UNIVERSITY OF LJUBLJANA
FACULTY OF MATHEMATICS AND PHYSICS
DEPARTMENT OF PHYSICS

Darius Alexander Faroughy

Non-Resonant Phenomena Beyond the Standard Model in Top-Quark Processes

Doctoral thesis

ADVISER: Dr. Jernej F. Kamenik

Ljubljana, 2019

UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO
ODDELEK ZA FIZIKO

Darius Alexander Faroughy

Neresonančni pojavi izven standardnega modela v procesih s kvarkom t

Doktorska disertacija

MENTOR: Dr. Jernej F. Kamenik

Ljubljana, 2019

This is dedicated to my parents.

To my father. For sharing your curiosity and deep admiration for nature. I can't think of anybody more passionate about the mysteries of the universe than you.

To my mother. For the life you've built for us. I am very lucky to be your son. The years in Santa Sofia are the happiest moments in my life.

The scientist does not study nature because it is useful to do so. He studies it because he takes pleasure in it, and he takes pleasure in it because it is beautiful [...]. I am not speaking, of course, of the beauty which strikes the senses, of the beauty of qualities and appearances. I am far from despising this, but it has nothing to do with science. What I mean is that more intimate beauty which comes from the harmonious order of its parts, and which a pure intelligence can grasp.

Henri Poincaré

Acknowledgements

First and foremost, I would like to acknowledge Jernej F. Kamenik for being my mentor during this special journey. I am grateful for his guidance, his patience and for always being available for questions no matter the circumstances. These last four years have been a very rewarding and enriching experience. I want to also thank everyone at the IJS theory department for contributing to such a stimulating and open environment for research. I am in particular grateful to the head of the department, Svjetlana Fajfer, who always treated me very kindly and invited me into her circle of collaborators.

During my graduate years, I had the privilege of learning from very talented researchers. I want to thank Admir Greljo, Ezequiel Alvarez, Roberto Morales, Alejandro Szynkman, Jure Zupan, Svjetlana Fajfer, Nejc Košnik, Ilja Doršner, Damir Bečirević, Olcyr Sumensari, Andrei Angelescu, David Marzocca, Alejandro Celis, Jose E. Camargo-Molina, Monalisa Patra and Barry Dillon for collaborations in many interesting projects. I also want to thank Fagner Correia, Luis Valle Silva, Miha Nemevšek, Borut Bajc, and Saša Prelovšek for interesting physics discussions during coffee or during lunch, and my Ph.D. colleagues at F1: Vasja, Labrini, Luka, Zala, Urška, Urša, Victor, Aleks, and Jan. I am very much in debt with Allen Horvat for helping me multiple times with the cluster, and with Matic Lubej for the Slovenian translation of the thesis summary and abstract.

I am also grateful to many people that helped me before coming to Ljubljana. If it were not for them I would not be writing this thesis today. I would like to thank my professors at ICTP, in particular, Alexei Yu. Smirnov and Goran Senjanović for inspiring me to become a particle physicist and for their huge support during periods of uncertainty before coming to Ljubljana. I am also in debt with Fernando Quevedo and Narain Kumar for their important support during my Diploma years at ICTP. I want to also thank my colleagues in Trieste: Juan Carlos, Javier, Barbara, Alejandro Cabo, Kelvin, Victor, Gabin, Daniel, and Gabriel. I have very fond memories from my ICTP years because of you.

I was fortunate enough to receive my undergraduate degree at the Universidad Simón Bolívar in Caracas, right before the massive exodus of professors, academics and professionals fleeing the country, victims of a shameful regime. I am forever grateful for this (crucial) incubation period where we were protected from an otherwise adverse environment outside the walls of the university. Like everything that is universal, I have no doubt that La Simón will one day rise and shine again as it once did. I want to thank my advisor Isbelia Martin, my professors, and my dear friends from the 2004/2005 cohort: Cesar Ali, Eliel, Marco, Paul, Roger, Irene del Carmen, Alejandro, Javier, Anais, Juan, Fran, Ramses, Elram, el Chino, Pinky, and many others with whom I shared unique moments in SALTE, COF, and el Amper.

Last but not the least, I want to thank my parents and brother for their unconditional support during all these years, and Dijana for her love, companionship and patience.

Neresonančni pojavi izven standardnega modela v procesih s kvarkom t

Izvleček

V tej disertaciji analiziramo različne ne-resonantne manifestacije fizike zunaj Standardnega Modela, natančneje, v sektorju kvarka t. Tega se lotimo iz smeri od konca proti začetku, tako da predlagamo nove opazljivke in nove iskalne strategije za iskanje procesov s kvarki t. Z izkoriščanjem različnih kinematičnih območij več-resonančnih stanj, ki so ustvarjena pri hadronskih trkalnihih, uvedemo množico *idealiziranih sipalnih presekov*, zato da opišemo ne-resonančne procese nove fizike. Te opazljivke uporabimo za študijo produkcije $pp \to t\bar{t}h$ na trkalniku LHC in da pokažemo kako se lahko različni efektivni operatorji dimenzije 6 razpletejo pri HL-LHC. Fiziko izven Standardnega modela, ki se sklaplja neposredno s kvarki t, je težko izmeriti na trkalniku LHC, še posebej, če efekti niso resonantni. Za obravnavo tega predlagamo kompletno strategijo iskanja za meritev redkih SM procesov $pp \to t\bar{t}t\bar{t}$ pri energiji 13 TeV na trkalniku LHC v enako-nabitem dileptonskem in trileptonskem kanalu. Pokažemo omejitve, ki nakazujejo, da je na LHC Run-II možno dobiti dokaz s signifikanco 3σ za meritev $\sigma_{\rm SM}(pp \to t\bar{t}t\bar{t})$ in da je možno odkriti ta razpadv SM s signifikanco 5σ. Prav tako nastavimo limite na ne-resonantno fiziko zunaj SM, natančneje za sklopitve t-filičnih mediatorjev z maso pod pragom produkcije para kvarkov t. Trdimo, da je možnmo imeti UV kompletnost, za katero je proces $pp \to t\bar{t}t\bar{t}$ zelo močno orodje. Nenazadnje raziščemo možnost t-filične nove fizike, da povzroči deviacije pri razpadih mezonov B. V luči trenutnih deviacij od leptonske univerzalnosti v procesih $b \to s\ell^+\ell^-$ pri LHCb, predstavimo razlago, ki bazira na t-filični novi fiziki. Po analizi anomalij pri mezonih B z uporabo efektivnih teorij, klasificiramo vse potencialne t-filične mediatorje in izmed njih izločimo Z' in dvo-leptokvarkovski model ($\widetilde{U}_1 \sim (\mathbf{3}, \mathbf{1}, 5/3)$ in $R_2 \sim (\mathbf{3}, \mathbf{2}, 7/6)$). Te poenostavljene modele analiziramo na LHC in pokažemo, da lahko testirajo relevantna območja faznega prostora parametrov. Na koncu skonstruiramo enostaven razred UV modelov, ki živijo na skali reda velikosti TeV in predstavijo vektorske leptokvarke (ki so $\mathrm{SU}(2)_L$ singleti) kot umeritvene bozone SU(4) razširitve SM. Ta konstrukt uporabimo zato, da priskrbimo kompleten UV model za leptokvarke $\widetilde{U}_1$ in R_2, ki vstopajo v prehode $b \to s$ na eno-zančnem nivoju.

