlet hadrons and resonances that form the observable final state. The resulting events, which include a full list of generator-level particles, are saved in the HepMC<sub>2</sub> format.

Detector effects are simulated using DELPHES (v3.4.2) [65], a fast parametric detector simulation framework. The `delphes_card_CMS.tcl` configuration card is used to emulate the response of the CMS detector, including the geometric acceptance, tracking efficiency, calorimeter energy resolution and segmentation, and magnetic field. Key reconstruction algorithms are applied within DELPHES:

- Jets are clustered from calorimeter towers using the anti- $k_t$  algorithm [85] with a distance parameter of  $R = 0.4$ , and b-tagging is simulated based on the efficiency and mis-tag rate of the CMS performance.
- Muons and electrons are identified with efficiency maps that are functions of  $p_T$  and  $\eta$ .
- The missing transverse energy (MET) is calculated from the negative vector sum of all reconstructed particle momenta.

The final output, containing reconstructed physics objects (jets, leptons, MET), is stored in ROOT format [122].

At this stage, the analysis of the simulated samples converges with the methodology applied to real collider data. The subsequent steps involve applying event selection criteria, calibrating and scaling the reconstructed objects (e.g., applying Jet Energy Corrections), and performing statistical interpretation. The reliability of the simulation is validated by comparing the modeling of well-known Standard Model processes (e.g., Drell-Yan,

$t\bar{t}$  production) against published results and data-driven control regions. Dominant theoretical systematic uncertainties, such as those arising from the choice of factorization and renormalization scales, PDF variations, and parton shower modeling, are evaluated and propagated through the analysis.

## 2.5 MEASUREMENT OF THE POWER OF AN ANALYSIS

---

In high-energy physics experiments, data is often discretized into bins (e.g., histograms of collision events versus energy or momentum) to test competing hypotheses [123]. The fundamental framework compares two scenarios: the *null hypothesis* ( $H_0$ ), representing background-only processes ( $b_i$  in each bin  $i$ ), and the *alternative hypothesis* ( $H_1$ ), including both signal and background ( $s_i + b_i$ ) [124]. Given the Poissonian nature of event counts  $n_i$ , the likelihood for observing the data under each hypothesis is the product of Poisson probabilities per bin and is therefore written as a binned Poisson likelihood [95, 123]:

$$\mathcal{L}(n_i | \lambda_i) = \frac{e^{-\lambda_i} \lambda_i^{n_i}}{n_i!}, \quad \text{where } \lambda_i = \begin{cases} b_i & \text{for } H_0, \\ s_i + b_i & \text{for } H_1. \end{cases} \quad (2.7)$$

The Neyman-Pearson lemma [124] provides a rigorous framework for hypothesis testing by establishing that the *likelihood ratio*  $Q = \mathcal{L}(\text{data} | H_1)/\mathcal{L}(\text{data} | H_0)$  is the most powerful test statistic for distinguishing between two simple hypotheses,  $H_0$  and  $H_1$  [95, 124]. This forms the basis for quantifying the evidence for new physics signals against known backgrounds [95, 96]. For binned analyses in particle physics, we define the likelihood ratio  $Q_i$  for each bin  $i$  as [95, 123],

$$Q_i = \frac{\mathcal{L}(n_i | s_i + b_i)}{\mathcal{L}(n_i | b_i)} = e^{-s_i} \left(1 + \frac{s_i}{b_i}\right)^{n_i}, \quad (2.8)$$

where  $n_i$  is the observed event count,  $s_i$  the expected signal, and  $b_i$  the expected background in bin  $i$  [95, 123].

The test for the full analysis is constructed as the product of individual bin likelihood ratios [95, 123]:

$$Q = \prod_{i=1}^N Q_i, \quad (2.9)$$

where  $N$  is the total number of bins [123]. Under this formulation, each bin is treated as an independent experiment, allowing us to analyze the data in a modular way; this is convenient when combining results from multiple search channels or energy ranges [95, 96].

For convenience and to connect with asymptotic results, one commonly works with the log-likelihood ratio:

$$-2 \ln Q = 2 \sum_{i=1}^N \left[ s_i - n_i \ln \left(1 + \frac{s_i}{b_i}\right) \right], \quad (2.10)$$

and, by Wilks' theorem, its asymptotic distribution under the null hypothesis is chi-square distributed in regular cases [95, 125].

In practice, the Neyman-Pearson lemma motivates the use of a test statistic  $t$  that quantifies the evidence for a signal against the background-only hypothesis, which can be written as

$$t = -2 \ln Q = \sum_{i=1}^N [2s_i - 2n_i w_i], \quad (2.11)$$

with the optimal weight of each bin given by  $w_i = \ln\left(1 + \frac{s_i}{b_i}\right)$ .

The discovery significance  $\kappa$  quantifies the statistical separation of  $t$  if  $n$  is distributed according to the background-only hypothesis ( $H_0$ ) versus the signal-plus-background hypothesis ( $H_1$ ), normalized by the standard deviation of the  $H_1$  distribution,

$$\kappa = \frac{\langle t \rangle_{H_0} - \langle t \rangle_{H_1}}{\sigma_{H_1}}. \quad (2.12)$$

The expected behavior differs under the signal-plus-background ( $H_1$ ) and background-only ( $H_0$ ) hypotheses:

- **Under  $H_1$**  we expect that the  $n_i$  data distribution follows  $\text{Pois}(s_i + b_i)$ :

$$\langle -2 \ln Q \rangle_{s+b} = \sum_i [2s_i - 2(s_i + b_i)w_i] \implies \sigma_{s+b}^2 = 4 \sum_i (s_i + b_i)w_i^2. \quad (2.13)$$

- **Under  $H_0$**  we expect that the  $n_i$  data distribution follows  $\text{Pois}(b_i)$

$$\langle -2 \ln Q \rangle_b = \sum_i [2s_i - 2b_i w_i] \implies \sigma_b^2 = 4 \sum_i b_i w_i^2 \quad (2.14)$$

Substituting in  $\kappa$  gives a useful expression for the discovery significance,

$$\kappa = \frac{\sum s_i w_i}{\sqrt{\sum (s_i + b_i) w_i^2}} \quad (2.15)$$

It quantifies the separation between the signal+background ( $s + b$ ) and background-only hypotheses in units of standard deviations ( $\sigma$ ), where  $\kappa = 5$  corresponds to the traditional  $5\sigma$  discovery threshold,  $\kappa = 3$  to a  $3\sigma$  evidence to the traditional anomaly detection threshold, and  $\kappa = 1.69$  to the 95% confidence level (CL) exclusion limit.

This figure of merit automatically optimizes sensitivity through the logarithmic weights  $w_i = \ln(1 + s_i/b_i)$ , which naturally emphasize bins with either high signal-to-background ratios ( $s_i/b_i$ ) or large absolute signal contributions ( $s_i$ ). In asymptotic limits,  $\kappa$  simplifies to intuitive forms: for dominant signals ( $s_i \gg b_i$ ), it approaches  $\sqrt{\sum s_i}$  (Poisson counting),

while in background-dominated regimes ( $s_i \ll b_i$ ), it reduces to an inverse-variance-weighted sum  $\sum s_i / \sqrt{\sum b_i (s_i/b_i)^2}$ . This dual behavior ensures optimal discrimination power across all signal regimes.

In practice, we must take into account systematic effects by incorporating nuisance parameters into the likelihood and profiling over uncertainty ranges. The power calculation can be extended to include systematic uncertainties by modifying the denominator as,

$$\kappa_{\text{sys}} = \frac{\sum_i s_i w_i}{\sqrt{\sum_i [(s_i + b_i) + \delta_{\text{sys,signal},i}^2 + \delta_{\text{sys,bkg},i}^2] w_i^2}}, \quad (2.16)$$

where  $\delta_{\text{sys}}$  terms represent the systematic uncertainties on signal and background predictions.

This framework not only provides a figure of merit for an analysis but also serves as a roadmap for experimental optimization. The expected signal and background in each bin,  $s_i$  and  $b_i$ , are not fundamental inputs but are themselves products of the experimental setup and analysis choices. They can be expressed in terms of more fundamental experimental parameters (with acceptance absorbed into the selection efficiencies):

$$s_i = \sigma_{s,i} \cdot \mathcal{L} \cdot \epsilon_{s,i}, \\ b_i = \sigma_{b,i} \cdot \mathcal{L} \cdot \epsilon_{b,i},$$

where  $\sigma_{s,i}$  and  $\sigma_{b,i}$  are the fiducial cross-sections for signal and background processes in bin  $i$ ,  $\mathcal{L}$  is the integrated luminosity, and  $\epsilon_{s,i}$  and  $\epsilon_{b,i}$  are the effective efficiencies (selection efficiency combined with detector acceptance and reconstruction effects).

Substituting these expressions into the significance  $\kappa$  reveals the multi-dimensional parameter space available for optimization:

$$\kappa = \frac{\sum_i \sigma_{s,i} \cdot \epsilon_{s,i} \cdot w_i}{\sqrt{\sum_i [(\sigma_{s,i} \epsilon_{s,i} + \sigma_{b,i} \epsilon_{b,i}) + \delta_{\text{sys}}^2] \cdot w_i^2}} \cdot \sqrt{\mathcal{L}}.$$

This decomposition shows that the discovery significance can be enhanced through several distinct strategies. The primary handles are:

- **Increasing integrated luminosity ( $\mathcal{L}$ ):** The  $\sqrt{\mathcal{L}}$  scaling represents the fundamental statistical limit - doubling sensitivity requires quadrupling data collection time. This drives the construction of higher-luminosity colliders and longer data-taking campaigns.
- **Reducing systematic uncertainties:** The  $\delta_{\text{sys}}$  terms encompass uncertainties from theoretical predictions, detector calibration, background estimation methods, and luminosity measurement. Their reduction requires dedicated calibration measurements, improved Monte Carlo simulations, and sophisticated data-driven background estimation techniques.

- **Improving detector performance:** Effective efficiencies  $\epsilon_{s,i}$  and  $\epsilon_{b,i}$  can be improved through better detector design, increased coverage, and enhanced reconstruction and calibration algorithms that recover and correctly identify more signal events while controlling backgrounds.
- **Choosing optimal observables:** The weights  $w_i = \ln(1 + s_i/b_i)$  are maximized when the analysis uses variables that provide the best separation between signal and background. This motivates the development of advanced feature engineering and the use of multivariate methods that automatically learn the most discriminating variables.
- **Optimizing selection criteria:** Signal efficiency  $\epsilon_{s,i}$  can be maximized while background efficiency  $\epsilon_{b,i}$  is minimized through sophisticated trigger algorithms, multivariate analysis techniques, and machine learning classifiers that exploit subtle differences between signal and background event features.

Therefore, the power of an analysis, quantified by  $\kappa$ , is the result of a concerted effort across accelerator operation, detector performance, and analysis strategy.

The key limitation of the binned formulation in Eq. 2.16 is its treatment of bins as independent experiments, which discards valuable information from inter-bin correlations. This approximation becomes particularly evident in regions of high sensitivity, where the shape information of distributions becomes crucial. In such cases, multivariate methods that exploit the full correlation structure—such as matrix element methods, deep learning classifiers, or template fits—typically outperform simple binned significance estimates.

However, the formalism presented here provides theoretical insight and a useful approximation for quick sensitivity estimates. In the asymptotic limit and for counting experiments, this approach yields results consistent with statistical packages commonly used in high-energy physics, such as RooStats and RooFit. These frameworks implement more rigorous statistical procedures that fully account for the likelihood structure, parameter correlations, and systematic uncertainties through nuisance parameters.

Despite this limitation, the  $\kappa$  metric remains invaluable for establishing *experimental sensitivity*, which defined as the minimum signal strength required to achieve a certain significance level (e.g., 95% CL exclusion or  $5\sigma$  discovery potential). It provides a practical tool for guiding analysis design, optimizing selection criteria, and prioritizing experimental efforts.

For experimental final results and interpretation, full statistical treatments using profile likelihood methods within frameworks like RooStats remain the gold standard, as they properly account for all correlations and systematic uncertainties. In this work, we are not interested in the final statistical interpretation of data, but rather in understanding and optimizing the experimental sensitivity to new physics signals. Therefore,

the  $\kappa$  metric serves as a practical and insightful tool for guiding analysis design and experimental strategy.

## 2.6 ML ENHANCED SIGNAL-BACKGROUND DISCRIMINATION

---

As shown in Section 2.5, the sensitivity of a search depends on optimally separating signal and background processes [95, 126]. Traditional cut-based analyses, which apply sequential selection criteria to individual observables, cannot fully exploit the discriminatory information contained in the high-dimensional feature space of collision events [126, 127]. In this approach, for an event described by a feature vector  $\mathbf{x} = (x_1, x_2, \dots, x_N)$ , cuts are applied in the form [128]

$$\text{Selection: } x_1 > c_1 \text{ AND } x_2 > c_2 \text{ AND } \dots \text{ AND } x_N > c_N. \quad (2.17)$$

This method has several limitations that become apparent when considering correlations among kinematic variables [126, 127]. A typical LHC event contains a large number of observables, and an optimal discriminator must consider these variables and their relationships simultaneously [75, 128]. Hard cut boundaries discard events that are signal-like in multivariate space but fall just outside univariate cuts [126, 127]. The challenge of dimensionality also arises, as optimizing many cuts becomes unstable and prone to statistical fluctuations [127, 129]. Finally, sequential cuts cannot capture non-linear relationships and complex decision boundaries that often provide the strongest discrimination [126, 127].

Supervised learning addresses these limitations directly [126, 127]. It learns a function  $f(\mathbf{x})$  that maps the high-dimensional input space to a continuous score approximating the posterior probability [130, 131]:

$$f : \mathbb{R}^N \rightarrow [0, 1], \quad f(\mathbf{x}) \approx P(\text{signal} | \mathbf{x}). \quad (2.18)$$

This score incorporates correlations and non-linearities present in the training data, providing a powerful, continuous discriminant [126, 127].

Formally, the classification problem can be stated as follows [127]. Each collision event is represented by a feature vector  $\mathbf{x} = (x_1, x_2, \dots, x_N)$ , where the components correspond to reconstructed kinematic variables or high-level observables [75, 87]. The task is to assign a label

$$y = \begin{cases} 1 & \text{if the event originates from the signal process,} \\ 0 & \text{if the event originates from the background.} \end{cases} \quad (2.19)$$

Rather than predicting a hard label, modern classifiers return a continuous score  $f(\mathbf{x}) \in [0, 1]$  that estimates the probability of an event being signal given its features. This formulation allows the classifier to exploit multi-dimensional correlations and complex decision boundaries that cut-based methods cannot capture [126, 130].

Logistic regression, SVMs [132], Random Forests [133] and boosted trees [129, 134] are common choices; see [127, 128] for comparisons and practical guidance.

Several algorithms are commonly employed in this context. Logistic regression is often used as a baseline due to its simplicity and transparency. It assumes a linear decision boundary in the feature space, with the discriminant

$$f(\mathbf{x}) = \sigma(\mathbf{w} \cdot \mathbf{x} + b), \quad \sigma(z) = \frac{1}{1 + e^{-z}}, \quad (2.20)$$

where  $\mathbf{w}$  are model parameters and  $b$  is a bias term. While it cannot capture complex non-linear relationships, logistic regression remains useful when signal and background are approximately linearly separable, and it provides a clear reference point against which more sophisticated methods can be compared.

Support Vector Machines (SVMs) instead seek to maximize the margin between signal and background classes. For non-linear problems, SVMs employ kernel functions  $K(\mathbf{x}_i, \mathbf{x})$  that implicitly map the inputs to a higher-dimensional space where linear separation becomes possible. The decision function is

$$f(\mathbf{x}) = \text{sign} \left( \sum_{i=1}^N \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + b \right), \quad (2.21)$$

where  $\alpha_i$  are Lagrange multipliers. SVMs are effective in high-dimensional spaces and have strong theoretical guarantees, but they scale poorly to large datasets and are highly sensitive to kernel choice and hyperparameter tuning.

Tree-based ensembles are among the most widely used methods. Random Forests combine multiple decision trees trained on bootstrap samples of the data, with random feature selection at each split:

$$F(\mathbf{x}) = \frac{1}{B} \sum_{b=1}^B T_b(\mathbf{x}), \quad (2.22)$$

where  $B$  is the number of trees and  $T_b$  is the prediction of the  $b$ -th tree. This approach reduces variance and mitigates overfitting through averaging. Random Forests are robust, parallelizable, and provide natural feature importance measures, though they typically require more trees than boosted methods and may yield slightly lower performance on well-tuned tasks.

Boosted Decision Trees (BDTs) have become a standard tool in particle physics because they balance interpretability and performance. A single decision tree partitions the feature space through binary splits (e.g., “Is  $x_i < \text{threshold?}$ ”) and assigns class probabilities to terminal nodes. On their own, trees are high-variance learners, but boosting combines many weak learners (typically shallow trees) into a strong ensemble:

$$F_M(\mathbf{x}) = \sum_{m=1}^M \gamma_m h_m(\mathbf{x}), \quad (2.23)$$

where each new tree  $h_m(\mathbf{x})$  corrects the errors of the current ensemble  $F_{m-1}(\mathbf{x})$ , and  $\gamma_m$  is its weight. This sequential correction process produces

a powerful classifier. BDTs are robust to outliers and non-Gaussian distributions, handle mixed variable types naturally, and perform implicit feature selection, often revealing which observables are most discriminating. These properties explain their widespread adoption in LHC analyses.

Finally, Deep Neural Networks (DNNs) represent the most expressive class of models, capable of learning highly complex, hierarchical representations of the input data. A deep network with multiple hidden layers can be written as

$$f(\mathbf{x}) = \sigma^{(L)} \left( W^{(L)} \sigma^{(L-1)} \left( \dots \sigma^{(1)} (W^{(1)} \mathbf{x} + b^{(1)}) \dots \right) + b^{(L)} \right), \quad (2.24)$$

where  $L$  is the number of layers. DNNs excel at capturing intricate patterns and correlations, but they require large training datasets, careful regularization, and extensive hyperparameter tuning. Their black-box nature also complicates interpretability and the assessment of systematic uncertainties.

In practice, Random Forests are often used as robust, low-maintenance alternatives, while SVMs have largely been superseded by tree-based methods and neural networks due to scalability issues. DNNs can outperform other methods on very complex problems with sufficient data, but they demand significantly more computational resources. For most LHC searches, BDTs—particularly implementations such as XGBoost—provide the best balance between performance, interpretability, and computational efficiency. In Sec. 3.4, and in Table 3.3, we compare the performance of different methods and show that the gain in accuracy from DNNs is marginal compared to BDTs, while the training time is substantially larger.

### 2.6.1 XGBOOST: OPTIMIZED GRADIENT BOOSTING

XGBoost (eXtreme Gradient Boosting) is a widely used implementation of gradient boosting that has become standard in high-energy physics due to its efficiency, predictive power, and robustness.

A single decision tree is a weak classifier: it partitions the feature space through binary splits but is highly sensitive to fluctuations in the training data. Boosting addresses this by constructing an ensemble of trees sequentially. At each step, a new tree is trained to predict the residual errors of the current ensemble. This iterative approach allows the classifier to improve gradually, with later trees focusing on events that were previously misclassified.

The algorithm minimizes a regularized objective function:

$$\mathcal{L} = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{m=1}^M \Omega(f_m), \quad (2.25)$$

where  $l(y_i, \hat{y}_i)$  measures the prediction error and  $\Omega(f_m)$  penalizes complex trees to prevent overfitting. The regularization term is:

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|\mathbf{w}\|^2 + \alpha \|\mathbf{w}\|_1, \quad (2.26)$$

where  $T$  is the number of leaves,  $\mathbf{w}$  are the leaf weights, and  $\gamma, \lambda, \alpha$  control tree complexity.

The model is built additively:

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + \eta f_t(\mathbf{x}_i), \quad (2.27)$$

where  $\eta$  is the learning rate. Each new tree is selected to minimize:

$$f_t = \arg \min_f \sum_{i=1}^n [g_i f(\mathbf{x}_i) + \frac{1}{2} h_i f^2(\mathbf{x}_i)] + \Omega(f), \quad (2.28)$$

with  $g_i$  and  $h_i$  the first and second derivatives of the loss function.

The performance of XGBoost depends on finding the right balance between underfitting and overfitting. Underfitting occurs when the model is too simple to capture the relevant patterns in the data, leading to poor performance on both training and test sets. In this case the model is characterized by high bias and low variance. Overfitting, on the other hand, arises when the model learns the training data too closely, including statistical fluctuations and noise. This yields very good performance on the training set but poor generalization to new data, with the model showing low bias and high variance.

Several hyperparameters play a central role in controlling this balance:

- **n\_estimators:** Number of boosting rounds. Too few trees lead to underfitting, while too many lead to overfitting. Early stopping is commonly used to determine the optimal number by monitoring validation performance.
- **Learning rate ( $\eta$ ):** Step size shrinkage applied at each boosting step. Small values (e.g. 0.01–0.1) improve generalization but require more trees; larger values speed up training but can overfit.
- **max\_depth:** Maximum depth of individual trees. Shallower trees (4–6) tend to be more stable, while deeper ones (7–10) can capture complex correlations but risk overfitting.
- **Regularization parameters:**
  - $\gamma$ : Minimum loss reduction required for a split. Larger values make the algorithm more conservative.
  - $\lambda$ : L2 regularization on leaf weights, which limits large values and stabilizes the model.
  - $\alpha$ : L1 regularization on leaf weights, which promotes sparsity and can serve as implicit feature selection.

Optimal hyperparameters are typically found through systematic search methods. The most common approach is *grid search with cross-validation* (GridSearchCV), which exhaustively tests all parameter combinations within

predefined ranges. The core of this method is *k-fold cross-validation*, a robust technique for assessing model generalization.

The k-fold cross-validation procedure consists of the following steps: First, the available training data is randomly shuffled and partitioned into k equal-sized subsets called folds. This partitioning is typically stratified to preserve the class distribution in each fold. Then, the model is trained and evaluated k times in a round-robin fashion. For each iteration i (where i = 1 to k), the i-th fold is held out as validation data, while the remaining k-1 folds are used for training. The model's performance metric (typically negative log-loss, accuracy, or area under the ROC curve) is computed on the validation fold. After all k iterations are completed, the performance scores from each validation fold are averaged to produce a single estimation of the model's generalization error. This approach ensures that every data point is used exactly once for validation while being used k-1 times for training, providing an unbiased estimate of model performance that is robust to the specific partitioning of the data.

For XGBoost in HEP applications, we typically use k= 5 folds as it offers a good balance between computational cost and reliable error estimation. Each parameter combination is evaluated through this cross-validation process, ensuring that selected parameters generalize well beyond the training data and are not overly tuned to specific statistical fluctuations.

The grid search tests all combinations in the parameter space defined by ranges such as: `learning_rate`  $\eta \in [0.01, 0.3]$ , `max_depth`  $\in [3, 10]$ , `n_estimators`  $\in [100, 1000]$ , with regularization parameters  $\gamma$ ,  $\lambda$ , and  $\alpha$  typically explored in logarithmic scales.

While grid search is thorough, it becomes computationally expensive for high-dimensional parameter spaces. In such cases, more efficient methods like *randomized search* (which samples parameter combinations randomly) or *Bayesian optimization* (which uses probabilistic models to guide the search toward promising regions) can be employed.

The optimization process is iterative: initial broad searches identify promising parameter regions, followed by finer-grained searches around the best-performing configurations. Early stopping during training—monitoring validation performance and halting when no improvement is observed for a specified number of rounds—prevents overfitting and significantly reduces computational cost, making the hyperparameter optimization feasible for large-scale HEP analyses.

### 2.6.2 STANDARD ML ANALYSIS WORKFLOW

The XGBoost output score  $f(\mathbf{x})$  transforms high-dimensional data into a single optimal discriminant. When binned, the resulting histogram gives expected yields:

$$s_i = \int_{\text{bin } i} \sigma_s \cdot \mathcal{L} \cdot \epsilon_s \cdot p_s(f) df, \quad (2.29)$$

$$b_i = \int_{\text{bin } i} \sigma_b \cdot \mathcal{L} \cdot \epsilon_b \cdot p_b(f) df, \quad (2.30)$$

where  $p_s(f)$  and  $p_b(f)$  are the output distributions.

Integrating machine learning into high-energy physics analysis follows a standardized workflow designed to maximize sensitivity while ensuring robustness against overfitting and systematic biases:

1. **Dataset Preparation and Balancing:** Monte Carlo simulations generate signal and background samples. The signal sample corresponds to the hypothetical new physics process, while background samples include all known Standard Model processes that can produce similar experimental signatures. To prevent classifier bias toward the typically dominant background, datasets are balanced through undersampling (selecting a subset of the majority class) or, more commonly, event weighting using  $w_i = \sigma \cdot \mathcal{L} \cdot \epsilon / N_{\text{gen}}$ . Equal numbers of signal and background events are often used during training to ensure the algorithm learns both classes effectively, though the final evaluation uses proper physics weights.
2. **Feature Preprocessing:** Input variables (kinematic observables such as  $p_T$ ,  $\eta$ ,  $\phi$ , invariant masses, and angular separations) are standardized using techniques like StandardScaler (transforming to zero mean and unit variance) or MinMaxScaler (scaling to a fixed range, typically  $[0, 1]$ ). While tree-based methods like XGBoost are theoretically scale-insensitive, preprocessing improves numerical stability and convergence speed. Dimensionality reduction techniques like Principal Component Analysis (PCA) may be used for visualization or to address severe multicollinearity, though trees naturally handle correlated features.
3. **Model Training and Hyperparameter Optimization:** The classifier is trained on the preprocessed data using the procedures described in Section 2.6.1. Key hyperparameters—including learning rate, maximum tree depth, L1/L2 regularization strengths, and minimum child weight—are optimized via grid search, random search, or Bayesian optimization as detailed in the hyperparameter optimization strategy. Performance is evaluated using k-fold cross-validation to ensure generalizability and avoid overfitting, with the optimal configuration selected based on the best cross-validated performance.

4. **Output Score Generation:** Instead of binary class assignments, the trained model's continuous output is obtained using `predict_proba()`, which provides a per-event probability score  $f(x) \in [0, 1]$  indicating the likelihood of belonging to the signal class. This score serves as a powerful discriminant variable that encapsulates the multidimensional separation power.
5. **Histogram Construction and Weighting:** Events are binned based on their classifier score to form a one-dimensional histogram. Each bin's content is calculated using the appropriate physics-level weights:

$$N_i^{\text{bin}} = \sum_{\text{events in bin } i} w_j = \sum_{\text{events in bin } i} (\sigma \cdot \mathcal{L} \cdot \epsilon / N_{\text{gen}})_j,$$

yielding the expected signal ( $s_i$ ) and background ( $b_i$ ) yields per bin. The binning is typically optimized to maximize the expected sensitivity, often with finer binning in regions of better signal-to-background ratio.

6. **Sensitivity Measurement:** The final histogram, incorporating all relevant systematic uncertainties as nuisance parameters, serves as input to the statistical model described in Section 2.5. The discovery significance  $\kappa$  (from Eq. 2.16) is computed, quantifying the analysis sensitivity and enabling comparison between different analysis strategies or machine learning approaches.

This end-to-end workflow seamlessly integrates machine learning into the established statistical framework of particle physics, transforming high-dimensional data into an optimized discriminant for sensitivity extraction.

# 3

---

## $U(1)_{T_R^3}$ GAUGE EXTENSION OF THE STANDARD MODEL

---

Based on [2]

Extensions of the Standard Model (SM) that introduce new  $U(1)$  gauge symmetries are among the most widely studied scenarios for physics beyond the SM. In particular, the  $U(1)_{T_R^3}$  symmetry is of interest. In this scenario, right-handed SM fermions and possible new states, such as right-handed neutrinos, are charged under  $U(1)_{T_R^3}$ . This symmetry has been explored in the context of left-right symmetric models [135–137].

In these frameworks,  $U(1)_{T_R^3}$  is identified as the diagonal, electrically neutral generator of  $SU(2)_R$ . It is often related to  $U(1)_{B-L}$  through the breaking pattern

$$U(1)_{B-L} \times U(1)_{T_R^3} \rightarrow U(1)_Y.$$

This motivates the existence of a new, massive, electrically neutral gauge boson associated with the extra  $U(1)$  symmetry [1, 138–141].

However, the Higgs doublet being a singlet under  $U(1)_{B-L}$ , acquires its hypercharge from  $U(1)_{T_R^3}$ . Its vacuum expectation value (VEV) links the symmetry-breaking scales of  $U(1)_Y$  and  $U(1)_{T_R^3}$ . Alternatively, these scales can be decoupled by introducing an additional  $U(1)_G$  group, under which SM fermions are singlets but the Higgs is charged. In this case [142]:

$$Y = Q_{T_R^3} + \frac{1}{2}Q_{B-L} + Q_G. \quad (3.1)$$

More generally, scenarios can be constructed where the hypercharge is not directly related to  $U(1)_{T_R^3}$ .

Recent theoretical and phenomenological work has focused on models where the low-energy gauge symmetry of the SM is extended by an Abelian  $U(1)_{T_R^3}$  group. In these models, the spontaneous breaking of  $U(1)_{T_R^3}$  is not tied to electroweak symmetry breaking [143–148].

The corresponding gauge boson of the extra  $U(1)$  is a neutral vector particle whose physical interpretation depends on its couplings and mass range. If the new boson couples directly to SM fermions with electroweak-strength interactions, it is often referred to as a  $Z'$ . If instead the new boson interacts only very weakly with the SM, typically through kinetic mixing with the hypercharge gauge boson, it is commonly called a dark photon  $A'$ . In either case, the gauge boson acquires mass through a Higgs-like mechanism. A complex scalar field  $\phi$ , singlet under the SM gauge group,

can provide the longitudinal degree of freedom. Its CP-odd component gives mass to the neutral vector boson, while its CP-even component can manifest as a dark Higgs,  $\phi'$ .

To ensure anomaly cancellation, a right-handed neutrino  $\nu_R$  is required for each SM generation that couples to  $U(1)_{T_R^3}$ . In addition, a set of new vector-like fermions ( $\chi_u, \chi_d, \chi_\ell, \chi_\nu$ ) is introduced to generate fermion masses in a UV-complete theory, following the universal see-saw mechanism [149–154]. This mechanism introduces a non-trivial  $\chi_u - t - \phi'$  coupling. This vertex allows for the production of  $t\chi_u\phi'$  final states via  $\chi_u - t$  fusion (see Figure 3.3). Since  $\chi_u$  couples to SM quarks and gluons, it can be copiously produced. Its energetic decay products, together with a  $\phi'$  mediator carrying significant transverse momentum, can be efficiently detected, especially if  $\phi'$  decays to visible SM particles in the central detector region.

This strategy is effective for reducing SM backgrounds and enhances the LHC discovery potential for heavy top partners and GeV-scale mediators, which are otherwise challenging to probe at hadron colliders. Moreover,  $t\chi_u\phi'$  final states can also arise from  $\chi_u\bar{\chi}_u$  production via QCD vertices, where one  $\chi_u$  decays to  $t\phi'$  (see Figure 3.4). The presence of energetic decay products and a mediator with substantial transverse momentum provides greater sensitivity than searches considering  $\chi_u$  or  $\phi'$  alone.

In this chapter, we perform a phenomenological study of search strategies at the LHC for a light (GeV-scale) scalar  $\phi'$  produced in association with a heavy (TeV-scale) top-partner  $\chi_u$ . We focus on the previously unexplored production channel  $pp \rightarrow t\chi_u\phi'$ , which contrasts with the more commonly studied  $pp \rightarrow TT \rightarrow t\phi't\phi'$  and di-photon  $\phi'$  decay channels, see Section 4.1 and [155–159].

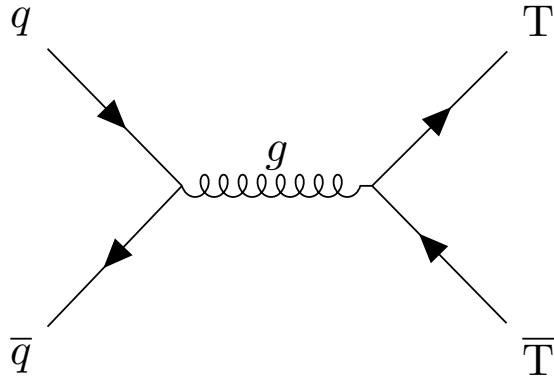
We consider the case where  $\phi'$  has family non-universal couplings to fermions, as proposed in [144]. Such couplings can address several open questions in the SM. Our analysis focuses on  $\phi' \rightarrow \mu^+\mu^-$  decays, as muons are efficiently reconstructed and identified. This allows for low  $p_T(\mu)$  triggers and provides a characteristically clean signature to suppress QCD multijet backgrounds.

To further maximize the sensitivity to this complex signal, a central component of our analysis is the use of ML. We employ an analysis based on Boosted Decision Trees (BDT) [129]. The BDT output is used in a profile-binned likelihood test to determine the signal significance for each model. The effectiveness of BDTs and other ML algorithms has been demonstrated in numerous experimental and phenomenological studies [1, 160–169]. Our results show that the BDT approach significantly improves sensitivity.

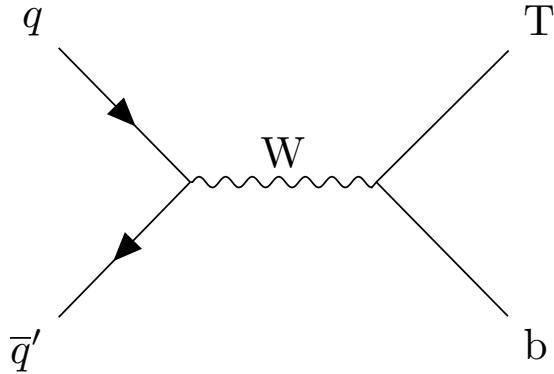
The remainder of this chapter is organized as follows. Section 4.1 describes the minimal  $U(1)_{T_R^3}$  model. Section 3.1 reviews current relevant LHC results. Section 3.3 details the Monte Carlo (MC) simulation samples used in this study. Section 3.4 discusses the motivation and implementation of the machine learning workflow, and Section 4.3 presents the main results.

### 3.1 CURRENT EXCLUSION LIMITS ON VECTOR-LIKE QUARKS

The ATLAS and CMS collaborations at CERN have conducted various searches for heavy vector-like quarks ( $T$ ). These searches utilized  $pp$  collisions at center-of-mass energies of  $\sqrt{s} = 8$  and 13 TeV. The studies primarily focused on  $T$  production through gluon-mediated QCD processes, either in pair production from quark-antiquark annihilation (Figure 3.1) or in single- $T$  production from electroweak processes involving associated quarks (Figure 3.2).



**Figure 3.1:** Representative Feynman diagram for  $T$  pair production via gluon-mediated QCD processes.



**Figure 3.2:** Representative Feynman diagram for single  $T$  production via electroweak processes.

In those studies,  $T$  decays into  $bW$ ,  $tZ$ , or  $tH$  have been considered. In the context of  $T$  pair production,  $T\bar{T}$ , via QCD processes, the cross sections are well-known and solely depend on the mass of the vector-like quark. Assuming a narrow  $T$  decay width ( $\Gamma/m(T) < 0.05$  or 0.1) and a 100% branching fraction to  $bW$ ,  $tZ$ , or  $tH$ , these searches have set stringent bounds on  $m(T)$ , excluding masses below almost 1.5 TeV at 95% confidence level [170–177]. The most recent analysis from the CMS collaboration probes  $T$ -quark production via  $pp \rightarrow Tqb$ , in final states with  $T \rightarrow tZ$  or  $T \rightarrow tH$ , considering scenarios with preferential couplings to third-generation fermions. The analysis sets 95% confidence level upper limits

of 68–1260 fb on the production cross section, for T masses ranging from 600–1200 GeV [171]. The latest studies from ATLAS probe vector-like quarks using the single-T production mode with the  $T \rightarrow tH$  decay channel leading to a fully hadronic final state [172], the single-T production mode with the  $T \rightarrow tZ$  decay channel leading to a multileptonic final state [173], the TT pair production mode with various T decay channels leading to multileptonic final states [174], and the TT pair production mode with various T decay channels leading to a single lepton plus missing momentum final state [175, 176]. The multilepton search offers the greatest sensitivity in most of the phase space, but the missing transverse energy based search has better sensitivity for low branching fraction  $\mathcal{B}(T \rightarrow Wb)$  and high  $\mathcal{B}(T \rightarrow Ht)$ . These searches have similar sensitivities for the singlet and doublet models, resulting in exclusion bounds for masses below about 1.25 TeV and 1.41 TeV, respectively.

A key consideration in the model interpretations summarized above is that the T branching fractions depend on the chosen model. The excluded mass range is less restrictive for specific branching fraction scenarios, such as  $\{\mathcal{B}(T \rightarrow tZ), \mathcal{B}(T \rightarrow bW), \mathcal{B}(T \rightarrow tH)\} = \{0.2, 0.6, 0.2\}$ , excluding masses below about 0.95 TeV. Moreover, if the  $T \rightarrow \phi't$  decay is allowed, or if the branching fractions  $\mathcal{B}(T \rightarrow tH/bW)$  are lower, the limits previously quoted must be re-evaluated. The authors of Ref. [178] emphasize that bounds on  $m(T)$  can be around 500 GeV when  $T \rightarrow t\phi'$  decays are permitted. Therefore, to facilitate a comprehensive study, benchmark scenarios in this paper are considered down to  $m(\chi_u) = 500$  GeV.

### 3.2 THE MINIMAL $U(1)_{T_R^3}$ MODEL

---

The model extends the SM by an Abelian gauge symmetry  $U(1)_{T_R^3}$ , under which only the right-handed fermions are charged. The symmetry breaking is achieved via two independent Higgs mechanisms: one with the SM Higgs doublet  $H$  for electroweak symmetry breaking, and another with a Higgs singlet  $\phi$  for breaking  $U(1)_{T_R^3}$ . These scalars acquire independent vacuum expectation values (VEVs),  $\langle H \rangle = v_h/\sqrt{2}$  and  $\langle \phi \rangle = v_\phi/\sqrt{2}$ . In the Kibble parametrization, the fields are written as:

$$H = \begin{pmatrix} G_+ \\ \frac{1}{\sqrt{2}}(v_h + \rho_0 + iG_0) \end{pmatrix}, \quad (3.2)$$

$$\phi = \frac{1}{\sqrt{2}}(v_\phi + \rho_\phi + iG_\phi). \quad (3.3)$$

In Eqs. (3.2) and (3.3),  $G_\pm$ ,  $G_0$ , and  $G_\phi$  are the Goldstone bosons absorbed by the SM  $W^\pm$  and  $Z$  bosons and the dark photon  $A'$  (associated with  $U(1)_{T_R^3}$ ) to acquire mass. The fields  $\rho_h$  and  $\rho_\phi$  mix to form the physical mass eigenstates, the SM-like Higgs boson  $h$  and a dark Higgs  $\phi'$ :

$$\begin{pmatrix} h \\ \phi' \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} \rho_0 \\ \rho_\phi \end{pmatrix}. \quad (3.4)$$

This mixing arises from diagonalizing the mass matrix derived from the gauge-invariant scalar potential:

$$\begin{aligned} V(H, \phi) = & \mu_H^2 H^\dagger H + \mu_\phi^2 \phi^* \phi \\ & + \lambda(H^\dagger H)(\phi^* \phi) + \lambda_H(H^\dagger H)^2 + \lambda_\phi(\phi^* \phi)^2. \end{aligned} \quad (3.5)$$

Minimizing the potential yields the tadpole equations:

$$\frac{\partial V}{\partial H} = \frac{v_h}{\sqrt{2}} \left( \mu_H^2 + \lambda_H v_h^2 + \frac{1}{2} \lambda v_\phi^2 \right) = 0, \quad (3.6)$$

$$\frac{\partial V}{\partial \phi} = \frac{v_\phi}{\sqrt{2}} \left( \mu_\phi^2 + \lambda_\phi v_\phi^2 + \frac{1}{2} \lambda v_h^2 \right) = 0. \quad (3.7)$$

The physical scalar masses are given by:

$$m_{h,\phi'}^2 = \frac{1}{2} (\lambda_H v_h^2 + \lambda_\phi v_\phi^2) \pm \sqrt{\lambda^2 v_h^2 v_\phi^2 + (\lambda_H v_h^2 - \lambda_\phi v_\phi^2)^2}, \quad (3.8)$$

and the mixing angle  $\alpha$  satisfies:

$$\tan 2\alpha = \frac{-\lambda v_h v_\phi}{\lambda_\phi v_\phi^2 - \lambda_H v_h^2}. \quad (3.9)$$

The quartic couplings can be expressed in terms of the physical parameters:

$$\lambda_H = \frac{m_{\phi'}^2 + m_h^2 + (m_{\phi'}^2 - m_h^2) \cos 2\alpha}{4v_h^2}, \quad (3.10)$$

$$\lambda_\phi = \frac{m_{\phi'}^2 + m_h^2 + (m_{\phi'}^2 - m_h^2) \cos 2\alpha}{4v_\phi^2}, \quad (3.11)$$

$$\lambda = \frac{m_{\phi'}^2 - m_h^2}{2v_h v_\phi} \sin 2\alpha. \quad (3.12)$$

Thus, the scalar sector has four free parameters: the masses  $m_h$  and  $m_{\phi'}$ , the mixing angle  $\alpha$ , and the dark Higgs VEV  $v_\phi$ . Similar to how  $v_h$  is fixed by the electroweak gauge boson masses,  $v_\phi$  is related to the dark photon mass by  $m_A^2 = g_{T_R^3}^2 v_\phi^2$ , where  $g_{T_R^3}$  is the  $U(1)_{T_R^3}$  gauge coupling. Depending on the value of  $g_{T_R^3}$ , this gauge boson can behave as a heavy  $Z'$  or a light dark photon. In this chapter, we assume  $g_{T_R^3}$  is sufficiently small such that  $A'$  can be treated as a dark photon.

### 3.2.1 THE UNIVERSAL SEESAW MECHANISM

In this model, the masses of the SM fermions are generated through a universal seesaw mechanism by mixing with vector-like fermions  $\chi_f$ . The relevant mass terms in the Lagrangian are:

$$-\mathcal{L} \supset Y_{f_L} \bar{f}'_L \chi'_{fR} H + Y_{f_R} \bar{\chi}'_{fL} f'_R \phi^* + m_{\chi'_f} \bar{\chi}'_{fL} \chi'_{fR} + \text{h.c.} \quad (3.13)$$

This leads to the mass matrix:

$$M_f = \begin{pmatrix} 0 & Y_{f_L} v_h / \sqrt{2} \\ Y_{f_R} v_\phi / \sqrt{2} & m_{\chi'_f} \end{pmatrix}. \quad (3.14)$$

The mass eigenstates  $(f, \chi_f)$  are obtained by rotating the gauge eigenstates:

$$\begin{pmatrix} f_{L,R} \\ \chi_{f_{L,R}} \end{pmatrix} = \begin{pmatrix} \pm \cos \theta_{f_{L,R}} & \mp \sin \theta_{f_{L,R}} \\ \sin \theta_{f_{L,R}} & \cos \theta_{f_{L,R}} \end{pmatrix} \begin{pmatrix} f'_{L,R} \\ \chi'_{f_{L,R}} \end{pmatrix}, \quad (3.15)$$

such that  $\mathcal{R}(\theta_{f_L}) M_f \mathcal{R}^{-1}(\theta_{f_R}) = \text{diag}(m_f, m_{\chi_f})$ . For real parameters, the physical masses and mixing angles are given by:

$$m_f m_{\chi_f} = \frac{Y_{f_L} v_h Y_{f_R} v_\phi}{2}, \quad (3.16)$$

$$m_f^2 + m_{\chi_f}^2 = m_{\chi'_f}^2 + \frac{1}{2} (Y_{f_L}^2 v_h^2 + Y_{f_R}^2 v_\phi^2), \quad (3.17)$$

$$\tan \theta_{f_{L,R}} = \frac{\sqrt{2}}{m_{\chi'_f}} \left( \frac{Y_{f_{L,R}} v_{h,\phi}}{2} - \frac{m_f^2}{Y_{f_{L,R}} v_{h,\phi}} \right). \quad (3.18)$$

The Yukawa interactions of the physical fermions with the scalars  $h$  and  $\phi'$  are:

$$-\mathcal{L}_{\text{yuk}} = h \bar{\psi}_{f_L} \gamma_h \psi_{f_R} + \phi' \bar{\psi}_{f_L} \gamma_\phi \psi_{f_R}, \quad (3.19)$$

where  $\psi_f = (f, \chi_f)^T$ . The Yukawa matrices are:

$$\gamma_h = \frac{1}{\sqrt{2}} \mathcal{R}(\theta_{f_L}) (Y_{f_L} \sigma_+ \cos \alpha - Y_{f_R} \sigma_- \sin \alpha) \mathcal{R}^{-1}(\theta_{f_R}), \quad (3.20)$$

$$\gamma_\phi = \frac{1}{\sqrt{2}} \mathcal{R}(\theta_{f_L}) (Y_{f_L} \sigma_+ \sin \alpha + Y_{f_R} \sigma_- \cos \alpha) \mathcal{R}^{-1}(\theta_{f_R}), \quad (3.21)$$

with  $\sigma_\pm = (\sigma_1 \pm i\sigma_2)/2$  being the ladder Pauli matrices.

The expressions above provide a simplified, one-generation view. The complete model involves a non-trivial flavor structure where the mass matrices are general  $3 \times 3$  matrices. The diagonalization of the full  $6 \times 6$  mass matrices, the procedure for absorbing unphysical unitary rotations, and the emergence of the CKM matrix are detailed in Appendix B. Furthermore, the appendix contains a rigorous treatment of the mass eigenvalue problem, deriving the exact relationship between the fundamental parameters  $(m_L, m_R, m_\chi)$  and the physical observables  $(m_f, m_F, \theta_L)$ , which leads to critical constraints on the model's parameter space to ensure perturbativity.

### 3.2.2 MINIMAL UV-COMPLETE THEORY

To generate non-zero masses for all SM fermions and ensure gauge anomaly cancellation, the model must include at least one full generation of vector-like fermions  $\{\chi_u, \chi_d, \chi_\ell, \chi\}$  and the right-handed neutrinos  $\nu_R$  for each SM

generation. Their quantum numbers are listed in Table 3.1. The Yukawa interactions in the UV-complete theory are:

$$\begin{aligned} -\mathcal{L} \supset & Y_{Lu}^{ij} \bar{q}'^i_L \chi'_{uR} \tilde{H} + Y_{Ru}^{ij} \bar{\chi}'^i_{uL} u'_R \phi^* + m_{\chi_u}^{ij} \bar{\chi}'^i_{uL} \chi'_{uR} \\ & + Y_{Ld}^{ij} \bar{q}'^i_L \chi'_{dR} H + Y_{Rd}^{ij} \bar{\chi}'^i_{dL} d'_R \phi + m_{\chi_d}^{ij} \bar{\chi}'^i_{dL} \chi'_{dR} \\ & + Y_{L\ell}^{ij} \bar{\ell}'^i_L \chi'_{\ell R} H + Y_{R\ell}^{ij} \bar{\chi}'^i_{\ell L} \ell'_R \phi + m_{\chi_\ell}^{ij} \bar{\chi}'^i_{\ell L} \chi'_{\ell R} \\ & + Y_{L\nu}^{ij} \bar{\ell}'^i_L \chi'_{\nu R} \tilde{H} + Y_{R\nu}^{ij} \bar{\chi}'^i_{\nu L} \nu'_R \phi^* + m_{\chi_\nu}^{ij} \bar{\chi}'^i_{\nu L} \chi'_{\nu R} + \text{h.c.} \end{aligned} \quad (3.22)$$

Here,  $i, j = 1, 2, 3$  are generation indices. The diagonalization of the mass matrices for each fermion type follows the structure outlined in Eqs. (3.16) and (3.17), while the Yukawa matrices generalize the structure of Eqs. (3.20) and (3.21), now encoding the CKM and PMNS mixing matrices. The neutrino sector has a more complex structure due to the possibility of a Majorana mass term for the vector-like neutrinos  $\chi'_\nu$ .

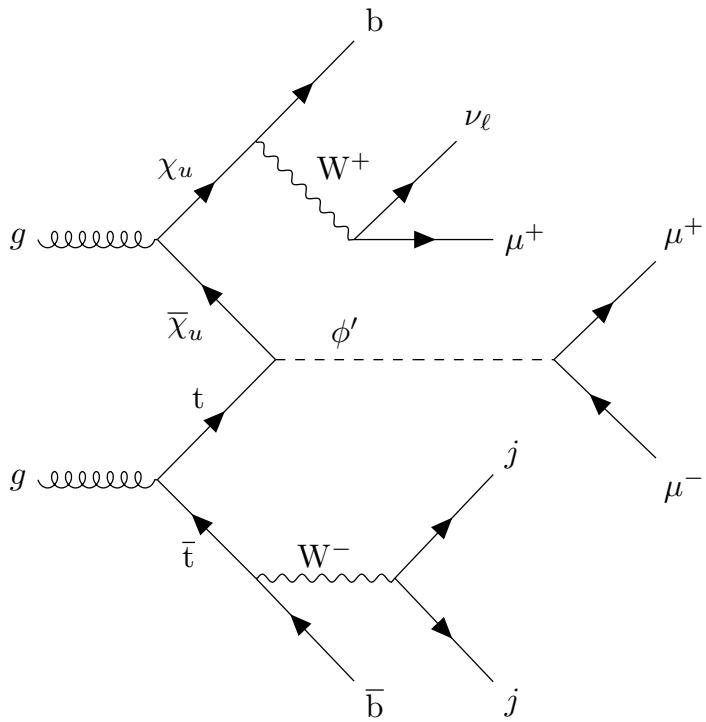
Field	SU(3) <sub>C</sub>	SU(2) <sub>L</sub>	U(1) <sub>Y</sub>	U(1) <sub>T<sub>R</sub><sup>3</sup></sub>
$q'_L$	<b>3</b>	<b>2</b>	1/6	0
$\ell'_L$	<b>1</b>	<b>2</b>	-1/2	0
H	<b>1</b>	<b>2</b>	1/2	0
$u'^c_R$	<b>3</b>	<b>1</b>	-2/3	-2
$d'^c_R$	<b>3</b>	<b>1</b>	1/3	2
$\ell'^c_R$	<b>1</b>	<b>1</b>	1	2
$\nu'^c_R$	<b>1</b>	<b>1</b>	0	-2
$\phi$	<b>1</b>	<b>1</b>	0	2
$\chi'_{uL}$	<b>3</b>	<b>1</b>	2/3	0
$\chi'_{uR}$	<b>3</b>	<b>1</b>	-2/3	0
$\chi'_{dL}$	<b>3</b>	<b>1</b>	-1/3	0
$\chi'_{dR}$	<b>3</b>	<b>1</b>	1/3	0
$\chi'_{\ell L}$	<b>1</b>	<b>1</b>	-1	0
$\chi'_{\ell R}$	<b>1</b>	<b>1</b>	1	0
$\chi'_{\nu L}$	<b>1</b>	<b>1</b>	0	0
$\chi'_{\nu R}$	<b>1</b>	<b>1</b>	0	0

**Table 3.1:** Minimal field content of the model and their representations under the SM and U(1)<sub>T<sub>R</sub><sup>3</sup></sub> gauge groups.

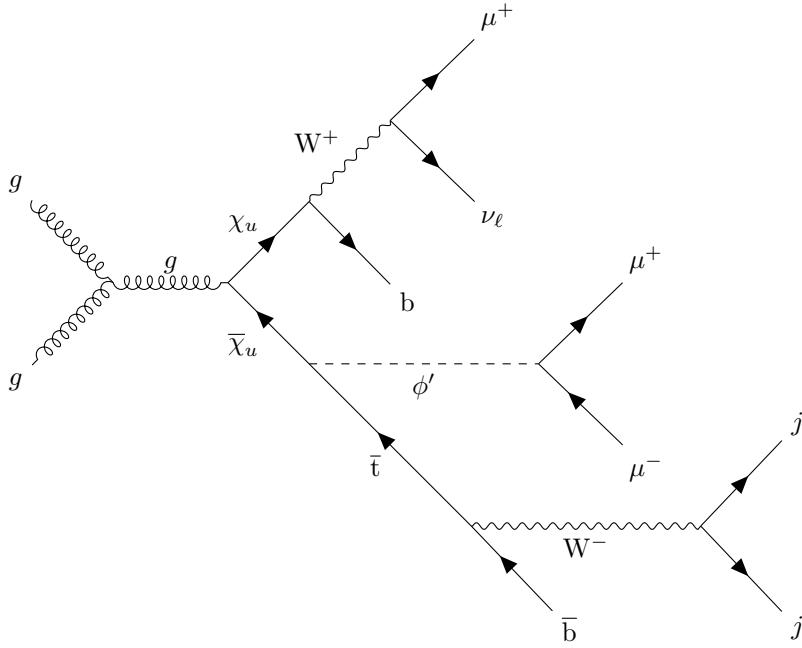
### 3.3 SAMPLES AND SIMULATION

The minimal U(1)<sub>T<sub>R</sub><sup>3</sup></sub> model described in Sec. 4.1 is implemented *at tree level* into the FeynRules package [64], which generates the Feynman rules and exports them into a Universal FeynRules Output (UFO) [117]. The resulting UFO is utilized as input for a generator to produce the MC samples. Both signal and background events are generated with the MadGraph5\_aMC@NLO v3.2.0 program [63, 118] at leading order (LO) in QCD, considering pp beams colliding with a center-of-mass energy of  $\sqrt{s} = 13.6$  TeV. Each

signal and background sample is generated separately, with no interference effects between the signal and background considered. The impact of these interference effects has been evaluated, and for all values of  $\chi_u$  and  $\phi'$  masses considered, the effect on the signal plus background cross section is found to be less than  $< 0.5\%$ . Additionally, the effect on the shape of the b-jet  $p_T$  distribution is less than 6% for  $p_T < 300$  GeV and less than 2% for b-jet  $p_T > 300$  GeV. We use the NNPDF3.0 NLO [71] set for parton distribution functions (PDFs) for all event generation. Parton-level events are then interfaced with PYTHIA (v8.2.44) [69] to account for parton showering and hadronization processes. Finally, we use DELPHES (v3.4.2) [65] to simulate smearing and other detector effects using the CMS detector geometric configurations and parameters for particle identification and reconstruction, using the CMS input card with 140 average pileup interactions.



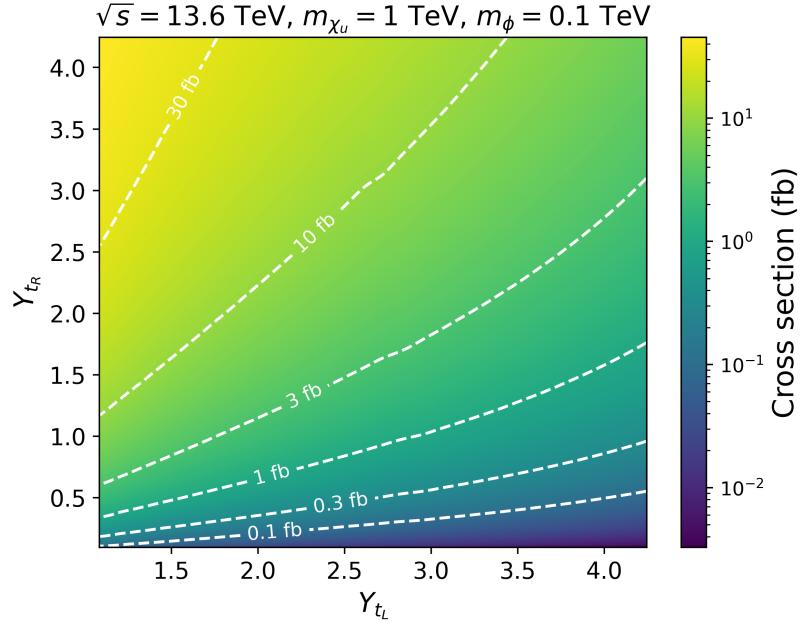
**Figure 3.3:** Representative Feynman diagram for the production of a  $\phi'$  boson in association with a  $\chi_u$  vector-like quark through the fusion of a top quark and  $\chi_u$  vector-like quark. Once again, the  $\phi'$  decays to a pair of muons, the top quark decays fully hadronically, and the  $\chi_u$  decays semi-leptonically to muons, neutrinos and b-jets.



**Figure 3.4:** Representative Feynman diagram for the production of a  $\phi'$  boson in association with a  $\chi_u$  vector-like quark through the fusion of a gluon pair from incoming protons. The  $\phi'$  decays to a pair of muons, the top quark that decays fully hadronically, and the  $\chi_u$  decay semi-leptonically to muons, neutrinos and jets.

All signal cross sections used in this analysis are obtained requiring the following kinematic criteria on leptons  $\ell$ , b quarks, and light-quark/gluon jets ( $j$ ) at parton level in MadGraph:  $p_T(\ell) > 35 \text{ GeV}$ ,  $|\eta(b)| < 2.5$ ,  $|\eta(\ell)| < 2.3$ ,  $p_T(j) > 20 \text{ GeV}$ , and  $|\eta(j)| < 5$ . These parton-level selections were applied exclusively to the signal processes to restrict event generation to the relevant phase space regions. For background processes, these default parton level requirements in MadGraph were imposed:  $p_T(\ell) > 10 \text{ GeV}$ ,  $|\eta(\ell)| < 2.5$ ,  $p_T(j) > 20 \text{ GeV}$ ,  $|\eta(j)| < 5$ , and  $|\eta(b)| < 5$ . This ensures that the phase space regions for the background near the analysis-level selection criteria are adequately described after parton showering since the pre-selections at the analysis level are more stringent than the parton-level requirements. Furthermore, we use the MLM algorithm for jet matching and jet merging. The parameters  $x_{\text{qcut}}$  and  $q_{\text{cut}}$  of the MLM algorithm are set to 30 and 45 respectively to ensure continuity of the differential jet rate as a function of jet multiplicity. Each simulated signal and background sample is produced separately at LO, with one million events at the generation level, neglecting potential interference effects between the signal and background due to the suppression caused by the different orders of magnitude in the coupling constants of the signal and background.

Signal samples are generated considering the production of a  $\phi'$  boson, an associated  $\chi_u$  vector-like quark, and a top quark ( $pp \rightarrow \chi_u t \phi'$ ), inclusive in both  $\alpha$  and  $\alpha_s$  (see Figures 3.3-3.4). We have used the implementation of the  $U(1)_{T_R^3}$  model in Ref. [148]. Signal samples were



**Figure 3.5:** Signal production cross section,  $pp \rightarrow \chi_u t \phi'$ , in the  $Y_{t_R}$  versus  $Y_{t_L}$  plane, for a benchmark point with  $m(\phi') = 100$  GeV and  $m(\chi_u) = 1.00$  TeV. The white-dashed contours show specific cross section values in the two dimensional plane.

created considering coupling values of  $Y_{t_R} = Y_{t_L} = 2\sqrt{2}$  in the range of masses  $m(\phi') \in \{5, 10, 50, 100, 325\}$  GeV for the dark higgs and  $m(\chi_u) \in \{0.50, 0.75, 1.0, 1.5, 2.0, 2.5\}$  TeV for the vector-like quark  $\chi_u$  [179]. The production cross section for  $pp \rightarrow \chi_u t \phi'$  is highly dependent on the choice of the Yukawa couplings in the Lagrangian. The  $\chi_u - t$  fusion process shown in Figure 3.3 is dominated by the  $Y_{t_R}$  coupling. However, the decay  $\chi_u \rightarrow t \phi'$  shown in Figure 3.4 is inversely proportional to the  $Y_{t_L}$  coupling. This effect is shown in Figure 3.5, which displays the total signal cross section, as a function of  $Y_{t_R}$  and  $Y_{t_L}$ , for a benchmark point with  $m(\phi') = 100$  GeV and  $m(\chi_u) = 1.0$  TeV.

We target signal events where the top quark decays hadronically into a bottom quark and two jets ( $t \rightarrow bW \rightarrow bq\bar{q}'$ ), the  $\chi_u$  decays semileptonically into a b quark, lepton, and neutrino (via  $\chi_u \rightarrow bW$  and  $W \rightarrow \mu\nu_\mu$ ), and the  $\phi'$  produces two muons. We note that the scalar  $\phi'$  particle could result from the mixture of the SM Higgs boson and additional scalar fields, and the Yukawas of the fermions could additionally arise from the mixing of the SM fermions with additional copies of the associated vector-like fermions. Therefore, the  $\phi'$  branching ratios are dependent on the chosen mechanism and model by which this mixture occurs, see for example, Refs. [180–183]. For the purpose of this work, and similar to Refs. [144, 148], the considered benchmark signal scenarios have  $\mathcal{B}(\chi_u \rightarrow b W)$  of about 0.5 and  $\mathcal{B}(\phi' \rightarrow \mu^+ \mu^-) = 1.00$ . Figure 3.6 shows the production cross section in fb, as a function of  $m(\phi')$  and  $m(\chi_u)$  masses, assuming the aforementioned decays, branching ratios, and couplings.

We note that for the parameter space of focus in this paper, the total mass of the  $t\chi_u$  system is larger than  $m(\phi')$ , thus the large rest energy of the  $t\chi_u$  system is converted into potentially large momentum values for the  $\phi'$ . Similarly, the  $t$ -quark produced through the  $\chi_u$ - $t$  fusion interaction can also have large momentum values, and thus in some cases the hadronic  $t$  decay products cannot be fully reconstructed independently of each other. This results in three possible  $t$  reconstruction scenarios: a fully merged scenario where the  $W \rightarrow jj$  system and the  $b$  quarks are very collimated and reconstructed as a single “fat jet” (henceforth referred to as a FatJet, FJ); a partially merged scenario, where the decay products of the  $W$  boson form a single FatJet but the  $b$  quark can still be separately identified; and an unmerged scenario where all decay products can be independently identified. Jets are clustered using the anti- $k_t$  algorithm [184] using the FastJet (v3.4.2) [185] package with a distance parameter of  $R = 0.4$  for standard jets and  $R = 0.8$  for fat jet objects. Each scenario has an associated identification efficiency and misidentification rate, which depends on the choice of the boosted  $t/W$  algorithm (our choice of efficiency and misidentification rates is described later).

Based on the above details, the final state of interest in this paper consists of three muons (two from the  $\phi'$  decay and one from the  $\chi_u$  decay), a (possibly boosted) top-tagged system, at least one  $b$ -tagged jet, and large missing transverse momentum ( $\vec{p}_T^{\text{miss}}$ ). For the partially merged and unmerged scenarios, there will be two  $b$  quarks present in the final state (one of which is part of the top tagged system).

We consider background sources from SM processes which can give similar objects in the final state as those expected for signal. Several background sources were considered and studied, such as QCD multijet events, production of vector boson pairs ( $VV$  :  $WW, ZZ, WZ$ ), vector boson triplets ( $VVV$  :  $WWZ, WZZ, ZZZ, WWW$ ), top-quark pairs in association with weak bosons ( $t\bar{t}X$ ), and  $t\bar{t}t\bar{t}$  processes. The dominant sources of SM background events are from the  $t\bar{t}X$ ,  $ZZW$ , and  $t\bar{t}t\bar{t}$  processes. The  $t\bar{t}X$  background is primarily associated production of a  $Z/\gamma^*$  from  $t\bar{t}$  fusion processes. The  $ZZW$  process becomes a background when one  $Z$  decays  $b\bar{b}$ , another  $Z$  decays to a pair of muons, and the  $W$  decays to a muon and a neutrino. Events from  $ZZW$  and  $t\bar{t}t\bar{t}$  have been combined, after being weighted by their corresponding production cross section. The combination is presented as the “ $b\bar{b}\mu\mu\nu$ ” background in the remainder of this paper. The  $t\bar{t}X$  process is presented as part of the “ $t\bar{t}\mu^+\mu^-$ ” background. Table 3.2 shows the production cross sections for the dominant background sources. The rest of the aforementioned background processes do not contribute meaningfully in our context, accounting for  $\ll 1\%$  of the total expected background yield.

The identification of leptons, boosted top quarks, and bottom quarks plays an important role in the ability to identify signal events, the ability to minimize the rate of SM backgrounds, and thus also the discovery reach in the high-luminosity environment of the LHC. It is worth noting that

Background Process	Cross-Section $\sigma$ [pb]
$pp \rightarrow t\bar{t} \mu^+ \mu^-$	$2.574 \times 10^{-3}$
$pp \rightarrow b\bar{b} \mu\mu\nu$	$4.692 \times 10^{-4}$

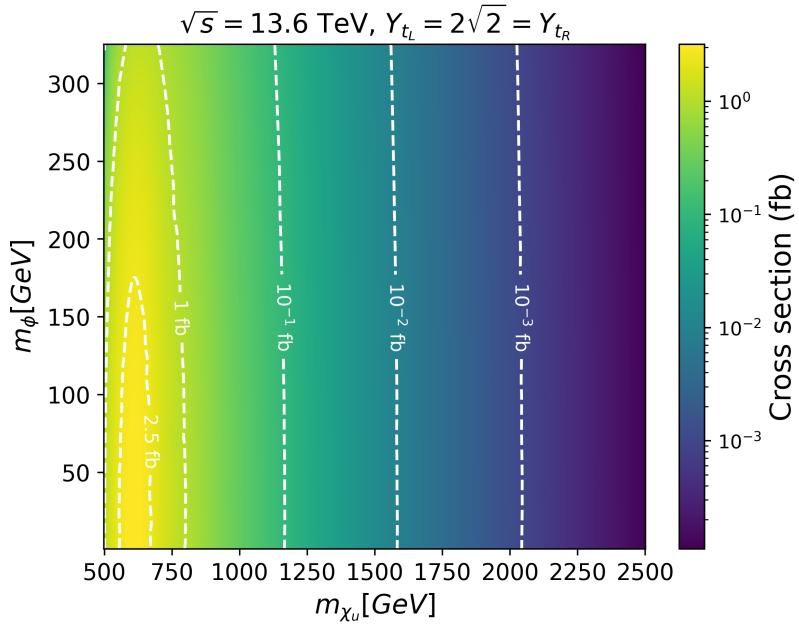
**Table 3.2:** A summary of dominant SM backgrounds produced by  $pp$  collisions and their cross sections in pb, as computed by MadGraph with  $n = 10^6$  events.

the reconstruction and identification of leptons and the decay products of the top/bottom quarks may be non-trivial at the High-Luminosity LHC (HL-LHC) due to the presence of a potentially large number of secondary  $pp$  interactions (pileup). The impact of pileup on the new physics discovery reach, and the importance of pileup mitigation at CMS and ATLAS has been outlined in many papers, for example in Ref. [82]. We note the expected performance of the upgraded ATLAS and CMS detectors for the HL-LHC is beyond the scope of this work; however, the studies presented here do attempt to provide reasonable expectations by conservatively assuming some degradation in lepton and hadron identification efficiencies, using Ref. [82] as a benchmark, and considering the case of 140 average pileup interactions.

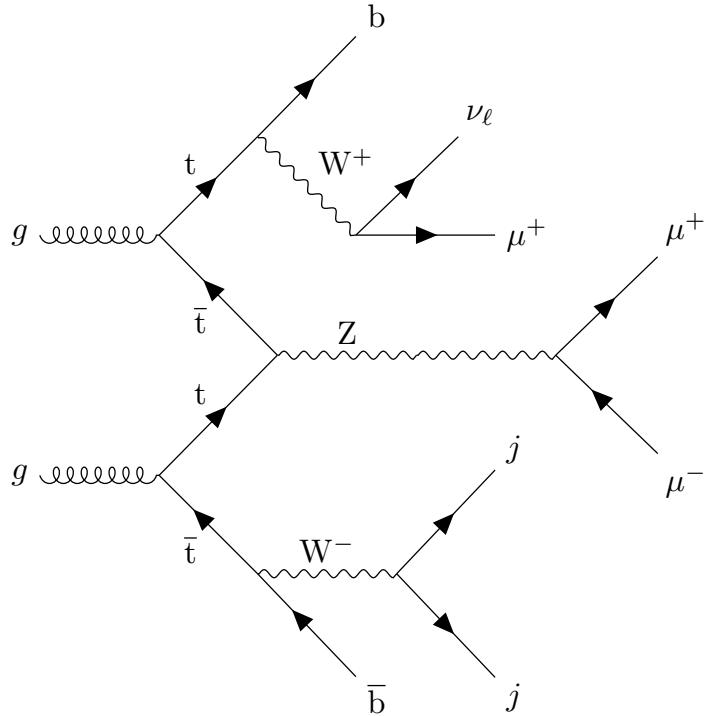
For muons with  $|\eta| < 1.5$ , the assumed identification efficiency is 95% with a 0.3% misidentification rate [82, 89]. The performance degrades linearly with  $\eta$  for  $1.5 < |\eta| < 2.5$ , and we assume an identification efficiency of 65% with a 0.5% misidentification rate at  $|\eta| = 2.5$ . Similarly, the charged hadron tracking efficiency, which contributes to the jet clustering algorithm and missing transverse momentum ( $\vec{p}_T^{\text{miss}}$ ) calculation, is 97% for  $1.5 < |\eta| < 2.5$ , and degrades to about 85% at  $|\eta| = 2.5$ . These potential inefficiencies due to the presence of secondary  $pp$  interactions contribute to how well the lepton and top kinematics can be reconstructed. Following Refs. [186, 187], we consider the “Loose” working point for the identification of the fully merged (partially merged)  $t$  decays, which results in 80-85% top (W) identification efficiency and 11-25% misidentification rate, depending on the FatJet transverse momentum ( $p_T^{\text{FJ}}$ ). Following Ref. [188], we consider the “Loose” working point of the DeepCSV algorithm [77], which gives a 70-80% b-tagging efficiency and 10% light quark mis-identification rate. The choice of boosted  $t/W$  and b-tagging working points is determined through an optimization process that maximizes discovery reach. It is noted the contribution from SM backgrounds with a misidentified boosted  $t/W$  is negligible, and thus our discovery projections are not sensitive to uncertainties related to the boosted  $t/W$  misidentification rates.

### 3.4 DATA ANALYSIS USING MACHINE LEARNING

The analysis of signal and background events is performed utilizing machine learning techniques. A machine learning-based approach offers sizeable advantages when compared to traditional event classification techniques. Unlike conventional methods, machine learning models have the



**Figure 3.6:** Projected cross section (fb) plot for  $\text{pp} \rightarrow t\chi_u\phi'$  and subsequent decay as a function of  $m(\chi_u)$  and  $m(\phi')$ .



**Figure 3.7:** Representative Feynman diagram for a background event. A Z boson is produced in association with a top quark through the fusion of a top, anti top pair from incoming protons. The Z boson subsequently decays to a pair of muons and the two spectator top quarks decay semi-leptonically and purely hadronically to muons, neutrinos and jets, resulting in the same final states as the signal event.

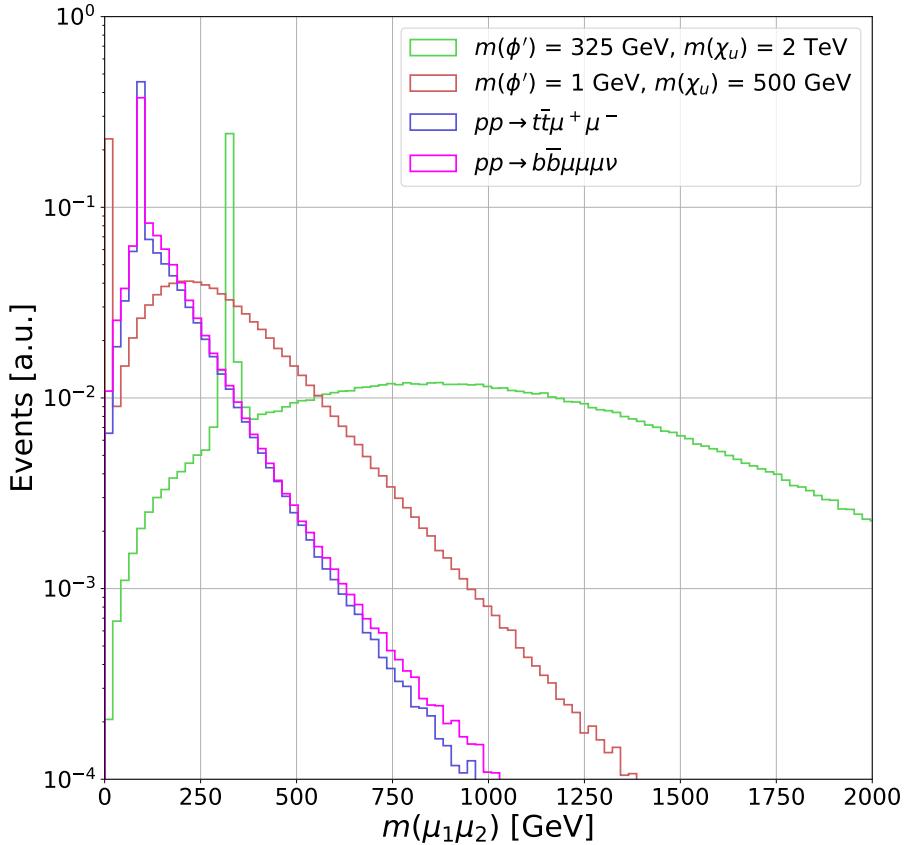
capability to simultaneously consider all kinematic variables, allowing them to efficiently navigate the complex and high-dimensional space of event kinematics. Consequently, machine learning models can effectively enact sophisticated selection criteria that take into account the entirety of this high-dimensional space. This makes them ideal for high-energy physics applications.

The BDT method is a powerful machine learning technique that has proven its effectiveness in various applications, particularly in the field of collider physics. In this method, decision trees are trained greedily in a sequential manner, with each tree focusing on learning the discrepancies or residuals between its predictions and the expected values obtained from the previously trained tree. This iterative process aims to progressively minimize errors, making BDTs a particularly effective approach for enhancing model performance.

In the context of collider physics, BDTs have demonstrated their utility in addressing classification problems. In particular, BDTs can effectively discriminate between signal and background events, enabling accurate and efficient event classification. Their ability to handle subtle non-linear relationships within the data with high interpretability makes BDTs a valuable tool to handle large amounts of data with a large number of parameters for each event.

The first step in our workflow involves the use of a specialized `MadAnalysis Expert Mode C++` script [189]. This script extracts essential kinematic and topological information from the simulated samples. The script will process the aforementioned variables contained within these files and transform them into a structured and informative CSV (Comma-Separated Values) format that can be used to train our machine learning models. These kinematic variables include crucial details about the events, such as particle momenta, energies, and topologies, providing the fundamental building blocks for our machine learning analysis. Figure 3.13 shows the features that are used for training the machine learning models and their importance for a benchmark point.

To account for the differential significance of various events, we apply cross-section weighting. This ensures that the relative importance of signal and background events is appropriately balanced in the dataset. This weighting is crucial for addressing the varying likelihood of observing different types of events in high-energy physics experiments. The prepared and weighted datasets are then passed to our `MadAnalysis Expert Mode C++` script, where the simulated signal and background events are initially filtered, before being passed to the CSV file for use by the machine learning algorithm. The filtering process requires at least one well-reconstructed and identified b-jet candidate, at least one jet (regular or FJ) not tagged as a b jet, and exactly three identified muons. The filtering selections are motivated by experimental constraints, such as the geometric constraints of the CMS/ATLAS detectors, the typical kinematic thresholds for the reconstruction of particle objects, and the available lepton triggers which

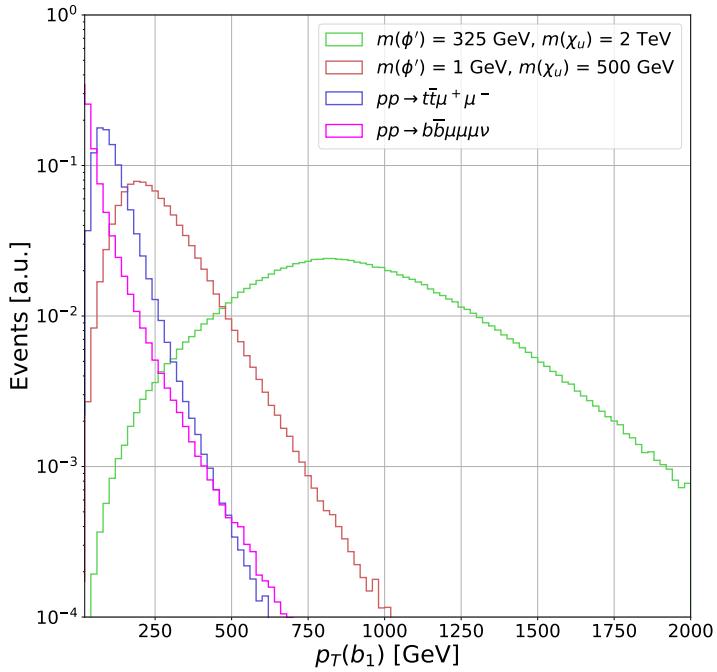


**Figure 3.8:** Invariant mass distribution of the muon pair with the highest and second highest transverse momentum. The distributions are shown for the two main SM background processes and two signal benchmark points.

also drive the minimal kinematic thresholds. Selected jets must have  $p_T > 30$  GeV and  $|\eta(j)| < 5.0$ , while b-jet candidates with  $p_T > 20$  GeV and  $|\eta(b)| < 2.5$  are chosen. The  $\mu$  object must pass a  $p_T > 35$  GeV threshold and be within a  $|\eta(\ell)| < 2.3$ . We will refer to this filtering criteria as pre-selections. The efficiency of the pre-selections depends on  $m(\phi')$  and  $m(\chi_u)$ , but is typically about 25-30% for the signal samples. Events passing this pre-selection are used as input for the machine learning algorithm, which classifies them as signal or background, using a probability factor.

We explore the performance of a diverse set of machine learning models, specifically three neural networks of differing architectures and a BDT algorithm. To ensure robust model assessment, we employed a standard 90-10 train-test split of the dataset, partitioning it into a 90% portion for training and a 10% portion for testing. This division allows us to gauge the generalization capabilities of our models on unseen data.

The training and evaluation of the BDT were carried out in a high-performance computing environment. Specifically, an Nvidia A100 GPU was used. The canonical PyTorch [190] deep learning framework was employed for configuring, training, and evaluating the neural networks.



**Figure 3.9:** Transverse momentum distribution of the leading b-quark jet candidate. The distributions are shown for the two main SM background processes and two signal benchmark points.

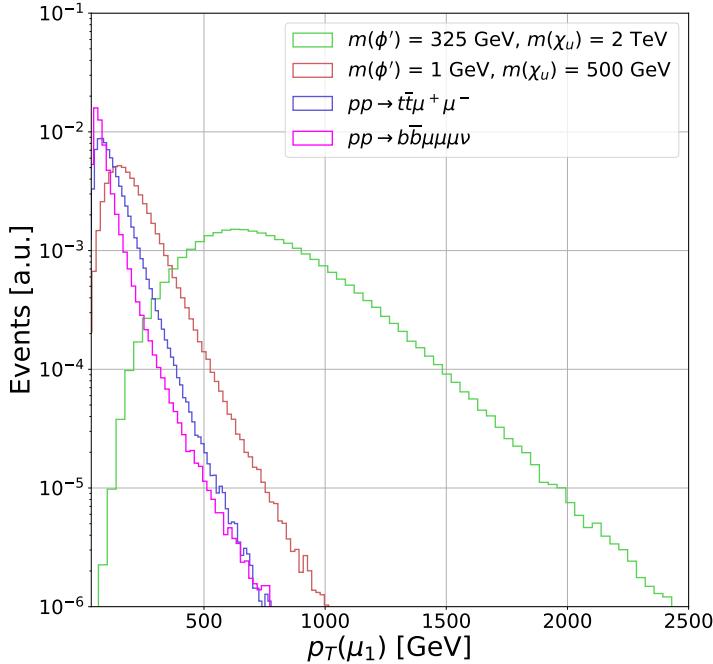
Model	Train/Test Acc.	Training Time
BDT	N.A./0.9993	6s
Neural Network 1	0.9999/0.9997	1h 58m
Neural Network 2	0.9999/0.9998	2h 12m
Neural Network 3	0.9999/0.9998	2h 32m

**Table 3.3:** Train/test results for the ML models.

PyTorch is well-regarded for its flexibility and performance in deep learning applications.

For the BDT algorithm, we used hyperparameters  $\eta = 0.3$ ,  $\gamma = 0$ , and `max_depth = 6`. The XGBoost [134] library was used for the implementation of the Boosted Decision Tree algorithm. It offers high efficiency, optimization, and interpretability, making it a suitable choice for this particular task.

It is worth mentioning that we experimented with deep neural networks of various architectures. Although we found that they yield similar signal sensitivity to the BDT, the complex nature of the studies in this work (particle objects considered, experimental constraints in a high luminosity LHC, etc.) motivates the use of a BDT over a deep neural network because of its usefulness, efficiency, and simplicity in understanding the machine learning output in addition to significantly shorter training times. Therefore, we perform our proceeding analysis using the BDT. The outcomes of our model training and evaluation are presented in Table 3.



**Figure 3.10:** Transverse momentum distribution of the leading muon candidate. The distributions are shown for the two main SM background processes and two signal benchmark points.

### 3.5 RESULTS

---

Figures 3.8, 3.9, and 3.10 show relevant kinematic distributions for two benchmark signal points and the dominant SM backgrounds, using the subset of events passing the pre-selections defined above. The signal benchmark points in these figures are  $m(\phi') = 325 \text{ GeV}$ ,  $m(\chi_u) = 2 \text{ TeV}$ , and  $m(\phi') = 1 \text{ GeV}$ ,  $m(\chi_u) = 500 \text{ GeV}$ . The distributions are normalized such that the area under the curve is unity. These distributions correspond to the reconstructed mass,  $m(\mu_1, \mu_2)$ , between the two muon candidates with the highest transverse momentum ( $\mu_1$  and  $\mu_2$ ), the transverse momentum of the b-jet candidate with the highest transverse momentum  $p_T(b_1)$ , and the muon candidate with the highest transverse momentum  $p_T(\mu_1)$ , respectively. These distributions are among the variables identified by the BDT algorithm with the highest signal to background discrimination power (see Figure 3.13).

As can be seen from Figure 3.8, the  $\phi'$  mass can be reconstructed through its associated muon decay pair, which is observed as a peak in the  $m(\mu_1, \mu_2)$  distribution around the expected  $m(\phi')$  value, and has low- and high-mass tails which are a consequence of cases where the leading and/or subleading muon is not from the  $\phi'$  decay, but rather from the associated W boson from the  $\chi_u$  decay. For the backgrounds, muons come from Z (W) decays. Therefore, the  $m(\mu_1, \mu_2)$  background distributions show a peak near  $m_{W/Z}$ , combined with a broad distribution indicative of the combination of two muon candidates from different decay vertices.

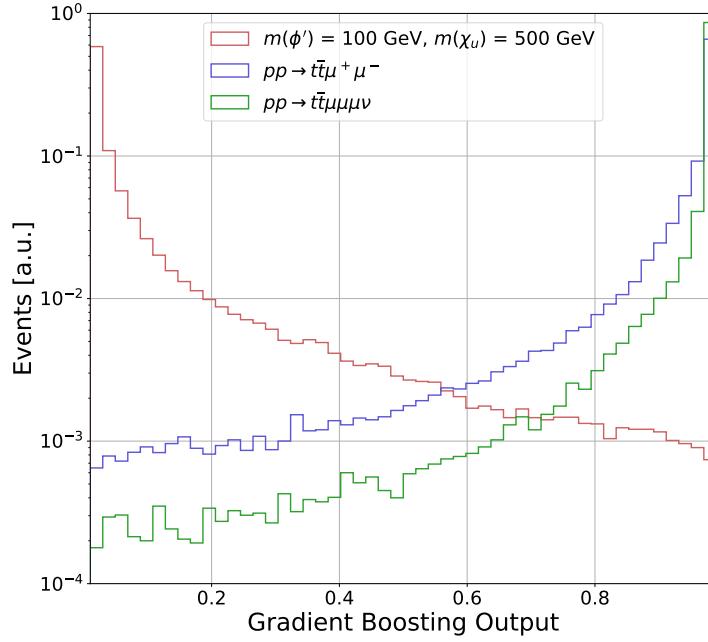
We note that the  $\phi' \rightarrow \mu^+ \mu^-$  decay width depends on the square of the  $\phi' \rightarrow \mu^+ \mu^-$  coupling and  $\frac{m_\mu^2}{m_{\phi'}^2}$  and is thus suppressed by the relatively small muon mass. For the new physics phase space considered in this paper, the  $\phi'$  decay width is less than 1% of the  $\phi'$  resonant mass. Furthermore, as indicated previously, the signal/background interference effects are small and negligible compared to effects from experimental resolution. Therefore, the width of the  $m(\mu_1, \mu_2)$  signal distributions is driven by the experimental resolution in the reconstruction of the muon momenta, as well as the probability that the two leading muons are the correct pair from the  $\phi'$  decay. Since the probability that the two highest- $p_T$  muons are the correct pair from the  $\phi' \rightarrow \mu^+ \mu^-$  decay depends on  $m(\phi')$  and  $m(\chi_u)$ , it is important to include all possible combinations of dimuon pairs (i.e.,  $m(\mu_1, \mu_3)$  and  $m(\mu_2, \mu_3)$ ) in the training of the BDT.

Figure 3.9 shows the distribution for the b-jet candidate with the highest  $p_T$ ,  $p_T(b_1)$ , for the same simulated samples shown in Figure 3.8. Based on the signal topology and our choice of parameter space (i.e.,  $m_{\chi_u} > m_t$ ), it is expected that the leading b-jet candidate comes from the  $\chi_u$  decay, with an average  $p_T$  close to  $\frac{m_{\chi_u} - m_W}{2}$ , as observed in Figure 3.9. For the  $t\bar{t}\mu^+ \mu^-$  background, the b-jet candidates come from top-quark decays. Therefore, their average transverse momentum is expected to be  $\frac{m_t - m_W}{2} \approx 45$  GeV, as observed in Figure 3.9. On the other hand, the b-jet candidates for the  $b\bar{b}\mu^+ \mu^-$  background can come from off-mass-shell  $Z^*/\gamma^*$ , and thus typically have an even softer spectrum in comparison to the  $t\bar{t}\mu^+ \mu^-$  background.

Figure 3.10 shows the distribution for the muon candidate with the highest  $p_T$ ,  $p_T(\mu_1)$ . Similar to Figure 3.9, when  $m(\chi_u) > m_t$  it is expected that the leading muon candidate comes from the  $\chi_u$  decay, with an average  $p_T$  of approximately  $\frac{m(\chi_u) - m_W}{4}$ , as observed in Figure 3.10. For the major SM backgrounds, the muon candidates come from  $Z/W/\gamma^*$  decays. Therefore, their average transverse momentum is expected to be much lower,  $\frac{m_{Z/W}}{4} \approx 40 - 45$  GeV. This kinematic feature provides a nice handle to discriminate high  $m(\chi_u)$  signal events amongst the large SM backgrounds, which have lower average  $p_T(\mu)$  constrained by the SM weak boson masses.

In addition to these aforementioned variables in Figures 3.8-3.10, several other kinematic variables were included as inputs to the BDT algorithm. In particular, 27 such variables were used in total, and these included the momenta of b and muon candidates; invariant masses of pairs of muons; angular differences between b jets and between the muons.

As mentioned above, the variables  $m(\mu_i, \mu_j)$  for  $i, j \neq 1$  provide some additional discrimination between signal and background when the leading muons are not a  $\phi'$  decay candidate. The angular separation variables, such as  $\Delta R(\mu_i, \mu_j)$ , are designed to be sensitive to lower mass  $\phi'$ , since the low rest mass of those particles means they acquire more boost, and thus smaller angular separation  $\Delta R$  between the muon candidates. The trained BDT returns the discriminating power of each of its inputs, and the feature

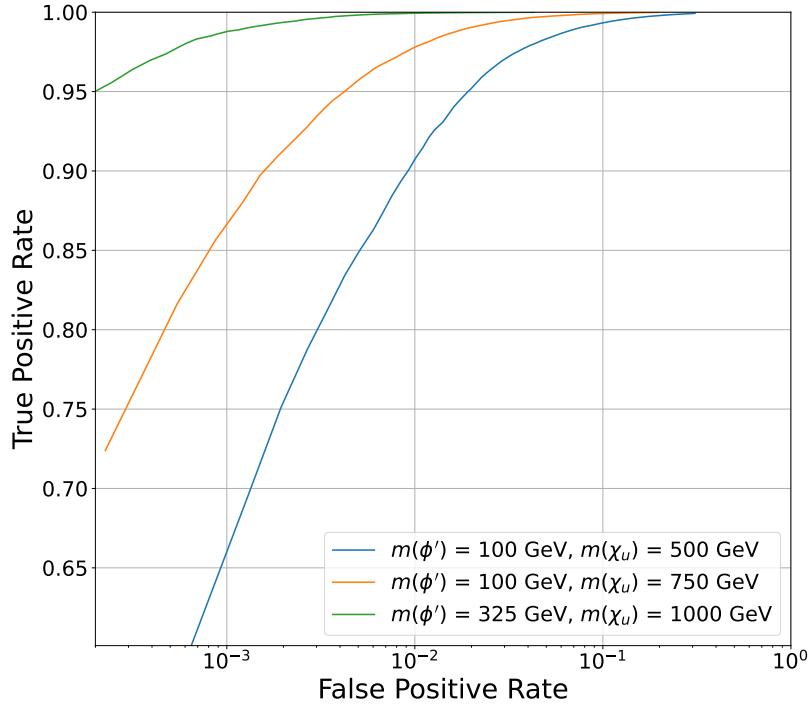


**Figure 3.11:** Output of the gradient boosting algorithm for a benchmark  $m(\phi') = 100$  GeV and  $m(\chi_u) = 500$  GeV signal, and dominant backgrounds. The distributions are normalized to unity.