Ključne besede: fizika izven standardnega modela, top kvarki, fizika trkalnikov, LHC, neresonančna nova fizika, nova fizika v fiziki top kvarkov, efektivna teorija polja, poenostavljeni modeli, B-anomalije, kršitev univerzalnosti leptonskega okusa, leptokvarki, UV modeli.

PACS: 14.65.Ha, 12.60.Cn, 14.80.Ec, 14.80.Sv, 14.65.Jk

Non-Resonant Phenomena Beyond the Standard Model in Top-Quark Processes

ABSTRACT

In this dissertation we analyze different non-resonant manifestations of physics beyond the Standard Model in the top-quark sector. We do this from a bottom up approach by proposing new observables and new search strategies for top-quark processes. By exploiting the different kinematical regimes of multi-resonance production at hadron colliders, we introduce a set of *idealized cross-sections* to characterize non-resonant New Physics (NP). We use these observables to study $pp \to t\bar{t}h$ production at the LHC and show how different effective operators of dimension-6 can be disentangled at the HL-LHC. Physics beyond the SM coupling exclusively to top-quarks is challenging to measure at the LHC, in particular, if the effects are not resonant. To address this we propose a complete search strategy for measuring SM rare $pp \to t\bar{t}t\bar{t}$ production at the 13 TeV LHC in the same-sign dilepton and the trilepton decay channels. We produce bounds demonstrating that a 3σ evidence measurement of $\sigma_{\rm SM}(pp \to t\bar{t}t\bar{t})$ is accessible during Run-II and that a 5σ discovery of this process in the SM is possible. We also set limits on non-resonant BSM physics, in particular on the couplings of a top-philic mediator with masses below the top-pair production threshold. We argue that it is possible to have a UV completion for which $pp \to t\bar{t}t\bar{t}$ is a very powerful probe. Finally, we explore the possibility of top-philic NP producing deviations in B-meson decays. Inspired by the current deviations from lepton flavor universality in $b \to s\ell^+\ell^-$ at LHCb, we present an explanation based on top-philic NP. After analyzing the B-anomaly using effective theories, we classify all potential top-philic mediators and single out a Z' and a two-leptoquark model ($\widetilde{U}_1 \sim (\mathbf{3}, \mathbf{1}, 5/3)$ and $R_2 \sim (\mathbf{3}, \mathbf{2}, 7/6)$). We analyze these simplified models at the LHC and show that it can probe the relevant regions of parameter space. In the end, we construct a simple class of UV model living at the TeV scale that give rise to vector leptoquarks (that are $\mathrm{SU}(2)_L$ singlets) as gauge bosons of SU(4) extensions of the SM. We use this construction to provide a UV complete model for the $\widetilde{U}_1$ and R_2 leptoquarks entering $b \to s$ transitions at the one-loop level.

Keywords: beyond the standard model, top-quarks, collider physics, LHC, non-resonant new physics, new physics in top quark physics, effective field theory, simplified models, B-anomalies, lepton flavor universality violation, leptoquarks, UV models.

PACS: 14.65.Ha, 12.60.Cn, 14.80.Ec, 14.80.Sv, 14.65.Jk

Contents

List of figures

List of tables

Chapter 1

Introduction

In the year 2012, the last particle predicted by the Standard Model (SM) was discovered simultaneously by the ATLAS and CMS experiments at the world largest accelerator, the *Large Hadron Collider* (LHC), at CERN [10, 11]. The discovery of the Higgs Boson at the LHC marked the end of a six decades quest and the beginning of a new era of precision measurements at the high energy frontier. In the upcoming years, the LHC will be capable of probing energies above the electroweak scale with unprecedented precision in search for hints of physics beyond the SM (BSM). Besides the Higgs and gauge bosons, the only particle with a mass at the electroweak scale that could potentially play an important role in this search is the top-quark. The reason for this is threefold: first, the LHC is basically a top-quark factory that has already produced a huge amount of top-quarks with the aim of performing precision tests of the SM. Second, the top-quark is the only free quark that has a short enough lifetime to decay before hadronization kicks in and also the only fermion that interacts with the Higgs boson with a coupling of order $\mathcal{O}(1)$. Third, the current bounds from LHC Run-I and Run-II, Tevatron, electroweak precision data and low-energy flavor constraints are still the weakest for third generation quarks. This allows for new phenomena beyond the SM in top-quarks to be hidden inside regions of parameter space that still remain unexplored at energy scales accessible at the LHC. These are good enough reasons for directing our experimental and theoretical efforts into the top-quark sector for the remaining life span of the LHC.