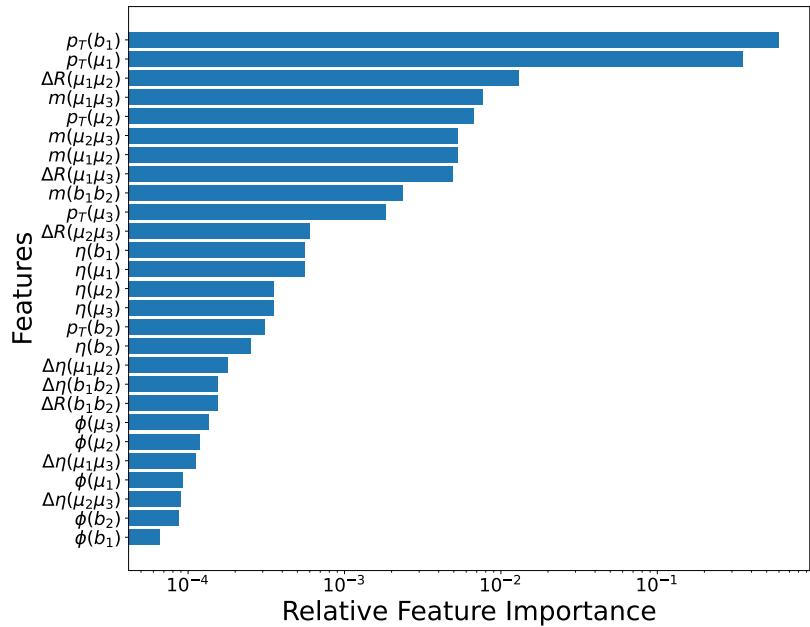
importance for each variable is shown in Figure 3.13 for a signal benchmark point with  $m(\phi') = 325$  GeV and  $m(\chi_u) = 2000$  GeV.

Figure 3.11 shows the distributions for the output of the BDT algorithm, normalized to unity, for the representative signal benchmark point of  $m(\phi') = 1$  GeV,  $m(\chi_u) = 0.5$  TeV and the two dominant backgrounds. The output of the BDT algorithm is a value between 0 and 1, which quantifies the likelihood that an event is either background-like (BDT output near 1) or signal-like (BDT output near 0). Figure 3.12 illustrates the true positive rate (TPR), defined as the probability of correctly selecting signal events using the BDT output, plotted against the false positive rate (FPR), defined as the probability of incorrectly selecting background events. For example, for  $m(\phi') = 100$  GeV and  $m(\chi_u) = 500$  GeV, when signal events are selected at 65% probability, the background is selected at about  $10^{-3}$  probability. We note that the primary discriminating feature between the signal and background is the boosted b-jet  $p_T$  coming from the  $\chi_u$  vector-like quark. The  $p_T$  of said b jet increases with  $m(\chi_u)$ , peaking at around  $[m(\chi_u) - m(W)]/2$ . This enhanced boost increases the separation between signal and background, improving the performance of the BDT algorithm as  $m(\chi_u)$  increases.

The outputs from the BDT machine learning algorithm are used to perform a profile-bin likelihood analysis to estimate the signal significance for a luminosity of  $3000 \text{ fb}^{-1}$ , corresponding to the expected amount of collected data by the end of the LHC era. For this purpose, the BDT distributions are normalized to cross section times pre-selection efficiency times luminosity for the different signal models. The significance is then calcu-



**Figure 3.12:** Receiver operating characteristic curve of the BDT algorithm for three different signal benchmark scenarios.



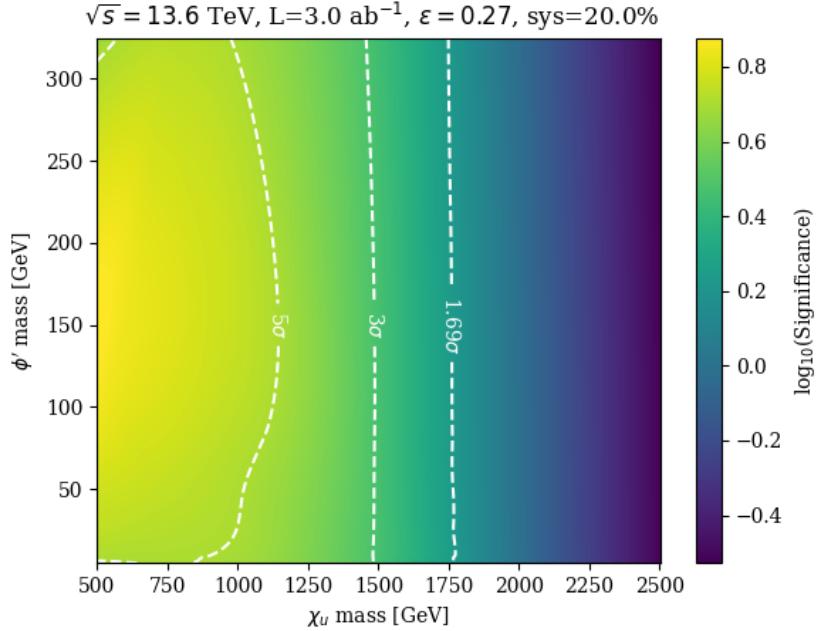
**Figure 3.13:** Relative importance of features in training for a benchmark signal scenario with  $m(\phi') = 325 \text{ GeV}$  and  $m(\chi_u) = 2000 \text{ GeV}$ .

lated using the expected bin-by-bin yields of the BDT output distribution in a profile likelihood fit, using the ROOTFit [92] package developed by CERN. The expected signal significance  $Z_{\text{sig}}$  is calculated using the probability of obtaining the same test statistic for the signal plus background and the signal-null hypotheses, defined as the local p-value. Similar to Refs. [191–197], the significance corresponds to the point where the integral of a Gaussian distribution between  $Z_{\text{sig}}$  and  $\infty$  results in a value equal to the local p-value. The estimation of  $Z_{\text{sig}}$  incorporates systematic uncertainties. The uncertainty values have been included as nuisance parameters, considering lognormal priors for normalization and Gaussian priors for uncertainties associated with the modeling of the shapes similar to Refs. [198, 199].

The systematic uncertainties that have been included result from experimental and theoretical constraints. A 1-5% systematic uncertainty, depending on the simulated MC sample, has been included to account for the choice of Parton Distribution Function (PDF) set. The systematic uncertainty effect was incorporated following the PDF4LHC [92] recommendations. This systematic uncertainty has a small impact on the expected event yields for signal and background, but it does not affect the shape of the BDT output distribution. We additionally considered theoretical uncertainties related to the absence of higher-order contributions to the signal cross sections, which can change the pre-selection efficiencies and the shapes of kinematic variables used as inputs to the BDT algorithm. This uncertainty was calculated by varying the renormalization and factorization scales by  $\times 2$ , and studying the resulting change in the bin-by-bin yields of the BDT distributions. They are found to be at most 2% in a given bin.

Regarding experimental uncertainties, following experimental measurements from CMS on the estimation of the integrated luminosity, a conservative 3% effect has been included [74]. A 5% systematic uncertainty associated with the reconstruction and identification of b-quark jets has been included, independent of  $p_T$  and  $\eta$  of the b-jet candidates. According to Ref. [188], this uncertainty is correlated between signal and background processes with genuine b-jets and is also correlated across BDT bins for each process. For muons, we include a 2% uncertainty associated with the reconstruction, identification, and isolation requirements, and a 3% systematic uncertainty to account for scale and resolution effects on the momentum and energy measurement. We consider jet energy scale uncertainties ranging from 2-5%, contingent on  $\eta$  and  $p_T$ , resulting in shape-based uncertainties on the BDT output distribution. Jet energy scale uncertainties were assumed to range from 1-5%, contingent on  $\eta$  and  $p_T$ . These assumptions lead to shape-based uncertainties on the BDT output distribution, varying from 1-2%. Additionally, we include a 10% systematic uncertainty to account for errors in the signal and background predictions. Considering all the various sources of systematic uncertainties, our conservative estimate yields a total effect of about 20%.

Figure 3.14 shows the expected signal significance considering an integrated luminosity of  $3000 \text{ fb}^{-1}$ . The significance is shown as a heat map



**Figure 3.14:** Signal significance for the high luminosity LHC era, considering with  $3000 \text{ fb}^{-1}$  of collected data.

in a two-dimensional plane for different  $\phi'$  and  $\chi_u$  masses. The x-axis corresponds to  $m(\chi_u)$ , the y-axis to  $m(\phi')$ , and the heat map to  $\log_{10}(Z_{\text{sig}})$ . The white dashed lines are contours of constant signal significances of  $1.69\sigma$ ,  $3\sigma$  and  $5\sigma$  to represent regions of possible exclusion, evidence of new physics, and discovery, respectively. Under these conditions,  $\phi'$  ( $\chi_u$ ) masses ranging from 1 to 325 GeV (500 to 1800 GeV) can be probed. The range for a discovery with  $5\sigma$  signal significance varies from  $\chi_u$  masses from  $m(\chi_u) = 770\text{-}1100 \text{ GeV}$ , depending  $m(\phi')$ . For large  $m(\chi_u)$ , the significance is almost independent of  $m(\phi')$  because the primary discriminating feature—the boosted b-quark originating from  $\phi'$ —is driven predominantly by the large  $m(\chi_u)$ , with the kinematic impact of  $m(\phi')$  being relatively negligible.

### 3.6 DISCUSSION

---

The LHC will continue to run with pp collisions at  $\sqrt{s} = 13.6 \text{ TeV}$  for the next decade. Given the increase in the integrated luminosity expected from the high-luminosity program, it is important to consider unexplored new physics phase space that diverges from the conventional assumptions made in many BSM theories, and which could have remained hidden in processes that have not yet been thoroughly examined. It is additionally crucial to explore advanced analysis techniques, in particular the use of artificial intelligence algorithms, to enhance the probability of detecting these rare

corners where production cross sections are lower and discrimination from SM backgrounds is difficult.

In this work, we examine a model based on a  $U(1)_{T_R^3}$  extension of the SM, which can address various conceptual and experimental issues with the SM, including the mass hierarchy between generations of fermions, the thermal dark matter abundance, and the muon  $g - 2$ ,  $R_{(D)}$ , and  $R_{(D^*)}$  anomalies. This model contains a light scalar boson  $\phi'$ , with potential masses below the electroweak scale, and TeV-scale vector-like quarks  $\chi_u$ . We consider the scenario where the scalar  $\phi'$  has family non-universal fermion couplings and  $m(\phi') \geq 1$  GeV, as was suggested in Ref. [144], and thus the  $\phi'$  can primarily decay to a pair of muons. Previous works in Refs. [148, 158] considered scenarios motivating a search methodology with a merged diphoton system from  $\phi' \rightarrow \gamma\gamma$  decays. The authors of Ref [148], in which  $m(\phi') < 1$  GeV, indeed pointed out that if the  $\phi'$  is heavier than about 1 GeV, then decays to  $\mu^+\mu^-$  can become the preferable mode for discovery, which is the basis for the work presented in this paper. We further note that the final state topology studied in this paper would represent the most important mode for discovery at  $m(\phi') < 2m_t$  where the  $\phi' \rightarrow t\bar{t}$  decay is kinematically forbidden.

The main result of this paper is that we have shown that the LHC can probe the visible decays of new bosons with masses below the electroweak scale, down to the GeV-scale, by considering the simultaneous production of heavy QCD-coupled particles, which then decay to the SM particles that contain large momentum values and can be observed in the central regions of the CMS and ATLAS detectors. The boosted system combined with innovative machine learning algorithms allows for the signal extraction above the lower-energy SM background. The LHC search strategy described here can be used to discover the prompt decay of new light particles. An important conclusion from this paper is that the detection prospects for low-mass particles are enhanced when it is kinematically possible to simultaneously access the heavy degrees of freedom which arise in the UV completion of the low-energy model. This specific scenario in which the couplings of the light scalars are generationally dependent, with important coupling values to the top quark, is an ideal example which would be difficult to directly probe at low energy beam experiments.

The proposed data analysis represents a competitive alternative to complement searches already being conducted at the LHC, allowing us to probe  $\phi'$  masses from 1 to 325 GeV, for  $m(\chi_u)$  values up to almost 2 TeV, at the HL-LHC. Therefore, we strongly encourage the ATLAS and CMS Collaborations to consider the proposed analysis strategy in future new physics searches.



# 4

---

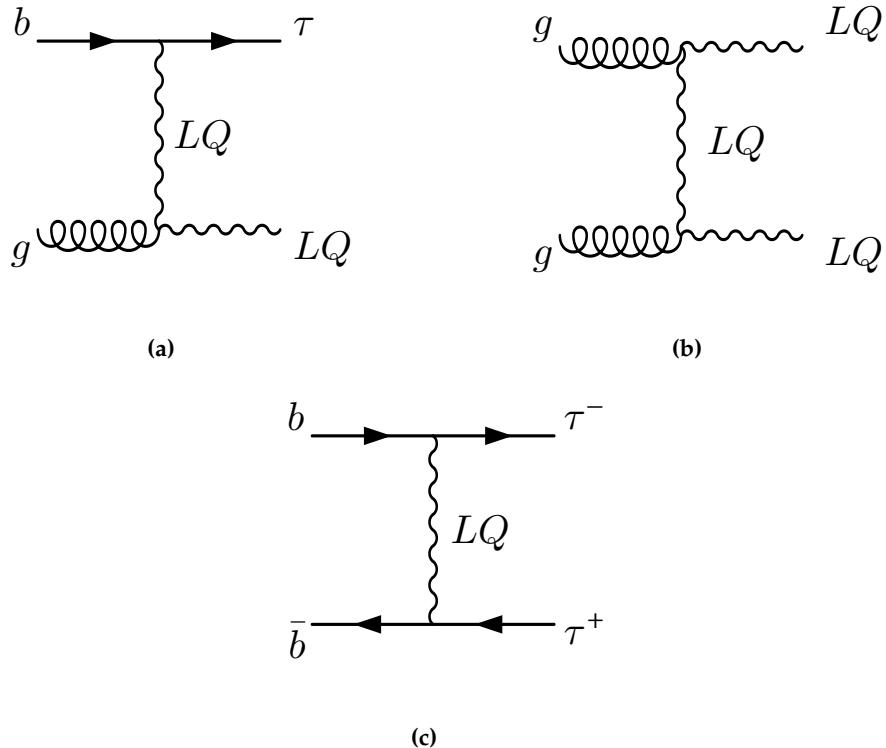
## ON VECTORIAL LEPTOQUARKS SENSITIVITY AT THE LHC

---

Leptoquarks (LQs) are hypothetical bosons carrying both baryon and lepton number, thus interacting jointly with a lepton and a quark. They are a common ingredient in SM extensions where quarks and leptons share the same multiplet. Typical examples of these can be found in the Pati-Salam [200] and SU(5) GUT [201] models. In addition, they can also be found in theories with strong interactions, such as compositeness [202]. Due to their exotic coupling which allows quark-lepton transitions, they have a diverse phenomenology, which naturally leads to several constraints. An important one comes from proton decay, which forces the LQ mass to values close to the Planck scale, unless baryon and lepton numbers are not violated. Furthermore, in models where the latter are conserved, the LQ can still be subject to a wide variety of bounds [54, 203–207]. Examples of these come from meson mixing, electric and magnetic dipole moments, atomic parity violation tests, rare decays, and direct searches. Nevertheless, the significance of each bound is a model dependent question.

In the last years, an increased interest in low scale LQs has emerged due to the anomalies in the precision measurements of the B-meson decay rates. As it is well known, these corresponded mainly to deviations in the  $R_{K^{(*)}}$  [59, 60, 208, 209] and  $R_{D^{(*)}}$  [61, 210–220] ratios, which measure the violation of lepton flavour universality (LFU). What followed was a very intense theoretical development, aiming to explain the anomalies by TeV scale LQ exchange at tree level [62, 221–235]. Before the end of 2022, it was generally agreed that, within proposed single LQ solutions, the only candidate capable of addressing all B-meson anomalies simultaneously and surviving all other constraints was a vector LQ ( $U_1$ ), transforming as  $(\mathbf{3}, \mathbf{1}, 2/3)$ , and coupling mainly to third-generation fermions via  $b\tau$  and  $t\nu_\tau$  vertices [62, 233]. In spite of a recent re-analysis of  $R_{K^{(*)}}$  data showing this ratio to be compatible with the SM prediction [68, 236–238], the solution to the  $R_{D^{(*)}}$  anomaly is still an open question and remains a valid motivation for the study of scenarios where new particles have preferential couplings to third-generation fermions. Thus, it is still of interest to continue exploring the possibility of observing the  $U_1$  LQ at the LHC [235].

As expected, the theoretical community has extensively participated in probing LQ models by scrutinizing search strategies, recasting LHC results,



**Figure 4.1:** Representative Feynman diagrams of single (a), pair (b), and non-resonant (c) production leptoquarks in proton-proton collision experiments. In single and pair production, the diagrams shown involve t-channel LQ exchange, dominant for lower LQ mass. However, for larger mass there exist s-channel diagrams featuring a virtual bottom quark and gluon, respectively.

and predicting the reach in the parameter space via different searches involving third-generation fermions (see for instance [162, 239–247]). In addition, several 13 TeV searches for LQs decaying into t/b and  $\tau/\nu$  final states have been performed by the CMS [248–256] and ATLAS [257–263] collaborations.

Of the searches above, we find [254] particularly interesting. Here, the CMS collaboration explores signals corresponding to  $t\nu b\tau$  and  $t\nu\tau$  final states, with  $137\text{ fb}^{-1}$  of proton-proton ( $p p$ ) collision data. The former is motivated by LQ pair production, with one LQ decaying into  $t\nu$  and the other into  $b\tau$ , while the latter arises from a single LQ produced in association with a  $\tau$ , with a subsequent LQ decay into  $t\nu$  (see Figure 4.1 for the corresponding diagrams). From the combination of both production channels, the search excludes  $U_1$  masses under  $1.3 - 1.7\text{ TeV}$ , with this range depending on the  $U_1$  coupling to gluons and on its coupling  $g_U$  in the  $b_L \tau_L$  vertex.

What makes this search particularly attractive is that, for the first time, an LHC collaboration directly places (mass dependent) bounds on  $g_U$ . This is important, since having information on this parameter is crucial in order to understand if the  $U_1$  is really responsible for the  $R_{D^{(*)}}$

anomaly. The inclusion of the single-LQ production mode is important, since its cross-section is directly proportional to  $g_U^2$ . However, as can be seen in Figure 6 of [254], the current constraints are dominated by pair production, with single-LQ production playing a subleading role. While this is expected [242], it still leads us to ponder the possibility of improving the sensitivity of LHC searches to single-LQ production, and thus on achieving better constraints on  $g_U$ . Other complementary and similar searches to [254] were carried out by both ATLAS [262] and CMS [256].

It is also well known, though, that searches for an excess in the high- $p_T$  tails of  $\tau$  lepton distributions can strongly probe  $g_U$ , up to very large LQ masses. Indeed, as shown in [235, 264], the new physics effective operators contributing to  $R_D^{(*)}$  also contribute to an enhancement in the  $p p \rightarrow \tau\tau$  production rates. This has motivated a large number of recasts [233, 235, 242, 243, 265–269], as well as a CMS search explicitly providing constraints in terms of  $U_1$  [255]. Nevertheless, it is important to note that for these  $p p \rightarrow \tau\tau$  processes, the LQ participates non-resonantly, so contributions to the  $p p \rightarrow \tau\tau$  rates and kinematic distributions from non-LQ BSM diagrams containing possible virtual particles, such as a heavy neutral vector boson  $Z'$ , could spoil a straightforward interpretation of any possible excess [243]. Thus, it is also necessary to understand how the presence of other virtual particles can affect the sensitivity of an analysis probing  $g_U$ .

In this work we study the projected LQ sensitivity at the LHC, considering already available  $p p$  data as well as the expected amount of data to be acquired during the High-Luminosity LHC (HL-LHC) runs. We explore a proposed analysis strategy which utilizes a combination of single-, double-, and non-resonant-LQ production, targeting final states with varying  $\tau$ -lepton and b-jet multiplicities. The studies are performed considering various benchmark scenarios for different LQ masses and couplings, also taking into account distinct chiralities for the third-generation fermions in the LQ vertex. We also assess the impact of a companion  $Z'$ , which is typical of gauge models, in non-resonant LQ probes, and find that interference effects can have a significant effect on the discovery reach. We consider this effect to be of high interest, given that non-resonant LQ production can have the largest cross-section, and thus could be an important channel in terms of discovery potential.

An important aspect of this work is that the analysis strategy is developed using a machine learning (ML) algorithm based on Boosted Decision Trees (BDT)[129]. The output of the event classifier is used to perform a profile-binned likelihood test to extract the overall signal significance for each model considered in the analysis. The advantage of using BDTs and other ML algorithms has been demonstrated in several experimental and phenomenological studies [160–166]. In our studies, we find that the BDT algorithm gives sizeable improvement in signal significance.

#### 4.1 A SIMPLIFIED MODEL FOR THE $U_1$ LEPTOQUARK

---

Extending the SM with a massive  $U_1$  vector LQ is not straightforward, as one has to ensure the renormalizability of the model. Most of the theoretical community has focused on extensions of the Pati-Salam (PS) models which avoid proton decay, such as the scenario found in [135]. Other examples include PS models with vector-like fermions [270–272], the so-called 4321 models [273–275], the twin PS<sup>2</sup> model [276, 277], the three-site PS<sup>3</sup> model [278–280], as well as composite PS models [281–283].

In what follows, we shall restrict ourselves to a simplified non-renormalizable lagrangian, understood to be embedded into a more complete model. The SM is thus extended by adding the following terms featuring the  $U_1$  LQ:

$$\begin{aligned} \mathcal{L}_{U_1} = & -\frac{1}{2} U_{\mu\nu}^\dagger U^{\mu\nu} + M_U^2 U_{1\mu}^\dagger U_1^\mu \\ & -ig_s U_{1\mu}^\dagger T^a U_{1\nu} G^{a\mu\nu} - i\frac{2}{3} g' U_{1\mu}^\dagger U_{1\nu} B^{\mu\nu} \\ & + \frac{g_u}{\sqrt{2}} [U_{1\mu} (\bar{Q}_3 \gamma^\mu L_3 + \beta_L^{s\tau} \bar{Q}_2 \gamma^\mu L_3 \\ & + \beta_R \bar{b}_R \gamma^\mu \tau_R) + h.c.] \end{aligned} \quad (4.1)$$

where  $U_{\mu\nu} \equiv \mathcal{D}_\mu U_{1\nu} - \mathcal{D}_\nu U_{1\mu}$ , and  $\mathcal{D}_\mu \equiv \partial_\mu + ig_s T^a G_\mu^a + i\frac{2}{3} g' B_\mu$ . As evidenced by the second line above, we assume that the LQ has a gauge origin <sup>1</sup>.

The third and fourth lines in Eq. (4.1) shows the LQ interactions with SM fermions, with coupling  $g_u$ , which we have chosen as preferring the third generation <sup>2</sup>. These are particularly relevant for the LQ decay probabilities, as well as for the single-LQ production cross-section. The  $\beta_L^{s\tau}$  parameter, which is the  $LQ \rightarrow s\tau$  coupling in the  $\beta_L$  matrix (see footnote), is chosen to be equal to 0.2, following the fit done in [267], in order to simultaneously solve the  $R_{D^{(*)}}$  anomaly and satisfy the  $p p \rightarrow \tau^+ \tau^-$  constraints. Although  $\beta_L^{s\tau}$  technically alters the single-LQ production cross-section and LQ branching fractions, we have confirmed that a value of  $\beta_L^{s\tau} = 0.2$  results in negligible impact on our collider results, and thus is ignored in our subsequent studies.

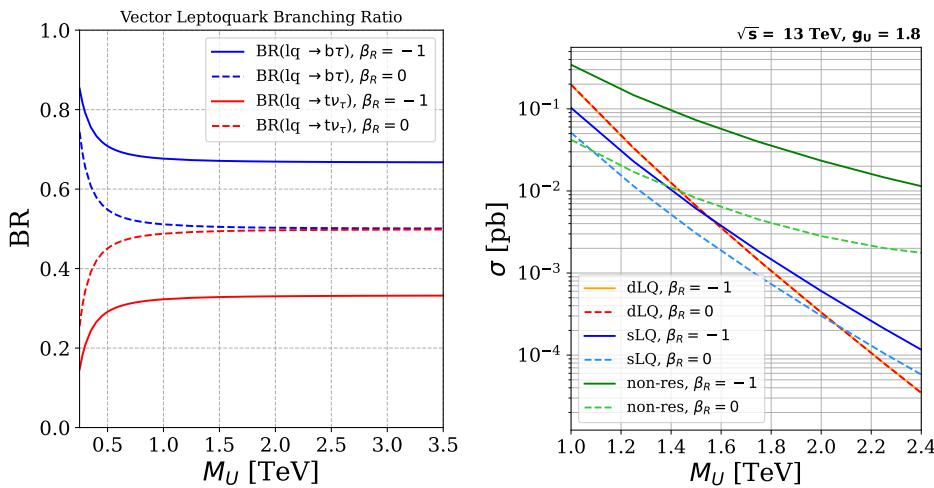
The LQ right-handed coupling is modulated with respect to the left-handed one by the  $\beta_R$  parameter. The choice of  $\beta_R$  is important phenomenologically, as it affects the LQ branching ratios <sup>3</sup>, as well as the single-LQ production cross-section. To illustrate the former, Figure 4.2 (top) shows the  $LQ \rightarrow b\tau$  and  $LQ \rightarrow t\nu$  branching ratios as functions of the LQ

<sup>1</sup> The couplings in the second line of Eq. (4.1) can be found in the literature as  $g_s \rightarrow g_s(1 - \kappa_U)$  and  $g' \rightarrow g'(1 - \tilde{\kappa}_U)$ , in order to take into account the possibility of an underlying strong interaction.

<sup>2</sup> Before the demise of the  $R_{K^{(*)}}$  anomaly [68, 236–238], a  $3 \times 3$   $\beta_L$  matrix would be used instead, with values fitted to solve all B meson anomalies.

<sup>3</sup> Having  $\beta_L^{s\tau}$  different from zero also opens new decay channels. These, however, are either suppressed by  $\beta_L^{s\tau}$  and powers of  $\lambda_{CKM}$ . In any case, this effect would decrease  $BR(LQ \rightarrow b\tau)$  and  $BR(LQ \rightarrow t\nu)$  by less than 3%.

mass, for two values of  $\beta_R$ . For large LQ masses, we confirm that with  $\beta_R = 0$  then  $\text{BR}(\text{LQ} \rightarrow b\tau) \approx \text{BR}(\text{LQ} \rightarrow t\nu) \approx \frac{1}{2}$ . However, for  $\beta_R = -1$ , as was chosen in [232], the additional coupling adds a new term to the total amplitude, leading to  $\text{BR}(\text{LQ} \rightarrow b\tau) \approx \frac{2}{3}$ . The increase in this branching ratio can thus weaken bounds from LQ searches targeting decays into  $t\nu$  final states, which motivates exploring the sensitivity in  $b\tau$  final states exclusively. Note that although a  $\text{BR}(\text{LQ} \rightarrow b\tau) \approx 1$  scenario is possible by having the LQ couple exclusively to right-handed currents (i.e.,  $g_U \rightarrow 0$ , but  $g_U \beta_R \neq 0$ ), it does not solve the observed anomalies in the  $R_{D^{(*)}}$  ratios. Therefore, although some LHC searches assume  $\text{BR}(\text{LQ} \rightarrow b\tau) = 1$ , we stress that in our studies we assume values of the model parameters and branching ratios that solve the  $R_{D^{(*)}}$  ratios.



**Figure 4.2:** Left: The  $\text{LQ} \rightarrow b\tau$  and  $\text{LQ} \rightarrow t\nu$  branching ratios for  $\beta_R = 0$  (solid lines) and  $\beta_R = -1$  (dashed lines). Right: Signal cross-section as a function of the LQ mass, for  $\sqrt{s} = 13$  TeV, with  $g_U = 1.8$ . We show single, pair, and non-resonant production, for  $\beta_R = -1, 0$  in solid and dashed lines, respectively.

To further understand the role of  $\beta_R$  at colliders, Figure 4.2 (bottom) shows the cross-section for single-LQ (sLQ), double-LQ (dLQ), and non-resonant (non-res) production, as a function of mass and for a fixed coupling  $g_U = 1.8$ , assuming  $p p$  collisions at  $\sqrt{s} = 13$  TeV. We note that this benchmark scenario with  $g_U = 1.8$  results in a  $\text{LQ} \rightarrow b\tau$  decay width that is  $< 5\%$  of the LQ mass, for mass values from 250 GeV to 2.5 TeV. In the Figure, we observe that, since dLQ production is mainly mediated by events from quantum chromodynamic processes, the choice of  $\beta_R$  does not affect the cross-section. However, for sLQ production, a non-zero value for  $\beta_R$  increases the cross-section by about a factor of 2 and by almost one order of magnitude in the case of non-res production. These results shown in Figure 4.2 are easily understood by considering the diagrams shown in Figure 4.1. The LQ mass value where the sLQ production cross-section exceeds the dLQ cross-section depends on the choice of  $g_U$ .

We also note that to solve the  $R_{D^{(*)}}$  anomaly, the authors of [267] point out that the wilson coefficient  $C_U \equiv g_U^2 v_{SM}^2 / (4 M_U^2)$  is constrained to a

specific range of values, and this range depends on the value of the  $\beta_R$  parameter. Therefore, the allowed values of the coupling  $g_U$  depend on  $M_U$  and  $\beta_R$ , and thus our studies are performed in this multi-dimensional phase space.

As noted in section ??, we study the role of a  $Z'$  boson in  $p p \rightarrow \tau\tau$  production. The presence of a  $Z'$  boson in LQ models has been justified in various papers, for example, in [243]. The argument is that minimal extensions of the SM which include a massive gauge  $U_1$  LQ, uses the gauge group  $SU(4) \times SU(3)' \times SU(2)_L \times U(1)_{T_R^3}$ . Such an extension implies the presence of an additional massive boson,  $Z'$ , and a color-octet vector,  $G'$ , arising from the spontaneous symmetry breaking into the SM <sup>4</sup>. The  $Z'$  in particular can play an important role in the projected LQ discovery reach, as it can participate in  $p p \rightarrow \tau\tau$  production by s-channel exchange, both resonantly and as a virtual mediator. To study the effect of a  $Z'$  on the  $p p \rightarrow \tau\tau$  production cross-sections and kinematics, we extend our benchmark Lagrangian in Eq. (4.1) with further non-renormalizable terms involving the  $Z'$ . Accordingly, we assume the  $Z'$  only couples to third-generation fermions. Our simplified model is thus extended by:

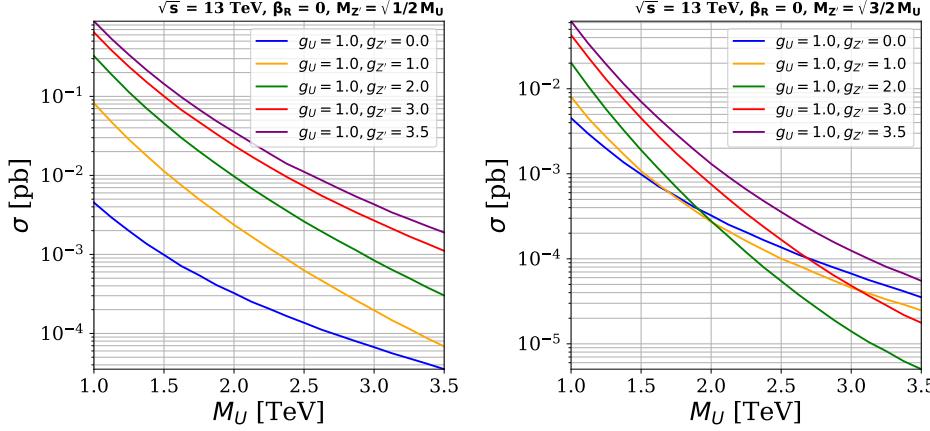
$$\begin{aligned} \mathcal{L}_{Z'} = & -\frac{1}{4} Z'_{\mu\nu} Z'^{\mu\nu} + \frac{1}{2} M_{Z'}^2 Z'_\mu Z'^\mu \\ & + \frac{9 Z'}{2\sqrt{6}} Z'^\mu (\zeta_q \bar{Q}_3 \gamma_\mu Q_3 + \zeta_t \bar{t}_R \gamma_\mu t_R \\ & + \zeta_b \bar{b}_R \gamma_\mu b_R - 3\zeta_\ell \bar{L}_3 \gamma_\mu L_3 - 3\zeta_\tau \bar{\tau}_R \gamma_\mu \tau_R) \end{aligned} \quad (4.2)$$

where the constants  $M_{Z'}$ ,  $g_{Z'}$ ,  $\zeta_q$ ,  $\zeta_t$ ,  $\zeta_b$ ,  $\zeta_\ell$ ,  $\zeta_\tau$ , are model dependent.

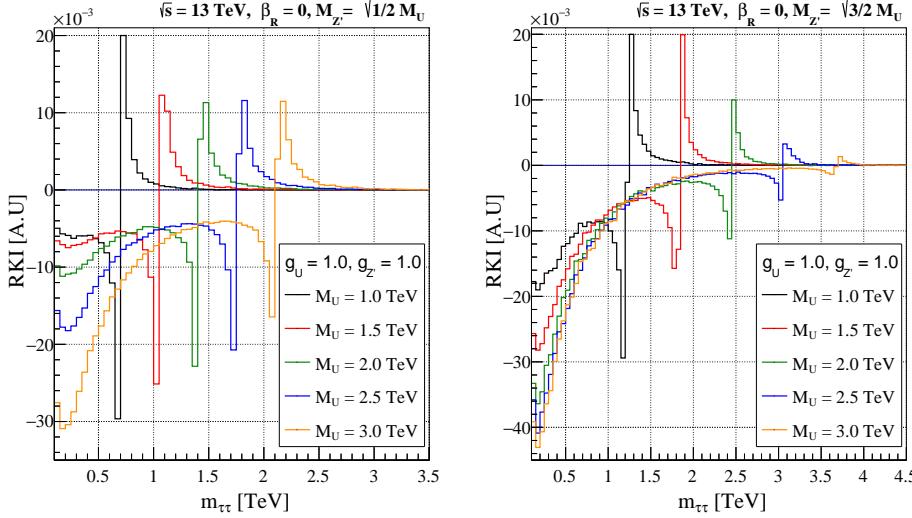
We study two extreme cases for the  $Z'$  mass, following [284], namely  $M_{Z'} = \sqrt{\frac{1}{2}} M_U < M_U$  and  $M_{Z'} = \sqrt{\frac{3}{2}} M_U > M_U$ . We also assume the LQ and  $Z'$  are uniquely coupled to left-handed currents, i.e.  $\zeta_q = \zeta_\ell = 1$  and  $\zeta_t = \zeta_b = \zeta_\tau = 0$ . With these definitions, Figure 4.3 shows the effect of the  $Z'$  on the  $\tau\tau$  production cross-section, considering  $g_U = 1$ ,  $\beta_R = 0$ , and different  $g_{Z'}$  couplings. On the left, the cross-sections corresponding to the cases where  $M_{Z'} = \sqrt{\frac{1}{2}} M_U$  are shown. As expected, the  $\tau\tau$  production cross-section for the inclusive case (i.e.,  $g_{Z'} \neq 0$ ) is larger than that for the LQ-only non-res process ( $g_{Z'} = 0$ , depicted in blue). This effect increases with  $g_{Z'}$  and, within the evaluated values, can exceed the LQ-only cross-section by up to two orders of magnitude. In contrast, a more intricate behaviour can be seen on the right of Figure 4.3, which corresponds to  $M_{Z'} = \sqrt{\frac{3}{2}} M_U$ . Here, for low values of  $M_U$ , a similar increase in the cross-section is observed. However, for higher values of  $M_U$ , the inclusive  $p p \rightarrow \tau\tau$  cross-section is smaller than the LQ-only  $\tau\tau$  cross-section. This behaviour suggests the presence of a dominant destructive interference at high masses, leaving its imprint on the results.

---

<sup>4</sup> Naively, the LQs are associated to the breaking of  $SU(4) \rightarrow SU(3)[4] \times U(1)_{B-L}$ , the  $G'$  arises from  $SU(3)[4] \times SU(3)' \rightarrow SU(3)_c$ , and the  $Z'$  comes from the breaking of  $U(1)_{B-L} \times U(1)_{T_R^3} \rightarrow U(1)_Y$ . Notice that the specific pattern of breaking, and the relations between the masses and couplings, are connected to the specific scalar potential used.



**Figure 4.3:**  $\tau\tau$  cross-section as a function of the LQ mass for different values of  $g_U$  and  $g_{Z'}$ . The estimates are performed at  $\sqrt{s} = 13$  TeV,  $\beta_R = 0$ ,  $M_{Z'} = \sqrt{1/2}M_U$  (left), and  $M_{Z'} = \sqrt{3/2}M_U$  (right).



**Figure 4.4:** The relative kinematic interference (RKI), as a function of the reconstructed mass of two taus, for different LQ masses. The studies are performed assuming  $\sqrt{s} = 13$  TeV,  $\beta_R = 0$ ,  $g_U = 1.0$ ,  $g_{Z'} = 1.0$ ,  $M_{Z'} = \sqrt{1/2}M_U$  (left), and  $M_{Z'} = \sqrt{3/2}M_U$  (right).

In order to further illustrate the effect, Figure 4.4 shows the relative kinematic interference (RKI) as a function of the reconstructed invariant mass  $m_{\tau\tau}$ , for  $g_{Z'} = 1$  and varying values of  $M_U$ . The RKI parameter is defined as

$$RKI(m_{\tau\tau}) = \frac{1}{\sigma_{LQ+Z'}} \left[ \frac{d\sigma_{LQ+Z'}}{dm_{\tau\tau}} - \left( \frac{d\sigma_{LQ}}{dm_{\tau\tau}} + \frac{d\sigma_{Z'}}{dm_{\tau\tau}} \right) \right], \quad (4.3)$$

where  $\sigma_X$  is the production cross-section arising due to contributions from  $X$  particles. For example,  $\sigma_{LQ+Z'}$  represents the inclusive cross-section where both virtual LQ and s-channel  $Z'$  exchange contribute. For both cases, we can observe the presence of deep valleys in the RKI curves when  $m_{\tau\tau} \rightarrow 0$ , indicating destructive interference between the LQ and the  $Z'$

contributions. This interference generates a suppression of the differential cross-section for lower values of  $m_{\tau\tau}$  and, therefore, in the integrated cross-section.

The observed interference effects are consistent with detailed studies on resonant and non-res  $p p \rightarrow t\bar{t}$  production, performed in reference [285].

## 4.2 SEARCH STRATEGY AND SIMULATION

---

Our proposed analysis strategy utilizes single-LQ (i.e.  $p p \rightarrow \tau LQ$ ), double-LQ (i.e.  $p p \rightarrow LQ LQ$ ), and non-resonant LQ production (i.e.  $p p \rightarrow \tau\tau$ ) as shown in Figure 4.1. At leading order in  $\alpha_s$ , since we focus on  $U_1 \rightarrow b\tau$  decays, the sLQ process results in the  $b\tau\tau$  mode, the dLQ process results in the  $bb\tau\tau$  mode, and the non-res process results in the  $\tau\tau$  mode. Therefore, in all cases we obtain two  $\tau$  leptons, with either 0, 1, or 2  $b$  jets. The  $\tau$  leptons decay to hadrons ( $\tau_h$ ) or semi-leptonically to electrons or muons ( $\tau_\ell$ ,  $\ell = e$  or  $\mu$ ). To this end, we study six final states:  $\tau_h\tau_{h/\ell}$ ,  $b\tau_h\tau_{h/\ell}$ , and  $bb\tau_h\tau_{h/\ell}$ , which can be naively associated to non-res, sLQ and dLQ production, respectively. Nevertheless, experimentally it is possible for  $b$  jets to not be properly identified or reconstructed, leading, for instance, to a fraction of dLQ signal events falling into the  $b\tau_h\tau_{h/\ell}$  and  $\tau_h\tau_{h/\ell}$  categories. Similarly, soft jets can fake  $b$  jets, such that non-res processes can contribute to the  $b\tau_h\tau_{h/\ell}$  and  $bb\tau_h\tau_{h/\ell}$  final states. This kind of signal loss and mixing is taken into account in our analysis<sup>5</sup>.

The contributions of signal and background events are estimated using Monte Carlo (MC) simulations. We implemented the  $U_1$  model from [243], adjusted to describe the lagrangian in Equations (4.1) and (4.2), using FeynRules (v2.3.43) [64, 116]. The branching ratios and cross-sections have been calculated using MadGraph5\_aMC (v3.1.0) [63, 118], the latter at leading order in  $\alpha_s$ . The corresponding samples are generated considering  $p p$  collisions at  $\sqrt{s} = 13$  TeV and  $\sqrt{s} = 13.6$  TeV. All samples are generated using the NNPDF3.0 NLO [71] set for parton distribution functions (PDFs) and using the full amplitude square SDE strategy for the phase-space optimization due to strong interference effects with the  $Z'$  boson. Parton level events are then interfaced with the PYTHIA (v8.2.44) [69] package to include parton fragmentation and hadronization processes, while DELPHES (v3.4.2) [65] is used to simulate detector effects, using the input card for the CMS detector geometric configurations, and for the performance of particle reconstruction and identification.

At parton level, jets and leptons are required to have a minimum transverse momentum ( $p_T$ ) of 20 GeV, while  $b$  jets are required to have a minimum  $p_T$  of 30 GeV. Additionally, we constrain the pseudorapidity ( $\eta$ )

---

<sup>5</sup> Note that further signal mixing can also occur at the event generation level by including terms at larger order in  $\alpha_s$ . For example, in the non-res diagram in Figure 4.1, one of the initial  $b$  could come from a  $g \rightarrow b\bar{b}$  splitting, leading to non resonant production of  $b\tau_h\tau_{h/\ell}$ . Simulating and studying the role of such NLO contributions is outside the scope of this work.

Variable	Threshold							
	$\tau_h \tau_h$	$b\tau_h \tau_h$	$bb\tau_h \tau_h$	$\tau_h \tau_\ell$	$b\tau_h \tau_\ell$	$bb\tau_h \tau_\ell$		
$N(b)$	= 0	= 1	$\geq 2$	= 0	= 1	$\geq 2$		
$p_T(b)$	-		$\geq 30 \text{ GeV}$	-		$\geq 30 \text{ GeV}$		
$ \eta(b) $	-		$\leq 2.4$	-		$\leq 2.4$		
$N(\ell)$		$= 0$			$= 1$			
$p_T(e)$		-			$\geq 35 \text{ GeV}$			
$p_T(\mu)$		-			$\geq 30 \text{ GeV}$			
$ \eta(\ell) $		-			$\leq 2.4$			
$N(\tau_h)$		$= 2$			$= 1$			
$p_T(\tau_h)$		$\geq 50 \text{ GeV}$						
$ \eta(\tau_h) $		$\leq 2.3$						
$\Delta R(p_i, p_j)$		$\geq 0.3$						

**Table 4.1:** Preliminary event selection criteria used to filter events before feeding them to the BDT algorithm. A  $\Delta R(p_i, p_j) > 0.3$  requirement is imposed between all pairs of reconstructed particle candidates  $p_i, p_j$ .

to  $|\eta| < 2.5$  for b jets and leptons, and  $|\eta| < 5.0$  for jets. The production cross-sections shown in the bottom panel of Figures 4.2 and 4.3 are obtained with the aforementioned selection criteria.