With LHC Run-II coming to an end at a luminosity of 150 fb^{-1} of gathered data, no conclusive sign of New Physics (NP) has been announced by the LHC collaborations. This *null result* teaches us that all current measurements performed both by ATLAS and CMS are consistent with the Standard Model. As a consequence, the well-motivated theories from the past, like e.g. the minimal Supersymmetric SM (MSSM) or Large Extra Dimensions (LED), have been essentially ruled out in most regions of parameter space between the electroweak scale and the $\mathcal{O}(1)$ TeV energy scale. This has profound implications for theoretical particle physics, since the main theoretical justifications for pushing experimentally the energy frontier up to and beyond to the TeV scale may have been incorrect. However, we know that the SM cannot be the ultimate description of our physical reality. We have plenty of reasons to believe that a deeper theory must exist. The discovery of neutrino masses at the turn of the millennium and the observation of gravitationally interacting Dark Matter permeating our Universe are some of the best hints pointing towards physics beyond the SM. Unfortunately, none of these provide a compelling

indication of the precise scale at which NP should lie and does not justify searches for new fundamental degrees of freedom at high-energy colliders. This leaves BSM collider physics in a tight spot: for theorists, there is no theoretical motivation for building BSM models at low energy scales and for LHC experimentalists there isn't any well-motived NP model to probe, nor any input from phenomenologists that guarantees a discovery in the foreseeable future. BSM physics must be sincere with this situation and adapt accordingly. At the present moment, a bottom-up approach towards BSM is probably the best strategy for the LHC. In this dissertation, we follow this approach by proposing new observables and new search strategies for rare SM processes at the LHC. Our aim is to guide in an efficient yet model-independent way the experimental collaborations at the LHC towards a discovery.

One of the main discovery mechanisms at particle colliders are resonance searches. Massive particles produced on-shell in the s-channel that decay in two or more states near the collision vertex will exhibit in the final state invariant mass spectrum a distinctive bump or Breit-Wigner peak. The location of this peak is centered at the mass of the mother particle and the width of the peak is related to the lifetime. Some of the most important SM particles have been discovered in invariant mass *bump hunts*, like e.g. the Z boson discovered in dijet and dilepton spectrums, or more recently the Higgs boson in the diphoton and four-lepton spectrum. At the present moment, none of the standard BSM resonance searches performed by ATLAS and CMS have found any significant bump in the data at invariant masses above the top-quark boson mass. This situation leaves us with the following scenarios for BSM physics:

- (i) The new particles are too heavy to be directly produced on-shell at current LHC energies.
- (ii) The masses of the new particles are within reach but these have a non-trivial flavor structure that only allows for couplings to SM particles for which resonant production is kinematically forbidden.
- (iii) The new particles are produced on-shell at the LHC but are not possible to reconstruct from the decay products because of large combinatorics, missing energy carried by invisible final states, a too large resonance width or a combination of these three.

If any of these possibilities are realized in nature, then the standard bump hunts currently performed by ATLAS and CMS will fail to uncover *non-resonant* NP phenomena.

In this thesis, we investigate the collider phenomenology of BSM physics coupling predominantly to top-quarks in the quark sector, that give rise to non-resonant effects at the LHC. Our work is divided into three main parts, each related to one of the three scenarios described above. We mostly focus on a model-independent approach to BSM physics and employ effective field theories and simplified model descriptions.

If the new particles are too heavy to be produced at the LHC, this would imply a significant mass-gap between the electroweak scale and the NP scale. In this case, the effects of BSM physics can be described by an effective field theory where the NP effects at the LHC can be systematically incorporated into an infinite tower

of non-renormalizable operators and a set of corresponding Wilson coefficients that parametrize our ignorance of the underlying theory in the UV regime. In this case, even if no bump is present in the data, there will still be featureless deviations from the SM predictions in some of the observables. These deviations can potentially produce measurable effects in cross-sections or in the tails of the differential distributions of processes involving top-quarks and other massive particles like the Higgs boson. One interesting process, recently discovered by ATLAS and CMS is $pp \to t\bar{t}h$ production [12, 13]. This process is paramount for characterizing the Higgs boson properties and is also sensitive to many generic extensions of the SM. It is the most relevant tree-level process that serves as a direct probe of the top-quark Yukawa coupling, as well as other important properties, like e.g. the CP nature of the Higgs boson. Current methods of measuring $pp \to t\bar{t}h$ consist of combining multiple decay channels with very different systematic and statistical uncertainties. Improving the current results is a challenging task that may require a new approach. We devise a novel way of approaching $pp \to t\bar{t}h$ production or any process with a large multiplicity of massive SM resonances (t, h, $W^{\pm}$, Z) in the final state. For this we introduce a set of *idealized cross-sections* based on a very natural partition of the production process phase space into different kinematical regions. The partition is in fact inspired by boosted fat jet substructure techniques that can be easily used in an experimental setup. These idealized observables have most of the theoretical and experimental benefits of the so-called pseudo-observables currently used to study both electroweak production and decay of the Higgs boson at the LHC. As a proof of principle, we perform a computation of the idealized cross-sections for $pp \to t\bar{t}h$ and demonstrate how these are highly sensitive to different directions in BSM parameter space. We explicitly show this for a representative set of dimension-6 effective operators entering $pp \to t\bar{t}h$ production. This alternative way of characterizing top-quark and Higgs properties could be used during the High Luminosity (HL) LHC phase or at next-generation hadron colliders.

One of the aims of this dissertation is to show that NP could also be relatively light and hidden in exotic signals at the LHC. In particular, new light states that couple exclusively to top-quarks will not affect the traditional top-quark production processes or decay mechanisms at the LHC at tree-level. For concreteness, one can envisage a new *top-philic* force mediated by a light Z' boson with mass below the top pair production threshold. This scenario (corresponding to item (ii) described above) will not produce any bump in top-quark data given that the decay channel $Z' \to t\bar{t}$ is kinematically closed. One can argue that this specific type of NP in tops is one of the most challenging to attack experimentally since the new particles will only contribute to deviations in the cross-section of four-top production $pp \to t\bar{t}t\bar{t}$ at tree-level or in $t\bar{t}$ production at loop-level. The simultaneous production of four on-shell top-quarks at a hadron collider is one of the ultimate top-physics challenges for the LHC in the years to come. With a gargantuan production threshold close to the TeV scale (at $4m_t \approx 700$ GeV to be precise) this process would correspond to the largest SM feature in the invariant mass distributions potentially accessible at the LHC. Furthermore, it has been pointed out in several studies that $pp \to t\bar{t}t\bar{t}$ is a very good channel to probe for BSM theories like top-color models, partial top compositeness, etc. In the SM, four-top production is also a sensitive probe of the top-quark Yukawa coupling. Measuring the cross-section precisely would provide an independent handle for this quantity. Up to date, very preliminary attempts have