Table 4.1 shows the preliminary event selection criteria for each channel at analysis level. The channels are divided based on the multiplicity of b jets,  $N(b)$ , number of light leptons,  $N(\ell)$ , number of hadronic tau leptons,  $N(\tau_h)$ , and kinematic criteria based on  $\eta$ ,  $p_T$  and spatial separation of particles in the detector volume ( $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ ). The minimum  $p_T$  thresholds for leptons are chosen following references [254, 255, 259], based on experimental constraints associated to trigger performance. Following reference [114], we use a flat identification efficiency for b jets of 70% across the entire  $p_T$  spectrum with misidentification rate of 1%. These values correspond with the “medium working point” of the CMS algorithm to identify b jets, known as DeepCSV. We also explored the “Loose” (“Tight”) working point using an efficiency of 85% (45%) and mis-identification rate of 10% (0.1%). The “medium working point” was selected as it gives the best signal significance for the analysis.

For the performance of  $\tau_h$  identification in DELPHES, we consider the latest technique described in [78], which is based on a deep neural network (i.e. DeepTau) that combines variables related to isolation and  $\tau$ -lepton lifetime as input to identify different  $\tau_h$  decay modes. Following [78], we consider three possible DeepTau “working points”: (i) the “Medium” working point of the algorithm, which gives a 70%  $\tau_h$ -tagging efficiency and 0.5% light-quark and gluon jet mis-identification rate; (ii) the “Tight”

Sample	t <bar>t</bar>	single t	VV	V+jets	signals
$N_{\text{events}} \times 10^{-6}$	24.31	11.50	32.35	39.45	0.60

**Table 4.2:** The number of simulated events for the signal and background samples.

working point, which gives a 60%  $\tau_h$ -tagging efficiency and 0.2% light-quark and gluon jet mis-identification rate; and (iii) the “VTight” working point, which gives a 50%  $\tau_h$ -tagging efficiency and 0.1% light-quark and gluon jet mis-identification rate. Similar to the choice of b-tagging working point, the choice of  $\tau_h$ -tagging working point is determined through an optimization process which maximizes discovery reach. The “Medium” working point was ultimately shown to provide the best sensitivity and therefore chosen for this study. For muons (electrons), the assumed identification efficiency is 95% (85%), with a 0.3% (0.6%) mis-identification rate [82, 88, 89].

After applying the preliminary selection criteria, the primary sources of background are production of top quark pairs ( $t\bar{t}$ ), and single-top quark processes (single t), followed by production of vector bosons with associated jets from initial or final state radiation (V+jets), and pair production of vector bosons (VV). The number of simulated MC events used for each sample is shown in Table 4.2.

We use two different sets of signal samples. The first set includes various  $\{M_U, g_U\}$  scenarios, for two different values of  $\beta_R \in \{0, -1\}$ . We generate signal samples for  $M_U$  values between 250 GeV and 5000 GeV, in steps of 250 GeV. The considered  $g_U$  coupling values are between 0.25 and 3.5, in steps of 0.25. Although the signal cross-sections depend on both  $M_U$  and  $g_U$ , the efficiencies of our selections only depend on  $M_U$  (for all practical purposes) since the decay widths are relatively small compared to the mass of  $M_U$  ( $\frac{\Gamma_U}{M_U} < 5\%$ ), and thus more sensitive to experimental resolution. In total there are 280  $\{M_U, g_U, \beta_R\}$  scenarios simulated for this first set of signal samples, and for each of these scenarios two subsets of samples are generated, which are used separately for the training and testing of the machine learning algorithm. The second set of signal samples is used to evaluate interference effects between LQs and the  $Z'$  bosons in non-res production. Using benchmark values  $g_U = 1.8$  and  $\beta_R = 0$ , we consider various  $\{M_U, g_{Z'}\}$  scenarios for two different  $Z'$  mass hypotheses,  $(M_{Z'}/M_U)^2 \in \{\frac{1}{2}, \frac{3}{2}\}$ . The  $M_U$  values vary between 500 GeV and 5000 GeV, in steps of 250 GeV. The  $g_{Z'}$  coupling values are between 0.25 and 3.5, in steps of 0.25. Therefore, in total there are 280  $\{M_U, g_{Z'}, (M_{Z'}/M_U)^2\}$  scenarios simulated for this second set of signal samples, and for each of these scenarios a total of  $6.0 \times 10^5$  MC events are generated.

As noted previously, the simulated signal and background events are initially filtered using selections which are motivated by experimental constraints, such as the geometric constraints of the CMS detector, the typical kinematic thresholds for reconstruction of particle objects, and the available triggers. The remaining events after the preliminary event selection criteria are used to train and execute a BDT algorithm for each signal point in

the  $\{M_U, g_U\}$  space, in order to maximize the probability to detect signal amongst background events. The BDT algorithm is implemented using the scikit-learn [75] and xgboost (XGB) [134] python libraries. We use the XGBClassifier class from the xgboost library, a 10-fold cross validation using the scikit-learn method (GridCV<sup>6</sup>) for a grid in a hyperparameter space with 75, 125, 250, and 500 estimators, maximum depth in 3, 5, 7, 9, as well as learning rates of 0.01, 0.1, 1, and 10. For the cost function, we utilize the default mean square error (MSE). Additionally, we use the tree method based on the approximate greedy algorithm (histogram-optimized), referred to as hist, with a uniform sample method. These choices allow us to maximize the detection capability of the BDT algorithm by carefully tuning the hyperparameters, selecting an appropriate cost function, and utilizing an optimized tree construction method.

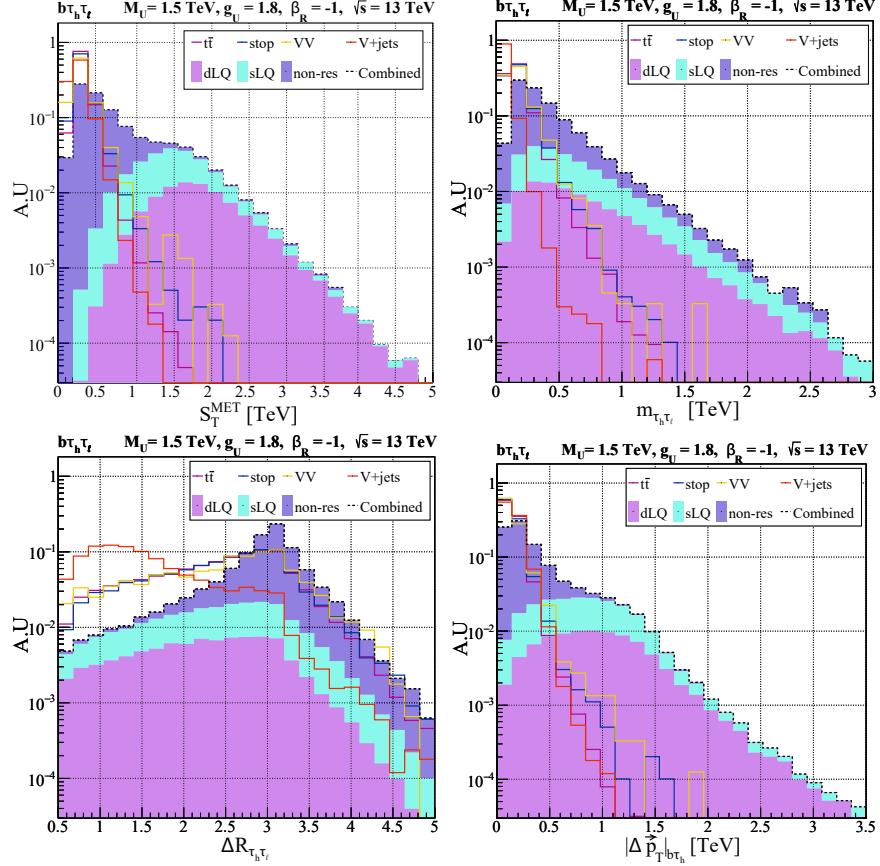
For each of the six analysis channels and  $\{M_U, g_U\}$  signal point, the binary XGB classifier was trained (tested) with 20% (80%) of the simulated events, for each signal and background MC sample. Over forty kinematic and topological variables were studied as input for the XGB. These included the momenta of b jets and  $\tau_{h,\ell}$  candidates; both invariant and transverse masses of pairs of  $\tau$  objects and of b $\tau$  combinations; angular differences between b jets, between  $\tau$  objects, and between the  $\tau_{h,\ell}$  and b jets; and additional variables derived from the missing momentum in the events. After studying correlations between variables and their impact on the performance of the BDT, we found that only eight variables were necessary and responsible for the majority of the sensitivity of the analysis. The variable that provides the best signal to background separation is the scalar sum of the  $p_T$  of the final state objects ( $\tau_h$ ,  $\tau_{h/\ell}$ , and b jets) and the missing transverse momentum, referred to as  $S_T^{\text{MET}}$ :

$$S_T^{\text{MET}} = |\vec{p}_T^{\text{miss}}| + \sum_{\tau_h, \tau_{h/\ell}, b} |\vec{p}_T| \quad (4.4)$$

The  $S_T^{\text{MET}}$  variable has been successfully used in LQ searches at the LHC, since it probes the mass scale of resonant particles involved in the production processes. Other relevant variables include the magnitude of the vectorial difference in  $p_T$  between the two lepton candidates ( $|\Delta\vec{p}_T|_{\tau_h \tau_{h/\ell}}$ ), the  $\Delta R_{\tau_h \tau_{h/\ell}}$  separation between them, the reconstructed dilepton mass  $m_{\tau_h \tau_{h/\ell}}$ , and the product of their electric charges ( $Q_{\tau_h} \times Q_{\tau_{h/\ell}}$ ). We also use the  $|\Delta\vec{p}_T|$  between the  $\tau_h$  candidate and  $\vec{p}_T^{\text{miss}}$ , and (if applicable) the  $|\Delta\vec{p}_T|$  between the  $\tau_h$  candidate and the leading b jet. For the final states including two  $\tau_h$  candidates, the one with the highest  $p_T$  is used.

---

<sup>6</sup> GridCV is a method that allows to find the best combination of hyperparameter values for the model, as this choice is crucial to achieve an optimal performance.



**Figure 4.5:**  $S_T^{\text{MET}}$ ,  $m_{\tau_h \tau_\ell}$ ,  $\Delta R_{\tau_h \tau_\ell}$ ,  $|\Delta \vec{p}_T|_{b\tau_h}$  signal and background distributions for the  $b\tau_h \tau_\ell$  channel. The signal distributions are generated for a benchmark sample with LQ mass of 1.5 TeV maximally coupled to right-handed currents. The combined distribution (shown as a stacked histogram) is the sum of the distributions, correctly weighted according to their respective cross-sections, assuming a coupling  $g_U = 1.8$ .

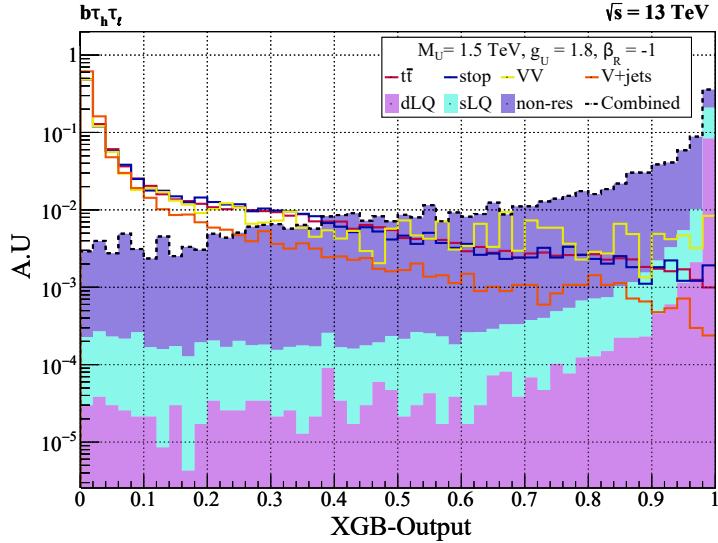
Figure 4.5 shows some relevant topological distributions, including  $S_T^{\text{MET}}$  on the top, for the  $b\tau_h \tau_\ell$  category. In the Figure we include all signal production modes to this channel, with each component weighted with respect to their total contribution to the combined signal. The combined signal distribution is normalised to unity. We also show all background processes contributing to this channel, each of them individually normalised to unity. We find that the combined signal is dominated by sLQ production for large values of  $S_T^{\text{MET}}$ , while non-res production dominates for small  $S_T^{\text{MET}}$ . Interestingly, the backgrounds also sit at low  $S_T^{\text{MET}}$  values, since  $S_T^{\text{MET}}$  is driven by the mass scale of the SM particles being produced, in this case top quarks and Z/W bosons. This suggest that the sLQ and dLQ signals can indeed be separated from the SM background. As expected, the  $S_T^{\text{MET}}$  sLQ and dLQ signal distributions have a mean near  $M_U$ , representative of resonant production, and a broad width as expected for large mass  $M_U$  hypotheses when information about the z-components of the momenta of objects is not utilised in the  $S_T^{\text{MET}}$  calculation.

Figure 4.5 (second from the top) shows the reconstructed mass of the ditau system, for the  $b\tau_h\tau_\ell$  search channel. Since the two  $\tau$  candidates in signal events arise from different production vertices (e.g., each  $\tau$  candidate in dLQ production comes from a different LQ decay chain), the ditau mass distribution for signal scales as  $m_{\tau_h\tau_\ell} \sim p_T(\tau_h) + p_T(\tau_\ell)$ , and thus has a tail which depends on  $M_U$  and sits above the expected SM spectrum. On the other hand, the SM  $m_{\tau_h\tau_\ell}$  distributions sit near  $m_{Z/W}$  since the  $\tau$  candidates in SM events arise from  $Z/W$  decays.

Figure 4.5 (third from the top) shows the  $\Delta R_{\tau_h\tau_\ell}$  distribution for the  $b\tau_h\tau_\ell$  channel. In the case of the  $p p \rightarrow \tau\tau$  non-res signal distribution, the two  $\tau$  leptons must be back-to-back to preserve conservation of momentum. Therefore, the visible  $\tau$  candidates,  $\tau_h$  and  $\tau_\ell$ , give rise to a  $\Delta R_{\tau_h\tau_\ell}$  distribution that peaks near  $\pi$  radians. In the case of sLQ production, although the LQ and associated  $\tau$  candidate must be back-to-back, the second  $\tau$  candidate arising directly from the decay of the LQ does not necessarily move along the direction of the LQ (since the LQ also decays to a b quark). As a result, the  $\Delta R_{\tau_h\tau_\ell}$  distribution for the sLQ signal process is smeared out, is broader, and has a mean below  $\pi$  radians. On the other hand, the  $\tau_h$  candidate in  $t\bar{t}$  events is often a jet being misidentified as a genuine  $\tau_h$ . When this occurs, the fake  $\tau_h$  candidate can arise from the same top quark decay chain as the  $\tau_\ell$  candidate, thus giving rise to small  $\Delta R_{\tau_h\tau_\ell}$  values. This difference in the signal and background distributions provides a nice way for the ML algorithm to help decipher signal and background processes.

As noted above, the  $|\Delta\vec{p}_T|$  distribution between the visible  $\tau$  candidates and the b-quark jets is an important variable to help the BDT distinguish between signal and background processes. The discriminating power can be seen in Figure 4.5 (bottom), which shows the  $|\Delta\vec{p}_T|$  between the  $\tau_h$  and b-jet candidate of the  $b\tau_h\tau_\ell$  channel. In the case of dLQ production, the b quarks and  $\tau$  leptons from the  $LQ \rightarrow b\tau$  decay acquire transverse momentum of  $p_T \sim M_U/2$ . However, when the  $\tau$  lepton decays hadronically (i.e.  $\tau \rightarrow \tau_h\nu$ ), a large fraction of the momentum is lost to the neutrino. Therefore, the  $|\Delta\vec{p}_T|_{b\tau_h}$  distribution for the dLQ (and sLQ) process peaks below  $M_U/2$ . On the other hand, for a background process such as V+jets, the b jet arises due to initial state radiation, and thus must balance the momentum of the associated vector boson (i.e.  $p_T(b) \sim p_T(V) \sim m_V$ ). Since the visible  $\tau$  candidate is typically produced from the V boson decay chain, its momentum (on average) is approximately  $p_T(\tau_h) \sim p_T(V)/4 \sim m_V/4$ . Therefore, to first order, the  $|\Delta\vec{p}_T|$  distribution for the V+jets background is expected to peak below the  $m_V$  mass.

Lets us turn to the results of the  $b\tau_h\tau_\ell$  BDT classifier, which is shown in Figure 4.6 for the different signal production modes and backgrounds. Similar to Figure 4.5, the distribution for each individual signal production mode is weighted with respect to their total contribution to the combined signal. The background distributions and combined signal distribution are normalized to an area under the curve of unity. Figure 4.6 shows the XGB



**Figure 4.6:** Postfit XGB-output normalised distribution in the  $b\tau_h\tau_\ell$  channel, for LQ mass of 1.5 TeV, constant coupling  $g_U = 1.8$ , and maximally coupled to right-handed currents.

distributions for a signal benchmark point with  $M_U = 1.5$  TeV,  $g_U = 1.8$ , and  $\beta_R = -1$ . The XGB output is a value between 0 and 1, which quantifies the likelihood that an event is either signal-like (XGB output near 1) or background-like (XGB output near 0). We see that the presence of the sLQ and dLQ production modes is observed as an enhancement near a XGB output of unity, while the backgrounds dominate over the low end of the XGB output spectrum, especially near zero. In fact, over eighty percent of the sLQ and dLQ distributions reside in the last two bins, XGB output greater than 0.96, while more than sixty percent of the backgrounds fall in the first two bins, XGB output less than 0.04. It is also interesting to note that in comparison to the sLQ and dLQ distributions in Figure 4.6, non-res is broader and not as narrowly peaked near XGB output of 1, which is expected due to the differences in kinematics described above. Overall, if we focus on the last bin in this distribution, we find approximately 0.2% of the background, in contrast to 22% of the non-res, 78% of the sLQ, and 91% of the dLQ signal distributions. These numbers highlight the effectiveness of the XGB output in reducing the background in the region where the signal is expected.

The output signal and background distributions of the XGB classifier, normalised to their cross section times pre-selection efficiency times luminosity, are used to perform a profile binned likelihood statistical test in order to determine the expected signal significance. The estimation is performed using the RooFit [286] package, following the same methodology as in Refs. [191–196, 287–296]. The value of the significance ( $Z_{sig}$ ) is measured using the probability to obtain the same outcome from the test statistic in the background-only hypothesis, with respect to the signal plus

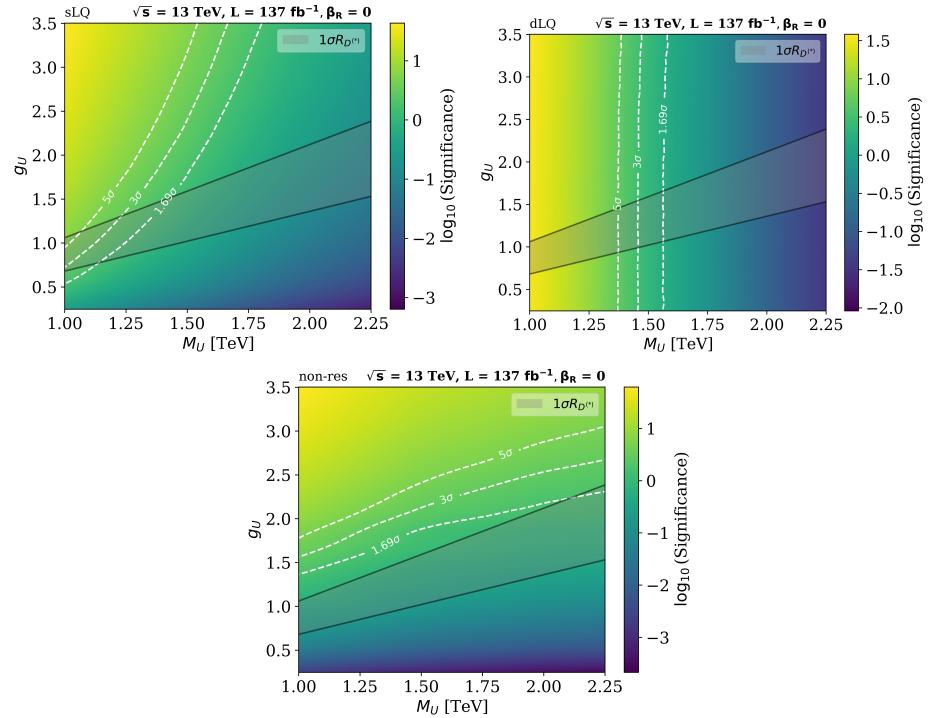
background hypothesis. This allows for the determination of the local p-value and thus the calculation of the signal significance, which corresponds to the point where the integral of a Gaussian distribution between  $Z_{\text{sig}}$  and  $\infty$  results in a value equal to the local p-value.

Systematic uncertainties are incorporated as nuisance parameters, considering log-priors for normalization and Gaussian priors for shape uncertainties. Our consideration of systematic uncertainties includes both experimental and theoretical effects, focusing on the dominant sources of uncertainty. Following [74], we consider a 3% systematic uncertainty on the measurement of the integrated luminosity at the LHC. A 5% uncertainty arises due to the choice of the parton distribution function used for the MC production, following the PDF4LHC prescription [92]. The chosen PDF set only has an effect on the overall expected signal and background yields, but the effect on the shape of the XGB output distribution is negligible. Reference [78] reports a systematic uncertainty of 2-5%, depending on the  $p_T$  and  $\eta$  of the  $\tau_h$  candidate. Therefore, we utilize a conservative 5% uncertainty per  $\tau_h$  candidate, independent of  $p_T$  and  $\eta$ , which is correlated between signal and background processes with genuine  $\tau_h$  candidates, and correlated across XGB bins for each process. We assumed a 5%  $\tau_h$  energy scale uncertainty, independent of  $p_T$  and  $\eta$ , following the CMS measurements described in [78]. Finally, we assume a conservative 3% uncertainty per b-jet candidate, following reference [188], and an additional 10% uncertainty in all the background predictions to account for possible mismodeling by the simulated samples. The uncertainties on the background estimates are typically derived from collision data using dedicated control samples that have negligible signal contamination and are enriched with events from the specific targeted background. The systematic uncertainties on the background estimates are treated as uncorrelated between background processes.

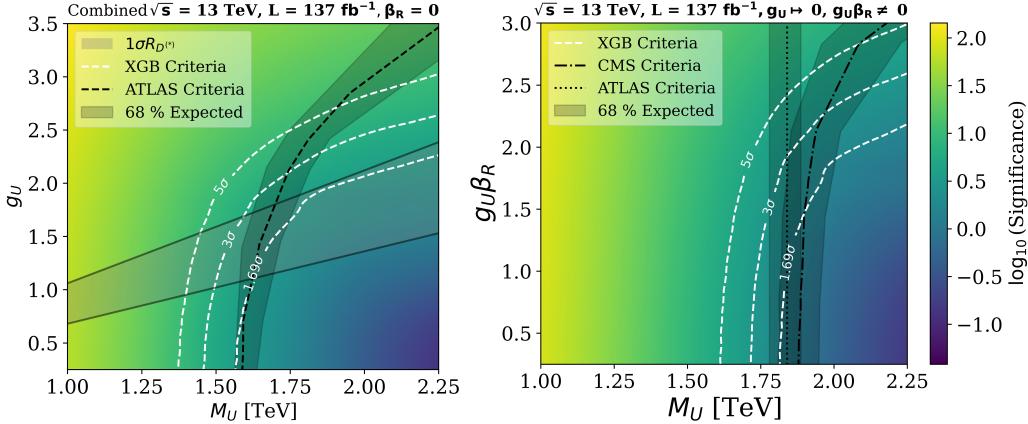
### 4.3 RESULTS

---

The expected signal significance for sLQ, dLQ and non-res production, and their combination, is presented in Figure 4.7. Here, the significance is shown as a heat map in a two dimensional plane of  $g_U$  versus  $M_U$ , considering exclusive couplings to left-handed currents, *i.e.*  $\text{BR}(\text{LQ} \rightarrow b\tau) = \frac{1}{2}$ . The dashed lines show the contours of constant signal significance. The  $1.69\sigma$  contour represents exclusion at 95% confidence level, and the  $3-5\sigma$  contours represent potential discovery. The grey band defines the set of  $\{M_U, g_U\}$  values that can explain the B-meson anomalies,  $C_U \sim 0.01$  for this scenario. The estimates are performed under the conditions for the second run, RUN-II, of the LHC ( $\sqrt{s} = 13 \text{ TeV}$  and  $L = 137 \text{ fb}^{-1}$ ). We find that the dLQ interpretation plot (Figure 4.7 second from the top) does not depend on  $g_U$ , which is expected due to dLQ production arising exclusively from interactions with gluons. For this reason, the dLQ production process provides the best mode for discovery when  $g_U$  is small. On the other



**Figure 4.7:** Signal significance for different coupling scenarios and LQ masses, without right-handed currents, using the combination of all search channels. The results pertaining to sLQ, dLQ and non-res production are displayed respectively from the top. These results are for  $\sqrt{s} = 13 \text{ TeV}$  and  $137 \text{ fb}^{-1}$ .

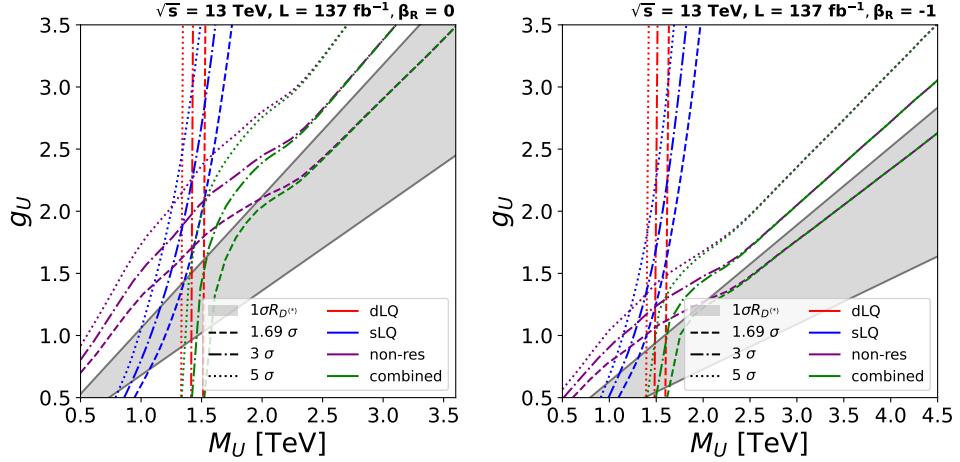


**Figure 4.8:** The top (bottom) panel shows signal significance comparison with ATLAS [262] (CMS and ATLAS [256, 263]) background only hypothesis, for the combination of all channels, with uniquely coupling to left-handed (right-handed) currents. The estimates are performed at  $\sqrt{s} = 13 \text{ TeV}$  and  $137 \text{ fb}^{-1}$ .

hand, the non-res channel is more sensitive to changes in the coupling parameter  $g_U$ , as its production cross-section depends on  $g_U^4$ . Therefore, the non-res production process provides the best mode for discovery when  $g_U$  is large. These results confirm the expectations from previous analyses (see for instance [242]), in the sense that the dLQ and non-res processes complement each other nicely at low and high  $g_U$  scenarios. The sLQ channel combines features from both the dLQ and non-res channels, in principle making it an interesting option to explore different scenarios and gain a better understanding of LQ properties, but the evolution of the signal significance in the full phase space is more complicated as it involves resonant LQ production with a cross-section that depends non-trivially on  $M_U$ ,  $g_U$ , and the LQ coupling to gluons. However, Figure 4.7 shows that the sLQ production process can provide complementary and competitive sensitivity to the non-res and dLQ processes, in certain parts of the phase space.

The top panel of Figure 4.8 presents the sensitivity of all signal production processes combined, and compares our expected exclusion region with the latest one from the ATLAS Collaboration [262]. The comparison suggests that our proposed analysis strategy provides better sensitivity than current methods being carried out at ATLAS, especially at large values of  $g_U$ . In particular, we find that with the pp data already available from RUN-II, our expected exclusion curves begin to probe solutions to the B-anomalies for LQ masses up to 2.25 TeV.

Figure 4.8 shows the expected signal significance considering  $\text{BR}(\text{LQ} \rightarrow b\tau) = 1$ , in order to compare our analysis with the corresponding results from the CMS [256] and ATLAS [263] Collaborations. Let us emphasize again that  $\text{BR}(\text{LQ} \rightarrow b\tau)$  depends on  $\beta_R$ , as illustrated on the top panel of Figure 4.2. Thus, although the  $\text{BR}(\text{LQ} \rightarrow b\tau) = 1$  scenario is a possible physical case, it does not solve the observed anomalies in the  $R_{D^{(*)}}$  ratios,



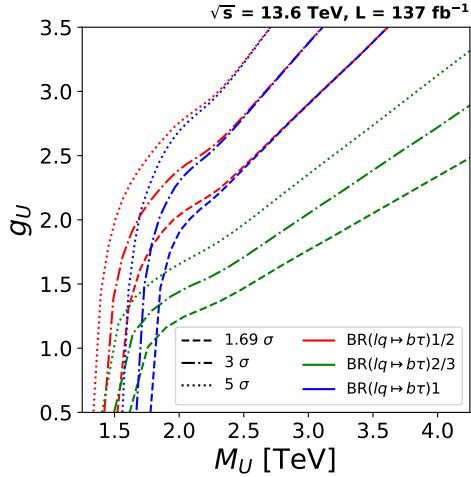
**Figure 4.9:** Signal significance for different coupling scenarios and LQ masses for all channels. This plot summarizes our results with  $\beta_R = 0$  (without right-handed currents) and  $\beta_R = -1$  (maximally coupled to right-handed currents). The estimates are performed at  $\sqrt{s} = 13 \text{ TeV}$  and  $137 \text{ fb}^{-1}$ .

as it corresponds to the case where LQs couple exclusively to right-handed currents.

With this in mind, the scenario studied by CMS in [256] considers couplings only to left-handed currents, setting artificially the condition  $\text{BR}(\text{LQ} \rightarrow b\tau) = 1$ . In order to compare, we scale the efficiency  $\times$  acceptance of our selection criteria for  $\beta_R = 0$ , by a factor of 2.0 for sLQ and 4.0 for dLQ. According to Figure 4.8, the ML approach that we have followed again suggests an optimisation of the signal and background separation, having the potential of improving the regions of exclusion ( $1.69\sigma$ ) with respect to that of CMS. In the bottom panel of the Figure we have also included a similar exclusion by ATLAS [263]. However, since ATLAS only considers dLQ production in the analysis, the results are not entirely comparable, so are included only as a reference.

We now turn to the role of  $\beta_R$ , and our capacity of probing the regions solving the B-meson anomalies. Figure 4.9 shows the maximum significant contours, under LHC RUN-II conditions, for the different LQ production mechanisms and their combination, considering scenarios with only left-handed currents ( $\beta_R = 0$ , top) and with maximal right-handed currents ( $\beta_R = -1$ , bottom). We find a noticeable improvement in signal significance in all channels when taking  $\beta_R = -1$ , as is expected from the increase in  $\text{BR}(\text{LQ} \rightarrow b\tau)$  branching ratio and production cross-sections (see Figure 4.2). However, the region solving the B-meson anomalies also changes, preferring lower values of  $g_U$ , such that in both cases we find ourselves just starting to probe this region at large  $M_U$ .

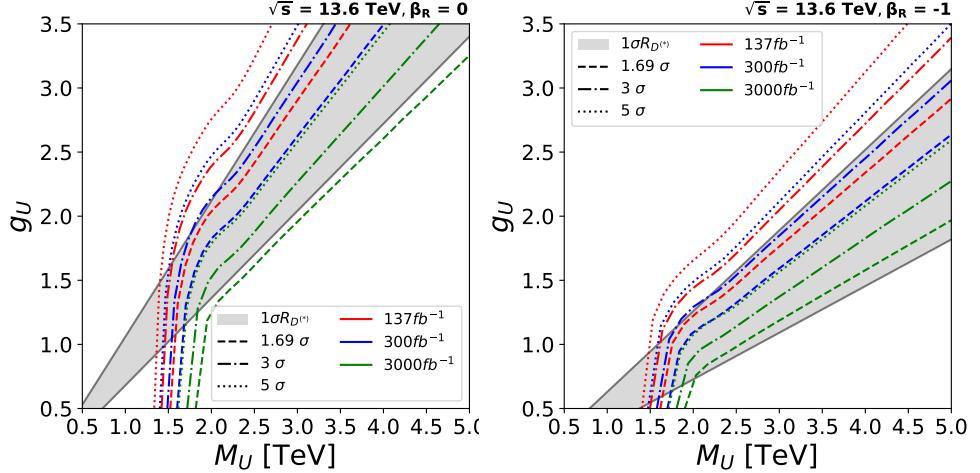
The combined significance contours for the different BR scenarios that have been considered is presented in Figure 4.10. These contours illustrate the regions of exclusion for the three cases of interest, namely exclusive left-handed currents ( $\text{BR}(\text{LQ} \rightarrow b\tau) = \frac{1}{2}$ ,  $\beta_R = 0$ ), maximal left and right



**Figure 4.10:** Signal significance for different coupling scenarios and LQ masses, considering the case without coupling to right-handed currents  $\text{BR}(\text{LQ} \rightarrow b\tau) = \frac{1}{2}$ , the case maximally coupled to right- and left-handed currents  $\text{BR}(\text{LQ} \rightarrow b\tau) = \frac{2}{3}$ , and the case uniquely coupled to right-handed currents  $\text{BR}(\text{LQ} \rightarrow b\tau) = 1$ . The estimates are performed at  $\sqrt{s} = 13 \text{ TeV}$  and  $137 \text{ fb}^{-1}$ .

couplings ( $\text{BR}(\text{LQ} \rightarrow b\tau) = \frac{2}{3}$ ,  $\beta_R = -1$ ), and exclusive right-handed currents ( $\text{BR}(\text{LQ} \rightarrow b\tau) = 1$ ,  $g_U \rightarrow 0$ ,  $g_U \beta_R = 1$ ). For small  $g_U$ , we find that the exclusive right-handed scenario is most sensitive, while the exclusive left-handed case is the worst. The reason for this is that this region is excluded principally by dLQ production, such that having the largest branching ratio is crucial in order to have a large number of events. For larger couplings, both exclusive scenarios end up having similar exclusion regions, with the  $\beta_R = -1$  case being significantly more sensitive. The reason in this case is that the exclusion is dominated by non-res, which has a much larger production cross-section if both currents are turned on.

In order to finalise our analysis of the LQ-only model, we show in Figure 4.11 the expected combined significance in the relatively near future. For this, considering  $\sqrt{s} = 13.6 \text{ TeV}$ , we show contours for the sensitivity corresponding to integrated luminosities of  $137 \text{ fb}^{-1}$ ,  $300 \text{ fb}^{-1}$ , and  $3000 \text{ fb}^{-1}$ , for scenarios with only left-handed currents (top) and with maximal coupling to right-handed currents (bottom). Note that for  $\beta_R = 0$  ( $\beta_R = -1$ ), couplings  $g_U$  close to 3.18 (1.85) and  $M_U = 5.0 \text{ TeV}$  can be excluded with 1.69 $\sigma$  significance for the high luminosity LHC era, allowing us to probe the practically the entirety of the B-meson anomaly favored region. Note that the background yields for the high luminosity LHC might be larger due to pileup effects. Nevertheless, as it was mentioned in Section 4.2, we have included a conservative 10% systematic uncertainty associated with possible fluctuations on the background estimations. Although effects from larger pileup might be significant, they can be mitigated by improvements in the algorithms for particle reconstruction and identification, and also on the data-analysis techniques.

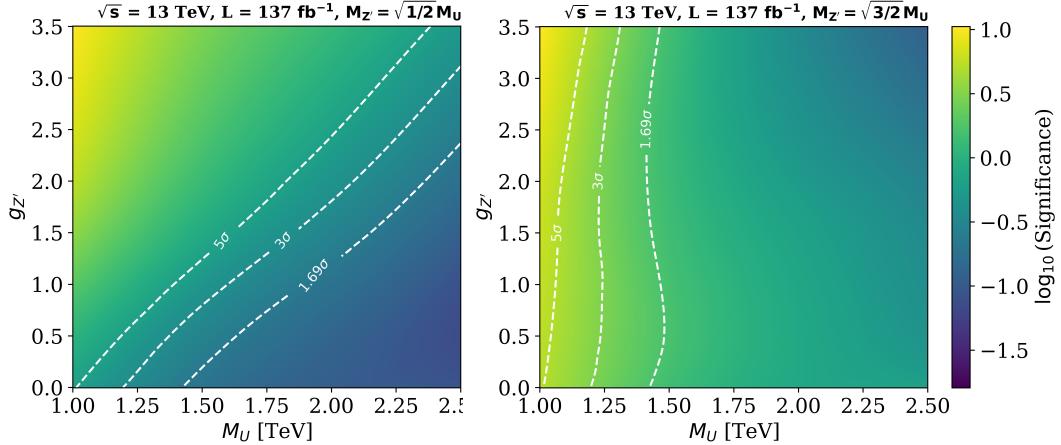


**Figure 4.11:** Projected signal significance for different coupling scenarios and LQ masses maximally coupled to right-handed currents. The estimates are performed at  $\sqrt{s} = 13.6 \text{ TeV}$ ,  $137 \text{ fb}^{-1}$ ,  $300 \text{ fb}^{-1}$  and  $3000 \text{ fb}^{-1}$ .

As commented on the Introduction, non-res production can be significantly affected by the presence of a companion  $Z'$ , which provides additional s-channel diagrams that add to the total cross-section and can interfere destructively with the LQ t-channel process (see Figures 4.3 and 4.4). From our previous results, we see that non-res always is of high importance in determining the exclusion region, particularly at large  $M_U$  and  $g_U$ , meaning it is crucial to understand how this role is affected in front of a  $Z'$  with similar mass.

The change in sensitivity on the non-res signal significance due this interference effect with the  $Z'$  boson is shown in Figure 4.12. We consider two opposite cases for the  $Z'$  mass:  $M_{Z'}^2 = M_U^2/2$  (top) and  $M_{Z'}^2 = 3M_U^2/2$  (bottom). Our results are shown on the  $g_{Z'} - M_U$  plane, for a fixed  $g_U = 1.8$  and  $\beta_R = 0$ . For the  $M_{Z'}^2 = M_U^2/2$  scenario, there is an overall increase in the total cross-section, with a larger  $g_{Z'}$  implying a larger sensitivity. This means that our ability to probe smaller values of  $g_U$  could be enhanced, as a given observation would be reproduced with both a specific  $g_U$  and vanishing  $g_{Z'}$ , or a smaller  $g_U$  with large  $g_{Z'}$ . Thus, for a large enough  $g_{Z'}$ , it could be possible to enhance non-res to the point that the entire region favoured by B-anomalies could be ruled out. In contrast, for  $M_{Z'}^2 = 3M_U^2/2$  the cross-section is strongly affected by the large destructive interference, such that a larger  $g_{Z'}$  does not necessarily imply an increase in sensitivity. In fact, as can be seen in the bottom panel, for large  $M_U$  the significance is reduced as  $g_{Z'}$  increases, leading to the opposite conclusion than above, namely, that a large  $g_{Z'}$  could reduce the effectiveness of non-res.

The impact of the above can be seen in Figure 4.13, which shows our previous sensitivity curves on the  $M_U - g_U$  plane, but this time with a  $Z'$  contribution to non-res. We use the same values of  $M_{Z'}$  as before, but fix  $g_{Z'} = 3.5$ . For smaller  $M_{Z'}$  (top), the non-res contribution is enhanced so much, that both sLQ and dLQ play no role whatsoever in determining the



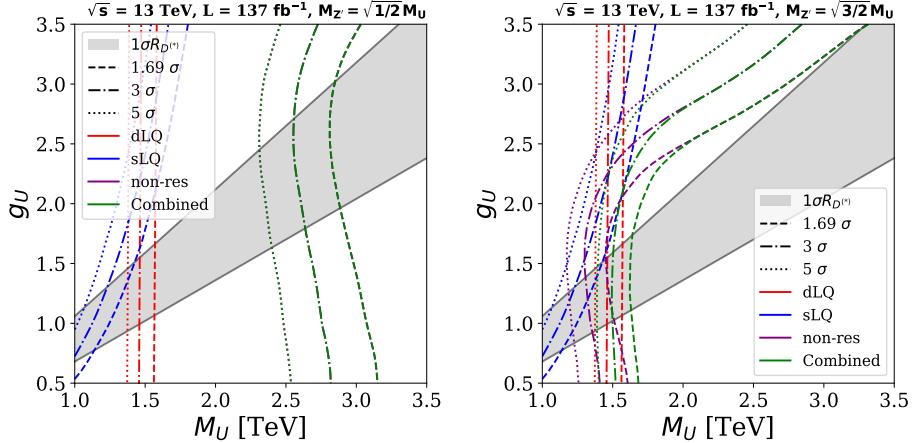
**Figure 4.12:** Change on the non-res signal significance for different  $Z'$  coupling scenarios and LQ masses. The estimates are performed at  $\sqrt{s} = 13.0$  TeV,  $\beta_R = 0$ ,  $g_U = 1.8$ ,  $M_{Z'} = \sqrt{1/2} M_U$  (top), and  $M_{Z'} = \sqrt{3/2} M_U$  (bottom).

exclusion region. We find that, for small  $g_U$ , the sensitivity is dominated by  $Z'$  production such that, since  $M_U$  is related to  $M_{Z'}$ , LQ masses up to  $\sim 3$  TeV are excluded. This bound is slightly relaxed for larger values of  $g_U$ , which is attributed to destructive interference effects due to an increased LQ contribution.

The bottom panel of Figure 4.13 shows that case where  $M_{Z'}$  is larger than  $M_U$ . As expected from our previous discussion, the behaviour and impact of non-res is modified. For small  $g_U$ , we again have the pure  $Z'$  production dominating the non-res cross-section, leading to a null sensitivity on  $g_U$ , similar to what happens in dLQ. In contrast, for very large  $g_U$ , we find that the pure LQ non-res production is the one that dominates, and we recover sensitivity regions with a slope similar to those shown in Figures 4.7-4.11, shifted towards larger values of  $g_U$ . For intermediate values of this coupling, the destructive interference have an important effect again, twisting the exclusion region slightly towards the left. Still, even in this case, we find that sLQ plays a marginal role in defining the combined exclusion region, and that the final result again depends primarily on dLQ and non-res production.

#### 4.4 DISCUSSION AND CONCLUSIONS

Experimental searches for LQs with preferential couplings to third generation fermions are currently of great interest due to their potential to explain observed tensions in the  $R(D)$  and  $R(D^*)$  decay ratios of B mesons with respect to the SM predictions. Although the LHC has a broad physics program on searches for LQs, it is very important to consider the impact of each search within wide range of different theoretical assumptions within a specific model. In addition, in order to improve the sensitivity to detect possible signs of physics beyond the SM, it is also important to strongly



**Figure 4.13:** Signal significance for different coupling scenarios and LQ masses, for all channels, with an additional  $Z'$  contribution to non-res production. We set  $\beta_R = 0$  and  $g_{Z'} = 3.5$ , taking  $M_{Z'}^2$  equal to  $M_U^2/2$  ( $3M_U^2/2$ ) on the top (bottom) panel.

consider new computational techniques based on machine learning (ML). Therefore, we have studied the production of  $U_1$  LQs with preferential couplings to third generation fermions, considering different couplings, masses and chiral currents. These studies have been performed considering  $p p$  collisions at  $\sqrt{s} = 13$  TeV and 13.6 TeV and different luminosity scenarios, including projections for the high luminosity LHC. A ML algorithm based on boosted decision trees is used to maximize the signal significance. The signal to background discrimination output of the algorithm is taken as input to perform a profile binned-likelihood test statistic to extract the expected signal significance.

The expected signal significance for sLQ, dLQ and non-res production, and their combination, is presented as contours on a two dimensional plane of  $g_U$  versus  $M_U$ . We present results for the case of exclusive couplings to left-handed, mixed, and exclusive right-handed currents. For the first two, the region of the phase space that could explain the B meson anomalies is also presented. We confirm the findings of previous works that the largest production cross-section and best overall significance comes from the combination of dLQ and non-res production channels. We also find that the sensitivity to probe the parameter space of the model is highly dependent on the chirality of the couplings. Nevertheless, the region solving the B-meson anomalies also changes with each choice, such that in all evaluated cases we find ourselves just starting to probe this region at large  $M_U$ .

Our studies compare our exclusion regions with respect to the latest reported results from the ATLAS and CMS Collaborations. The comparison suggests that our ML approach has a better sensitivity than the standard cut-based analyses, especially at large values of  $g_U$ . In addition, our projections for the HL-LHC cover the whole region solving the B-anomalies, for masses up to 5.00 TeV.

Finally, we consider the effects of a companion  $Z'$  boson on non-res production. We find that such a contribution can have a considerable impact on the LQ sensitivity regions, depending on the specific masses and couplings. In spite of this, we still consider non-res production as an essential channel for probing LQs in the future.



# A

---

## SUPERVISED LEARNING

---

Supervised learning represents one of the most fundamental paradigms in machine learning and artificial intelligence. In this framework, algorithms learn to make predictions by examining labeled training data—examples where both the input features and the desired output are known. This approach mirrors, in many ways, how humans learn: by observing patterns, receiving feedback, and gradually improving performance through experience.

The power of supervised learning lies in its ability to discover complex relationships between input variables and target outcomes that may not be immediately apparent to human observers. In the context of high-energy physics, this capability has proven invaluable for tasks ranging from particle identification and signal-background discrimination to the reconstruction of physical parameters from detector measurements [[guest2018machine](#)]. The sophisticated detectors at the Large Hadron Collider (LHC), for instance, generate petabytes of data annually, making supervised learning techniques essential tools for extracting meaningful physics insights from this vast information landscape.

Historically, the foundations of supervised learning can be traced back to the early developments in statistics and pattern recognition in the mid-20th century. The perceptron algorithm, introduced by Frank Rosenblatt in 1957 [[rosenblatt1958perceptron](#)], marked one of the first attempts to create machines capable of learning from examples. However, it was not until the advent of more sophisticated algorithms and the exponential growth in computational power that supervised learning truly came into its own as a practical tool for scientific discovery.

In this chapter, we provide a comprehensive introduction to supervised learning, emphasizing both the theoretical foundations and practical applications relevant to particle physics research. We begin by formalizing the supervised learning problem within a rigorous mathematical framework, drawing parallels to the optimization principles familiar from field theory. We then explore the two primary categories of supervised learning: classification, where the goal is to assign discrete labels to inputs, and regression, where we seek to predict continuous-valued outputs.

## A.1 MATHEMATICAL FRAMEWORK OF SUPERVISED LEARNING

---

To establish a solid foundation, we begin by formalizing the supervised learning problem in mathematical terms. Let  $\mathcal{X}$  denote the input space (also called the feature space) and  $\mathcal{Y}$  the output space (or target space). In supervised learning, we assume the existence of an unknown joint probability distribution  $P(X, Y)$  over  $\mathcal{X} \times \mathcal{Y}$ , from which our training data is drawn.

### A.1.1 THE LEARNING PROBLEM

Given a training dataset  $\mathcal{D} = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$  consisting of  $n$  independent and identically distributed (i.i.d.) samples from  $P(X, Y)$ , the goal of supervised learning is to find a function  $f : \mathcal{X} \rightarrow \mathcal{Y}$  that can accurately predict the output  $y$  for any given input  $x$ .

More precisely, we seek to minimize the *expected risk* (also known as the *generalization error*):

$$R(f) = \mathbb{E}_{(X, Y) \sim P} [\ell(f(X), Y)] = \int_{\mathcal{X} \times \mathcal{Y}} \ell(f(x), y) dP(x, y), \quad (\text{A.1})$$

where  $\ell : \mathcal{Y} \times \mathcal{Y} \rightarrow \mathbb{R}_+$  is a *loss function* that quantifies the cost of making prediction  $f(x)$  when the true output is  $y$ .

Since the true distribution  $P(X, Y)$  is unknown, we cannot directly minimize the expected risk. Instead, we resort to *empirical risk minimization* (ERM), where we minimize the *empirical risk*:

$$\hat{R}_n(f) = \frac{1}{n} \sum_{i=1}^n \ell(f(x_i), y_i). \quad (\text{A.2})$$

This approach is analogous to the principle of least action in classical mechanics, where the true dynamics emerge from minimizing an action functional over all possible paths. Similarly, in supervised learning, the optimal predictor emerges from minimizing the empirical risk over all functions in our hypothesis class.

### A.1.2 HYPOTHESIS CLASSES AND MODEL COMPLEXITY

In practice, we cannot search over all possible functions  $f : \mathcal{X} \rightarrow \mathcal{Y}$ , as this would be computationally intractable and would likely lead to overfitting. Instead, we restrict our search to a specific *hypothesis class*  $\mathcal{H} = \{h_\theta : \theta \in \Theta\}$ , where  $\Theta$  is the parameter space.

The choice of hypothesis class represents a fundamental trade-off between expressiveness and complexity. A more expressive class can potentially capture more complex patterns in the data, but it also has a higher capacity to memorize noise, leading to poor generalization. This trade-off is formalized in the *bias-variance decomposition*.

For a given input  $x$ , if we denote the true relationship as  $y = f^*(x) + \epsilon$ , where  $\epsilon$  is irreducible noise with  $\mathbb{E}[\epsilon] = 0$  and  $\text{Var}(\epsilon) = \sigma^2$ , then the expected squared error of any predictor  $\hat{f}$  can be decomposed as:

$$\mathbb{E}[(y - \hat{f}(x))^2] = \underbrace{(\hat{f}(x) - \mathbb{E}[\hat{f}(x)])^2}_{\text{Bias}^2} + \underbrace{\text{Var}(\hat{f}(x))}_{\text{Variance}} + \underbrace{\sigma^2}_{\text{Irreducible Error}}. \quad (\text{A.3})$$

The **bias** measures how far the expected prediction is from the true value, capturing systematic errors due to oversimplified assumptions. The **variance** measures how much the predictions vary across different training sets, reflecting the model's sensitivity to fluctuations in the data.

*The bias-variance trade-off is reminiscent of the uncertainty principle in quantum mechanics: reducing bias (systematic error) typically increases variance (statistical fluctuations), and vice versa.*

### A.1.3 REGULARIZATION AND OCCAM'S RAZOR

To control model complexity and achieve good generalization, we often employ *regularization* techniques. Instead of minimizing only the empirical risk, we minimize a regularized objective:

$$\hat{R}_{\text{reg}}(f) = \hat{R}_n(f) + \lambda \Omega(f), \quad (\text{A.4})$$

where  $\Omega(f)$  is a *regularization term* that penalizes complex functions, and  $\lambda > 0$  is the *regularization parameter* that controls the strength of the penalty.

This approach embodies a mathematical formalization of Occam's razor: among competing hypotheses that explain the data equally well, the simplest one is preferred. Common regularization terms include:

- **L<sub>1</sub> regularization (Lasso):**  $\Omega(f) = \|\theta\|_1 = \sum_i |\theta_i|$ , which promotes sparsity by driving some parameters to exactly zero.
- **L<sub>2</sub> regularization (Ridge):**  $\Omega(f) = \|\theta\|_2^2 = \sum_i \theta_i^2$ , which shrinks parameters toward zero without necessarily making them exactly zero.
- **Elastic Net:** A combination of L<sub>1</sub> and L<sub>2</sub> regularization.

## A.2 THE CLASSIFICATION PROBLEM

---

Classification represents one of the two fundamental categories of supervised learning, where the goal is to assign discrete labels to input examples. In the context of particle physics, classification problems are ubiquitous: distinguishing signal from background events, identifying particle types based on detector signatures, or determining the decay modes of reconstructed particles.

### A.2.1 BINARY CLASSIFICATION

We begin with the simplest case: binary classification, where  $\mathcal{Y} = \{0, 1\}$  or  $\mathcal{Y} = \{-1, +1\}$ . The goal is to learn a decision function  $f : \mathcal{X} \rightarrow \mathcal{Y}$  that can accurately classify new examples.

A natural approach is to model the conditional probability  $P(Y = 1|X = x)$  and classify based on whether this probability exceeds a threshold (typically 0.5). This leads us to consider the *logistic regression* model:

$$P(Y = 1|X = x) = \frac{1}{1 + e^{-\theta^T x}} = \sigma(\theta^T x), \quad (\text{A.5})$$

where  $\sigma(z) = 1/(1 + e^{-z})$  is the *sigmoid function* and  $\theta$  are the model parameters.

The sigmoid function has several desirable properties: it maps any real number to the interval  $(0, 1)$ , it is smooth and differentiable everywhere, and it has a natural interpretation as a probability. The corresponding loss function is the *cross-entropy* or *log-likelihood*:

$$\ell(f(x), y) = -y \log f(x) - (1 - y) \log(1 - f(x)). \quad (\text{A.6})$$

This loss function has a profound connection to information theory: it measures the number of bits needed to encode the true label given the predicted probability distribution. Minimizing cross-entropy is equivalent to maximizing the likelihood of the observed data under the model.

### A.2.2 SUPPORT VECTOR MACHINES

An alternative approach to classification is the *Support Vector Machine* (SVM), which seeks to find the optimal separating hyperplane between classes. For linearly separable data, the SVM finds the hyperplane that maximizes the *margin*—the distance between the hyperplane and the nearest data points from each class.

Mathematically, for binary classification with labels  $y_i \in \{-1, +1\}$ , the SVM optimization problem is:

$$\min_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}\|^2 \quad (\text{A.7})$$

$$\text{subject to } y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1, \quad i = 1, \dots, n. \quad (\text{A.8})$$

For non-separable data, we introduce *slack variables*  $\xi_i \geq 0$  and the *soft-margin* SVM:

$$\min_{\mathbf{w}, b, \xi} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i \quad (\text{A.9})$$

$$\text{subject to } y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0. \quad (\text{A.10})$$

The parameter  $C$  controls the trade-off between margin maximization and classification error. The beauty of SVMs lies in their ability to handle

non-linear decision boundaries through the *kernel trick*, which implicitly maps the input space to a higher-dimensional feature space where linear separation may be possible.

### A.2.3 MULTI-CLASS CLASSIFICATION

Real-world classification problems often involve more than two classes. For  $K > 2$  classes, several strategies can be employed:

- **One-vs-Rest (OvR):** Train  $K$  binary classifiers, each distinguishing one class from all others.
- **One-vs-One (OvO):** Train  $\binom{K}{2}$  binary classifiers for each pair of classes.
- **Direct extension:** Extend binary algorithms directly to handle multiple classes (e.g., multinomial logistic regression).

*The kernel trick allows SVMs to learn non-linear decision boundaries without explicitly computing the high-dimensional feature representations, making them computationally efficient for complex classification tasks.*

For multinomial logistic regression, we use the *softmax function*:

$$P(Y = k|X = x) = \frac{e^{\theta_k^T x}}{\sum_{j=1}^K e^{\theta_j^T x}}, \quad (\text{A.11})$$

which generalizes the sigmoid function to multiple classes while ensuring that the predicted probabilities sum to unity.

---

## A.3 THE REGRESSION PROBLEM

Regression problems involve predicting continuous-valued outputs, making them particularly relevant for estimating physical parameters from experimental measurements. In particle physics, regression tasks include determining particle energies, reconstructing invariant masses, or estimating detector calibration constants.

### A.3.1 LINEAR REGRESSION

The simplest regression model assumes a linear relationship between inputs and outputs:

$$f(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_d x_d = \theta^T x, \quad (\text{A.12})$$

where we have included a bias term  $\theta_0$  by augmenting the input vector with a constant feature.

The most common loss function for regression is the *squared error*:

$$\ell(f(x), y) = (f(x) - y)^2. \quad (\text{A.13})$$

Minimizing the empirical risk with squared error loss leads to the famous *normal equations*:

$$\theta^* = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}, \quad (\text{A.14})$$

where  $\mathbf{X}$  is the design matrix with training examples as rows and  $\mathbf{y}$  is the vector of target values.

The squared error loss has a deep statistical interpretation: it corresponds to maximum likelihood estimation under the assumption that the noise follows a Gaussian distribution. If we assume  $y = \theta^\top x + \epsilon$  where  $\epsilon \sim \mathcal{N}(0, \sigma^2)$ , then minimizing squared error is equivalent to maximizing the likelihood of the observed data.

### A.3.2 POLYNOMIAL REGRESSION AND FEATURE ENGINEERING

Linear regression can be extended to capture non-linear relationships through *polynomial regression*. By transforming the input features, we can fit polynomial functions while maintaining linearity in the parameters:

$$f(x) = \theta_0 + \theta_1 x + \theta_2 x^2 + \cdots + \theta_k x^k. \quad (\text{A.15})$$

More generally, we can apply any non-linear transformation  $\phi : \mathcal{X} \rightarrow \mathbb{R}^p$  to create new features, leading to the model:

$$f(x) = \theta^\top \phi(x). \quad (\text{A.16})$$

This approach, known as *feature engineering*, allows linear models to capture complex non-linear patterns. Common transformations include polynomial features, trigonometric functions, radial basis functions, and domain-specific features derived from physical principles.

### A.3.3 ROBUST REGRESSION

The squared error loss is sensitive to outliers—data points that deviate significantly from the general pattern. In experimental physics, outliers can arise from detector malfunctions, rare background processes, or measurement errors. To mitigate their impact, we can employ robust loss functions:

- **Absolute error (L1 loss):**  $\ell(f(x), y) = |f(x) - y|$
- **Huber loss:** A hybrid that behaves quadratically for small errors and linearly for large errors:

$$\ell_\delta(f(x), y) = \begin{cases} \frac{1}{2}(f(x) - y)^2 & \text{if } |f(x) - y| \leq \delta \\ \delta(|f(x) - y| - \frac{1}{2}\delta) & \text{otherwise} \end{cases} \quad (\text{A.17})$$

These robust loss functions provide better performance when the dataset contains outliers or when the noise distribution has heavy tails.

### A.3.4 PERFORMANCE EVALUATION

Evaluating the performance of supervised learning models requires careful consideration of appropriate metrics. For classification problems, common evaluation metrics include:

- **Accuracy:** The fraction of correctly classified examples.
- **Precision:** The fraction of positive predictions that are actually positive.
- **Recall (Sensitivity):** The fraction of actual positive examples that are correctly identified.
- **F1-score:** The harmonic mean of precision and recall.
- **Area Under the ROC Curve (AUC-ROC):** A threshold-independent measure of classification performance.

For regression problems, typical metrics include:

- **Mean Squared Error (MSE):**  $\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$
- **Root Mean Squared Error (RMSE):**  $\sqrt{\text{MSE}}$
- **Mean Absolute Error (MAE):**  $\frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$
- **Coefficient of Determination (R<sup>2</sup>):** The proportion of variance explained by the model.

The selection of appropriate evaluation metrics depends on the specific requirements of the physics analysis. For signal detection in rare processes, maximizing sensitivity (recall) might be more important than overall accuracy. For precision measurements, minimizing systematic uncertainties requires careful attention to model bias and calibration.

In conclusion, supervised learning provides a powerful framework for extracting patterns from labeled data and making predictions on new examples. The mathematical foundations we have established—from empirical risk minimization to the bias-variance trade-off—provide the theoretical underpinning for understanding when and why these methods work. As we progress to more advanced topics, these fundamental concepts will serve as the building blocks for more sophisticated algorithms and applications in particle physics research.

*In particle physics, the choice of evaluation metric should align with the physics goals. For rare signal detection, recall might be more important than precision, while for precision measurements, minimizing systematic biases (related to model bias) is crucial.*

## A.4 BOOSTED DECISION TREES

---

Boosted Decision Trees (BDT) represent one of the most powerful and widely-used machine learning techniques in high-energy physics, particularly for signal-background discrimination tasks. The success of BDTs stems from their ability to combine the simplicity and interpretability of decision trees with the enhanced predictive power achieved through ensemble methods.

#### A.4.1 DECISION TREES: THE FOUNDATION

A decision tree is a hierarchical model that makes predictions by recursively splitting the input space based on feature values. Each internal node of the tree represents a decision based on a single feature, each branch represents the outcome of that decision, and each leaf node represents a prediction.

Mathematically, a decision tree can be viewed as a piecewise constant function that partitions the input space  $\mathcal{X}$  into disjoint regions  $\{R_1, R_2, \dots, R_M\}$ . The prediction for any input  $x$  falling into region  $R_m$  is:

$$f(x) = \sum_{m=1}^M c_m \mathbb{I}(x \in R_m), \quad (\text{A.18})$$

where  $c_m$  is the prediction for region  $R_m$  and  $\mathbb{I}(\cdot)$  is the indicator function.

The construction of a decision tree involves finding the optimal splits that minimize a chosen impurity measure. For classification problems, common impurity measures include:

- **Gini impurity:**  $G = \sum_{k=1}^K p_k(1 - p_k)$ , where  $p_k$  is the proportion of class  $k$  samples in the node.
- **Cross-entropy:**  $H = -\sum_{k=1}^K p_k \log p_k$
- **Misclassification error:**  $E = 1 - \max_k p_k$

For regression problems, the mean squared error is typically used:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2, \quad (\text{A.19})$$

where  $\bar{y}$  is the mean of the target values in the node.

#### A.4.2 THE BOOSTING PARADIGM

While individual decision trees are interpretable and fast to train, they often suffer from high variance and limited expressiveness. Boosting addresses these limitations by combining multiple weak learners (typically shallow trees) in a way that focuses on the most challenging examples.

The fundamental idea behind boosting is to train a sequence of models, where each subsequent model focuses on correcting the errors made by the previous ensemble. This is achieved through a weighted training procedure where examples that are misclassified by the current ensemble receive higher weights in the training of the next model.

#### *AdaBoost Algorithm*

The original AdaBoost (Adaptive Boosting) algorithm, introduced by Freund and Schapire, provides the foundational framework for understanding boosting. For binary classification with labels  $y_i \in \{-1, +1\}$ , the algorithm proceeds as follows:

1. Initialize uniform weights:  $w_i^{(1)} = \frac{1}{n}$  for  $i = 1, \dots, n$
2. For  $t = 1, 2, \dots, T$ :
  - (a) Train weak learner  $h_t$  using weights  $w^{(t)}$
  - (b) Compute weighted error:  $\epsilon_t = \sum_{i:h_t(x_i) \neq y_i} w_i^{(t)}$
  - (c) Compute learner weight:  $\alpha_t = \frac{1}{2} \ln \left( \frac{1-\epsilon_t}{\epsilon_t} \right)$
  - (d) Update example weights:  $w_i^{(t+1)} = w_i^{(t)} e^{-\alpha_t y_i h_t(x_i)} / Z_t$
3. Output final classifier:  $H(x) = \text{sign} \left( \sum_{t=1}^T \alpha_t h_t(x) \right)$

where  $Z_t$  is a normalization constant ensuring that the weights sum to unity.

#### A.4.3 GRADIENT BOOSTING FRAMEWORK

Modern BDT implementations, such as XGBoost used in particle physics analyses, are based on the gradient boosting framework, which provides a more general and flexible approach to boosting.

In gradient boosting, we view boosting as a functional gradient descent procedure in the space of functions. Given a differentiable loss function  $L(y, f(x))$ , the goal is to find a function  $f^*$  that minimizes:

$$f^* = \arg \min_f \mathbb{E}[L(Y, f(X))]. \quad (\text{A.20})$$

The gradient boosting algorithm approximates this optimization by building an additive model:

$$f_M(x) = \sum_{m=0}^M \gamma_m h_m(x), \quad (\text{A.21})$$

where  $h_0(x)$  is an initial guess (often a constant), and  $h_m(x)$  are the weak learners added at each iteration.

At each step  $m$ , we fit a new weak learner  $h_m$  to the negative gradient (pseudo-residuals) of the loss function:

$$r_{im} = - \left[ \frac{\partial L(y_i, f(x_i))}{\partial f(x_i)} \right]_{f=f_{m-1}}. \quad (\text{A.22})$$

#### A.4.4 XGBOOST: EXTREME GRADIENT BOOSTING

XGBoost (Extreme Gradient Boosting) represents the state-of-the-art implementation of gradient boosting, incorporating several key innovations that make it particularly effective for particle physics applications:

### *Regularized Objective Function*

XGBoost uses a regularized objective function that includes both the training loss and a regularization term:

$$\mathcal{L}(\phi) = \sum_i l(y_i, \hat{y}_i) + \sum_k \Omega(f_k), \quad (\text{A.23})$$

where  $l$  is the differentiable convex loss function,  $\hat{y}_i$  is the prediction for the  $i$ -th instance, and  $\Omega(f_k)$  is the regularization term for the  $k$ -th tree:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^T w_j^2, \quad (\text{A.24})$$

where  $T$  is the number of leaves,  $w_j$  is the score on leaf  $j$ ,  $\gamma$  controls the complexity penalty, and  $\lambda$  is the L2 regularization parameter.

### *Second-Order Taylor Expansion*

Instead of using only first-order gradients, XGBoost employs a second-order Taylor expansion of the loss function, leading to more accurate approximations and faster convergence:

$$\mathcal{L}^{(t)} \approx \sum_{i=1}^n [l(y_i, \hat{y}_i^{(t-1)}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t), \quad (\text{A.25})$$

where  $g_i = \partial_{\hat{y}_i^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)})$  and  $h_i = \partial_{\hat{y}_i^{(t-1)}}^2 l(y_i, \hat{y}_i^{(t-1)})$  are the first and second-order gradients.

### *Key Hyperparameters for Particle Physics*

In particle physics applications, several hyperparameters are crucial for optimizing BDT performance:

- **Learning rate ( $\eta$ ):** Controls the step size in the gradient descent procedure. Typical values in particle physics range from 0.1 to 0.3.
- **Maximum depth ( $\text{max\_depth}$ ):** Limits the depth of individual trees. Values of 4-8 are common for particle physics, balancing expressiveness with overfitting prevention.
- **Minimum child weight ( $\text{min\_child\_weight}$ ):** Controls the minimum sum of instance weights needed in a child node.
- **Regularization parameters ( $\gamma, \lambda, \alpha$ ):** Control model complexity and prevent overfitting.
- **Subsample ratio:** Determines the fraction of training samples used for each tree, introducing stochasticity to improve generalization.

#### A.4.5 FEATURE IMPORTANCE AND INTERPRETABILITY

One of the key advantages of BDTs in particle physics is their interpretability. XGBoost provides several measures of feature importance:

- **Gain:** The average improvement in accuracy brought by a feature to the branches it is on.
- **Cover:** The average coverage of a feature when it is used in trees.
- **Frequency:** The percentage of times a feature is used in all trees.

In the context of particle physics analyses, feature importance helps physicists understand which kinematic variables are most discriminating for signal-background separation, providing physical insights into the underlying processes.

## A.5 NEURAL NETWORKS

---

Neural networks represent one of the most powerful and flexible approaches to machine learning, offering the ability to learn complex non-linear relationships between inputs and outputs. In particle physics, neural networks have found applications ranging from event classification to detector simulation and beyond.

### A.5.1 THE PERCEPTRON: BUILDING BLOCK OF NEURAL NETWORKS

The fundamental unit of a neural network is the perceptron, which can be viewed as a generalization of linear regression with a non-linear activation function. A single perceptron computes:

$$y = \sigma \left( \sum_{i=1}^d w_i x_i + b \right) = \sigma(\mathbf{w}^\top \mathbf{x} + b), \quad (\text{A.26})$$

where  $\mathbf{w}$  are the weights,  $b$  is the bias term, and  $\sigma$  is the activation function.

Common activation functions include:

- **Sigmoid:**  $\sigma(z) = \frac{1}{1+e^{-z}}$
- **Hyperbolic tangent:**  $\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$
- **Rectified Linear Unit (ReLU):**  $\text{ReLU}(z) = \max(0, z)$
- **Leaky ReLU:**  $\text{LeakyReLU}(z) = \max(\alpha z, z)$  where  $\alpha$  is a small positive constant

### A.5.2 MULTI-LAYER PERCEPTRONS

A multi-layer perceptron (MLP) consists of multiple layers of perceptrons, allowing the network to learn complex non-linear mappings. For a network with  $L$  layers, the forward propagation can be described as:

$$\mathbf{z}^{(1)} = \mathbf{W}^{(1)}\mathbf{x} + \mathbf{b}^{(1)} \quad (\text{A.27})$$

$$\mathbf{a}^{(1)} = \sigma^{(1)}(\mathbf{z}^{(1)}) \quad (\text{A.28})$$

$$\vdots \quad (\text{A.29})$$

$$\mathbf{z}^{(L)} = \mathbf{W}^{(L)}\mathbf{a}^{(L-1)} + \mathbf{b}^{(L)} \quad (\text{A.30})$$

$$\mathbf{y} = \sigma^{(L)}(\mathbf{z}^{(L)}) \quad (\text{A.31})$$

where  $\mathbf{W}^{(l)}$  and  $\mathbf{b}^{(l)}$  are the weight matrix and bias vector for layer  $l$ , and  $\sigma^{(l)}$  is the activation function for layer  $l$ .

### A.5.3 BACKPROPAGATION ALGORITHM

Training neural networks requires computing gradients of the loss function with respect to all parameters. The backpropagation algorithm efficiently computes these gradients using the chain rule of calculus.

For a loss function  $L(\mathbf{y}, \hat{\mathbf{y}})$ , the gradients are computed by propagating errors backward through the network:

$$\frac{\partial L}{\partial \mathbf{W}^{(l)}} = \frac{\partial L}{\partial \mathbf{z}^{(l)}} \frac{\partial \mathbf{z}^{(l)}}{\partial \mathbf{W}^{(l)}} = \boldsymbol{\delta}^{(l)} (\mathbf{a}^{(l-1)})^T \quad (\text{A.32})$$

$$\frac{\partial L}{\partial \mathbf{b}^{(l)}} = \boldsymbol{\delta}^{(l)} \quad (\text{A.33})$$

$$\boldsymbol{\delta}^{(l-1)} = (\mathbf{W}^{(l)})^T \boldsymbol{\delta}^{(l)} \odot \sigma'^{(l-1)}(\mathbf{z}^{(l-1)}) \quad (\text{A.34})$$

where  $\boldsymbol{\delta}^{(l)} = \frac{\partial L}{\partial \mathbf{z}^{(l)}}$  is the error signal for layer  $l$ , and  $\odot$  denotes element-wise multiplication.

### A.5.4 OPTIMIZATION ALGORITHMS

Training neural networks involves optimizing the loss function using gradient-based methods. Several sophisticated optimization algorithms have been developed:

*Stochastic Gradient Descent (SGD)*

The basic SGD algorithm updates parameters using:

$$\theta_{t+1} = \theta_t - \eta \nabla_{\theta} L(\theta_t), \quad (\text{A.35})$$

where  $\eta$  is the learning rate.

### Momentum

Momentum accelerates SGD by accumulating gradients from previous steps:

$$\mathbf{v}_t = \beta \mathbf{v}_{t-1} + (1 - \beta) \nabla_{\theta} L(\theta_t) \quad (\text{A.36})$$

$$\theta_{t+1} = \theta_t - \eta \mathbf{v}_t \quad (\text{A.37})$$

### Adam Optimizer

Adam (Adaptive Moment Estimation) combines the benefits of momentum and adaptive learning rates:

$$\mathbf{m}_t = \beta_1 \mathbf{m}_{t-1} + (1 - \beta_1) \nabla_{\theta} L(\theta_t) \quad (\text{A.38})$$

$$\mathbf{v}_t = \beta_2 \mathbf{v}_{t-1} + (1 - \beta_2) (\nabla_{\theta} L(\theta_t))^2 \quad (\text{A.39})$$

$$\hat{\mathbf{m}}_t = \frac{\mathbf{m}_t}{1 - \beta_1^t} \quad (\text{A.40})$$

$$\hat{\mathbf{v}}_t = \frac{\mathbf{v}_t}{1 - \beta_2^t} \quad (\text{A.41})$$

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{\mathbf{v}}_t} + \epsilon} \hat{\mathbf{m}}_t \quad (\text{A.42})$$

## A.5.5 REGULARIZATION TECHNIQUES

Neural networks are prone to overfitting, especially when dealing with limited training data or high-dimensional feature spaces. Several regularization techniques are employed:

### Dropout

Dropout randomly sets a fraction of input units to zero during training, preventing co-adaptation of neurons:

$$\mathbf{a}^{(l)} = \mathbf{m}^{(l)} \odot \sigma^{(l)}(\mathbf{z}^{(l)}), \quad (\text{A.43})$$

where  $\mathbf{m}^{(l)}$  is a binary mask with elements drawn from a Bernoulli distribution.

### Batch Normalization

Batch normalization normalizes the inputs to each layer, accelerating training and improving generalization:

$$\hat{\mathbf{z}}^{(l)} = \frac{\mathbf{z}^{(l)} - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}, \quad (\text{A.44})$$

where  $\mu_B$  and  $\sigma_B^2$  are the mean and variance of the mini-batch.

### A.5.6 NEURAL NETWORKS IN PARTICLE PHYSICS

In particle physics applications, neural networks offer several advantages:

- **High expressiveness:** Neural networks can learn complex non-linear relationships between kinematic variables and physics outcomes.
- **Automatic feature learning:** Deep networks can automatically discover relevant feature combinations without manual feature engineering.
- **Scalability:** Neural networks can handle high-dimensional input spaces and large datasets effectively.

However, they also present some challenges:

- **Computational cost:** Neural networks require significantly more computational resources for training compared to BDTs.
- **Interpretability:** Understanding which features contribute to the network's decisions is more difficult than with tree-based methods.
- **Hyperparameter sensitivity:** Neural networks have many hyperparameters that require careful tuning.

### A.5.7 NETWORK ARCHITECTURES FOR PARTICLE PHYSICS

Different network architectures are suitable for different particle physics tasks:

#### *Dense Networks*

Standard fully-connected networks are appropriate for classification tasks with tabular data (kinematic variables):

- Input layer with the number of kinematic features
- Multiple hidden layers with ReLU activation
- Output layer with sigmoid (binary) or softmax (multi-class) activation

#### *Convolutional Neural Networks (CNNs)*

CNNs are particularly effective for image-like data such as calorimeter deposits or detector hit patterns.

#### *Recurrent Neural Networks (RNNs)*

RNNs and their variants (LSTM, GRU) are useful for sequential data such as particle tracks or time-series detector readouts.

## A.6 COMPARATIVE ANALYSIS: BDTS VS NEURAL NETWORKS

---

The choice between BDTS and neural networks in particle physics analyses depends on several factors:

### A.6.1 PERFORMANCE COMPARISON

In many particle physics applications, both BDTS and neural networks can achieve similar classification performance. The choice often depends on:

- **Dataset size:** Neural networks typically require larger datasets to reach their full potential
- **Feature complexity:** Neural networks excel when complex feature interactions are present
- **Training time:** BDTS generally train much faster than neural networks

### A.6.2 INTERPRETABILITY

BDTS offer superior interpretability through:

- Feature importance rankings
- Visualization of decision trees
- Clear understanding of decision boundaries

Neural networks are less interpretable but techniques like attention mechanisms and gradient-based attribution methods can provide insights.

### A.6.3 ROBUSTNESS AND GENERALIZATION

Both methods require careful attention to overfitting, but they differ in their susceptibility:

- BDTS are generally more robust to outliers
- Neural networks may be more sensitive to distribution shifts
- Both benefit from proper regularization and validation procedures

### A.6.4 COMPUTATIONAL CONSIDERATIONS

- **Training time:** BDTS typically train in minutes to hours, while neural networks may require hours to days
- **Inference speed:** Both can achieve fast inference, but BDTS generally have lower latency

- **Memory requirements:** Neural networks typically require more memory during training

In conclusion, both BDTs and neural networks are powerful tools for particle physics analyses. The choice between them should be based on the specific requirements of the analysis, including dataset characteristics, computational constraints, and interpretability needs. Many successful particle physics analyses have demonstrated that when properly implemented and tuned, both approaches can significantly enhance the sensitivity to new physics signals.

# B

---

## ON THE UNIVERSAL SEESAW MECHANISM

---

The masses of the SM fermions are generated not through direct Yukawa couplings to the Higgs, but via a universal seesaw mechanism by mixing with new vector-like fermions  $\chi_f$ . The most general gauge-invariant and renormalizable Lagrangian for the quark sector is given by:

$$\begin{aligned} \mathcal{L}_{\text{mass}}^{\text{quark}} \supset & \bar{Q}_L^i (Y_{uL})_{ij} \chi_{uR}^j \tilde{H} + \bar{Q}_L^i (Y_{dL})_{ij} \chi_{dR}^j H \\ & + \bar{\chi}_{uL}^i (Y_{uR})_{ij} u_R^j \phi^* + \bar{\chi}_{dL}^i (Y_{dR})_{ij} d_R^j \phi \\ & + \bar{\chi}_{uL}^i (m_{\chi u})_{ij} \chi_{uR}^j + \bar{\chi}_{dL}^i (m_{\chi d})_{ij} \chi_{dR}^j + \text{h.c.}, \end{aligned} \quad (\text{B.1})$$

where  $i, j = 1, 2, 3$  are flavor indices. An entirely analogous set of terms exists for the lepton sector. The Yukawa matrices  $Y_{uL}, Y_{dL}, Y_{uR}, Y_{dR}$  and the vector-like mass matrices  $m_{\chi u}, m_{\chi d}$  are general complex  $3 \times 3$  matrices, making the flavor structure highly non-trivial.

### B.1 DIAGONALIZATION AND FIELD REDEFINITIONS

---

To make physical predictions, we must diagonalize these matrices. We express them in terms of their singular value decompositions (i.e., their diagonal forms) and the associated unitary mixing matrices:

$$\begin{aligned} Y_{uL} &= U_{LL}^\dagger Y_{uL}^d U_{LR}, & Y_{uR} &= U_{RL}^\dagger Y_{uR}^d U_{RR}, \\ Y_{dL} &= V_{LL}^\dagger Y_{dL}^d V_{LR}, & Y_{dR} &= V_{RL}^\dagger Y_{dR}^d V_{RR}, \\ m_{\chi u} &= W_{uL}^\dagger m_{\chi u}^d W_{uR}, & m_{\chi d} &= W_{dL}^\dagger m_{\chi d}^d W_{dR}. \end{aligned}$$

Here, the matrices  $Y^d$  and  $m^d$  are real, diagonal, and non-negative. The unitary matrices  $U, V, W$  are not physical by themselves but encode the mixing between flavor states.

We now perform a series of field redefinitions to absorb the maximal number of these unitary matrices into the definitions of the fermion fields. The goal is to make as many mass parameters diagonal as possible. The redefinitions are:

$$\begin{aligned} Q_L &\rightarrow U_{LL} Q_L, & \chi_{uR} &\rightarrow W_{uR} \chi_{uR}, & \chi_{uL} &\rightarrow W_{uL} \chi_{uL}, & u_R &\rightarrow U_{RR} u_R, \\ \chi_{dR} &\rightarrow W_{dR} \chi_{dR}, & \chi_{dL} &\rightarrow W_{dL} \chi_{dL}, & d_R &\rightarrow V_{RR} d_R. \end{aligned}$$

Applying these transformations to the Lagrangian (B.1) and using the definitions above, we obtain the simplified form:

$$\begin{aligned}\mathcal{L}_{\text{Yuk}} = & \bar{Q}_L Y_{uL}^d (U_{LR} W_{uR}^\dagger) \chi_{uR} \tilde{H} + \bar{Q}_L (U_{LL} V_{LL}^\dagger) Y_{dL}^d (V_{LR} W_{dR}^\dagger) \chi_{dR} H \\ & + \bar{\chi}_{uL} (W_{uL} U_{RL}^\dagger) Y_{uR}^d u_R \phi^* + \bar{\chi}_{dL} (W_{dL} V_{RL}^\dagger) Y_{dR}^d d_R \phi \\ & + \bar{\chi}_{uL} m_{\chi u}^d \chi_{uR} + \bar{\chi}_{dL} m_{\chi d}^d \chi_{dR} + \text{h.c.}\end{aligned}$$

The matrix  $\tilde{V}_{\text{CKM}} \equiv U_{LL} V_{LL}^\dagger$  is identified as the unitary matrix that will yield the observed Cabibbo-Kobayashi-Maskawa (CKM) quark mixing. For simplicity, and to focus on the essential mass generation mechanism, we adopt a *flavor-aligned* scenario. This assumes that all other unitary matrices ( $U_{LR}$ ,  $W_{uR}$ ,  $W_{uL}$ ,  $U_{RL}$ , etc.) are equal to the identity matrix. This is a strong assumption that minimizes new sources of flavor violation beyond the SM. Under this assumption, the Lagrangian simplifies dramatically to:

$$\begin{aligned}\mathcal{L}_{\text{Yuk}} = & \bar{Q}_L Y_{uL}^d \chi_{uR} \tilde{H} + \bar{Q}_L \tilde{V}_{\text{CKM}} Y_{dL}^d \chi_{dR} H \\ & + \bar{\chi}_{uL} Y_{uR}^d u_R \phi^* + \bar{\chi}_{dL} Y_{dR}^d d_R \phi \\ & + \bar{\chi}_{uL} m_{\chi u}^d \chi_{uR} + \bar{\chi}_{dL} m_{\chi d}^d \chi_{dR} + \text{h.c.}\end{aligned}$$

All matrices  $Y^d$  and  $m^d$  are now diagonal. The only remaining off-diagonal flavor structure is in  $\tilde{V}_{\text{CKM}}$ .

## B.2 SYMMETRY BREAKING AND THE MASS MATRIX

---

After the electroweak symmetry breaking ( $\langle H \rangle = v_h/\sqrt{2}$ ) and the  $U(1)_{T_R^3}$  breaking ( $\langle \phi \rangle = v_\phi/\sqrt{2}$ ), the mass terms for the up-type quarks (and analogously for down-type and leptons) are generated. For a single generation, the mass terms in the basis  $(\bar{u}_L, \bar{\chi}_{uL}), (u_R, \chi_{uR})^\top$  form a  $2 \times 2$  matrix:

$$\mathcal{L}_{\text{mass}} = - (\bar{u}_L \quad \bar{\chi}_{uL}) \begin{pmatrix} 0 & m_L \\ m_R & m_\chi \end{pmatrix} \begin{pmatrix} u_R \\ \chi_{uR} \end{pmatrix} + \text{h.c.}, \quad (\text{B.2})$$

where the Dirac masses are:

$$\begin{aligned}m_L &= \frac{v_h}{\sqrt{2}} Y_{uL}, \\ m_R &= \frac{v_\phi}{\sqrt{2}} Y_{uR}.\end{aligned}$$

The entry  $m_\chi$  is the vector-like mass. For three generations, this generalizes to a  $6 \times 6$  matrix:

$$M_f = \begin{pmatrix} 0 & m_L \\ m_R & m_\chi \end{pmatrix}, \quad (\text{B.3})$$

where each entry is now a  $3 \times 3$  matrix:  $m_L = \frac{v_h}{\sqrt{2}} Y_{fL}^d$ ,  $m_R = \frac{v_\phi}{\sqrt{2}} Y_{fR}^d$ , and  $m_\chi = m_{\chi f}^d$ .

### B.3 BI-UNITARY TRANSFORMATION AND MASS EIGENVALUES

---

The general mass matrix  $M_f$  is diagonalized by a bi-unitary transformation:

$$U_R^\dagger M_f U_L = M_f^d = \text{diag}(m_{f_1}, m_{f_2}, m_{f_3}, m_{F_1}, m_{F_2}, m_{F_3}), \quad (\text{B.4})$$

where  $U_L$  and  $U_R$  are  $6 \times 6$  unitary matrices. The physical masses are found by solving the eigenvalues of the Hermitian matrices  $H_L = M_f M_f^\dagger$  and  $H_R = M_f^\dagger M_f$ , as  $U_L$  diagonalizes  $H_L$  and  $U_R$  diagonalizes  $H_R$ .

For the one-generation case, these matrices are:

$$\begin{aligned} H_L &= M_f M_f^\dagger = \begin{pmatrix} m_L m_L^\dagger & m_L m_X^\dagger \\ m_X m_L^\dagger & m_R m_R^\dagger + m_X m_X^\dagger \end{pmatrix} = \begin{pmatrix} |m_L|^2 & m_L m_X^* \\ m_X m_L^* & |m_R|^2 + |m_X|^2 \end{pmatrix}, \\ H_R &= M_f^\dagger M_f = \begin{pmatrix} m_R m_R^\dagger & m_R m_X^\dagger \\ m_X m_R^\dagger & m_L m_L^\dagger + m_X m_X^\dagger \end{pmatrix} = \begin{pmatrix} |m_R|^2 & m_R m_X^* \\ m_X m_R^* & |m_L|^2 + |m_X|^2 \end{pmatrix}. \end{aligned}$$

The eigenvalues  $\lambda$  of  $H_L$  (and  $H_R$ ) are found from the characteristic equation  $\det(H_L - \lambda I) = 0$ :

$$\begin{aligned} &| |m_L|^2 - \lambda | |m_R|^2 + |m_X|^2 - \lambda | - |m_L|^2 |m_X|^2 = 0 \\ &\Rightarrow \lambda^2 - \lambda(|m_L|^2 + |m_R|^2 + |m_X|^2) + |m_L|^2 |m_R|^2 = 0. \end{aligned}$$

The solutions to this quadratic equation are the squared masses of the two mass eigenstates:

$$m_f^2 = \frac{1}{2} \left( m_X^2 + m_L^2 + m_R^2 - \sqrt{(m_X^2 + m_L^2 + m_R^2)^2 - 4m_L^2 m_R^2} \right), \quad (\text{B.5})$$

$$m_F^2 = \frac{1}{2} \left( m_X^2 + m_L^2 + m_R^2 + \sqrt{(m_X^2 + m_L^2 + m_R^2)^2 - 4m_L^2 m_R^2} \right), \quad (\text{B.6})$$

where we have now assumed all parameters are real for clarity.  $m_f$  is the light SM-like fermion mass, and  $m_F$  is the heavy vector-like partner mass.

### B.4 YUKAWA COUPLING ENHANCEMENT AND PERTURBATIVITY

---

Equation (B.5) is fundamental. It shows that the light mass  $m_f$  is not simply proportional to  $m_L$  (the SM Higgs VEV). We can solve Eq. (B.5) for  $m_L^2$ :

$$\begin{aligned} m_f^2(m_X^2 + m_L^2 + m_R^2 - m_f^2) &= m_L^2 m_R^2 \quad (\text{from the exact seesaw relation}) \\ m_L^2(m_R^2 - m_f^2) &= m_f^2(m_X^2 + m_R^2 - m_f^2) \\ m_L^2 &= m_f^2 \left( \frac{m_X^2 + m_R^2 - m_f^2}{m_R^2 - m_f^2} \right) = m_f^2 \left( 1 + \frac{m_X^2}{m_R^2 - m_f^2} \right). \end{aligned}$$

Expressing this in terms of the original Yukawa couplings, where  $m_L = \frac{v_h}{\sqrt{2}} Y_{fL}$  and the SM Yukawa is defined by  $m_f = \frac{v_h}{\sqrt{2}} Y_f^{\text{SM}}$ , we find:

$$Y_{fL}^2 = (Y_f^{\text{SM}})^2 \left( 1 + \frac{m_X^2}{m_R^2 - m_f^2} \right). \quad (\text{B.7})$$

This relation reveals the core of the universal seesaw mechanism: the Yukawa coupling  $Y_{fL}$  that couples the SM fermions to the Higgs is *enhanced* compared to the standard model value  $Y_f^{\text{SM}}$ . The enhancement factor is  $\sqrt{1 + m_\chi^2/(m_R^2 - m_f^2)}$ .

This has profound implications:

- **Light Fermions (e.g., electron, u-quark):** Here,  $Y_f^{\text{SM}} \ll 1$ . A large hierarchy  $m_\chi^2 \gg m_R^2 \gg m_f^2$  can generate this tiny mass from a more “natural”  $Y_{fL} \sim \mathcal{O}(0.1 - 1)$ .
- **Top Quark:** Here,  $Y_t^{\text{SM}} \approx 1$  is already large. An enhancement could easily push  $Y_{tL}$  into the non-perturbative regime ( $Y_{tL}^2/4\pi > 1$ ). To avoid this, we must require the enhancement factor to be  $\mathcal{O}(1)$ , which implies  $m_\chi^2 \lesssim m_R^2 - m_t^2$ . Since  $m_R = \frac{v_\phi}{\sqrt{2}} Y_{fR}$ , this suggests  $v_\phi > m_\chi$  is a natural condition.