been made by phenomenologists to analyze the feasibility of measuring $pp \to t\bar{t}t\bar{t}$ at the LHC both in the SM and beyond. Experimental attempts have also been made, but have not been very successful. Until recently, the best limits on the SM cross-section were around $\sim \mathcal{O}(10)$ times above the SM prediction in decay channels with overwhelming backgrounds that will hardly improve with more data. With the ultimate aim of probing top-philic NP, we propose in this work a very simple cut-based search strategy for the SM $pp \to t\bar{t}t\bar{t}$ production based on the most sensitive decay channels, namely, the same-sign dilepton and the trilepton decay modes. We use our results to set the first limits on non-resonant top-philic Z' and top-philic scalar models. We also demonstrate that the LHC will be able to comfortably measure the SM four-top production cross-section in the near feature using our multi-lepton search strategy. Of course, the search proposed here can also be used to constrain heavier Z' or scalars for which the $t\bar{t}$ decay mode is kinematically accessible, or to set a limit on four-top contact interaction operators as well. The possibility of detecting top-philic mediators raises the question whether in nature a UV-complete theory for top-philic particles can exist without first being indirectly probed by experiments other than $pp \to t\bar{t}t\bar{t}$. To answer this we describe a fairly simple UV extension of the SM that give rise to a massive vector bosons and a scalar coupling dominantly to right-handed top-quarks.

In the final part of this thesis we take a leap into the realm of low-energy physics. The signs of NP coupling to top-quarks could make a first appearance as small discrepancies in precision observables measured in rare processes at scales much lower than the EW scale. If this is the case, then the well-known flavor pattern of the SM suggests that processes involving third-family quarks, namely the b-quark, are the most likely to exhibit the largest deviations from the SM. Indeed, rare B-meson decays with new heavy degrees of freedom and top-quarks running inside the loop are good candidates to search for the indirect effects of top-philic NP in low energy data. Furthermore, if the new heavy states also happen to couple to SM leptons with a non-trivial flavor structure, this might give rise to violations of lepton flavor universality (LFU) in such decays. The possibility of generating LFU violation in B-physics from beyond the SM physics in top-quarks is exciting because it would bridge two of the main theoretical and experimental pillars of particle physics: flavor physics at the high intensity frontier and collider physics at the high-energy frontier. Interestingly, several experimental indications of LFU violation beyond the SM in B-decays at B-factories, each ranging between 2 and 3 standard deviations from the SM prediction, have been reported in several different low-energy experiments in the last decade. To this date, whether these *B-physics anomalies* correspond to the first signs of NP or to statistical fluctuations is still an open question that will not be settled until more data is gathered in the next couple of years. Nonetheless, the persistence of these experimental discrepancies across completely different experiments during the years has given rise to many theoretical proposals to explain them with BSM physics. Of interest to us, are the discrepancies reported in $b \to s$ flavor-changing transitions measured by the LHCb collaboration at CERN in the semi-leptonic decays $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$ [14, 15, 16, 17, 18]. In recent years, the LHCb experiment has produced a series of precise measurements of LFU ratios comparing the muonic and electronic decay modes:

$$R_{K^{(*)}} \equiv \frac{\mathrm{Br}(B \to K^{(*)}\mu^+\mu^-)}{\mathrm{Br}(B \to K^{(*)}e^+e^-)}\,. \tag{1.1}$$

These two ratios have been measured to be $\approx 2.5\,\sigma$ away from the SM prediction in two different dilepton invariant mass bins. In addition, angular observables in the muonic channel have also been measured and confronted by the SM prediction. These also have been reported to be away from SM prediction at approximately $2\,\sigma$ and consistent with the $R_{K^{(*)}}$ anomaly. In our work, we explore the possibility of explaining these kinds of deviations with NP coupling predominantly to top-quarks and muons in the lepton sector, both within an effective field theory description and simplified dynamical models. In contrast to tree-level NP, one interesting consequence of accommodating $b \to s\ell^+\ell^-$ at the one-loop level is that the new states must not be too heavy if the theory is to be perturbative. In fact, loop-level top-philic explanations of the LHCb anomalies necessarily imply new particles with masses at the TeV scale. This opens the door wide open for probing these models in top-quark processes at the LHC. In this thesis we analyze the LHC phenomenology of different top-philic solutions of the $R_{K^{(*)}}$ anomalies and propose new search channels for the LHC, in particular, we demonstrate the capacity of LHC searches in $pp \to t\bar{t}t\bar{t}$ and $pp \to t\bar{t}\mu^+\mu^-$ to set relevant limits on Z' and leptoquark solutions of the LHCb anomalies. While resonant production is a possibility for TeV scale mediators in some of these processes, the presence of leptonic top decays in, e.g. the $pp \to t\bar{t}t\bar{t}$ search strategy, forbids an efficient reconstruction of the NP states due to large combinatorics in the final states and large missing energy carried away by several neutrinos. On a final note, having massive vector bosons that explain $R_{K^{(*)}}$ implies that more degrees of freedom must exist in order to have a consistent description of nature. This is even more important for leptoquark models because these will tend to arise from models where color is partially embedded in a larger gauge group. Interestingly enough, we show that it is possible to write down gauge theories based on SU(4) where any vector leptoquark representation that is a singlet under $\mathrm{SU}(2)_L$ is a gauge boson. Motivated by the B-physics anomalies we construct a top-philic UV model where the gauge leptoquark and a scalar give rise to LFU violation in $b \to s$ transitions at low energies.

This thesis is organized as follows: In Chapter 2 we provide a brief introduction to the SM physics followed by discussions about physics beyond the SM in Chapter 4, where we provide a general defintion of the top-philic flavor structure in the SM effective field theory. In Chapter 3 we introduce a new set of observables, the idealized cross-sections, for analyzing multi-resonance production processes like $pp \to t\bar{t}h$ in the SM and beyond. Next, in Chapter 5 we describe a complete search strategy for $pp \to t\bar{t}t\bar{t}$ within the SM and beyond, focusing on light top-philic BSM physics that can not be produced resonantly. In Chapter 6 we analyze the impact of top-philic NP in semi-leptonic B-decays within the EFT and simplified model framework and study the LHC phenomenology of these loop models. We end by proposing a UV complete model for gauge leptoquarks that can be used to explain the anomalies. In Chapter 6 we conclude.