## B.5 EXACT DIAGONALIZATION AND MIXING ANGLES

---

The bi-unitary transformation (B.4) is performed by matrices that can be parameterized by a mixing angle. For one generation, the left-handed mixing matrix is:

$$U_L = \begin{pmatrix} \cos \theta_L & \sin \theta_L \\ -\sin \theta_L & \cos \theta_L \end{pmatrix}. \quad (\text{B.8})$$

The angle  $\theta_L$  quantifies the mixing between the SM fermion and its vector-like partner. The exact expressions for the fundamental parameters  $m_L, m_R, m_\chi$  in terms of the physical masses  $m_f, m_F$  and the mixing angle  $\theta_L$  can be found by equating  $U_L^\dagger H_L U_L = \text{diag}(m_f^2, m_F^2)$ . This yields the system of equations:

$$\begin{aligned} m_L^2 &= m_f^2 \cos^2 \theta_L + m_F^2 \sin^2 \theta_L, \\ m_R^2 + m_\chi^2 &= m_f^2 \sin^2 \theta_L + m_F^2 \cos^2 \theta_L, \\ m_L m_\chi &= (m_F^2 - m_f^2) \sin \theta_L \cos \theta_L. \end{aligned}$$

Solving this system (and a similar one from  $H_R$  for  $\theta_R$ ) gives:

$$m_L^2 = \frac{1}{2} (m_f^2 + m_F^2 - (m_F^2 - m_f^2) \cos 2\theta_L), \quad (\text{B.9})$$

$$m_R^2 = \frac{m_f^2 m_F^2}{m_L^2} = \frac{2m_f^2 m_F^2}{m_f^2 + m_F^2 - (m_F^2 - m_f^2) \cos 2\theta_L}, \quad (\text{B.10})$$

$$m_\chi^2 = m_R^2 + m_F^2 + m_f^2 - m_L^2 - \frac{m_f^2 m_F^2}{m_L^2} = \frac{(m_F^2 - m_f^2)^2 \sin^2 2\theta_L}{4m_L^2}. \quad (\text{B.11})$$

Substituting Eq. (B.9) into the expression for  $m_\chi^2$  yields the form shown in the original text.

The critical constraint to keep the top Yukawa perturbative is  $m_\chi^2 < m_R^2$ . Using Eqs. (B.10) and (B.11), the ratio is:

$$\frac{m_\chi^2}{m_R^2} = \frac{(m_F^2 - m_f^2)^2 \sin^2 2\theta_L}{4m_f^2 m_F^2} < 1. \quad (\text{B.12})$$

For the top quark with  $m_f = m_t \approx 173$  GeV and assuming a heavy partner  $m_F \gg m_t$ , this simplifies to:

$$\frac{m_F^4 \sin^2 2\theta_L}{4m_t^2 m_F^2} \approx \frac{m_F^2}{4m_t^2} \sin^2 2\theta_L < 1 \Rightarrow \sin^2 2\theta_L < \frac{4m_t^2}{m_F^2}. \quad (\text{B.13})$$

This is a very strong constraint. For example, if  $m_F = 1$  TeV, then  $\sin^2 2\theta_L < 0.12$ , meaning  $\theta_L < 10^\circ$ . In the small  $\theta_L$  limit,  $\cos 2\theta_L \approx 1 - 2\theta_L^2$  and  $\sin^2 2\theta_L \approx 4\theta_L^2$ . Substituting this into Eq. (B.9):

$$\begin{aligned} m_L^2 &\approx \frac{1}{2} (m_t^2 + m_F^2 - (m_F^2 - m_t^2)(1 - 2\theta_L^2)) \\ &= \frac{1}{2} (m_t^2 + m_F^2 - m_F^2 + m_t^2 + 2(m_F^2 - m_t^2)\theta_L^2) \\ &= \frac{1}{2} (2m_t^2 + 2(m_F^2 - m_t^2)\theta_L^2) = m_t^2 + (m_F^2 - m_t^2)\theta_L^2. \end{aligned}$$

From the constraint (B.13),  $\theta_L^2 < m_t^2/m_F^2$ . Therefore:

$$m_L^2 < m_t^2 + (m_F^2 - m_t^2) \frac{m_t^2}{m_F^2} = m_t^2 + m_t^2 - \frac{m_t^4}{m_F^2} = 2m_t^2 - \frac{m_t^4}{m_F^2}. \quad (\text{B.14})$$

Converting back to Yukawa couplings:

$$Y_{tL}^2 \lesssim (Y_t^{\text{SM}})^2 \left( 2 - \frac{m_t^2}{m_F^2} \right). \quad (\text{B.15})$$

This shows that the maximum enhancement for the top Yukawa is less than a factor of  $\sqrt{2}$ , which is perfectly perturbative.

The generalization to three generations involves the diagonalization of the full  $6 \times 6$  matrices. The matrix  $\tilde{V}_{\text{CKM}}$  introduced during field redefinition will manifest in the charged current weak interactions of the mass eigenstates. After diagonalization, the SM  $W$  boson will couple not only to the three light quarks but also to the heavy vector-like quarks, with couplings suppressed by the mixing angles  $\theta_L^i$ . The observed  $3 \times 3$  CKM matrix emerges as the effective mixing matrix among the three light quarks when the heavy states are integrated out.

The lepton sector follows an identical procedure for the charged leptons. The neutrino sector, however, offers further richness. The right-handed neutrinos  $\nu_R$  can possess both Dirac masses ( $m_R$ ) from coupling to  $\phi$  and Majorana mass terms  $M_R \bar{\nu}_R^c \nu_R$ , which are allowed by the gauge symmetry. The vector-like neutrinos  $\chi_\nu$  can also have Majorana masses. This combination of Dirac and Majorana masses for both  $\nu_R$  and  $\chi_\nu$  can generate a double or triple seesaw mechanism, providing a natural explanation for the tiny masses of the observed light neutrinos. The diagonalization of this extended neutrino mass matrix also generates the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) mixing matrix.



# C

---

## THE 4321 MODEL

---

This appendix summarises the main features of the 4321 model presented in [273], based on the construction showed in [138]. The model is built upon the extended gauge group

$$\mathcal{G}_{4321} \equiv \mathrm{SU}(4) \times \mathrm{SU}(3)' \times \mathrm{SU}(2)_L \times \mathrm{U}(1)'.$$

The Standard Model (SM) gauge group,  $\mathcal{G}_{321} \equiv \mathrm{SU}(3)_c \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ , is embedded into  $\mathcal{G}_{4321}$  through two key identifications.

First, the SM strong force is identified with the diagonal subgroup of the two  $\mathrm{SU}(3)$  factors:

$$\mathrm{SU}(3)_c = (\mathrm{SU}(3)_{[4]} \times \mathrm{SU}(3)')_{\text{diag}}, \quad (\text{C.1})$$

where  $\mathrm{SU}(3)_{[4]} \subset \mathrm{SU}(4)$ . Second, and more crucially, the SM hypercharge is a linear combination of charges from the  $\mathrm{SU}(4)$  and  $\mathrm{U}(1)'$  sectors:

$$Y = Q_{B-L} + Y'. \quad (\text{C.2})$$

Here, the baryon minus lepton number ( $Q_{B-L}$ ) is generated by a diagonal  $\mathrm{SU}(4)$  generator,  $Q_{B-L} = 2\sqrt{6}T^{15}/3$ ,

The spontaneous breaking of the full  $\mathcal{G}_{4321}$  symmetry down to the SM  $\mathcal{G}_{321}$  gives mass to the gauge bosons associated with the broken generators. The spectrum of these new massive vectors and their quantum numbers under the SM group are:

- A vector leptoquark,  $U \sim (\mathbf{3}, \mathbf{1}, 2/3)$ ,
- A coloron,  $g' \sim (\mathbf{8}, \mathbf{1}, 0)$ ,
- A massive neutral boson,  $Z' \sim (\mathbf{1}, \mathbf{1}, 0)$ .

Heuristically, each of these bosons originates from a distinct part of the symmetry breaking pattern: the leptoquark ( $U$ ) emerges from the breaking  $\mathrm{SU}(4) \rightarrow \mathrm{SU}(3)_{[4]} \times \mathrm{U}(1)_{B-L}$ , the coloron ( $g'$ ) from  $\mathrm{SU}(3)_{[4]} \times \mathrm{SU}(3)' \rightarrow \mathrm{SU}(3)_c$ , and the  $Z'$  from  $\mathrm{U}(1)_{B-L} \times \mathrm{U}(1)_{T_R^3} \rightarrow \mathrm{U}(1)_Y$ .

The spontaneous breaking of the  $\mathcal{G}_{4321}$  symmetry down to the Standard Model  $\mathcal{G}_{321}$  and subsequently to electromagnetism is achieved through a scalar sector comprising four multiplets. The primary breaking  $\mathcal{G}_{4321} \rightarrow \mathcal{G}_{321}$  is induced by the vacuum expectation values (vevs) of three scalar fields:

*As we see in chapter 3, the  $\mathrm{U}(1)'$  charge could be identified with twice the third component of right-handed isospin,  $Y' \equiv 2Q_{T_R^3}$ . This specific embedding reveals the model's left-right symmetric foundation; the SM electric charge operator can now be expressed in the manifestly left-right symmetric form:*

$$Q = Q_{T_L^3} + Q_{T_R^3} + \frac{1}{2}Q_{B-L}.$$

- $\Omega_1 \sim (\bar{\mathbf{4}}, \mathbf{1}, \mathbf{1}, -1/2)$ ,
- $\Omega_3 \sim (\bar{\mathbf{4}}, \mathbf{3}, \mathbf{1}, 1/6)$ ,
- $\Omega_{15} \sim (\mathbf{15}, \mathbf{1}, \mathbf{1}, 0)$  (taken to be a real field).

The final electroweak symmetry breaking,  $\mathcal{G}_{321} \rightarrow U(1)_{EM}$ , is triggered by the Higgs doublet  $H \sim (\mathbf{1}, \mathbf{1}, \mathbf{2}, 1/2)$ .

A suitable scalar potential (analysed in detail in Section ??) allows for a vev configuration that ensures this breaking pattern. Phenomenological constraints suggest a clear hierarchy between these scales:

$$\langle \Omega_3 \rangle > \langle \Omega_1 \rangle \gg \langle \Omega_{15} \rangle \gg \langle H \rangle. \quad (C.3)$$

Given this hierarchy, we simplify the analysis by first considering the  $\Omega_3$  and  $\Omega_1$  system in isolation to understand the primary TeV-scale breaking. The effects of incorporating the smaller vevs of  $\Omega_{15}$  and  $H$  will be discussed subsequently.

To analyze the  $\Omega_3$ - $\Omega_1$  subsystem, we represent these fields as a  $4 \times 3$  matrix and a 4-vector, transforming as  $\Omega_3 \rightarrow U_4^* \Omega_3 U_3^T$ , and  $\Omega_1 \rightarrow U_4^* \Omega_1$  under  $SU(4) \times SU(3)'$ , respectively. The desired vacuum configuration that breaks  $\mathcal{G}_{4321}$  to  $\mathcal{G}_{321}$  is:

$$\langle \Omega_3 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} v_3 & 0 & 0 \\ 0 & v_3 & 0 \\ 0 & 0 & v_3 \\ 0 & 0 & 0 \end{pmatrix}, \quad \langle \Omega_1 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 0 \\ v_1 \end{pmatrix}. \quad (C.4)$$

The most general renormalizable scalar potential that admits this vacuum as a stationary point, and in the limit where the bare masses vanish ( $\mu_3 = \mu_1 = 0$ ) and the cubic coupling is absent ( $\lambda_6 = 0$ ), can be written as:

$$\begin{aligned} V_{\Omega_3, \Omega_1} = & \mu_1^2 |\Omega_1|^2 + \mu_3^2 \text{Tr}(\Omega_3^\dagger \Omega_3) \\ & + \lambda_1 \left( \text{Tr}(\Omega_3^\dagger \Omega_3) - \frac{3}{2} v_3^2 \right)^2 + \lambda_2 \text{Tr} \left( \Omega_3^\dagger \Omega_3 - \frac{1}{2} v_3^2 \mathbb{1}_3 \right)^2 \\ & + \lambda_3 \left( |\Omega_1|^2 - \frac{1}{2} v_1^2 \right)^2 + \lambda_4 \left( \text{Tr}(\Omega_3^\dagger \Omega_3) - \frac{3}{2} v_3^2 \right) \left( |\Omega_1|^2 - \frac{1}{2} v_1^2 \right) \\ & + \lambda_5 \Omega_1^\dagger \Omega_3 \Omega_3^\dagger \Omega_1 + \lambda_6 ([\Omega_3 \Omega_3 \Omega_3 \Omega_1]_1 + \text{h.c.}). \end{aligned} \quad (C.5)$$

Here,  $\mathbb{1}_3$  denotes the  $3 \times 3$  identity matrix. We have used a relative rephasing between the fields  $\Omega_1$  and  $\Omega_3$  to remove the phase of  $\lambda_6$ . The unique quartic term,

$$[\Omega_3 \Omega_3 \Omega_3 \Omega_1]_1 \equiv \epsilon_{\alpha \beta \gamma \delta} \epsilon^{abc} (\Omega_3)_a^\alpha (\Omega_3)_b^\beta (\Omega_3)_c^\gamma (\Omega_1)^\delta, \quad (C.6)$$

is required to avoid accidental global symmetries in the scalar potential that would lead to unwanted massless Goldstone bosons.

The inclusion of the other two representations,  $\Omega_{15}$  and  $H$ , in the scalar potential can be safely considered as a perturbation. They are assumed to take the vevs  $\langle \Omega_{15} \rangle = T_{15} v_{15}$  and  $\langle H \rangle = \frac{1}{\sqrt{2}} (0, v)^T$ , with  $v = 246$  GeV. This

treatment is justified because their vevs are subleading for phenomenological reasons and they do not alter the pattern of global symmetries of the  $\Omega_3 - \Omega_1$  potential. Finally, the decomposition of  $\Omega_{15}$  under  $\mathcal{G}_{321}$  is  $\Omega_{15} \rightarrow (\mathbf{1}, \mathbf{1}, 0) \oplus (\mathbf{3}, \mathbf{1}, 2/3) \oplus (\bar{\mathbf{3}}, \mathbf{1}, -2/3) \oplus (\mathbf{8}, \mathbf{1}, 0)$ . The mixing of these states with those contained in  $\Omega_{3,1}$  is parametrically suppressed by the ratio  $v_{15}^2/v_{3,1}^2$ , hence they play a subleading role in phenomenology.

Given the extended gauge group  $\mathcal{G}_{4321}$ , we denote the gauge fields by  $H_\mu^\alpha, G'_\mu^a, W_\mu^i, B'_\mu$ ; the gauge couplings by  $g_4, g_3, g_2, g_1$ ; and the generators by  $T^\alpha, T^a, T^i, Y'$  (with indices  $\alpha = 1, \dots, 15$ ,  $a = 1, \dots, 8$ ,  $i = 1, 2, 3$ ).

To determine the gauge boson spectrum, we start from the covariant derivatives acting on the scalar fields  $\Omega_{3,1,15}$ :

$$\begin{aligned} D_\mu \Omega_1 &= (\partial_\mu + ig_4 H_\mu^\alpha T^{\alpha*} - \frac{1}{2}ig_1 B'_\mu) \Omega_1, \\ D_\mu \Omega_3 &= (\partial_\mu + ig_4 H_\mu^\alpha T^{\alpha*} - ig_3 G'_\mu^a T^a + \frac{1}{6}ig_1 B'_\mu) \Omega_3, \\ D_\mu \Omega_{15} &= \partial_\mu \Omega_{15} - ig_4 [T^\alpha, \Omega_{15}] H_\mu^\alpha. \end{aligned}$$

We define the index  $A = 9, \dots, 14$  to span the  $SU(4)/(SU(3)_4 \times U(1)_4)$  coset. Neglecting electroweak symmetry breaking effects, the gauge boson masses are extracted from the canonically normalized kinetic terms of the scalar fields:

$$\begin{aligned} \mathcal{L} \supset &+ \frac{1}{2} \left( g_4^2 v_1^2 + g_4^2 v_3^2 + \frac{4}{3} g_4^2 v_{15}^2 \right) H_\mu^A H^{\mu A} \\ &+ \frac{v_3^2}{4} (H_\mu^a \quad G'_\mu^a) \begin{pmatrix} g_4^2 & -g_4 g_3 \\ -g_4 g_3 & g_3^2 \end{pmatrix} (H^{b\mu} \quad G'^{b\mu}) \\ &+ \frac{3v_1^2 + v_3^2}{4} (H_\mu^{15} \quad B'_\mu) \begin{pmatrix} \frac{g_4^2}{4} & -\frac{g_4 g_1}{2\sqrt{6}} \\ -\frac{g_4 g_1}{2\sqrt{6}} & \frac{g_1^2}{6} \end{pmatrix} (H^{15\mu} \quad B'^{\mu}). \end{aligned} \tag{C.7}$$

Diagonalizing these mass matrices, we obtain the massive gauge boson spectrum:

$$U_\mu^{1,2,3} = \frac{1}{\sqrt{2}} (H_\mu^{9,11,13} - i H_\mu^{10,12,14}), \quad M_U^2 = \frac{1}{4} g_4^2 \left( v_1^2 + v_3^2 + \frac{4}{3} v_{15}^2 \right), \tag{C.8}$$

$$g'_\mu^a = \frac{g_4 H_\mu^a - g_3 G'_\mu^a}{\sqrt{g_4^2 + g_3^2}}, \quad M_g^2 = \frac{1}{2} (g_4^2 + g_3^2) v_3^2, \tag{C.9}$$

$$Z'_\mu = \frac{g_4 H_\mu^{15} - \sqrt{\frac{2}{3}} g_1 B'_\mu}{\sqrt{g_4^2 + \frac{2}{3} g_1^2}}, \quad M_Z^2 = \frac{1}{4} \left( g_4^2 + \frac{2}{3} g_1^2 \right) \left( v_1^2 + \frac{1}{3} v_3^2 \right). \tag{C.10}$$

The combinations orthogonal to (C.9) and (C.10) correspond to the massless  $SU(3)_c \times U(1)_Y$  gauge bosons of  $\mathcal{G}_{321}$  prior to electroweak symmetry breaking:

$$g_\mu^a = \frac{g_3 H_\mu^a + g_4 G'_\mu^a}{\sqrt{g_4^2 + g_3^2}}, \quad (C.11)$$

$$B_\mu = \frac{\sqrt{\frac{2}{3}} g_1 H_\mu^{15} + g_4 B'_\mu}{\sqrt{g_4^2 + \frac{2}{3} g_1^2}}. \quad (C.12)$$

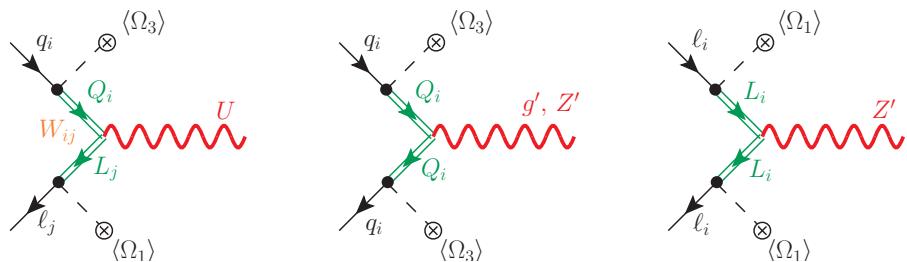
The matching between the fundamental couplings  $g_4, g_3, g_1$  and the SM couplings  $g_s, g_Y$  is readily obtained by acting with the covariant derivative on a field which transforms trivially under  $SU(4)$ . This yields:

$$g_s = \frac{g_4 g_3}{\sqrt{g_4^2 + g_3^2}}, \quad (C.13)$$

$$g_Y = \frac{g_4 g_1}{\sqrt{g_4^2 + \frac{2}{3} g_1^2}}. \quad (C.14)$$

Evolving the SM gauge couplings up to  $\mu = 2$  TeV, we obtain  $g_s = 1.02$  and  $g_Y = 0.363$ . Since  $g_s \leq g_{4,3}$  and  $g_Y \leq \sqrt{\frac{3}{2}} g_4, g_1$ , the hierarchy  $g_s \gg g_Y$  also implies  $g_{4,3} \gg g_Y \simeq g_1$ . In the limit  $v_3 \gg v_1 \gg v_{15}$ , the mass spectrum simplifies. For example, if the gauge couplings also satisfy  $g_4 \sim g_3$ , one finds  $M_{g'} \simeq \sqrt{2} M_U$  and  $M_{Z'} \simeq \frac{1}{\sqrt{2}} M_U$ .

In the 4321 model, the observed SM fermion masses and mixings arise from the mixing between elementary chiral fermions—charged under  $SU(3)' \times SU(2)_L \times U(1)'$  with SM-like quantum numbers—and three generations of vector-like fermions transforming as fundamentals of  $SU(4)$ . This mixing is triggered once the scalars  $\Omega_1$  and  $\Omega_3$  acquire VEVs (see Figure C.1). The full matter content of the model is summarized in Table C.1.



**Figure C.1:** Diagrammatic representation of the interactions between the SM fermions and the heavy vector-like fermions  $\Psi$ , induced by the Yukawa couplings to  $\Omega_1$  and  $\Omega_3$  after symmetry breaking.

Field	SU(4)	SU(3)'	SU(2) <sub>L</sub>	U(1)'
q_L <sup>i</sup>	<b>1</b>	<b>3</b>	<b>2</b>	1/6
u_R <sup>i</sup>	<b>1</b>	<b>3</b>	<b>1</b>	2/3
d_R <sup>i</sup>	<b>1</b>	<b>3</b>	<b>1</b>	-1/3
l_L <sup>i</sup>	<b>1</b>	<b>1</b>	<b>2</b>	-1/2
e_R <sup>i</sup>	<b>1</b>	<b>1</b>	<b>1</b>	-1
Ψ_L <sup>i</sup>	<b>4</b>	<b>1</b>	<b>2</b>	0
Ψ_R <sup>i</sup>	<b>4</b>	<b>1</b>	<b>2</b>	0
H	<b>1</b>	<b>1</b>	<b>2</b>	1/2
Ω <sub>1</sub>	<b>4</b>	<b>1</b>	<b>1</b>	-1/2
Ω <sub>3</sub>	<b>4</b>	<b>3</b>	<b>1</b>	1/6
Ω <sub>15</sub>	<b>15</b>	<b>1</b>	<b>1</b>	0

**Table C.1:** Field content of the 4321 model. The index  $i = 1, 2, 3$  runs over generations.

The mixing between the elementary fermions and the vector-like fermions is described by the Yukawa Lagrangian  $\mathcal{L}_Y = \mathcal{L}_{\text{SM-like}} + \mathcal{L}_{\text{mix}}$ , where

$$\mathcal{L}_{\text{SM-like}} = -\bar{q}'_L Y_d H d'_R - \bar{q}'_L Y_u \tilde{H} u'_R - \bar{l}'_L Y_e H e'_R + \text{h.c.}, \quad (\text{C.15})$$

$$\mathcal{L}_{\text{mix}} = -\bar{q}'_L \lambda_q \Omega_3^T \Psi_R - \bar{l}'_L \lambda_\ell \Omega_1^T \Psi_R - \bar{\Psi}_L (M + \lambda_{15} \Omega_{15}) \Psi_R + \text{h.c..} \quad (\text{C.16})$$

Here,  $\tilde{H} = i\sigma_2 H^*$ , and  $Y_{u,d,e}, \lambda_{q,\ell,15}, M$  are  $3 \times 3$  matrices in flavour space.

The vector-like fermions transform under  $\mathcal{G}_{4321}$  as

$$\Psi_{L,R} = \begin{pmatrix} Q'_{L,R} \\ L'_{L,R} \end{pmatrix} \sim (\mathbf{4}, \mathbf{1}, \mathbf{2}, 0). \quad (\text{C.17})$$

Under the breaking  $\text{SU}(4) \rightarrow \text{SU}(3)_{[4]} \times \text{U}(1)_{B-L}$ , they decompose as  $Q'_{L,R} \sim (\mathbf{3}, \mathbf{2}, 1/6)$  and  $L'_{L,R} \sim (\mathbf{1}, \mathbf{2}, -1/2)$ . Their vector-like masses are generated by the  $M$  term and are split by the VEV of  $\Omega_{15}$ :

$$M_Q = M + \frac{\lambda_{15} v_{15}}{2\sqrt{6}}, \quad M_L = M - \frac{3\lambda_{15} v_{15}}{2\sqrt{6}}. \quad (\text{C.18})$$

To comply with flavour constraints, the authors on [138] employ the following Yukawa textures as a starting point

$$\lambda_q = \hat{\lambda}_q \equiv \text{diag}(\lambda_{12}^q, \lambda_{12}^q, \lambda_3^q),$$

$$\lambda_\ell = \hat{\lambda}_\ell W^\dagger \equiv \text{diag}(\lambda_1^\ell, \lambda_2^\ell, \lambda_3^\ell) \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_{LQ} & -\sin \theta_{LQ} \\ 0 & \sin \theta_{LQ} & \cos \theta_{LQ} \end{pmatrix}, \quad (\text{C.19})$$

$$\lambda_{15} \propto \hat{M} \propto \mathbb{1}.$$

After the  $\text{SU}(3)_{[4]} \times \text{SU}(3)' \rightarrow \text{SU}(3)_c$  symmetry breaking, the  $6 \times 6$  fermion mass matrices for the quarks read:

$$\mathcal{M}_u = \begin{pmatrix} V^\dagger \hat{Y}_u \frac{v}{\sqrt{2}} & \hat{\lambda}_q \frac{v_3}{\sqrt{2}} \\ 0 & \hat{M}_Q \end{pmatrix}, \quad \mathcal{M}_d = \begin{pmatrix} \hat{Y}_d \frac{v}{\sqrt{2}} & \hat{\lambda}_q \frac{v_3}{\sqrt{2}} \\ 0 & \hat{M}_Q \end{pmatrix}. \quad (\text{C.20})$$

Similarly, after the  $U(1)_{B-L} \times U(1)_{T_R^3} \rightarrow U(1)_Y$  symmetry breaking, the  $6 \times 6$  fermion mass matrices for the leptons read:

$$\mathcal{M}_N = \begin{pmatrix} 0 & \hat{\lambda}_\ell \frac{v_1}{\sqrt{2}} \\ 0 & \hat{M}_L \end{pmatrix}, \quad \mathcal{M}_e = \begin{pmatrix} \hat{Y}_e \frac{v}{\sqrt{2}} & \hat{\lambda}_\ell W^\dagger \frac{v_1}{\sqrt{2}} \\ 0 & \hat{M}_L \end{pmatrix}. \quad (C.21)$$

Here,  $\hat{Y}_{u,d,e}$  and  $\hat{\lambda}_{q,\ell}$  are diagonal matrices,  $V$  and  $W$  are unitary Cabibbo-like mixing matrices, and  $M_Q, M_L$  are proportional to the identity matrix.

The structure of the mass matrices in Eqs. (C.21) allows them to be diagonalized by unitary transformations of the form  $\psi'_x = U_x \psi_x$ , where  $\psi_x$  ( $x = q, u, d, \ell, e, N$ ) denotes a 6-dimensional vector containing both the chiral and vector-like fermions, and the unprimed fields represent the mass eigenstates.

The chosen flavour structure in Eq. (C.19) ensures that in the limit  $W \rightarrow \mathbb{1}$ , the mixing is family-specific: each vector-like fermion generation mixes predominantly with only one generation of chiral fermions (up to CKM rotations). At leading order, the unitary mixing matrices are given by:

$$\begin{aligned} U_q &\approx \mathcal{R}_{14}(\theta_{q_1}) \mathcal{R}_{25}(\theta_{q_2}) \mathcal{R}_{36}(\theta_{q_3}), & U_\ell &\approx \mathcal{R}_{14}(\theta_{\ell_1}) \mathcal{R}_{25}(\theta_{\ell_2}) \mathcal{R}_{36}(\theta_{\ell_3}), \\ U_u &\approx \mathcal{R}_{14}(\theta_{u_R}) \mathcal{R}_{25}(\theta_{c_R}) \mathcal{R}_{36}(\theta_{t_R}), & U_e &\approx \begin{pmatrix} 1 & 0 \\ 0 & W \end{pmatrix} \mathcal{R}_{14}(\theta_{e_R}) \mathcal{R}_{25}(\theta_{\mu_R}) \mathcal{R}_{36}(\theta_{\tau_R}), \\ U_d &\approx \mathcal{R}_{14}(\theta_{d_R}) \mathcal{R}_{25}(\theta_{s_R}) \mathcal{R}_{36}(\theta_{b_R}), & U_N &\approx \begin{pmatrix} 0 & 0 \\ 0 & W \end{pmatrix}. \end{aligned}$$

Here, we have adopted a flavour basis for the SM  $SU(2)_L$  fermion multiplets defined by:

$$q^i = \begin{pmatrix} V_{ji}^* u_L^j \\ d_L^i \end{pmatrix}, \quad \ell^\alpha = \begin{pmatrix} v_L^\alpha \\ e_L^\alpha \end{pmatrix}, \quad (C.22)$$

where  $V$  is the CKM matrix. The mixing angles are related to the Lagrangian parameters by:

$$\begin{aligned} \sin \theta_{q_i} &= \frac{\lambda_i^q v_3}{\sqrt{|\lambda_i^q|^2 v_3^2 + 2\hat{M}_Q^2}}, & \cos \theta_{q_i} &= \frac{\sqrt{2}\hat{M}_Q}{\sqrt{|\lambda_i^q|^2 v_3^2 + 2\hat{M}_Q^2}}, \\ \sin \theta_{\ell_i} &= \frac{\lambda_i^\ell v_1}{\sqrt{|\lambda_i^\ell|^2 v_1^2 + 2\hat{M}_L^2}}, & \cos \theta_{\ell_i} &= \frac{\sqrt{2}\hat{M}_L}{\sqrt{|\lambda_i^\ell|^2 v_1^2 + 2\hat{M}_L^2}}, \\ \sin \theta_{u_R^i} &= \frac{m_{u_i}}{M_{Q_i}} \tan \theta_{q_i}, & \sin \theta_{d_R^i} &= \frac{m_{d_i}}{M_{Q_i}} \tan \theta_{q_i}, \\ \sin \theta_{e_R^i} &= \frac{m_{e_i}}{M_{L_i}} \tan \theta_{\ell_i}, & \cos \theta_{f_R^i} &= 1 \quad (f = u, d, e). \end{aligned} \quad (C.23)$$

In these expressions,  $m_i$  and  $M_i$  denote the physical fermion masses. Up to corrections of  $\mathcal{O}(m_i^2/M_i^2)$ , these are given by:

$$\begin{aligned} M_{L_i} &= \sqrt{\frac{|\lambda_i^\ell|^2 v_1^2}{2} + \hat{M}_L^2}, & M_{Q_i} &= \sqrt{\frac{|\lambda_i^q|^2 v_3^2}{2} + \hat{M}_Q^2}, \\ m_{f_i} &\approx |\hat{Y}_f| \cos \theta_{f_i} \frac{v}{\sqrt{2}} \quad (f = u, d, e). \end{aligned} \quad (C.24)$$

The interaction terms of the massive gauge bosons with the fermions in the interaction basis are derived from the covariant derivative. For the left-handed fields, we find:

$$\begin{aligned} \mathcal{L}_L &= \frac{g_4}{\sqrt{2}} \bar{Q}'_L \gamma^\mu L'_L U_\mu + \text{h.c.} \\ &+ g_s \left( \frac{g_4}{g_3} \bar{Q}'_L \gamma^\mu T^a Q'_L - \frac{g_3}{g_4} \bar{q}'_L \gamma^\mu T^a q'_L \right) g'_\mu^a \\ &+ g_Y \left( \sqrt{\frac{3}{2}} \frac{g_4}{g_1} Y(Q'_L) \bar{Q}'_L \gamma^\mu Q'_L - \sqrt{\frac{2}{3}} \frac{g_1}{g_4} Y(q'_L) \bar{q}'_L \gamma^\mu q'_L \right) Z'_\mu \\ &+ g_Y \left( \sqrt{\frac{3}{2}} \frac{g_4}{g_1} Y(L'_L) \bar{L}'_L \gamma^\mu L'_L - \sqrt{\frac{2}{3}} \frac{g_1}{g_4} Y(\ell'_L) \bar{\ell}'_L \gamma^\mu \ell'_L \right) Z'_\mu, \end{aligned} \quad (\text{C.25})$$

and for the right-handed fields:

$$\begin{aligned} \mathcal{L}_R &= \frac{g_4}{\sqrt{2}} \bar{Q}'_R \gamma^\mu L'_R U_\mu + \text{h.c.} \\ &+ g_s \left( \frac{g_4}{g_3} \bar{Q}'_R \gamma^\mu T^a Q'_R - \frac{g_3}{g_4} \bar{u}'_R \gamma^\mu T^a u'_R - \frac{g_3}{g_4} \bar{d}'_R \gamma^\mu T^a d'_R \right) g'_\mu^a \\ &+ g_Y \left( \sqrt{\frac{3}{2}} \frac{g_4}{g_1} Y(Q'_R) \bar{Q}'_R \gamma^\mu Q'_R - \sqrt{\frac{2}{3}} \frac{g_1}{g_4} Y(u'_R) \bar{u}'_R \gamma^\mu u'_R - \sqrt{\frac{2}{3}} \frac{g_1}{g_4} Y(d'_R) \bar{d}'_R \gamma^\mu d'_R \right) Z'_\mu \\ &+ g_Y \left( \sqrt{\frac{3}{2}} \frac{g_4}{g_1} Y(L'_R) \bar{L}'_R \gamma^\mu L'_R - \sqrt{\frac{2}{3}} \frac{g_1}{g_4} Y(\ell'_R) \bar{\ell}'_R \gamma^\mu \ell'_R \right) Z'_\mu. \end{aligned} \quad (\text{C.26})$$

The SM hypercharges are:  $Y(Q'_L) = Y(Q'_R) = Y(q'_L) = \frac{1}{6}$ ,  $Y(u'_R) = \frac{2}{3}$ ,  $Y(d'_R) = -\frac{1}{3}$ ,  $Y(L'_L) = Y(L'_R) = Y(\ell'_L) = -\frac{1}{2}$ , and  $Y(e'_R) = -1$ .

To express the interactions in Eqs. (C.25) and (C.26) in the fermion mass basis, we collect the fields into 6-dimensional multiplets,  $\psi_x$  ( $x = q, u, d, \ell, e$ ), and apply the corresponding unitary transformations  $U_x$ . Neglecting the right-handed rotations—which are suppressed by the small masses of the SM fermions—we obtain:

$$\begin{aligned} \mathcal{L}_U &= \frac{g_4}{\sqrt{2}} U_\mu [\beta \bar{\Psi}_q \gamma^\mu \psi_\ell + W \bar{Q}_R \gamma^\mu L_R] + \text{h.c.}, \\ \mathcal{L}_{g'} &= g_s \frac{g_4}{g_3} g'_\mu^a [\kappa_q \bar{\Psi}_q \gamma^\mu T^a \psi_q + \kappa_u \bar{\Psi}_u \gamma^\mu T^a \psi_u + \kappa_d \bar{\Psi}_d \gamma^\mu T^a \psi_d], \\ \mathcal{L}_{Z'} &= \frac{g_Y}{2\sqrt{6}} \frac{g_4}{g_1} Z'_\mu [\zeta_q \bar{\Psi}_q \gamma^\mu \psi_q + \zeta_u \bar{\Psi}_u \gamma^\mu \psi_u \\ &\quad + \zeta_d \bar{\Psi}_d \gamma^\mu \psi_d - 3\zeta_\ell \bar{\Psi}_\ell \gamma^\mu \psi_\ell - 3\zeta_e \bar{\Psi}_e \gamma^\mu \psi_e]. \end{aligned} \quad (\text{C.27})$$

The coupling matrices are defined as follows (with indices  $A, B = 4, 5, 6$  spanning the heavy vector-like states, and  $\alpha, \beta = 1, \dots, 6$  spanning the full 6-dimensional space):

$$\begin{aligned} \beta^{\alpha\beta} &= [U_q]_{A\alpha}^* [W]_{AB} [U_\ell]_{B\beta}, \\ \kappa_q^{\alpha\beta} &= [U_q]_{A\alpha}^* [U_q]_{A\beta} - \frac{g_3^2}{g_4^2} \delta_{\alpha\beta}, \quad \kappa_u \approx \kappa_d \approx \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{1}_{3\times 3} \end{pmatrix} - \frac{g_3^2}{g_4^2} \mathbb{1}_{6\times 6}, \\ \zeta_q^{\alpha\beta} &= [U_q]_{A\alpha}^* [U_q]_{A\beta} - \frac{2g_1^2}{3g_4^2} \delta_{\alpha\beta}, \quad \zeta_u \approx \zeta_d \approx \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{1}_{3\times 3} \end{pmatrix} - \frac{2g_1^2}{3g_4^2} \mathbb{1}_{6\times 6}, \\ \zeta_\ell^{\alpha\beta} &= [U_\ell]_{A\alpha}^* [U_\ell]_{A\beta} - \frac{2g_1^2}{3g_4^2} \delta_{\alpha\beta}, \quad \zeta_e \approx \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{1}_{3\times 3} \end{pmatrix} - \frac{2g_1^2}{3g_4^2} \mathbb{1}_{6\times 6}. \end{aligned} \tag{C.28}$$

A key result of the assumed flavour structure is that the matrix  $W$  cancels due to unitarity in the  $Z'$  and  $g'$  interactions. This cancellation is crucial for suppressing unwanted flavour-changing neutral currents (FCNCs) in these sectors.

Assuming  $W = \mathcal{R}_{56}(\theta_{LQ})$  and no CP violation in the mixing angles, the left-handed coupling matrices can be explicitly written as:

$$\beta \approx \begin{pmatrix} \text{diag}(s_{q_1}s_{\ell_1}, s_{q_2}s_{\ell_2}, s_{q_3}s_{\ell_3})W & -\text{diag}(s_{q_1}c_{\ell_1}, s_{q_2}c_{\ell_2}, s_{q_3}c_{\ell_3})W \\ \text{diag}(c_{q_1}s_{\ell_1}, c_{q_2}s_{\ell_2}, c_{q_3}s_{\ell_3})W & -\text{diag}(c_{q_1}c_{\ell_1}, c_{q_2}c_{\ell_2}, c_{q_3}c_{\ell_3})W \end{pmatrix}, \tag{C.29}$$

$$\kappa_q \approx \begin{pmatrix} \text{diag}(s_{q_1}^2, s_{q_2}^2, s_{q_3}^2) & -\frac{1}{2}\text{diag}(s_{2q_1}, s_{2q_2}, s_{2q_3}) \\ -\frac{1}{2}\text{diag}(s_{2q_1}, s_{2q_2}, s_{2q_3}) & \text{diag}(c_{q_1}^2, c_{q_2}^2, c_{q_3}^2) \end{pmatrix} - \frac{g_3^2}{g_4^2} \mathbb{1}_{6\times 6}, \tag{C.30}$$

$$\zeta_q \approx \begin{pmatrix} \text{diag}(s_{q_1}^2, s_{q_2}^2, s_{q_3}^2) & -\frac{1}{2}\text{diag}(s_{2q_1}, s_{2q_2}, s_{2q_3}) \\ -\frac{1}{2}\text{diag}(s_{2q_1}, s_{2q_2}, s_{2q_3}) & \text{diag}(c_{q_1}^2, c_{q_2}^2, c_{q_3}^2) \end{pmatrix} - \frac{2g_1^2}{3g_4^2} \mathbb{1}_{6\times 6}, \tag{C.31}$$

$$\zeta_\ell \approx \begin{pmatrix} \text{diag}(s_{\ell_1}^2, s_{\ell_2}^2, s_{\ell_3}^2) & -\frac{1}{2}\text{diag}(s_{2\ell_1}, s_{2\ell_2}, s_{2\ell_3}) \\ -\frac{1}{2}\text{diag}(s_{2\ell_1}, s_{2\ell_2}, s_{2\ell_3}) & \text{diag}(c_{\ell_1}^2, c_{\ell_2}^2, c_{\ell_3}^2) \end{pmatrix} - \frac{2g_1^2}{3g_4^2} \mathbb{1}_{6\times 6}. \tag{C.32}$$

Following the flavour structure in [62], the assumption of a single  $U(2)_q$  breaking spurion in both the leptoquark and SM Yukawa couplings implies the relation  $\beta_L^{13} = V_{td}^*/V_{ts}^* \beta_L^{23}$ . More generally, from  $U(2)$  symmetries acting on both quark and lepton sectors, we expect the hierarchy:

$$|\beta_L^{31}| \ll |\beta_L^{23}|, |\beta_L^{32}| \ll |\beta_R^{33}|, |\beta_L^{33}| = \mathcal{O}(1),$$

and analogously for the  $\zeta_{\ell,e,Q}^{ij}$  and  $\kappa_Q^{ij}$  couplings.

This structure can be achieved with a specific choice of the  $\lambda$  parameters in the potential, leading to particular values for the mixing angles  $\theta_{q_i}$  and  $\theta_{\ell_i}$ . To explain the B-physics anomalies, a large mixing angle  $\theta_{\ell_3}$  is required, while the other two angles  $\theta_{\ell_{1,2}}$  must be small to avoid large contributions to muon and electron observables. The quark mixing angles

$\theta_{q_i}$  should be small for the first two generations to avoid large contributions to meson mixing observables, while a moderate value of  $\theta_{q_3}$  is needed to explain the B-physics anomalies.



---

## BIBLIOGRAPHY

---

- [1] A. Flórez et al. "On the sensitivity reach of LQ production with preferential couplings to third generation fermions at the LHC". In: *The European Physical Journal C* 83.11 (Nov. 2023), p. 1023. ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-023-12177-4](https://doi.org/10.1140/epjc/s10052-023-12177-4).
- [2] Umar Sohail Qureshi et al. "Probing light scalars and vector-like quarks at the high-luminosity LHC". In: *Eur. Phys. J. C* 85.4 (2025), p. 379. DOI: [10.1140/epjc/s10052-025-14085-1](https://doi.org/10.1140/epjc/s10052-025-14085-1). arXiv: [2410.17854 \[hep-ph\]](https://arxiv.org/abs/2410.17854).
- [3] W Greiner. *Relativistic Quantum Mechanics. Wave Equations*. Springer, 2000. ISBN: 9783540674573.
- [4] Stefan Pokorski. *Gauge field theories*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2000. ISBN: 9780521478168.
- [5] David Tong. *Quantum Field Theory University of Cambridge Part III Mathematical Tripos*. Vol. 45. 12. 1995.
- [6] Cristian Rodriguez. "Evaluacion del Lagrangiano de Supergravedad  $\mathcal{N} = 1$  utilizando el formalismo Superconforme." BA thesis. UPTC, 2018.
- [7] Herbert Goldstein, Charles P Poole, and John L Safko. *Classical mechanics*. 3rd ed. Addison Wesley, 2001. ISBN: 9780201657029,0201657023.
- [8] J V José and E J Saletan. *Classical Dynamics: A Contemporary Approach*. Cambridge University Press, 1998. ISBN: 9780521636360.
- [9] Roberto Martínez. *Teoría cuántica de campos*. Colección Textos. Universidad Nacional de Colombia, Facultad de Ciencias, Departamento de Física, 2002. ISBN: 9789587011203.
- [10] Steven Weinberg. *The Quantum Theory of Fields, Volume 3: Supersymmetry*. Cambridge University Press, 2005. ISBN: 0521670551,9780521670555.
- [11] D.A. Bromley Walter Greiner Joachim Reinhardt. *Field quantization*. 1st ed. Springer, 1996. ISBN: 3540591796,9783540591795.
- [12] Dan V. Schroeder Michael E. Peskin. *An Introduction To Quantum Field Theory*. Frontiers in Physics. Westview Press, 1995. ISBN: 9780201503975.
- [13] P C W Davies N. D. Birrell. *Quantum fields in curved space*. CUP. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1984. ISBN: 9780521278584,0521278589.
- [14] Amitabha Lahiri and Palash B Pal. *A First Book of Quantum Field Theory*. Alpha Science International, 2006, p. 380. ISBN: 1842652494.

- [15] Reinhold A. Bertlmann. *Anomalies in Quantum Field Theory*. Oxford University Press, 1996. ISBN: 9780198520476.
- [16] Daniel Z. Freedman and Antoine Van Proeyen. *Supergravity*. Cambridge University Press, 2012. ISBN: 9780521194013.
- [17] Diego Gallego. *Estudio de uplifts efectivos en compactificaciones de supercuerdas tipo-IIB*. Tech. rep. Tunja: Universidad Pedagógica y Tecnológica de Colombia, 2016, p. 96.
- [18] Antoine Van Proeyen. “Tools for Supersymmetry”. In: *Physics AUC* 9.I (1999), pp. 1–49. arXiv: 9910030 [hep-th].
- [19] Stephen P. Martin. “TASI 2011 lectures notes: two-component fermion notation and supersymmetry”. In: *arXiv* (2012). arXiv: 1205.4076.
- [20] M Robinson. *Symmetry and the Standard Model: Mathematics and Particle Physics*. 1st ed. SpringerLink : Bucher. New York: Springer-Verlag, 2011. ISBN: 9781441982674. DOI: 10.1007/978-1-4419-8267-4.
- [21] “Physics: 1963–1970”. In: Nobel lectures, including presentation speeches and laureates’ biographies. Elsevier Publishing Company, 1972.
- [22] F. J. Dyson. “The Radiation Theories of Tomonaga, Schwinger, and Feynman”. In: *Phys. Rev.* 75 (3 1949), pp. 486–502. DOI: 10.1103/PhysRev.75.486.
- [23] P.A. Zyla et al. “Review of Particle Physics”. In: *PTEP* 2020.8 (2020), p. 083C01. DOI: 10.1093/ptep/ptaa104.
- [24] C. S. Wu et al. “Experimental Test of Parity Conservation in Beta Decay”. In: *Phys. Rev.* 105 (4 1957), pp. 1413–1415. DOI: 10.1103/PhysRev.105.1413.
- [25] Tadao Nakano and Kazuhiko Nishijima. “Charge Independence for V-particles\*”. In: *Progress of Theoretical Physics* 10.5 (Nov. 1953), pp. 581–582. ISSN: 0033-068X. DOI: 10.1143/PTP.10.581.
- [26] B. Odom et al. “New Measurement of the Electron Magnetic Moment Using a One-Electron Quantum Cyclotron”. In: *Phys. Rev. Lett.* 97 (3 2006), p. 030801. DOI: 10.1103/PhysRevLett.97.030801.
- [27] Steven Weinberg. “A Model of Leptons”. In: *Phys. Rev. Lett.* 19 (21 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.
- [28] Sheldon L. Glashow. “Partial-symmetries of weak interactions”. In: *Nuclear Physics* 22.4 (1961), pp. 579–588. ISSN: 0029-5582. DOI: [https://doi.org/10.1016/0029-5582\(61\)90469-2](https://doi.org/10.1016/0029-5582(61)90469-2).
- [29] Alejandro Rivero and Andre Gsponer. *The strange formula of Dr. Koide*. 2005. DOI: 10.48550/ARXIV.HEP-PH/0505220.
- [30] Carlo Rovelli. *Quantum gravity*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2007. ISBN: 9780521715966,0521715962.

- [31] Ronald J. Adler, Brendan Casey, and Ovid C. Jacob. "Vacuum catastrophe: An elementary exposition of the cosmological constant problem". In: *American Journal of Physics* 63 (7 July 1995), pp. 620–626. ISSN: 0002-9505. DOI: [10.1119/1.17850](https://doi.org/10.1119/1.17850).
- [32] A. N. Lasenby M. P. Hobson G. P. Efstathiou. *General relativity: an introduction for physicists*. Cambridge University Press, 2005. ISBN: 9780521536394.
- [33] A.D. Sakharov. "Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe". In: *Soviet Journal of Experimental and Theoretical Physics Letters* 5 (Jan. 1967), p. 24.
- [34] G. 't Hooft. "Symmetry Breaking through Bell-Jackiw Anomalies". In: *Phys. Rev. Lett.* 37 (1 1976), pp. 8–11. DOI: [10.1103/PhysRevLett.37.8](https://doi.org/10.1103/PhysRevLett.37.8).
- [35] James M. Cline. *Baryogenesis*. 2006. DOI: [10.48550 / ARXIV .HEP-PH/0609145](https://arxiv.org/abs/0609145).
- [36] S. Abe et al. "Precision Measurement of Neutrino Oscillation Parameters with KamLAND". In: *Physical Review Letters* 100.22 (2008). DOI: [10.1103/physrevlett.100.221803](https://doi.org/10.1103/physrevlett.100.221803).
- [37] J. P. Lees et al. "Evidence for an Excess of  $\bar{B} \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$  Decays". In: *Phys. Rev. Lett.* 109 (10 2012), p. 101802. DOI: [10.1103/PhysRevLett.109.101802](https://doi.org/10.1103/PhysRevLett.109.101802).
- [38] R. Aaij et al. "Measurement of the Ratio of Branching Fractions  $\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\mu^-\bar{\nu}_\mu)$ ". In: *Phys. Rev. Lett.* 115 (11 2015), p. 111803. DOI: [10.1103/PhysRevLett.115.111803](https://doi.org/10.1103/PhysRevLett.115.111803).
- [39] Wolfgang Altmannshofer and David M. Straub. "New physics in  $b \rightarrow s$  transitions after LHC run 1". In: *The European Physical Journal C* 75.8 (2015). DOI: [10.1140/epjc/s10052-015-3602-7](https://doi.org/10.1140/epjc/s10052-015-3602-7).
- [40] T. Hurth, F. Mahmoudi, and S. Neshatpour. "On the anomalies in the latest LHCb data". In: *Nuclear Physics B* 909 (2016), pp. 737–777. DOI: [10.1016/j.nuclphysb.2016.05.022](https://doi.org/10.1016/j.nuclphysb.2016.05.022).
- [41] LHCb Collaboration et al. *Test of lepton universality in beauty-quark decays*. 2021. DOI: [10.48550/ARXIV.2103.11769](https://arxiv.org/abs/2103.11769).
- [42] Bernat Capdevila et al. "Patterns of New Physics in  $b \rightarrow s\ell^+\ell^-$  transitions in the light of recent data". In: *Journal of High Energy Physics* 2018.1 (2018). DOI: [10.1007/jhep01\(2018\)093](https://doi.org/10.1007/jhep01(2018)093).
- [43] Thomas Blum et al. *The Muon ( $g-2$ ) Theory Value: Present and Future*. 2013. DOI: [10.48550/ARXIV.1311.2198](https://arxiv.org/abs/1311.2198).
- [44] B. Abi et al. "Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm". In: *Physical Review Letters* 126.14 (2021). DOI: [10.1103/physrevlett.126.141801](https://doi.org/10.1103/physrevlett.126.141801).
- [45] CDF Collaboration et al. "High-precision measurement of the  $W$  boson mass with the CDF II detector". In: *Science* 376.6589 (2022), pp. 170–176. DOI: [10.1126/science.abk1781](https://doi.org/10.1126/science.abk1781).

- [46] M. Aaboud et al. "Measurement of the W-boson mass in pp collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector". In: *The European Physical Journal C* 78.2 (2018). doi: [10.1140/epjc/s10052-017-5475-4](https://doi.org/10.1140/epjc/s10052-017-5475-4).
- [47] and R. Aaij et al. "Measurement of the W boson mass". In: *Journal of High Energy Physics* 2022.1 (2022). doi: [10.1007/jhep01\(2022\)036](https://doi.org/10.1007/jhep01(2022)036).
- [48] S. Schael et al. "Measurement of the W boson mass and width in e+e- collisions at LEP". In: *The European Physical Journal C* 47.2 (2006), pp. 309–335. doi: [10.1140/epjc/s2006-02576-8](https://doi.org/10.1140/epjc/s2006-02576-8).
- [49] V. M. Abazov et al. "Measurement of the W Boson Mass with the Do Detector". In: *Physical Review Letters* 108.15 (2012). doi: [10.1103/physrevlett.108.151804](https://doi.org/10.1103/physrevlett.108.151804).
- [50] J. C. Hardy and I. S. Towner. "Superallowed  $0^+ \rightarrow 0^+$  nuclear  $\beta$  decays: 2020 critical survey, with implications for  $V_{ud}$  and CKM unitarity". In: *Phys. Rev. C* 102 (4 2020), p. 045501. doi: [10.1103/PhysRevC.102.045501](https://doi.org/10.1103/PhysRevC.102.045501).
- [51] Andreas Crivellin and Martin Hoferichter. " $\beta$  Decays as Sensitive Probes of Lepton Flavor Universality". In: *Phys. Rev. Lett.* 125 (11 2020), p. 111801. doi: [10.1103/PhysRevLett.125.111801](https://doi.org/10.1103/PhysRevLett.125.111801).
- [52] The CMS collaboration. "Search for resonant and nonresonant new phenomena in high-mass dilepton final states at  $\sqrt{s}=13$  TeV". In: *Journal of High Energy Physics* 2021.7 (2021), p. 208. ISSN: 1029-8479. doi: [10.1007/JHEP07\(2021\)208](https://doi.org/10.1007/JHEP07(2021)208).
- [53] H.P. P Nilles. "Supersymmetry, supergravity and particle physics". In: *Physics Reports* 110.1-2 (1984), pp. 1–162. ISSN: 03701573. doi: [10.1016/0370-1573\(84\)90008-5](https://doi.org/10.1016/0370-1573(84)90008-5).
- [54] I. Doršner et al. "Physics of leptoquarks in precision experiments and at particle colliders". In: *Phys. Rept.* 641 (2016), pp. 1–68. doi: [10.1016/j.physrep.2016.06.001](https://doi.org/10.1016/j.physrep.2016.06.001). arXiv: [1603.04993 \[hep-ph\]](https://arxiv.org/abs/1603.04993).
- [55] Lisa Randall and Raman Sundrum. "A Large mass hierarchy from a small extra dimension". In: *Phys. Rev. Lett.* 83 (1999), pp. 3370–3373. doi: [10.1103/PhysRevLett.83.3370](https://doi.org/10.1103/PhysRevLett.83.3370). arXiv: [hep-ph/9905221](https://arxiv.org/abs/hep-ph/9905221).
- [56] G. F. Giudice. "Naturalness after LHC". In: *arXiv e-prints* (2013). arXiv: [1307.7879 \[hep-ph\]](https://arxiv.org/abs/1307.7879).
- [57] Gianfranco Bertone, Dan Hooper, and Joseph Silk. "Particle dark matter: evidence, candidates and constraints". In: *Physics Reports* 405 (2005), pp. 279–390. doi: [10.1016/j.physrep.2004.08.031](https://doi.org/10.1016/j.physrep.2004.08.031). arXiv: [hep-ph/0404175 \[hep-ph\]](https://arxiv.org/abs/hep-ph/0404175).
- [58] P. K. Mohapatra, R. N. Mohapatra, and P. B. Pal. "Implications of E6 Grand Unification". In: *Phys. Rev. D* 33 (1986), p. 2010. doi: [10.1103/PhysRevD.33.2010](https://doi.org/10.1103/PhysRevD.33.2010).

- [59] LHCb Collaboration. “Test of lepton universality using  $B^+ \rightarrow K^+\ell^+\ell^-$  decays”. In: *Phys. Rev. Lett.* 113 (2014), p. 151601. doi: [10.1103/PhysRevLett.113.151601](https://doi.org/10.1103/PhysRevLett.113.151601). arXiv: [1406.6482 \[hep-ex\]](https://arxiv.org/abs/1406.6482).
- [60] LHCb Collaboration. “Test of lepton universality with  $B^0 \rightarrow K^{*0}\ell^+\ell^-$  decays”. In: *JHEP* 08 (2017), p. 055. doi: [10.1007/JHEP08\(2017\)055](https://doi.org/10.1007/JHEP08(2017)055). arXiv: [1705.05802 \[hep-ex\]](https://arxiv.org/abs/1705.05802).
- [61] BaBar Collaboration. “Evidence for an excess of  $\bar{B} \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$  decays”. In: *Phys. Rev. Lett.* 109 (2012), p. 101802. doi: [10.1103/PhysRevLett.109.101802](https://doi.org/10.1103/PhysRevLett.109.101802). arXiv: [1205.5442 \[hep-ex\]](https://arxiv.org/abs/1205.5442).
- [62] D. Buttazzo et al. “B-physics anomalies: a guide to combined explanations”. In: *JHEP* 11 (2017), p. 044. doi: [10.1007/JHEP11\(2017\)044](https://doi.org/10.1007/JHEP11(2017)044). arXiv: [1706.07808 \[hep-ph\]](https://arxiv.org/abs/1706.07808).
- [63] J. Alwall et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In: *JHEP* 07 (2014), p. 079. doi: [10.1007/JHEP07\(2014\)079](https://doi.org/10.1007/JHEP07(2014)079). arXiv: [1405.0301 \[hep-ph\]](https://arxiv.org/abs/1405.0301).
- [64] A. Alloul et al. “FeynRules 2.0 - A complete toolbox for tree-level phenomenology”. In: *Comput. Phys. Commun.* 185 (2014), pp. 2250–2300. doi: [10.1016/j.cpc.2014.04.012](https://doi.org/10.1016/j.cpc.2014.04.012). arXiv: [1310.1921 \[hep-ph\]](https://arxiv.org/abs/1310.1921).
- [65] DELPHES 3 Collaboration. “DELPHES 3, A modular framework for fast simulation of a generic collider experiment”. In: *JHEP* 02 (2014), p. 057. doi: [10.1007/JHEP02\(2014\)057](https://doi.org/10.1007/JHEP02(2014)057). arXiv: [1307.6346 \[hep-ex\]](https://arxiv.org/abs/1307.6346).
- [66] Georges Aad et al. “Search for high-mass dilepton resonances using  $139 \text{ fb}^{-1}$  of  $pp$  collision data collected at  $\sqrt{s} = 13 \text{ TeV}$  with the ATLAS detector”. In: *Phys. Lett. B* 796 (2019), pp. 68–87. doi: [10.1016/j.physletb.2019.07.016](https://doi.org/10.1016/j.physletb.2019.07.016). arXiv: [1903.06248 \[hep-ex\]](https://arxiv.org/abs/1903.06248).
- [67] Albert M Sirunyan et al. “Search for resonant and nonresonant new phenomena in high-mass dilepton final states at  $\sqrt{s} = 13 \text{ TeV}$ ”. In: *JHEP* 07 (2021), p. 208. doi: [10.1007/JHEP07\(2021\)208](https://doi.org/10.1007/JHEP07(2021)208). arXiv: [2103.02708 \[hep-ex\]](https://arxiv.org/abs/2103.02708).
- [68] M. Ciuchini et al. “Constraints on lepton universality violation from rare B decays”. In: *Phys. Rev. D* 107 (5 2023), p. 055036. doi: [10.1103/PhysRevD.107.055036](https://doi.org/10.1103/PhysRevD.107.055036).
- [69] T. Sjöstrand et al. “An introduction to PYTHIA 8.2”. In: *Comput. Phys. Commun.* 191 (2015), pp. 159–177. doi: [10.1016/j.cpc.2015.01.024](https://doi.org/10.1016/j.cpc.2015.01.024). arXiv: [1410.3012 \[hep-ph\]](https://arxiv.org/abs/1410.3012).
- [70] M. Cacciari, G. P. Salam, and G. Soyez. “FastJet User Manual”. In: *Eur. Phys. J. C* 72 (2012), p. 1896. doi: [10.1140/epjc/s10052-012-1896-2](https://doi.org/10.1140/epjc/s10052-012-1896-2). arXiv: [1111.6097 \[hep-ph\]](https://arxiv.org/abs/1111.6097).
- [71] NNPDF Collaboration. “Parton distributions for the LHC Run II”. In: *JHEP* 04 (2015), p. 040. doi: [10.1007/JHEP04\(2015\)040](https://doi.org/10.1007/JHEP04(2015)040). arXiv: [1410.8849 \[hep-ph\]](https://arxiv.org/abs/1410.8849).

- [72] G. Aad et al. “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *JINST* 3 (2008), S08003. doi: [10.1088/1748-0221/3/08/S08003](https://doi.org/10.1088/1748-0221/3/08/S08003).
- [73] S. Chatrchyan et al. “The CMS Experiment at the CERN LHC”. In: *JINST* 3 (2008), S08004. doi: [10.1088/1748-0221/3/08/S08004](https://doi.org/10.1088/1748-0221/3/08/S08004).
- [74] Paul Lujan et al. *The Pixel Luminosity Telescope: A detector for luminosity measurement at CMS using silicon pixel sensors*. Tech. rep. CMS-DN-21-008. Accepted by Eur. Phys. J. C. 2022. arXiv: 2206.08870 [physics.ins-det]. URL: <https://arxiv.org/abs/2206.08870>.
- [75] F. Pedregosa et al. “Scikit-learn: Machine Learning in Python”. In: *J. Mach. Learn. Res.* 12 (2011), pp. 2825–2830. ISSN: 1532-4435. doi: [10.48550/arXiv.1201.0490](https://doi.org/10.48550/arXiv.1201.0490).
- [76] T. Chen and C. Guestrin. “XGBoost: A Scalable Tree Boosting System”. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD ’16. San Francisco, California, USA: ACM, 2016, pp. 785–794. ISBN: 978-1-4503-4232-2. doi: [10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785).
- [77] E. Bols et al. “Jet flavour classification using DeepJet”. In: *Journal of Instrumentation* 15.12 (2020), P12012. doi: [10.1088/1748-0221/15/12/P12012](https://doi.org/10.1088/1748-0221/15/12/P12012).
- [78] CMS Collaboration. “Identification of hadronic tau lepton decays using a deep neural network”. In: *JINST* 17.07 (2022), Po7023. doi: [10.1088/1748-0221/17/07/P07023](https://doi.org/10.1088/1748-0221/17/07/P07023). arXiv: 2201.08458 [hep-ex].
- [79] Lyndon Evans and Philip Bryant. “LHC Machine”. In: *Journal of Instrumentation* 3.08 (2008), S08001–S08001. doi: [10.1088/1748-0221/3/08/S08001](https://doi.org/10.1088/1748-0221/3/08/S08001).
- [80] ATLAS Collaboration. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Physics Letters B* 716 (2012). ATLAS Higgs-discovery paper; verify metadata if needed, pp. 1–29. doi: [10.1016/j.physletb.2012.08.020](https://doi.org/10.1016/j.physletb.2012.08.020).
- [81] CMS Collaboration. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: *Physics Letters B* 716 (2012). CMS Higgs-discovery paper; verify metadata if needed, pp. 30–61. doi: [10.1016/j.physletb.2012.08.021](https://doi.org/10.1016/j.physletb.2012.08.021).
- [82] CMS Collaboration. *Study of the Discovery Reach in Searches for Supersymmetry at CMS with 3000/fb*. CMS Physics Analysis Summary CMS-PAS-FTR-13-014. Accessed on January 07, 2023. CERN, 2013. URL: <https://cds.cern.ch/record/1607141>.
- [83] G. Apollinari et al. *High-Luminosity Large Hadron Collider (HL-LHC): Technical and performance projections*. Tech. rep. CERN Yellow Report / TBD. HL-LHC technical design / projection report; please verify exact report number/title and replace placeholder fields as desired. CERN, 2017.

- [84] D. Bertolini et al. "Pileup Per Particle Identification". In: *JHEP* 10 (2014). Introduces the PUPPI algorithm for per-particle pileup mitigation; ampliamente usado en ATLAS/CMS analyses., p. 59. arXiv: 1407.6013 [hep-ph].
- [85] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti-kt jet clustering algorithm". In: *Journal of High Energy Physics* 2008.04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063.
- [86] Izaak Neutelings. *CMS coordinate system*. URL: [https://tikz.net/axis3d\\_cms](https://tikz.net/axis3d_cms).
- [87] CMS Collaboration. "Particle-flow reconstruction and global event description with the CMS detector". In: *JINST* 12 (2017). Resumen del algoritmo de reconstrucción particle-flow utilizado en CMS; útil para la discusión de PF en el texto., P10003.
- [88] A.M. Sirunyan et al. "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC". In: *JINST* 16.05 (2021), Po5014. DOI: 10.1088/1748-0221/16/05/P05014.
- [89] A.M. Sirunyan et al. "Performance of the reconstruction and identification of high-momentum muons in proton-proton collisions at  $\sqrt{s} = 13$  TeV". In: *JINST* 15.02 (2020), Po2027. DOI: 10.1088/1748-0221/15/02/P02027.
- [90] The CMS Collaboration. "Description and performance of track and primary-vertex reconstruction with the CMS tracker". In: *Journal of Instrumentation* 9.10 (Oct. 2014), P10009–P10009. ISSN: 1748-0221. DOI: 10.1088/1748-0221/9/10/p10009.
- [91] John C. Collins, Davison E. Soper, and George F. Sterman. "Factorization of Hard Processes in QCD". In: *Adv. Ser. Direct. High Energy Phys.* 5 (1989), pp. 1–91. DOI: 10.1142/9789814503266\_0001. arXiv: hep-ph/0409313.
- [92] J. Butterworth, S. Carrazza, A. Cooper-Sarkar, A. De Roeck, J. Feltesse, S. Forte, J. Gao, S. Glazov, J. Huston, Z. Kassabov, R. McNulty, A. Morsch, P. Nadolsky, V. Radescu, J. Rojo and R. Thorne. "PDF4LHC recommendations for LHC Run II". In: *J. Phys. G* 43 (2016), p. 023001. DOI: 10.1088/0954-3899/43/2/023001. arXiv: 1510.03865 [hep-ph].
- [93] Albert M Sirunyan et al. "Performance of missing transverse momentum reconstruction in proton-proton collisions at  $\sqrt{s} = 13$  TeV using the CMS detector". In: *JINST* 14.07 (2019), Po7004. DOI: 10.1088/1748-0221/14/07/P07004. arXiv: 1903.06078 [hep-ex].
- [94] William R. Leo. *Techniques for Nuclear and Particle Physics Experiments*. Springer, 1994. ISBN: 978-3-540-57280-0, 978-3-642-57920-2. DOI: 10.1007/978-3-642-57920-2.
- [95] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J. C* 71 (2011). [Erratum: Eur.Phys.J.C 73, 2501 (2013)], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [physics.data-an].

- [96] Alexander L. Read. "Presentation of search results: The  $CL_s$  technique". In: *J. Phys. G* 28 (2002). Ed. by M. R. Whalley and L. Lyons, pp. 2693–2704. doi: [10.1088/0954-3899/28/10/313](https://doi.org/10.1088/0954-3899/28/10/313).
- [97] Wolfgang A. Rolke, Angel M. López, and Jan Conrad. "Limits and confidence intervals in the presence of nuisance parameters". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 551.2–3 (Oct. 2005), 493–503. ISSN: 0168-9002. doi: [10.1016/j.nima.2005.05.068](https://doi.org/10.1016/j.nima.2005.05.068).
- [98] Gary J. Feldman and Robert D. Cousins. "Unified approach to the classical statistical analysis of small signals". In: *Physical Review D* 57.7 (Apr. 1998), 3873–3889. ISSN: 1089-4918. doi: [10.1103/physrevd.57.3873](https://doi.org/10.1103/physrevd.57.3873).
- [99] Bo Andersson et al. "Parton Fragmentation and String Dynamics". In: *Phys. Rept.* 97 (1983), pp. 31–145. doi: [10.1016/0370-1573\(83\)90080-7](https://doi.org/10.1016/0370-1573(83)90080-7).
- [100] B. R. Webber. "A QCD Model for Jet Fragmentation Including Soft Gluon Interference". In: *Nucl. Phys. B* 238 (1984), pp. 492–528. doi: [10.1016/0550-3213\(84\)90333-X](https://doi.org/10.1016/0550-3213(84)90333-X).
- [101] Gavin P. Salam. "Towards Jetography". In: *Eur. Phys. J. C* 67 (2010), pp. 637–686. doi: [10.1140/epjc/s10052-010-1314-6](https://doi.org/10.1140/epjc/s10052-010-1314-6). arXiv: [0906.1833 \[hep-ph\]](https://arxiv.org/abs/0906.1833).
- [102] S. Catani et al. "Longitudinally invariant  $K_t$  clustering algorithms for hadron hadron collisions". In: *Nucl. Phys. B* 406 (1993), pp. 187–224. doi: [10.1016/0550-3213\(93\)90166-M](https://doi.org/10.1016/0550-3213(93)90166-M).
- [103] Stephen D. Ellis and Davison E. Soper. "Successive combination jet algorithm for hadron collisions". In: *Phys. Rev. D* 48 (1993), pp. 3160–3166. doi: [10.1103/PhysRevD.48.3160](https://doi.org/10.1103/PhysRevD.48.3160). arXiv: [hep-ph/9305266](https://arxiv.org/abs/hep-ph/9305266).
- [104] Yuri L. Dokshitzer et al. "Better jet clustering algorithms". In: *JHEP* 08 (1997), p. 001. doi: [10.1088/1126-6708/1997/08/001](https://doi.org/10.1088/1126-6708/1997/08/001). arXiv: [hep-ph/9707323](https://arxiv.org/abs/hep-ph/9707323).
- [105] Jesse Thaler and Ken Van Tilburg. "Identifying Boosted Objects with N-subjettiness". In: *JHEP* 03 (2011), p. 015. doi: [10.1007/JHEP03\(2011\)015](https://doi.org/10.1007/JHEP03(2011)015). arXiv: [1011.2268 \[hep-ph\]](https://arxiv.org/abs/1011.2268).
- [106] M. Cacciari, G. P. Salam, and G. Soyez. "The catchment area of jets". In: *JHEP* 04 (2008). Detalla conceptos de área de jets y técnicas relacionadas con la resta de pile-up., p. 005. doi: [10.1088/1126-6708/2008/04/005](https://doi.org/10.1088/1126-6708/2008/04/005).
- [107] Garvita Agarwal. "Jet Energy Scale and Resolution Measurements in CMS". In: *PoS ICHEP2022* (2022), p. 652. doi: [10.22323/1.414.0652](https://doi.org/10.22323/1.414.0652). arXiv: [2301.02175 \[hep-ex\]](https://arxiv.org/abs/2301.02175).
- [108] Jonathan M. Butterworth et al. "Jet substructure as a new Higgs search channel at the LHC". In: *Phys. Rev. Lett.* 100 (2008), p. 242001. doi: [10.1103/PhysRevLett.100.242001](https://doi.org/10.1103/PhysRevLett.100.242001). arXiv: [0802.2470 \[hep-ph\]](https://arxiv.org/abs/0802.2470).

- [109] Andrew J. Larkoski et al. "Soft Drop". In: *JHEP* 05 (2014), p. 146. doi: [10.1007/JHEP05\(2014\)146](https://doi.org/10.1007/JHEP05(2014)146). arXiv: [1402.2657 \[hep-ph\]](https://arxiv.org/abs/1402.2657).
- [110] A. M. Sirunyan et al. "Performance of reconstruction and identification of  $\tau$  leptons decaying to hadrons and  $\nu_\tau$  in pp collisions at  $\sqrt{s} = 13$  TeV". In: *JINST* 13.10 (2018), P10005. doi: [10.1088/1748-0221/13/10/P10005](https://doi.org/10.1088/1748-0221/13/10/P10005). arXiv: [1809.02816 \[hep-ex\]](https://arxiv.org/abs/1809.02816).
- [111] Georges Aad et al. "Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at  $\sqrt{s}=8$  TeV". In: *Eur. Phys. J. C* 75.7 (2015), p. 303. doi: [10.1140/epjc/s10052-015-3500-z](https://doi.org/10.1140/epjc/s10052-015-3500-z). arXiv: [1412.7086 \[hep-ex\]](https://arxiv.org/abs/1412.7086).
- [112] CMS Collaboration. "Performance of the CMS Level-1 trigger in proton-proton collisions at  $\sqrt{s} = 13$  TeV". In: *J. Inst.* 17.05 (2022), Po5024. doi: [10.31526/lhep.2022.195](https://doi.org/10.31526/lhep.2022.195). arXiv: [2202.08678 \[hep-ex\]](https://arxiv.org/abs/2202.08678).
- [113] ATLAS Collaboration. "Measurement of the energy response of the ATLAS calorimeter to charged pions from  $W^\pm \rightarrow \tau^\pm (\rightarrow \pi^\pm \nu_\tau) \nu_\tau$  events in Run 2 data". In: *Eur. Phys. J. C* 82.3 (2022), p. 223. doi: [10.1140/epjc/s10052-022-10933-6](https://doi.org/10.1140/epjc/s10052-022-10933-6). arXiv: [2202.13948 \[hep-ex\]](https://arxiv.org/abs/2202.13948).
- [114] CMS Collaboration. "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV". In: *JINST* 13.04 (2018), P05011. doi: [10.1088/1748-0221/13/05/P05011](https://doi.org/10.1088/1748-0221/13/05/P05011). arXiv: [1712.07158 \[hep-ex\]](https://arxiv.org/abs/1712.07158).
- [115] The CMS Collaboration. "The CMS experiment at the CERN LHC". In: *Journal of Instrumentation* 3.08 (2008), S08004–S08004. doi: [10.1088/1748-0221/3/08/s08004](https://doi.org/10.1088/1748-0221/3/08/s08004).
- [116] N. D. Christensen and C. Duhr. "FeynRules - Feynman rules made easy". In: *Comput. Phys. Commun.* 180 (2009), pp. 1614–1641. doi: [10.1016/j.cpc.2009.02.018](https://doi.org/10.1016/j.cpc.2009.02.018). arXiv: [0806.4194 \[hep-ph\]](https://arxiv.org/abs/0806.4194).
- [117] Celine Degrande et al. "UFO - The Universal FeynRules Output". In: *Comput. Phys. Commun.* 183 (2012), pp. 1201–1214. doi: [10.1016/j.cpc.2012.01.022](https://doi.org/10.1016/j.cpc.2012.01.022). arXiv: [1108.2040 \[hep-ph\]](https://arxiv.org/abs/1108.2040).
- [118] J. Alwall et al. "Computing decay rates for new physics theories with FeynRules and MadGraph 5\_aMC@NLO". In: *Comput. Phys. Commun.* 197 (2015), pp. 312–323. doi: [10.1016/j.cpc.2015.08.031](https://doi.org/10.1016/j.cpc.2015.08.031). arXiv: [1402.1178 \[hep-ph\]](https://arxiv.org/abs/1402.1178).
- [119] Johan Alwall et al. "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions". In: *Eur. Phys. J. C* 53 (2008), pp. 473–500. doi: [10.1140/epjc/s10052-007-0490-5](https://doi.org/10.1140/epjc/s10052-007-0490-5). arXiv: [0706.2569 \[hep-ph\]](https://arxiv.org/abs/0706.2569).
- [120] Andy Buckley et al. "LHAPDF6: parton density access in the LHC precision era". In: *Eur. Phys. J. C* 75 (2015), p. 132. doi: [10.1140/epjc/s10052-015-3318-8](https://doi.org/10.1140/epjc/s10052-015-3318-8). arXiv: [1412.7420 \[hep-ph\]](https://arxiv.org/abs/1412.7420).

- [121] Matt Dobbs and Jorgen Beck Hansen. “The HepMC C++ Monte Carlo event record for High Energy Physics”. In: *Comput. Phys. Commun.* 134 (2001), pp. 41–46. doi: [10.1016/S0010-4655\(00\)00189-2](https://doi.org/10.1016/S0010-4655(00)00189-2).
- [122] Rene Brun and Fons Rademakers. “ROOT — An object oriented data analysis framework”. In: *Nucl. Instrum. Meth. A* 389.1-2 (1997). Ed. by M. Werlen and D. Perret-Gallix, pp. 81–86. doi: [10.1016/s0168-9002\(97\)00048-x](https://doi.org/10.1016/s0168-9002(97)00048-x).
- [123] Steve Baker and Robert D. Cousins. “Clarification of the Use of Chi Square and Likelihood Functions in Fits to Histograms”. In: *Nucl. Instrum. Meth.* 221 (1984), pp. 437–442. doi: [10.1016/0167-5087\(84\)90016-4](https://doi.org/10.1016/0167-5087(84)90016-4).
- [124] Jerzy Neyman and Egon Sharpe Pearson. “On the Problem of the Most Efficient Tests of Statistical Hypotheses”. In: *Phil. Trans. Roy. Soc. Lond. A* 231.694-706 (1933), pp. 289–337. doi: [10.1098/rsta.1933.0009](https://doi.org/10.1098/rsta.1933.0009).
- [125] S. S. Wilks. “The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses”. In: *Annals Math. Statist.* 9.1 (1938), pp. 60–62. doi: [10.1214/aoms/1177732360](https://doi.org/10.1214/aoms/1177732360).
- [126] Dan Guest, Kyle Cranmer, and Daniel Whiteson. “Deep Learning and its Application to LHC Physics”. In: *Ann. Rev. Nucl. Part. Sci.* 68 (2018), pp. 161–181. doi: [10.1146/annurev-nucl-101917-021019](https://doi.org/10.1146/annurev-nucl-101917-021019). arXiv: [1806.11484 \[hep-ex\]](https://arxiv.org/abs/1806.11484).
- [127] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer, 2009. ISBN: 978-0-387-84857-0, 978-0-387-84858-7. doi: [10.1007/978-0-387-84858-7](https://doi.org/10.1007/978-0-387-84858-7).
- [128] A. Hoecker et al. *TMVA - Toolkit for Multivariate Data Analysis*. 2009. arXiv: [physics/0703039 \[physics.data-an\]](https://arxiv.org/abs/physics/0703039).
- [129] J. H. Friedman. “Greedy function approximation: A gradient boosting machine.” In: *The Annals of Statistics* 29.5 (2001). Publisher: Institute of Mathematical Statistics, pp. 1189–1232. ISSN: 0090-5364, 2168-8966. doi: [10.1214/aos/1013203451](https://doi.org/10.1214/aos/1013203451). (Visited on 09/05/2022).
- [130] Kyle Cranmer, Juan Pavez, and Gilles Louppe. “Approximating Likelihood Ratios with Calibrated Discriminative Classifiers”. In: (June 2015). arXiv: [1506.02169 \[stat.AP\]](https://arxiv.org/abs/1506.02169).
- [131] John Platt. *Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods*. Tech. rep. Microsoft Research, 1999.
- [132] Corinna Cortes and Vladimir Vapnik. “Support-vector networks”. In: *Machine Learning* 20.3 (1995), pp. 273–297. doi: [10.1007/BF00994018](https://doi.org/10.1007/BF00994018).
- [133] Leo Breiman. “Random Forests”. In: *Machine Learning* 45.1 (2001), pp. 5–32. doi: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324).

- [134] T. Chen and C. Guestrin. "XGBoost: A scalable tree boosting system". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '16. New York, NY, USA: Association for Computing Machinery, 2016, pp. 785–794. ISBN: 978-1-4503-4232-2. DOI: 10.1145/2939672.2939785. (Visited on 09/05/2022).
- [135] N. Assad, B. Fornal, and B. Grinstein. "Baryon Number and Lepton Universality Violation in Leptoquark and Diquark Models". In: *Phys. Lett. B* 777 (2018), pp. 324–331. DOI: 10.1016/j.physletb.2017.12.042. arXiv: 1708.06350 [hep-ph].
- [136] R. N. Mohapatra and J. C. Pati. "'Natural' left-right symmetry". In: *Physical Review D* 11.9 (May 1975), pp. 2558–2561. DOI: 10.1103/PhysRevD.11.2558.
- [137] G. Senjanovic and R. N. Mohapatra. "Exact left-right symmetry and spontaneous violation of parity". In: *Physical Review D* 12.5 (Sept. 1975), pp. 1502–1505. DOI: 10.1103/PhysRevD.12.1502.
- [138] Luca Di Luzio et al. "Maximal flavour violation: a Cabibbo mechanism for leptoquarks". In: *Journal of High Energy Physics* 2018.11 (Nov. 13, 2018), p. 81. ISSN: 1029-8479. DOI: 10.1007/JHEP11(2018)081. URL: [https://doi.org/10.1007/JHEP11\(2018\)081](https://doi.org/10.1007/JHEP11(2018)081).
- [139] Michael J. Baker et al. "High-\$p\_T\$ signatures in vector-leptoquark models". In: *The European Physical Journal C* 79.4 (2019), p. 334. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-6853-x.
- [140] Lisa Michaels and Felix Yu. "Probing new U(1) gauge symmetries via exotic  $Z \rightarrow Z'\gamma$  decays". In: *JHEP* 03 (2021), p. 120. DOI: 10.1007/JHEP03(2021)120. arXiv: 2010.00021 [hep-ph].
- [141] P. S. Bhupal Dev et al. "Searching for  $Z'$  bosons at the P2 experiment". In: *JHEP* 06 (2021), p. 039. DOI: 10.1007/JHEP06(2021)039. arXiv: 2103.09067 [hep-ph].
- [142] Bhaskar Dutta, Sumit Ghosh, and Jason Kumar. " $U(1)_{T3R}$  Extension of Standard Model: A Sub-GeV Dark Matter Model". In: *Snowmass 2021*. Mar. 2022. arXiv: 2203.07786 [hep-ph].
- [143] Bhaskar Dutta, Sumit Ghosh, and Jason Kumar. "Sub-GeV dark matter model". In: *Physical Review D* 100.7 (Oct. 2019), p. 075028. DOI: 10.1103/PhysRevD.100.075028.
- [144] Bhaskar Dutta, Sumit Ghosh, and Jason Kumar. "Opportunities for probing  $U(1)_{T3R}$  with light mediators". In: *Physical Review D* 102.7 (Oct. 2020), p. 075041. DOI: 10.1103/PhysRevD.102.075041.
- [145] Bhaskar Dutta, Sumit Ghosh, and Jason Kumar. "Contributions to  $N_{\text{eff}}$  from the dark photon of  $U(1)_{T3R}$ ". In: *Physical Review D* 102.1 (July 2020), p. 015013. DOI: 10.1103/PhysRevD.102.015013.

- [146] Bhaskar Dutta et al. "Explaining  $g_\mu - 2$  and  $R_K^{(*)}$  using the light mediators of  $U(1)_{T3R}$ ". In: *Physical Review D* 105.1 (Jan. 2022), p. 015011. DOI: [10.1103/PhysRevD.105.015011](https://doi.org/10.1103/PhysRevD.105.015011).
- [147] Valentina De Romeri et al. "Dark matter in the scotogenic model with spontaneous lepton number violation". In: *Phys. Rev. D* 107 (9 2023), p. 095019. DOI: [10.1103/PhysRevD.107.095019](https://doi.org/10.1103/PhysRevD.107.095019).
- [148] Bhaskar Dutta et al. "Probing an MeV-scale scalar boson in association with a TeV-Scale top-quark partner at the LHC". In: *Journal of High Energy Physics* 2023.3 (Mar. 2023), p. 164. ISSN: 1029-8479. DOI: [10.1007/JHEP03\(2023\)164](https://doi.org/10.1007/JHEP03(2023)164).
- [149] Z.G. Berezhiani. "The weak mixing angles in gauge models with horizontal symmetry — A new approach to quark and lepton masses". In: *Physics Letters B* 129.1 (1983), pp. 99–102. ISSN: 0370-2693. DOI: [10.1016/0370-2693\(83\)90737-2](https://doi.org/10.1016/0370-2693(83)90737-2).
- [150] Darwin Chang and Rabindra N. Mohapatra. "Small and calculable Dirac neutrino mass". In: *Phys. Rev. Lett.* 58 (16 1987), pp. 1600–1603. DOI: [10.1103/PhysRevLett.58.1600](https://doi.org/10.1103/PhysRevLett.58.1600).
- [151] Aharon Davidson and Kameshwar C. Wali. "Universal seesaw mechanism?" In: *Phys. Rev. Lett.* 59 (4 1987), pp. 393–395. DOI: [10.1103/PhysRevLett.59.393](https://doi.org/10.1103/PhysRevLett.59.393).
- [152] S. Rajpoot. "See-Saw Masses for quarks and leptons in an ambidextrous electroweak interaction model". In: *Modern Physics Letters A* 02.05 (1987), pp. 307–315. DOI: [10.1142/S0217732387000422](https://doi.org/10.1142/S0217732387000422).
- [153] K. S. Babu and Rabindra N. Mohapatra. "CP violation in seesaw models of quark masses". In: *Phys. Rev. Lett.* 62 (10 1989), pp. 1079–1082. DOI: [10.1103/PhysRevLett.62.1079](https://doi.org/10.1103/PhysRevLett.62.1079).
- [154] K. S. Babu and Rabindra N. Mohapatra. "Solution to the strong CP problem without an axion". In: *Phys. Rev. D* 41 (4 1990), pp. 1286–1291. DOI: [10.1103/PhysRevD.41.1286](https://doi.org/10.1103/PhysRevD.41.1286).
- [155] Akanksha Bhardwaj et al. "Roadmap to explore vectorlike quarks decaying to a new scalar or pseudoscalar". In: *Physical Review D* 106.9 (Nov. 2022). ISSN: 2470-0029. DOI: [10.1103/physrevd.106.095014](https://doi.org/10.1103/physrevd.106.095014).
- [156] Akanksha Bhardwaj et al. "Discovery prospects of a vectorlike top partner decaying to a singlet boson". In: *Physical Review D* 106.7 (Oct. 2022). ISSN: 2470-0029. DOI: [10.1103/physrevd.106.075024](https://doi.org/10.1103/physrevd.106.075024).
- [157] Jai Bardhan et al. "Machine learning-enhanced search for a vectorlike singlet B quark decaying to a singlet scalar or pseudoscalar". In: *Physical Review D* 107.11 (June 2023). ISSN: 2470-0029. DOI: [10.1103/physrevd.107.115001](https://doi.org/10.1103/physrevd.107.115001).
- [158] Shankha Banerjee et al. "Implications of a high-mass diphoton resonance for heavy quark searches". In: *Journal of High Energy Physics* 2016.11 (Nov. 2016). ISSN: 1029-8479. DOI: [10.1007/jhep11\(2016\)154](https://doi.org/10.1007/jhep11(2016)154).

- [159] João M. Alves et al. “Vector-like singlet quarks: A roadmap”. In: *Physics Reports* 1057 (Mar. 2024), 1–69. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2023.12.004.
- [160] X. Ai et al. “Probing highly collimated photon-jets with deep learning”. In: *JPCS* 2438.1 (2023), p. 012114. DOI: 10.1088/1742-6596/2438/1/012114.
- [161] ATLAS Collaboration. “Search for the standard model Higgs boson produced in association with top quarks and decaying into a  $b\bar{b}$  pair in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector”. In: *Phys. Rev. D* 97.7 (2018), p. 072016. DOI: 10.1103/PhysRevD.97.072016. arXiv: 1712.08895 [hep-ex].
- [162] A. Biswas et al. “Collider signature of  $U_1$  leptoquark and constraints from  $b \rightarrow c$  observables”. In: *J. Phys. G* 47.4 (2020), p. 045005. DOI: 10.1088/1361-6471/ab6948. arXiv: 1808.04169 [hep-ph].
- [163] Y. L. Chung, S. C. Hsu, and B. Nachman. “Disentangling boosted Higgs boson production modes with machine learning”. In: *JINST* 16 (2021), Po7002. DOI: 10.1088/1748-0221/16/07/P07002. arXiv: 2009.05930 [hep-ph].
- [164] J. Feng et al. “Improving heavy Dirac neutrino prospects at future hadron colliders using machine learning”. In: *JHEP* 2022.9 (2022). DOI: 10.1007/jhep09(2022)141.
- [165] D. Barbosa et al. “Probing a  $Z'$  with non-universal fermion couplings through top quark fusion, decays to bottom quarks, and machine learning techniques”. In: *EPJC* 83.413 (2023). DOI: 10.1140/epjc/s10052-023-11506-x.
- [166] S. Chigusa et al. “Deeply learned preselection of Higgs dijet decays at future lepton colliders”. In: *Phys. Lett. B* 833 (2022), p. 137301. DOI: 10.1016/j.physletb.2022.137301. arXiv: 2202.02534 [hep-ph].
- [167] Ernesto Arganda et al. “LHC study of third-generation scalar leptoquarks with machine-learned likelihoods”. In: *Phys. Rev. D* 109 (5 2024), p. 055032. DOI: 10.1103/PhysRevD.109.055032.
- [168] S. Ajmal et al. “Searching for exclusive leptoquarks with the Nambu-Jona-Lasinio composite model at the LHC and HL-LHC”. In: *Journal of High Energy Physics* 2024.8 (Aug. 2024). ISSN: 1029-8479. DOI: 10.1007/jhep08(2024)176.
- [169] Bhaskar Dutta et al. “Probing compressed bottom squarks with boosted jets and shape analysis”. In: *Physical Review D* 92.9 (Nov. 2015). ISSN: 1550-2368. DOI: 10.1103/physrevd.92.095009.
- [170] CMS Collaboration. *Review of searches for vector-like quarks, vector-like leptons, and heavy neutral leptons in proton-proton collisions at  $\sqrt{s} = 13$  TeV at the CMS experiment*. 2024. arXiv: 2405.17605 [hep-ex].