Chapter 2

The Standard Model of Particle Physics

The Standard Model (SM) of elementary particle physics is one of the crowning scientific achievements of the 20$^{\text{th}}$ century. Developed in the 1960s and 1970s as a unified framework, the SM provides an accurate description of how the most fundamental building blocks in nature behave under all the fundamental forces except gravity. Mathematically the SM is based on a quantum field theory (QFT) exhibiting local gauge invariance. The gauge group describing the fundamental strong, weak and electromagnetic interactions is the Lie group

$$\mathcal{G}_{\text{SM}} \;=\; \text{SU}(3)_c \otimes \text{SU}(2)_L \otimes \text{U}(1)_Y \,, \tag{2.1}$$

where the first group SU(3) describes the theory of strong interactions, known as *Quantum Chromodynamics* (QCD), and the group $\text{SU}(2) \otimes \text{U}(1)$ describes the electromagnetic and weak interactions in a unified framework known as *electroweak theory*.

In the SM, the matter content is described by fermionic fields that can be categorized depending on how these transform under $\mathcal{G}_{\text{SM}}$. Fermions that transform non-trivially under the color group $\text{SU}(3)_c$ are known as *quarks*, while fields that are singlets under this group, i.e with no color charge, are known as *leptons*. In addition, one of the most consequential peculiarities of the SM is that the different chiralities of fermions transform differently under the weak isospin group $\text{SU}(2)_L$. For instance, the left-handed fields, ψ_L, transform non-trivially under weak isospin while the right-handed fields, ψ_R, are singlets of this group. This leads to the violation of discrete parity symmetry in weak decays [19]. The Lorentz structure of weak currents has the well known $V - A$ (vector minus axial) form [20]. On the other hand, strong and electromagnetic interactions are known to respect parity symmetry and are correctly described by vector currents. In table 2.1 we list the matter content of the SM.

2.1 Quantum Chromodynamics

In the SM, quarks are the fundamental constituents of hadrons such as the proton (baryons) or the pion (mesons). In particular, the baryon Δ^{++} with spin 3/2, is represented in the quark model as a bound state of three up-quarks $\Delta^{++} \sim uuu$ in

fields	$SU(3)_c$	$SU(2)_L$	$U(1)_Y$
Q_L^i	**3**	**2**	1/6
L_L^i	**1**	**2**	−1/2
u_R^i	**3**	**1**	2/3
d_R^i	**3**	**1**	−1/3
e_R^i	**1**	**1**	−1
φ	**1**	**2**	1/2

Table 2.1: Matter and scalar field content of the SM.

the ground state. Such a state is described by a totally symmetric wave function, leading to the violation of the spin-statistics theorem for fermions. Introducing a new hidden degree of freedom for quarks, *color* [21], resolved this paradox if quarks come in triplets of a fundamental color group $SU(3)_c$. In this case, $\Delta^{++} \sim \epsilon_{ijk} u^i u^j u^k$ is now described as a totally anti-symmetric field where ϵ is the totally antisymmetric tensor of the color group.

In this description all quarks, by definition, transform under the triplet (fundamental) representation of the color group $SU(3)_c$. This implies that the spectrum of hadrons is composed of color-singlet bound states: baryons of the form $\epsilon_{ijk} q^i q^j q^k$ and mesons of the form $q^i \bar{q}^i$. The three-fold nature of the color charge of quarks has been tested in several experiments, such as in e^+e^- colliders measuring the ratio of the cross-sections between hadronic and leptonic final states, π^0 decays to di-photons via the Adler-Bell-Jackiw anomaly, tau lepton decays, etc.

The QCD Lagrangian

In QCD, strong interactions are described by combining the $SU(3)_c$ Yang-Mills Lagrangian [22] with the Dirac Lagrangian (with partial derivative replaced by the covariant derivative):

$$\mathcal{L}_{\text{QCD}} \;=\; -\frac{1}{4} G^A_{\mu\nu} G^{A\,\mu\nu} + \overline{\psi^a_i}\,(i\gamma^\mu D^{ab}_\mu - m_i \delta^{ab})\psi^b. \tag{2.2}$$

Here the field strength tensor G and the covariant derivative D are defined as

$$G^A_{\mu\nu} \;\equiv\; \partial_\mu G^A_\nu - \partial_\nu G^A_\mu + i g_s f^{ABC} G^B_\mu G^C_\nu \tag{2.3}$$

$$D^{ab}_\mu \;\equiv\; \partial_\mu \delta^{ab} + i g_s G^A_\mu (T^A)^{ab}\,, \tag{2.4}$$

respectively, and ψ_i is the quark field with color index $a = 1, 2, 3$ and with flavor index i running over N_f different quarks each with mass m_i, G^A is the $SU(3)_c$ gauge field in the adjoint representation ($A = 1, ..., 8$), describing the eight gluon field mediating strong interactions, and g_s is the strong coupling constant. The matrices T^A are the eight standard generators of the SU(3) group satisfying the algebra and normalization

$$[T^A, T^B] \;=\; i f^{ABC} T^C\,, \qquad \text{Tr}(T^A T^B) = \delta^{AB}/2\,, \tag{2.5}$$

and f^{ABC} are the corresponding structure constants of SU(3). The most relevant prediction of the non-abelian Yang-Mills sector of QCD is the presence of gluonic

self-interactions originating from the non-linear terms in (2.3). These interaction terms are absent in abelian theories like QED and give rise to the triple-gluon and four-gluon interactions

$$\mathcal{L}_{\text{QCD}} \supset g_s f^{ABC}(\partial^\mu G^{A\,\nu}) G^B_\mu G^C_\nu - \frac{g_s^2}{4} f^{ABC} f^{ADE} G^{B\,\mu} G^{C\,\nu} G^D_\mu G^E_\nu . \tag{2.6}$$

One very important consequence of these non-linear self-interactions is *asymptotic freedom* [23, 24], which predicts that strong interactions become weaker in the Ultra-Violet (UV). In this regime, perturbative theory can be safely used to perform calculations and compared to experiments at high-energies. On the other hand, in the Infrared (IR) regime of the theory, QCD interactions become strong leading to a break down of perturbation theory at scales $\Lambda_{\text{QCD}} \sim 1$ GeV. In this regime, new phenomena arise, like quark confinement leading to the formation of bound states. Another important feature of unbroken Yang-Mills theories is that the vector gauge bosons (the gluons) need to be massless fields since a term of the form $m^2 G^A_\mu G^{A\,\mu}$ is not gauge invariant.

Hadron colliders

Since the central theme of this work is the study of New Physics at high-energy colliders we will now briefly describe in the remainder of this section the underlying physics behind hadron colliders like the LHC. Hadron colliders are factories of strong interactions. Once energies exchanged in the center-of-mass collisions of the incoming hadrons are sufficiently high, asymptotic freedom guarantees that the short-distance processes (the hard process) can be calculated in perturbation theory using the *parton model* [25]. On the other hand, the incoming hadrons involved in the collision have a complicated internal structure characterized by long-distance effects that cannot be described using a perturbative description. Fortunately, the unknown details of these long-range IR effects can be parametrized by a set of process-independent (universal) probability distributions, known as *parton distribution functions* (PDFs). Interestingly, because of the large energy scale separation, the hard partonic process and the long-range physics can disentangle from each other leading to a huge theoretical simplification of high-energy hadron collisions. This leads to the well-known *factorization theorem* in QCD [26]. For hadron colliders like the LHC, the (differential) cross-section of a proton-proton inelastic collision into N final states can be written using the master formula

$$d\sigma(pp \to X) = \sum_{a,b} \int_0^1 dx_a dx_b \int_V d\Phi_N \, f_a(x_a, \alpha_s, \mu_F) f_b(x_b, \alpha_s, \mu_F) \, d\hat{\sigma}(x_a, x_b, \mu_R, \mu_F) \,, \tag{2.7}$$

where a and b are the two colliding partons inside each proton, respectively, $\hat{\sigma}$ is the partonic cross-section of the hard process $ab \to 1, ..., N$ and $x_{a,b}$ are the fractions of longitudinal momentum of a and b with respect to each parent proton momentum. The functions $f_{a,b}$ each correspond to the PDFs describing each parton and depend on an unphysical factorization scale μ_F and the strong coupling $\alpha_s = \alpha_s(\mu_R)$ which itself depends on the renormalization scale μ_R. The partonic cross-section also depends on the renormalization and factorization scales in such a way that the

total integrand inside eq. (2.7) remains scale independent. The quantity $d\Phi_N$ is the differential phase space volume for the $2 \to N$ body process defined in the usual way by:

$$d\Phi_N \equiv \prod_{i=1}^{N} \frac{d^3 p_i}{2(2\pi)^3 E_i} (2\pi)^4 \delta^{(4)} \left(p_a + p_b - \sum_{i=1}^{N} p_i \right) \tag{2.8}$$

This quantiity is integrated over the *fiducial volume* V which does not cover the totality of the kinematic phase space. Instead, V is smaller regions delimited either by the physical volume of the detector or by a set of kinematic cuts purposefully designed to isolate certain subregions that highlight desired features of the underlying hard process. These experimental cuts are defined within a dedicated search strategy.

One of the aims of collider physics is to confront predictions from theory (e.g. the SM) with collider data. The quantity that experimentalists want to measure is exactly the hard process (differential) cross-section $\sigma(pp \to N)$. This can be achieved by performing careful event counting experiments at fixed integrated luminosities. The extracted measurements of the cross-sections have to then be compared to the theoretical computation of $\hat{\sigma}(ab \to N)$ from perturbative quantum field theory techniques (e.g. Feynman diagrams) up to a given order in perturbation theory. Notice that this comparison is only possible provided we have the information about the PDFs and the phase space integration region V. If this is so, then an unfolding procedure, or deconvolution, of the PDFs can be applied allowing for theory and experiment to be compared directly. Indeed, the PDFs $f_{a,b}$, while not computable from first principles, are instead available from many previous experiments. These have been carefully extracted experimentally during the years from fits to data in deep inelastic scattering and other experimental setups. The phase space integration volume V is also known a priori since it has been defined and optimized from the beginning after having specified a search strategy. The factorization theorem in (2.7) and the possibility of unfolding explains the success of hadron colliders like Tevatron and the LHC in recent years.

Of course going from real measurements down to the hard process at a proton-proton collider has many complications that originate in the very nature of QCD itself. For instance, a large amount of "excess" energy not involved in the hard collision $ab \to 1, ..., N$ goes into what is known as initial state radiation (ISR) and final state radiation (FSR). This happens when the colliding or the outgoing hard partons radiate one or more extra partons that are not directly involved in the hard process. This radiation can carry large amounts of transverse momentum and can complicate the idealized picture described above. Further complications arise from the underlying event (UE), i.e. the radiation from the remnant partons from the protons, or from pile-up (contamination from other colliding protons in the beams). Finally, the outgoing partons of the collisions, as well as the ISR and FSR are not the objects that are actually reconstructed and detected in the collider. Colored hard particles like gluons and quarks radiate many other partons that then condense (with the top-quark being the only exception) into colorless hadrons before entering the particle tracker. This implies that QCD has a natural "firewall" at the hadronization scale that prevents us to directly access the quantum numbers of any of the outgoing partons. To deal with this it is necessary to understand how partons evolve from very high energies down to the QCD hadronization scale and

then how hadronization takes place to produce bound states. Parton evolution is nicely explained in QCD using *parton showers* while hadronization, being a purely non-perturbative effect, must be described by fitting fragmentation models to data. Parton showers and hadronization are usually implemented in many Monte Carlo (MC) event generators as described later on.

Jets

One of the phenomenological trademark signatures of QCD at high-energy colliders is the presence of highly collimated sprays of hadrons, known as *jets* [27]. A jet emerges when one high-p_T initial state parton produced in the underlying collision, undergoes subsequent QCD splittings into a cascade of multiple intermediate gluons and quarks. Perturbative QCD predicts that these *parton showers* occur mostly inside a confined angular region along the direction of the initial parton seed ultimately giving rise to a jet once the radiated partons reach the non-perturbative hadronization scale around $\Lambda_{\text{QCD}} \sim 1$ GeV.

Reversing this picture provides the opportunity to access the initial parton seed and infer its properties if one can define and characterize jets in a systematic way. This is what jet physics is all about and what jet forming algorithms seek to achieve. The idea behind a jet algorithm is to cluster detector level objects (final state particle four-vectors, or calorimeter hits) in a way that is consistent with perturbative QCD, i.e. IR and collinearly safe (see [28] for a precise definition of IRC safety) and fast on the computational side. At the end of the clustering procedure, a 4-momentum is assigned to the emerging jet which now serves as a high-level representation of the underlying (inaccessible) hard parton[1]. In this sense, the study of jets (and their substructure) at colliders gives us crucial experimental access to the hadronic activity occurring at parton level inside the hard process.