- [171] CMS Collaboration. *Search for production of a single vector-like quark decaying to tH or tZ in the all-hadronic final state in pp collisions at  $\sqrt{s} = 13$  TeV*. 2024. arXiv: 2405.05071 [hep-ex].
- [172] Georges Aad et al. “Search for single production of a vectorlike T quark decaying into a Higgs boson and top quark with fully hadronic final states using the ATLAS detector”. In: *Phys. Rev. D* 105.9 (2022), p. 092012. DOI: 10.1103/PhysRevD.105.092012. arXiv: 2201.07045 [hep-ex].
- [173] Georges Aad et al. “Search for singly produced vectorlike top partners in multilepton final states with 139 fb<sup>-1</sup> of pp collision data at s=13 TeV with the ATLAS detector”. In: *Phys. Rev. D* 109.11 (2024), p. 112012. DOI: 10.1103/PhysRevD.109.112012. arXiv: 2307.07584 [hep-ex].
- [174] Georges Aad et al. “Search for pair-production of vector-like quarks in pp collision events at s=13 TeV with at least one leptonically decaying Z boson and a third-generation quark with the ATLAS detector”. In: *Phys. Lett. B* 843 (2023), p. 138019. DOI: 10.1016/j.physletb.2023.138019. arXiv: 2210.15413 [hep-ex].
- [175] Georges Aad et al. “Search for pair-produced vector-like top and bottom partners in events with large missing transverse momentum in pp collisions with the ATLAS detector”. In: *Eur. Phys. J. C* 83.8 (2023), p. 719. DOI: 10.1140/epjc/s10052-023-11790-7. arXiv: 2212.05263 [hep-ex].
- [176] Georges Aad et al. “Search for single production of vector-like T quarks decaying into Ht or Zt in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector”. In: *JHEP* 08 (2023), p. 153. DOI: 10.1007/JHEP08(2023)153. arXiv: 2305.03401 [hep-ex].
- [177] Georges Aad et al. *Exploration at the high-energy frontier: ATLAS Run 2 searches investigating the exotic jungle beyond the Standard Model*. Mar. 2024. arXiv: 2403.09292 [hep-ex].
- [178] Giacomo Cacciapaglia et al. “Exotic decays of top partners: mind the search gap”. In: *Phys. Lett. B* 798 (2019), p. 135015. DOI: 10.1016/j.physletb.2019.135015. arXiv: 1908.07524 [hep-ph].
- [179] A. C. Canbay and O. Cakir. “Investigating the single production of vectorlike quarks decaying into a top quark and W boson through hadronic channels at the HL-LHC”. In: *Phys. Rev. D* 108 (9 2023), p. 095006. DOI: 10.1103/PhysRevD.108.095006.
- [180] Giacomo Cacciapaglia, Aldo Deandrea, and Shahram Vatani. “Is the muon a third family lepton?” In: *Physical Review D* 108.1 (July 2023). ISSN: 2470-0029. DOI: 10.1103/physrevd.108.016010.
- [181] Gianluca Blankenburg, Gino Isidori, and Joel Jones-Perez. “Neutrino Masses and LFV from Minimal Breaking of U(3)<sup>5</sup> and U(2)<sup>5</sup> flavor Symmetries”. In: *Eur. Phys. J. C* 72 (2012), p. 2126. DOI: 10.1140/epjc/s10052-012-2126-7. arXiv: 1204.0688 [hep-ph].

- [182] Joel Jones-Perez. “RGE behaviour of  $U(2)^3$  symmetry in SUSY”. In: *J. Phys. Conf. Ser.* 447 (2013). Ed. by Gustavo C. Branco et al., p. 012060. DOI: [10.1088/1742-6596/447/1/012060](https://doi.org/10.1088/1742-6596/447/1/012060). arXiv: [1302.6133 \[hep-ph\]](https://arxiv.org/abs/1302.6133).
- [183] L. Calibbi et al. “ $SU(3)$  Flavour Symmetries and CP Violation”. In: *PoS EPS-HEP2009* (2009), p. 167. DOI: [10.22323/1.084.0167](https://doi.org/10.22323/1.084.0167). arXiv: [0909.2501 \[hep-ph\]](https://arxiv.org/abs/0909.2501).
- [184] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. “The anti-kt jet clustering algorithm”. In: *Journal of High Energy Physics* 2008.04 (Apr. 2008), 063–063. ISSN: [1029-8479](https://doi.org/10.1088/1126-6708/2008/04/063). DOI: [10.1088/1126-6708/2008/04/063](https://doi.org/10.1088/1126-6708/2008/04/063).
- [185] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “FastJet user manual: (for version 3.0.2)”. In: *The European Physical Journal C* 72.3 (Mar. 2012). ISSN: [1434-6052](https://doi.org/10.1140/epjc/s10052-012-1896-2). DOI: [10.1140/epjc/s10052-012-1896-2](https://doi.org/10.1140/epjc/s10052-012-1896-2).
- [186] Albert M Sirunyan et al. “Identification of heavy, energetic, hadronically decaying particles using machine-learning techniques”. In: *JINST* 15.06 (2020), Po6005. DOI: [10.1088/1748-0221/15/06/P06005](https://doi.org/10.1088/1748-0221/15/06/P06005). arXiv: [2004.08262 \[hep-ex\]](https://arxiv.org/abs/2004.08262).
- [187] Morad Aaboud et al. “Performance of top-quark and  $W$ -boson tagging with ATLAS in Run 2 of the LHC”. In: *Eur. Phys. J. C* 79.5 (2019), p. 375. DOI: [10.1140/epjc/s10052-019-6847-8](https://doi.org/10.1140/epjc/s10052-019-6847-8). arXiv: [1808.07858 \[hep-ex\]](https://arxiv.org/abs/1808.07858).
- [188] A.M. Sirunyan et al. “Identification of heavy-flavour jets with the CMS detector in  $pp$  collisions at 13 TeV”. In: *JINST* 13 (2018), P05011. DOI: [10.1088/1748-0221/13/05/P05011](https://doi.org/10.1088/1748-0221/13/05/P05011). arXiv: [1712.07158 \[hep-ex\]](https://arxiv.org/abs/1712.07158).
- [189] Eric Conte, Benjamin Fuks, and Guillaume Serret. “MadAnalysis 5, a user-friendly framework for collider phenomenology”. In: *Computer Physics Communications* 184.1 (2013), pp. 222–256. ISSN: [0010-4655](https://doi.org/10.1016/j.cpc.2012.09.009). DOI: [10.1016/j.cpc.2012.09.009](https://doi.org/10.1016/j.cpc.2012.09.009).
- [190] Adam Paszke et al. *PyTorch: An Imperative Style, High-Performance Deep Learning Library*. 2019. arXiv: [1912.01703 \[cs.LG\]](https://arxiv.org/abs/1912.01703).
- [191] A. Flórez et al. “Probing axionlike particles with  $\gamma\gamma$  final states from vector boson fusion processes at the LHC”. In: *Phys. Rev. D* 103.9 (2021), p. 095001. DOI: [10.1103/PhysRevD.103.095001](https://doi.org/10.1103/PhysRevD.103.095001). arXiv: [2101.11119 \[hep-ph\]](https://arxiv.org/abs/2101.11119).
- [192] A. Flórez et al. “Anapole dark matter via vector boson fusion processes at the LHC”. In: *Phys. Rev. D* 100.1 (2019), p. 016017. DOI: [10.1103/PhysRevD.100.016017](https://doi.org/10.1103/PhysRevD.100.016017). arXiv: [1902.01488 \[hep-ph\]](https://arxiv.org/abs/1902.01488).
- [193] A. Flórez et al. “Probing heavy spin-2 bosons with  $\gamma\gamma$  final states from vector boson fusion processes at the LHC”. In: *Phys. Rev. D* 99.3 (2019), p. 035034. DOI: [10.1103/PhysRevD.99.035034](https://doi.org/10.1103/PhysRevD.99.035034). arXiv: [1812.06824 \[hep-ph\]](https://arxiv.org/abs/1812.06824).

- [194] A. Flórez et al. "Expanding the reach of heavy neutrino searches at the LHC". In: *Phys. Lett. B* 778 (2018), pp. 94–100. doi: [10.1016/j.physletb.2018.01.009](https://doi.org/10.1016/j.physletb.2018.01.009). arXiv: [1708.03007 \[hep-ph\]](https://arxiv.org/abs/1708.03007).
- [195] A. Flórez et al. "Searching for new heavy neutral gauge bosons using vector boson fusion processes at the LHC". In: *Phys. Lett. B* 767 (2017), pp. 126–132. doi: [10.1016/j.physletb.2017.01.062](https://doi.org/10.1016/j.physletb.2017.01.062). arXiv: [1609.09765v2 \[hep-ph\]](https://arxiv.org/abs/1609.09765v2).
- [196] A. Flórez et al. "Probing the stau-neutralino coannihilation region at the LHC with a soft tau lepton and a jet from initial state radiation". In: *Phys. Rev. D* 94.7 (2016), p. 073007. doi: [10.1103/PhysRevD.94.073007](https://doi.org/10.1103/PhysRevD.94.073007). arXiv: [1606.08878 \[hep-ph\]](https://arxiv.org/abs/1606.08878).
- [197] R. Leonardi et al. "Phenomenology at the LHC of composite particles from strongly interacting Standard Model fermions via four-fermion operators of NJL type". In: *The European Physical Journal C* 80.4 (Apr. 2020). ISSN: [1434-6052](https://doi.org/10.1140/epjc/s10052-020-7822-0). doi: [10.1140/epjc/s10052-020-7822-0](https://doi.org/10.1140/epjc/s10052-020-7822-0).
- [198] Cardona Natalia et al. *Long-term LHC Discovery Reach for Compressed Higgsino-like Models using VBF Processes*. 2021. arXiv: [2102.10194 \[hep-ph\]](https://arxiv.org/abs/2102.10194).
- [199] Andrés Flórez et al. "Probing axionlike particles with  $\gamma\gamma$  final states from vector boson fusion processes at the LHC". In: *Phys. Rev. D* 103 (9 2021), p. 095001. doi: [10.1103/PhysRevD.103.095001](https://doi.org/10.1103/PhysRevD.103.095001).
- [200] J. C. Pati and A. Salam. "Lepton number as the fourth color". In: *Phys. Rev. D* 10 (1974). [Erratum: *Phys. Rev. D* 11, 703–703 (1975)], pp. 275–289. doi: [10.1103/PhysRevD.10.275](https://doi.org/10.1103/PhysRevD.10.275).
- [201] H. Georgi and S. L. Glashow. "Unity of all elementary particle forces". In: *Phys. Rev. Lett.* 32 (1974), pp. 438–441. doi: [10.1103/PhysRevLett.32.438](https://doi.org/10.1103/PhysRevLett.32.438).
- [202] B. Schrempp and F. Schrempp. "Light leptoquarks". In: *Phys. Lett. B* 153 (1985), pp. 101–107. doi: [10.1016/0370-2693\(85\)91450-9](https://doi.org/10.1016/0370-2693(85)91450-9).
- [203] M. Leurer. "A comprehensive study of leptoquark bounds". In: *Phys. Rev. D* 49 (1994), pp. 333–342. doi: [10.1103/PhysRevD.49.333](https://doi.org/10.1103/PhysRevD.49.333). arXiv: [hep-ph/9309266](https://arxiv.org/abs/hep-ph/9309266).
- [204] S. Davidson, D. C. Bailey, and B. A. Campbell. "Model independent constraints on leptoquarks from rare processes". In: *Z. Phys. C* 61 (1994), pp. 613–644. doi: [10.1007/BF01552629](https://doi.org/10.1007/BF01552629). arXiv: [hep-ph/9309310](https://arxiv.org/abs/hep-ph/9309310).
- [205] M. Leurer. "Bounds on vector leptoquarks". In: *Phys. Rev. D* 50 (1994), pp. 536–541. doi: [10.1103/PhysRevD.50.536](https://doi.org/10.1103/PhysRevD.50.536). arXiv: [hep-ph/9312341](https://arxiv.org/abs/hep-ph/9312341).
- [206] J. L. Hewett and T. G. Rizzo. "Much ado about leptoquarks: A comprehensive analysis". In: *Phys. Rev. D* 56 (1997), pp. 5709–5724. doi: [10.1103/PhysRevD.56.5709](https://doi.org/10.1103/PhysRevD.56.5709). arXiv: [hep-ph/9703337](https://arxiv.org/abs/hep-ph/9703337).

- [207] F. S. Queiroz, K. Sinha, and A. Strumia. "Leptoquarks, Dark Matter, and Anomalous LHC Events". In: *Phys. Rev. D* 91.3 (2015), p. 035006. doi: [10.1103/PhysRevD.91.035006](https://doi.org/10.1103/PhysRevD.91.035006). arXiv: [1409.6301 \[hep-ph\]](https://arxiv.org/abs/1409.6301).
- [208] LHCb Collaboration. "Search for lepton-universality violation in  $B^+ \rightarrow K^+ \ell^+ \ell^-$  decays". In: *Phys. Rev. Lett.* 122.19 (2019), p. 191801. doi: [10.1103/PhysRevLett.122.191801](https://doi.org/10.1103/PhysRevLett.122.191801). arXiv: [1903.09252 \[hep-ex\]](https://arxiv.org/abs/1903.09252).
- [209] LHCb Collaboration. "Test of lepton universality in beauty-quark decays". In: *Nature Phys.* 18.3 (2022), pp. 277–282. doi: [10.1038/s41567-021-01478-8](https://doi.org/10.1038/s41567-021-01478-8). arXiv: [2103.11769 \[hep-ex\]](https://arxiv.org/abs/2103.11769).
- [210] BaBar Collaboration. "Measurement of an excess of  $\bar{B} \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$  decays and implications for charged Higgs bosons". In: *Phys. Rev. D* 88.7 (2013), p. 072012. doi: [10.1103/PhysRevD.88.072012](https://doi.org/10.1103/PhysRevD.88.072012). arXiv: [1303.0571 \[hep-ex\]](https://arxiv.org/abs/1303.0571).
- [211] Belle Collaboration. "Measurement of  $\mathcal{R}(D)$  and  $\mathcal{R}(D^*)$  with a semileptonic tagging method". In: (2019). doi: [10.48550/arXiv.1904.08794](https://doi.org/10.48550/arXiv.1904.08794). arXiv: [1904.08794 \[hep-ex\]](https://arxiv.org/abs/1904.08794).
- [212] Belle Collaboration. "Measurement of the  $\tau$  lepton polarization and  $R(D^*)$  in the decay  $\bar{B} \rightarrow D^*\tau^-\bar{\nu}_\tau$  with one-prong hadronic  $\tau$  decays at Belle". In: *Phys. Rev. D* 97 (2018), p. 012004. doi: [10.1103/PhysRevD.97.012004](https://doi.org/10.1103/PhysRevD.97.012004). arXiv: [1709.00129 \[hep-ex\]](https://arxiv.org/abs/1709.00129).
- [213] Belle Collaboration. "Measurement of the branching ratio of  $\bar{B}^0 \rightarrow D^{*+}\tau^-\bar{\nu}_\tau$  relative to  $\bar{B}^0 \rightarrow D^{*+}\ell^-\bar{\nu}_\ell$  decays with a semileptonic tagging method". In: *Phys. Rev. D* 94 (2016), p. 072007. doi: [10.1103/PhysRevD.94.072007](https://doi.org/10.1103/PhysRevD.94.072007). arXiv: [1607.07923 \[hep-ex\]](https://arxiv.org/abs/1607.07923).
- [214] Belle Collaboration. "Measurement of the  $\tau$  lepton polarization and  $R(D^*)$  in the decay  $\bar{B} \rightarrow D^*\tau^-\bar{\nu}_\tau$ ". In: *Phys. Rev. Lett.* 118 (2017), p. 211801. doi: [10.1103/PhysRevLett.118.211801](https://doi.org/10.1103/PhysRevLett.118.211801). arXiv: [1612.00529 \[hep-ex\]](https://arxiv.org/abs/1612.00529).
- [215] Belle Collaboration. "Measurement of the branching ratio of  $\bar{B} \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$  relative to  $\bar{B} \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell$  decays with hadronic tagging at Belle". In: *Phys. Rev. D* 92 (2015), p. 072014. doi: [10.1103/PhysRevD.92.072014](https://doi.org/10.1103/PhysRevD.92.072014). arXiv: [1507.03233 \[hep-ex\]](https://arxiv.org/abs/1507.03233).
- [216] LHCb Collaboration. "Measurement of the ratio of branching fractions  $\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\mu^-\bar{\nu}_\mu)$ ". In: *Phys. Rev. Lett.* 115.11 (2015). [Erratum: *Phys. Rev. Lett.* 115, 159901 (2015)], p. 111803. doi: [10.1103/PhysRevLett.115.111803](https://doi.org/10.1103/PhysRevLett.115.111803). arXiv: [1506.08614 \[hep-ex\]](https://arxiv.org/abs/1506.08614).
- [217] LHCb Collaboration. "Measurement of the ratio of branching fractions  $\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\mu^-\bar{\nu}_\mu)$ ". In: *Phys. Rev. Lett.* 115 (2015). [Erratum: *Phys. Rev. Lett.* 115, 159901 (2015)], p. 111803. doi: [10.1103/PhysRevLett.115.111803](https://doi.org/10.1103/PhysRevLett.115.111803). arXiv: [1506.08614 \[hep-ex\]](https://arxiv.org/abs/1506.08614).
- [218] LHCb Collaboration. "Measurement of the ratio of the  $B^0 \rightarrow D^{*-}\tau^+\bar{\nu}_\tau$  and  $B^0 \rightarrow D^{*-}\mu^+\bar{\nu}_\mu$  branching fractions using three-prong  $\tau$ -lepton decays". In: *Phys. Rev. Lett.* 120 (2018), p. 171802. doi: [10.1103/PhysRevLett.120.171802](https://doi.org/10.1103/PhysRevLett.120.171802). arXiv: [1708.08856 \[hep-ex\]](https://arxiv.org/abs/1708.08856).

- [219] LHCb Collaboration. “Test of lepton flavor universality by the measurement of the  $B^0 \rightarrow D^{*-} \tau^+ \nu_\tau$  branching fraction using three-prong  $\tau$  decays”. In: *Phys. Rev. D* 97.7 (2018), p. 072013. doi: 10.1103/PhysRevD.97.072013. arXiv: 1711.02505 [hep-ex].
- [220] LHCb Collaboration. “Measurement of the ratios of branching fractions  $\mathcal{R}(D^*)$  and  $\mathcal{R}(D^0)$ ”. In: (2023). doi: 10.48550/arXiv.2302.02886. arXiv: 2302.02886 [hep-ex].
- [221] G. Hiller and M. Schmaltz. “ $R_K$  and future  $b \rightarrow s\ell\ell$  physics beyond the standard model opportunities”. In: *Phys. Rev. D* 90 (2014), p. 054014. doi: 10.1103/PhysRevD.90.054014. arXiv: 1408.1627 [hep-ph].
- [222] B. Gripaios, M. Nardecchia, and S. A. Renner. “Composite leptoquarks and anomalies in B-meson decays”. In: *JHEP* 05 (2015), p. 006. doi: 10.1007/JHEP05(2015)006. arXiv: 1412.1791 [hep-ph].
- [223] R. Alonso, B. Grinstein, and J. M. Camalich. “Lepton universality violation and lepton flavor conservation in B-meson decays”. In: *JHEP* 10 (2015), p. 184. doi: 10.1007/JHEP10(2015)184. arXiv: 1505.05164 [hep-ph].
- [224] L. Calibbi, A. Crivellin, and T. Ota. “Effective field theory approach to  $b \rightarrow s\ell\ell^{(')}$ ,  $B \rightarrow K^{(*)}\nu\bar{\nu}$  and  $B \rightarrow D^{(*)}\tau\nu$  with third generation couplings”. In: *Phys. Rev. Lett.* 115 (2015), p. 181801. doi: 10.1103/PhysRevLett.115.181801. arXiv: 1506.02661 [hep-ph].
- [225] S. Fajfer and N. Košnik. “Vector leptoquark resolution of  $R_K$  and  $R_{D^{(*)}}$  puzzles”. In: *Phys. Lett. B* 755 (2016), pp. 270–274. doi: 10.1016/j.physletb.2016.02.018. arXiv: 1511.06024 [hep-ph].
- [226] M. Bauer and M. Neubert. “Minimal leptoquark explanation for the  $R_{D^{(*)}}$ ,  $R_K$ , and  $(g-2)_\mu$  anomalies”. In: *Phys. Rev. Lett.* 116.14 (2016), p. 141802. doi: 10.1103/PhysRevLett.116.141802. arXiv: 1511.01900 [hep-ph].
- [227] D. Bećirević et al. “Palatable leptoquark scenarios for lepton flavor violation in exclusive  $b \rightarrow s\ell_1\ell_2$  modes”. In: *JHEP* 11 (2016), p. 035. doi: 10.1007/JHEP11(2016)035. arXiv: 1608.07583 [hep-ph].
- [228] A. Crivellin, D. Muller, and T. Ota. “Simultaneous explanation of  $R(D^*)$  and  $b \rightarrow \mu^+\mu^-$ :the last scalar leptoquark standing”. In: *JHEP* 09 (2017), p. 040. doi: 10.1007/JHEP09(2017)040. arXiv: 1703.09226 [hep-ph].
- [229] G. D’Amico et al. “Flavour anomalies after the  $R_{K^*}$  measurement”. In: *JHEP* 09 (2017), p. 010. doi: 10.1007/JHEP09(2017)010. arXiv: 1704.05438 [hep-ph].
- [230] G. Hiller and I. Nisandžić. “ $R_K$  and  $R_{K^*}$  beyond the standard model”. In: *Phys. Rev. D* 96.3 (2017), p. 035003. doi: 10.1103/PhysRevD.96.035003. arXiv: 1704.05444 [hep-ph].

- [231] D. Bećirević et al. “Scalar leptoquarks from grand unified theories to accommodate the B-physics anomalies”. In: *Phys. Rev. D* 98.5 (2018), p. 055003. doi: [10.1103/PhysRevD.98.055003](https://doi.org/10.1103/PhysRevD.98.055003). arXiv: [1806.05689](https://arxiv.org/abs/1806.05689) [hep-ph].
- [232] C. Cornellà, J. Fuentes-Martín, and G. Isidori. “Revisiting the vector leptoquark explanation of the B-physics anomalies”. In: *JHEP* 07 (2019), p. 168. doi: [10.1007/JHEP07\(2019\)168](https://doi.org/10.1007/JHEP07(2019)168). arXiv: [1903.11517](https://arxiv.org/abs/1903.11517) [hep-ph].
- [233] A. Angelescu et al. “Single leptoquark solutions to the B-physics anomalies”. In: *Phys. Rev. D* 104.5 (2021), p. 055017. doi: [10.1103/PhysRevD.104.055017](https://doi.org/10.1103/PhysRevD.104.055017). arXiv: [2103.12504](https://arxiv.org/abs/2103.12504) [hep-ph].
- [234] Geneviève Belanger et al. “Leptoquark manoeuvres in the dark: a simultaneous solution of the dark matter problem and the  $R_{D^{(*)}}$  anomalies”. In: *JHEP* 02 (2022), p. 042. doi: [10.1007/JHEP02\(2022\)042](https://doi.org/10.1007/JHEP02(2022)042). arXiv: [2111.08027](https://arxiv.org/abs/2111.08027) [hep-ph].
- [235] J. Aebrischer et al. “Confronting the vector leptoquark hypothesis with new low- and high-energy data”. In: *The European Physical Journal C* 83.2 (Feb. 16, 2023), p. 153. ISSN: 1434-6052. doi: [10.1140/epjc/s10052-023-11304-5](https://doi.org/10.1140/epjc/s10052-023-11304-5).
- [236] LHCb Collaboration. “Test of lepton universality in  $b \rightarrow \ell^+ \ell^-$  decays”. In: (2022). doi: [10.48550/arXiv.1705.05802](https://doi.org/10.48550/arXiv.1705.05802). arXiv: [2212.09152](https://arxiv.org/abs/2212.09152) [hep-ex].
- [237] LHCb Collaboration. “Measurement of lepton universality parameters in  $B^+ \rightarrow K^+ \ell^+ \ell^-$  and  $B^0 \rightarrow K^{*0} \ell^+ \ell^-$  decays”. In: (2022). doi: [10.48550/arXiv.2212.09153](https://doi.org/10.48550/arXiv.2212.09153). arXiv: [2212.09153](https://arxiv.org/abs/2212.09153) [hep-ex].
- [238] A. Greljo et al. “Rare b decays meet high-mass Drell-Yan”. In: *JHEP* 2023.5 (2023). doi: [10.1007/jhep05\(2023\)087](https://doi.org/10.1007/jhep05(2023)087).
- [239] B. Diaz, M. Schmaltz, and Y. Zhong. “The leptoquark Hunter’s guide: Pair production”. In: *JHEP* 10 (2017), p. 097. doi: [10.1007/JHEP10\(2017\)097](https://doi.org/10.1007/JHEP10(2017)097). arXiv: [1706.05033](https://arxiv.org/abs/1706.05033) [hep-ph].
- [240] I. Doršner and A. Greljo. “Leptoquark toolbox for precision collider studies”. In: *JHEP* 05 (2018), p. 126. doi: [10.1007/JHEP05\(2018\)126](https://doi.org/10.1007/JHEP05(2018)126). arXiv: [1801.07641](https://arxiv.org/abs/1801.07641) [hep-ph].
- [241] N Vignaroli. “Seeking leptoquarks in the  $t\bar{t}$  plus missing energy channel at the high-luminosity LHC”. In: *Phys. Rev. D* 99 (3 2019), p. 035021. doi: [10.1103/PhysRevD.99.035021](https://doi.org/10.1103/PhysRevD.99.035021). arXiv: [1808.10309](https://arxiv.org/abs/1808.10309) [hep-ph].
- [242] M. Schmaltz and Y. Zhong. “The leptoquark Hunter’s guide: large coupling”. In: *JHEP* 01 (2019), p. 132. doi: [10.1007/JHEP01\(2019\)132](https://doi.org/10.1007/JHEP01(2019)132). arXiv: [1810.10017](https://arxiv.org/abs/1810.10017) [hep-ph].
- [243] M. J. Baker et al. “High- $p_T$  signatures in vector-leptoquark models”. In: *Eur. Phys. J. C* 79.4 (2019), p. 334. doi: [10.1140/epjc/s10052-019-6853-x](https://doi.org/10.1140/epjc/s10052-019-6853-x). arXiv: [1901.10480](https://arxiv.org/abs/1901.10480) [hep-ph].

- [244] U. Haisch and G. Polesello. “Resonant third-generation leptoquark signatures at the Large Hadron Collider”. In: *JHEP* 05 (2021), p. 057. doi: [10.1007/JHEP05\(2021\)057](https://doi.org/10.1007/JHEP05(2021)057). arXiv: [2012.11474 \[hep-ph\]](https://arxiv.org/abs/2012.11474).
- [245] A. Bhaskar et al. “Improving third-generation leptoquark searches with combined signals and boosted top quarks”. In: *Phys. Rev. D* 104.7 (2021), p. 075037. doi: [10.1103/PhysRevD.104.075037](https://doi.org/10.1103/PhysRevD.104.075037). arXiv: [2106.07605 \[hep-ph\]](https://arxiv.org/abs/2106.07605).
- [246] J. Bernigaud et al. “LHC signatures of  $\tau$ -flavoured vector leptoquarks”. In: *JHEP* 08 (2022), p. 127. doi: [10.1007/JHEP08\(2022\)127](https://doi.org/10.1007/JHEP08(2022)127). arXiv: [2112.12129 \[hep-ph\]](https://arxiv.org/abs/2112.12129).
- [247] R. Leonardi et al. “Phenomenology at the LHC of composite particles from strongly interacting Standard Model fermions via four-fermion operators of NJL type”. In: *EPJC* 80.309 (2020). doi: [10.1140/epjc/s10052-020-7822-0](https://doi.org/10.1140/epjc/s10052-020-7822-0). arXiv: [1810.11420 \[hep-ph\]](https://arxiv.org/abs/1810.11420).
- [248] CMS Collaboration. “Search for heavy neutrinos or third-generation leptoquarks in final states with two hadronically decaying  $\tau$  leptons and two jets in proton-proton collisions at  $\sqrt{s} = 13$  TeV”. In: *JHEP* 03 (2017), p. 077. doi: [10.1007/JHEP03\(2017\)077](https://doi.org/10.1007/JHEP03(2017)077). arXiv: [1612.01190 \[hep-ex\]](https://arxiv.org/abs/1612.01190).
- [249] CMS Collaboration. “Search for third-generation scalar leptoquarks and heavy right-handed neutrinos in final states with two tau leptons and two jets in proton-proton collisions at  $\sqrt{s} = 13$  TeV”. In: *JHEP* 07 (2017), p. 121. doi: [10.1007/JHEP07\(2017\)121](https://doi.org/10.1007/JHEP07(2017)121). arXiv: [1703.03995 \[hep-ex\]](https://arxiv.org/abs/1703.03995).
- [250] CMS Collaboration. “Search for third-generation scalar leptoquarks decaying to a top quark and a  $\tau$  lepton at  $\sqrt{s} = 13$  TeV”. In: *Eur. Phys. J. C* 78 (2018), p. 707. doi: [10.1140/epjc/s10052-018-6143-z](https://doi.org/10.1140/epjc/s10052-018-6143-z). arXiv: [1803.02864 \[hep-ex\]](https://arxiv.org/abs/1803.02864).
- [251] CMS Collaboration. “Constraints on models of scalar and vector leptoquarks decaying to a quark and a neutrino at  $\sqrt{s} = 13$  TeV”. In: *Phys. Rev. D* 98.3 (2018), p. 032005. doi: [10.1103/PhysRevD.98.032005](https://doi.org/10.1103/PhysRevD.98.032005). arXiv: [1805.10228 \[hep-ex\]](https://arxiv.org/abs/1805.10228).
- [252] CMS Collaboration. “Search for a singly produced third-generation scalar leptoquark decaying to a  $\tau$  lepton and a bottom quark in proton-proton collisions at  $\sqrt{s} = 13$  TeV”. In: *JHEP* 07 (2018), p. 115. doi: [10.1007/JHEP07\(2018\)115](https://doi.org/10.1007/JHEP07(2018)115). arXiv: [1806.03472 \[hep-ex\]](https://arxiv.org/abs/1806.03472).
- [253] CMS Collaboration. “Search for heavy neutrinos and third-generation leptoquarks in hadronic states of two  $\tau$  leptons and two jets in proton-proton collisions at  $\sqrt{s} = 13$  TeV”. In: *JHEP* 03 (2019), p. 170. doi: [10.1007/JHEP03\(2019\)170](https://doi.org/10.1007/JHEP03(2019)170). arXiv: [1811.00806 \[hep-ex\]](https://arxiv.org/abs/1811.00806).
- [254] CMS Collaboration. “Search for singly and pair-produced leptoquarks coupling to third-generation fermions in proton-proton collisions at  $s=13$  TeV”. In: *Phys. Lett. B* 819 (2021), p. 136446. doi: [10.1016/j.physletb.2021.136446](https://doi.org/10.1016/j.physletb.2021.136446). arXiv: [2012.04178 \[hep-ex\]](https://arxiv.org/abs/2012.04178).

- [255] CMS Collaboration. "Searches for additional Higgs bosons and for vector leptoquarks in  $\tau\tau$  final states in proton-proton collisions at  $\sqrt{s} = 13$  TeV". In: (2022). doi: 10.48550/arXiv.2208.02717. arXiv: 2208.02717 [hep-ex].
- [256] CMS Collaboration. "The search for a third-generation leptoquark coupling to a  $\tau$  lepton and a b quark through single, pair and nonresonant production at  $\sqrt{s} = 13$  TeV". In: (2022). URL: <https://inspirehep.net/literature/2110188>.
- [257] ATLAS Collaboration. "Searches for third-generation scalar leptoquarks in  $\sqrt{s} = 13$  TeV pp collisions with the ATLAS detector". In: *JHEP* 06 (2019), p. 144. doi: 10.1007/JHEP06(2019)144. arXiv: 1902.08103 [hep-ex].
- [258] ATLAS Collaboration. "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". In: *Eur. Phys. J. C* 80.8 (2020), p. 737. doi: 10.1140/epjc/s10052-020-8102-8. arXiv: 2004.14060 [hep-ex].
- [259] ATLAS Collaboration. "Search for pair production of third-generation scalar leptoquarks decaying into a top quark and a  $\tau$ -lepton in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector". In: *JHEP* 06 (2021), p. 179. doi: 10.1007/JHEP06(2021)179. arXiv: 2101.11582 [hep-ex].
- [260] ATLAS Collaboration. "Search for new phenomena in final states with b-jets and missing transverse momentum in  $\sqrt{s} = 13$  TeV pp collisions with the ATLAS detector". In: *JHEP* 05 (2021), p. 093. doi: 10.1007/JHEP05(2021)093. arXiv: 2101.12527 [hep-ex].
- [261] ATLAS Collaboration. "Search for new phenomena in pp collisions in final states with tau leptons, b-jets, and missing transverse momentum with the ATLAS detector". In: *Phys. Rev. D* 104.11 (2021), p. 112005. doi: 10.1103/PhysRevD.104.112005. arXiv: 2108.07665 [hep-ex].
- [262] ATLAS Collaboration. "Search for leptoquarks decaying into the  $b\tau$  final state in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector". In: (2023). doi: 10.48550/arXiv.2305.15962. arXiv: 2305.15962 [hep-ex].
- [263] ATLAS Collaboration. *Search for pair production of third-generation leptoquarks decaying into a bottom quark and a  $\tau$ -lepton with the ATLAS detector*. 2023. doi: 10.48550/arXiv.2303.01294. arXiv: 2303.01294 [hep-ex].
- [264] D. A. Faroughy, A. Greljo, and J. F. Kamenik. "Confronting lepton flavor universality violation in B decays with high- $p_T$  tau lepton searches at LHC". In: *Phys. Lett. B* 764 (2017), pp. 126–134. doi: 10.1016/j.physletb.2016.11.011. arXiv: 1609.07138 [hep-ph].

- [265] A. Angelescu et al. “Closing the window on single leptoquark solutions to the B-physics anomalies”. In: *JHEP* 10 (2018), p. 183. doi: [10.1007/JHEP10\(2018\)183](https://doi.org/10.1007/JHEP10(2018)183). arXiv: [1808.08179 \[hep-ph\]](https://arxiv.org/abs/1808.08179).
- [266] A. Bhaskar et al. “Precise limits on the charge- $2/3$   $U_1$  vector leptoquark”. In: *Phys. Rev. D* 104.3 (2021), p. 035016. doi: [10.1103/PhysRevD.104.035016](https://doi.org/10.1103/PhysRevD.104.035016). arXiv: [2101.12069 \[hep-ph\]](https://arxiv.org/abs/2101.12069).
- [267] C. Cornellà et al. “Reading the footprints of the B-meson flavor anomalies”. In: *JHEP* 08 (2021), p. 050. doi: [10.1007/JHEP08\(2021\)050](https://doi.org/10.1007/JHEP08(2021)050). arXiv: [2103.16558 \[hep-ph\]](https://arxiv.org/abs/2103.16558).
- [268] L. Allwicher et al. “Drell-Yan tails beyond the standard model”. In: *JHEP* 2023.3 (2023). doi: [10.1007/jhep03\(2023\)064](https://doi.org/10.1007/jhep03(2023)064).
- [269] U. Haisch, L. Schnell, and S. Schulte. “Drell-Yan production in third-generation gauge vector leptoquark models at NLO+PS in QCD”. In: *JHEP* 2023.2 (2023). doi: [10.1007/jhep02\(2023\)070](https://doi.org/10.1007/jhep02(2023)070).
- [270] L. Calibbi, A. Crivellin, and T. Li. “Model of vector leptoquarks in view of the B-physics anomalies”. In: *Phys. Rev. D* 98.11 (2018), p. 115002. doi: [10.1103/PhysRevD.98.115002](https://doi.org/10.1103/PhysRevD.98.115002). arXiv: [1709.00692 \[hep-ph\]](https://arxiv.org/abs/1709.00692).
- [271] M. Blanke and A. Crivellin. “B Meson anomalies in a Pati-Salam model within the Randall-Sundrum background”. In: *Phys. Rev. Lett.* 121.1 (2018), p. 011801. doi: [10.1103/PhysRevLett.121.011801](https://doi.org/10.1103/PhysRevLett.121.011801). arXiv: [1801.07256 \[hep-ph\]](https://arxiv.org/abs/1801.07256).
- [272] S. Iguro et al. “TeV-scale vector leptoquark from Pati-Salam unification with vectorlike families”. In: *Phys. Rev. D* 104.7 (2021), p. 075008. doi: [10.1103/PhysRevD.104.075008](https://doi.org/10.1103/PhysRevD.104.075008). arXiv: [2103.11889 \[hep-ph\]](https://arxiv.org/abs/2103.11889).
- [273] L. Di Luzio, A. Greljo, and M. Nardecchia. “Gauge leptoquark as the origin of B-physics anomalies”. In: *Phys. Rev. D* 96.11 (2017), p. 115011. doi: [10.1103/PhysRevD.96.115011](https://doi.org/10.1103/PhysRevD.96.115011). arXiv: [1708.08450 \[hep-ph\]](https://arxiv.org/abs/1708.08450).
- [274] Admir Greljo and Ben A. Stefanek. “Third family quark-lepton unification at the TeV scale”. In: *Phys. Lett. B* 782 (2018), pp. 131–138. doi: [10.1016/j.physletb.2018.05.033](https://doi.org/10.1016/j.physletb.2018.05.033). arXiv: [1802.04274 \[hep-ph\]](https://arxiv.org/abs/1802.04274).
- [275] L. Di Luzio et al. “Maximal flavour violation: a Cabibbo mechanism for leptoquarks”. In: *JHEP* 11 (2018), p. 081. doi: [10.1007/JHEP11\(2018\)081](https://doi.org/10.1007/JHEP11(2018)081). arXiv: [1808.00942 \[hep-ph\]](https://arxiv.org/abs/1808.00942).
- [276] S. F. King. “Twin Pati-Salam theory of flavour with a TeV scale vector leptoquark”. In: *JHEP* 11 (2021), p. 161. doi: [10.1007/JHEP11\(2021\)161](https://doi.org/10.1007/JHEP11(2021)161). arXiv: [2106.03876 \[hep-ph\]](https://arxiv.org/abs/2106.03876).
- [277] M. Fernández-Navarro and S. F. King. “B-anomalies in a twin Pati-Salam theory of flavour including the 2022 LHCb  $R_{K^{(*)}}$  analysis”. In: *JHEP* 2023.2 (2023). doi: [10.1007/jhep02\(2023\)188](https://doi.org/10.1007/jhep02(2023)188).

- [278] M. Bordone et al. “A three-site gauge model for flavor hierarchies and flavor anomalies”. In: *Phys. Lett. B* 779 (2018), pp. 317–323. doi: 10.1016/j.physletb.2018.02.011. arXiv: 1712.01368 [hep-ph].
- [279] M. Bordone et al. “Low-energy signatures of the PS<sup>3</sup> model: from B-physics anomalies to LFV”. In: *JHEP* 10 (2018), p. 148. doi: 10.1007/JHEP10(2018)148. arXiv: 1805.09328 [hep-ph].
- [280] Javier Fuentes-Martín et al. “Flavor hierarchies, flavor anomalies, and Higgs mass from a warped extra dimension”. In: *Phys. Lett. B* 834 (2022), p. 137382. doi: 10.1016/j.physletb.2022.137382. arXiv: 2203.01952 [hep-ph].
- [281] B. Gripaios. “Composite leptoquarks at the LHC”. In: *JHEP* 02 (2010), p. 045. doi: 10.1007/JHEP02(2010)045. arXiv: 0910.1789 [hep-ph].
- [282] R. Barbieri, C. W. Murphy, and F. Senia. “B-decay anomalies in a composite leptoquark model”. In: *Eur. Phys. J. C* 77.1 (2017), p. 8. doi: 10.1140/epjc/s10052-016-4578-7. arXiv: 1611.04930 [hep-ph].
- [283] R. Barbieri and A. Tesi. “B-decay anomalies in Pati-Salam SU(4)”. In: *Eur. Phys. J. C* 78.3 (2018), p. 193. doi: 10.1140/epjc/s10052-018-5680-9. arXiv: 1712.06844 [hep-ph].
- [284] J. Fuentes-Martín et al. “Vector leptoquarks beyond tree level. III. Vectorlike fermions and flavor-changing transitions”. In: *Phys. Rev. D* 102 (11 2020), p. 115015. doi: 10.1103/PhysRevD.102.115015.
- [285] A. Djouadi et al. “Interference effects in t̄t production at the LHC as a window on new physics”. In: *JHEP* 03 (2019), p. 119. doi: 10.1007/JHEP03(2019)119. arXiv: 1901.03417 [hep-ph].
- [286] The ROOT team. “ROOT - An Object Oriented Data Analysis Framework”. In: *Nucl. Inst. & Meth. in Phys. Res. A* 389 (1997). Proceedings AIHENP’96 Workshop, Lausanne, Sep. 1996. See also “ROOT”[software], Release v6.18/02, and <https://root.cern/manual/rootfit/>, pp. 81–86. doi: 10.5281/zenodo.3895860.
- [287] D. Barbosa et al. “Probing a Z' with non-universal fermion couplings through top quark fusion, decays to bottom quarks, and machine learning techniques”. In: *Eur. Phys. J. C* 83.5 (2023), p. 413. doi: 10.1140/epjc/s10052-023-11506-x. arXiv: 2210.15813 [hep-ph].
- [288] B. Dutta et al. “Probing an MeV-Scale Scalar Boson in Association with a TeV-Scale Top-Quark Partner at the LHC”. In: *JHEP* 03.164 (2023). doi: 10.1007/JHEP03%282023%29164. arXiv: 2202.08234 [hep-ph].
- [289] R. Arnowitt et al. “Determining the Dark Matter Relic Density in the Minimal Supergravity Stau-Neutralino Coannihilation Region at the Large Hadron Collider”. In: *Phys. Rev. Lett.* 100 (2008), p. 231802. doi: 10.1103/PhysRevLett.100.231802. arXiv: 0802.2968 [hep-ph].

- [290] B. Dutta et al. “Supersymmetry Signals of Supercritical String Cosmology at the Large Hadron Collider”. In: *Phys. Rev. D* 79 (2009), p. 055002. DOI: [10.1103/PhysRevD.79.055002](https://doi.org/10.1103/PhysRevD.79.055002). arXiv: [0808.1372](https://arxiv.org/abs/0808.1372) [hep-ph].
- [291] C. Avila et al. “Connecting Particle Physics and Cosmology: Measuring the Dark Matter Relic Density in Compressed Supersymmetry at the LHC”. In: *Physics of the Dark Universe* 27 (2020), p. 100430. DOI: [10.1016/j.dark.2019.100430](https://doi.org/10.1016/j.dark.2019.100430). arXiv: [1801.03966](https://arxiv.org/abs/1801.03966) [hep-ph].
- [292] B. Dutta et al. “Vector Boson Fusion Processes as a Probe of Supersymmetric Electroweak Sectors at the LHC”. In: *Phys. Rev. D* 87 (2013), p. 035029. DOI: [10.1103/PhysRevD.87.035029](https://doi.org/10.1103/PhysRevD.87.035029). arXiv: [1210.0964](https://arxiv.org/abs/1210.0964) [hep-ph].
- [293] A. Delannoy et al. “Probing Dark Matter at the LHC using Vector Boson Fusion Processes”. In: *Phys. Rev. Lett.* 111 (2013), p. 061801. DOI: [10.1103/PhysRevLett.111.061801](https://doi.org/10.1103/PhysRevLett.111.061801). arXiv: [1304.7779](https://arxiv.org/abs/1304.7779) [hep-ph].
- [294] B. Dutta et al. “Probing Compressed Sleptons at the LHC using Vector Boson Fusion Processes”. In: *Phys. Rev. D* 91 (2015), p. 055025. DOI: [10.1103/PhysRevD.91.055025](https://doi.org/10.1103/PhysRevD.91.055025). arXiv: [1411.6043](https://arxiv.org/abs/1411.6043) [hep-ph].
- [295] B. Dutta et al. “Probing Compressed Top Squarks at the LHC at 14 TeV”. In: *Phys. Rev. D* 90 (2014), p. 095022. DOI: [10.1103/PhysRevD.90.095022](https://doi.org/10.1103/PhysRevD.90.095022). arXiv: [1312.1348](https://arxiv.org/abs/1312.1348) [hep-ph].
- [296] B. Dutta et al. “Probing Compressed Bottom Squarks with Boosted Jets and Shape Analysis”. In: *Phys. Rev. D* 92 (2015), p. 095009. DOI: [10.1103/PhysRevD.92.095009](https://doi.org/10.1103/PhysRevD.92.095009). arXiv: [1507.01001](https://arxiv.org/abs/1507.01001) [hep-ph].