The most reliable and successful jet clustering technique used at hadron colliders are the *jet sequential recombination* schemes. The widespread k_T [29], anti-k_T [30], and Cambridge/Aachen (C/A) [31, 32] jet algorithms used at the LHC belong to this category of jet algorithms. In the language of unsupervised learning, these algorithms fall under the umbrella of measure-based *agglomerative hierarchical clustering* models. The algorithm is defined as follows:

- Input: N data point p_i (particles) in the data sample $\mathcal{D} = \{p_1, p_2, ...p_N\}$ (event), a distance function $\langle \cdot, \cdot \rangle$ and the empty set $\mathcal{C} = \{\emptyset\}$.
- (i) select one data point $p_i \in \mathcal{D}$
- (ii) Find $p_j \in \mathcal{D}$ that minimizes $\langle p_i, p_j \rangle$.
- (iii) If a *merging criterion* is satisfied, then merge (cluster) $p_i \cup p_j \to p_k$ in $\mathcal{D}$.
- (iv) Else, remove p_i from $\mathcal{D}$ and include in $\mathcal{C}$.
- (v) Repeat until $\mathcal{D}$ is empty.

[1] In principle what is meant by "partons" is also ambiguous and must also be defined precisely, see [28].

After finalizing the run over all the data $\mathcal{D}$ in the sample, the algorithm outputs a set of clusters $\mathcal{C}$ with K elements such that $0 < K < N$. The generalized-k_T algorithm for jet clustering uses exactly this procedure where the data points p_i are hadrons or particles defined by their kinematics $(p_T^i, \eta^i, \varphi^i)$, i.e. transverse momentum, rapidity, and azimuthal angle. The distance function between two particles is defined as

$$\langle p_i, p_j \rangle \;=\; \min\left(p_{i_T}^{2\alpha}, p_{j_T}^{2\alpha}\right)\left(\delta_{ij} \;+\; \frac{\Delta R_{ij}^2}{R^2}\right) \tag{2.9}$$

where the planar distance is defined as $\Delta R_{ij}^2 \equiv (\varphi_i - \varphi_j)^2 + (y_i - y_j)^2$ and the three different choices $\alpha = \{1, -1, 0\}$ define the three common types of clustering procedures, k_T, anti-k_T and C/A, respectively. Here, R is a free parameter (or hyperparameter) of order 1 known as the clustering cone radius. The resulting clusters in $\mathcal{C}$ are the jets. The merging operation $p_i \cup p_j \to p_k$ in step (iii) is simply 4-momentum addition $p_k \equiv p_i + p_j$ and the merging criterion is given by the condition

$$\min(\langle p_i, p_i \rangle, \langle p_j, p_j \rangle) \geq \langle p_i, p_j \rangle \,. \tag{2.10}$$

Notice that the distance function $\langle \cdot, \cdot \rangle$ is not formally a metric since the reflexive axiom $\langle p, p \rangle = 0$ is not satisfied[2]. The clustering cone hyperparameter R used in the algorithm fixes the geometric size of the clustered jet and also controls the final number of jets in the event. R is an ad hoc parameter that the user must fix beforehand depending on the kinematics of the process he or she wishes to study at the collider. In LHC searches, the default clustering cone radius for non-boosted underlying partons is $R = 0.4$. When characterizing boosted resonances such as top-quarks or Higgs Bosons, the cone radius for the clustering algorithm is usually taken with a wider angle of $R \gtrsim 1$, leading to large area *fat jets*. These fat jets capture most of the overlapping (sub)jets generated by the daughter partons.

As stressed above, jets originate from the final state partons produced in the hard collision. Unfortunately, there is not a one-to-one mapping between the number of jets and the number of hard partons coming from the collision. The presence of ISR and FSR will usually give rise to additional jets. The number of jets is also dependant on the choice of the cone radius hyperparameter R and also on the kinematics of the initial partons. Besides this, jets being prolonged objects may also accidentally overlap inside the detector, complicating things further.

Simulation pipeline for collider studies

In order to perform reliable phenomenological collider studies one needs to rely on MC event generators and simulations that can model all the different physics involved in hadron collisions. We now introduce some of the computational tools that we have used extensively for the work presented in this thesis. These tools can be thought of as a semi-automatized pipeline. At one end of this pipeline, we input a Lagrangian $\mathcal{L}$ that potentially describes nature along with a cleverly chosen process $pp \to X$ that we believe highlights certain aspects of $\mathcal{L}$ that we desire to study at the collider. At the other end of the pipeline, an output is produced corresponding to the "response" of an idealized representation of the detector (e.g. ATLAS, CMS, LHCb) to the input process $pp \to X$ computed using the Lagrangian $\mathcal{L}$. Between input

[2]In mathematics metrics that do not satisfy the reflexive axion are known as *meta-metrics*

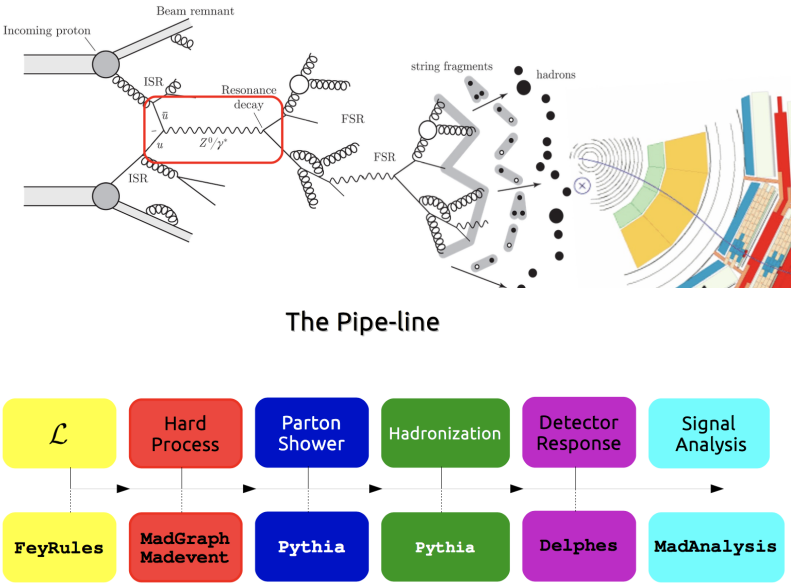


Figure 2.1: (Above) Schematic representation of the underlying physics at a proton-proton collider. The hard process ($q\bar{q} \to Z/\gamma^* \to q\bar{q}$) is shown inside the red rectangle. (Bellow) The simulation pipeline for collider studies used throughout this work.

and output, there is a sequence of tools feeding into each other that closely mirror the different physics present in (2.7) in high-energy collisions. This is illustrated in figure 2.1.

The target Lagrangian $\mathcal{L}$ is first implemented in a `Mathematica` package called `FeynRules` [33]. This program extracts the Feynman rules necessary for the matrix element computation $ab \to N$. The output of `FeynRules` in the UFO model format [34]) is then fed into `Madgraph` [35]. At this level, the total cross-section of the (PDF-convoluted) process $pp \to X$ is computed using $\mathcal{L}$ at leading order (LO) in QCD and MC events are generated. These partonic events are then taken as input for `Pythia` [36] where they are showered and then ultimately hadronized. The final state hadrons in each event are then clustered into jets using, typically, the anti-k_t jet algorithm implemented in `FastJet` [37]. The energy of each clustered jet and other final state objects (photons, leptons, missing energy) are then smeared and reconstructed using the fast detector simulator `Delphes` [38]. Finally, events are then filtered by applying a set of kinematic cuts and object selections defining a *signal region* with the aim of isolating the most interesting parts of phase space. These selection cuts, defined by the search strategy, are imposed on the event sets using software like `ROOT` [39] or `MadAnalysis` [40] in order to extract signal and background efficiencies.

2.2 Electroweak Theory

The starting point of electroweak theory can be traced back to Fermi's theory of β-decays [41]. After Pauli proposed the existence of the neutrino ν_e in order to save conservation of energy, Fermi devised a theory for the neutron β-decay $n \to pe^-\bar{\nu}_e$ based on the four-fermion interaction Lagrangian

$$\mathcal{L}_\beta \;=\; -\frac{G_F}{\sqrt{(2)}}\,(\bar{p}\,\Gamma n)\,(\bar{e}\,\Gamma\nu_e) + \text{h.c.} \tag{2.11}$$

where G_F is the well-known dimensionful Fermi constant $G_F \simeq 10^{-5}\,\text{GeV}^{-2}$ and $\Gamma = \sum_I c_I \Gamma^I$ is an element of the space-time Clifford algebra generated by the basis elements $\Gamma^I = \{1, \gamma^5, \gamma^\mu, \gamma^\mu\gamma^5, \sigma^{\mu\nu}\}$ corresponding, respectively, to the scalar, pseudoscalar, vector, pseudovector and tensor elements and γ^μ are the 4×4 Dirac matrices satisfying the Dirac algebra $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$. In the original theory of beta decay, the Lorentz structure of (2.11) was chosen to be vectorial $\Gamma = \gamma^\mu$ in analogy to electric currents. Later on, the discovery of parity violation in weak interactions led to the rethinking of the Lorentz structure of Γ^I and proposal of the $V - A$ structure of weak interactions with $\Gamma = \gamma^\mu(1 - \gamma^5)/2$. The Lagrangian (2.11) is of mass dimension-6, hence, non-renormalizable. This implies that in this theory is not well behaved in the UV, i.e. certain processes violate unitarity (e.g. cross-sections growing linearly with the center-of-mass energy) and also the divergent quantities appearing in higher-orders in perturbation theory cannot be reabsorbed in a finite number of bare parameters leading to untrustworthy predictions.

The pathologies present in the $V - A$ four-fermion Lagrangian described above can be removed if we think of this theory as an effective theory at low energies of a more fundamental theory that is renormalizable, much in the same spirit as electromagnetism. Oskar Klein postulated an intermediate massive vector boson W with the EM-like interaction

$$\mathcal{L}_W \;=\; -\frac{1}{4}W_{\mu\nu}W^{\mu\nu} \;+\; \frac{M_W^2}{2}W_\mu W^\mu \;+\; g\,(\bar{p}\,\gamma_\mu n + \bar{e}\,\gamma_\mu\nu_e)W^\mu + \text{h.c.} \tag{2.12}$$

with a dimensionless coupling g, vector boson mass M_W and field strength $W_{\mu\nu}$. Such an interaction, for momentum transfers $q \ll M_W$ leads to the effective Lagrangian (2.11) with $G_F = \sqrt{2}g^2/M_W^2$ if the vector boson propagator has the structure:

$$i\Delta_{\mu\nu} \;=\; -i\Big(g_{\mu\nu} - \frac{p_\mu p_\nu}{M_W^2}\Big)/(p^2 - M_W^2). \tag{2.13}$$

The asymptotic behavior of the second "longitudinal" term in the propagator (2.13) for large momenta saturates and becomes constant for $p \to \infty$ leading again to a non-renormalizable theory. This issue, originating from the massive nature of the W vector boson is elegantly solved by the Brout-Englert-Higgs mechanism for gauge theories.

$\text{SU}(2) \otimes \text{U}(1)$ Unification

The fact that both the electromagnetic and weak interactions can be described by intermediate vector bosons coupling to matter currents led to the proposal of electroweak unification. Glashow showed that the correct theory achieving unification while still incorporating the $V - A$ structure of weak interactions along with the vectorial nature of electromagnetism is given by a gauge theory based on the group $\text{SU}(2)_L \otimes \text{U}(1)_Y$, where $\text{SU}(2)_L$ is the *weak isospin* group and $\text{U}(1)_Y$ is the abelian *weak hypercharge* group [42]. In this theory, the electromagnetic group $\text{U}(1)_{\text{em}}$ is embedded as a diagonal subgroup of the electroweak group $\text{SU}(2)_L \otimes \text{U}(1)_Y$ with the electric charge operator given as a linear combination

$$Q \;=\; T^3 + \frac{1}{2}Y\,, \tag{2.14}$$

of the Cartan generator T^3 of $\mathrm{SU}(2)_L$ and the abelian hypercharge generator Y. This expression relates the electromagnetic charge to the weak isospin.

Since weak interactions have a $V-A$ Lorentz structure, this implies that the weak boson W only couples to the Left-Handed (LH) fermion doublets $\psi_L \equiv \frac{1}{2}(1-\gamma^5)\psi$, while Right-Handed (RH) fermions, defined as $\psi_R \equiv \frac{1}{2}(1+\gamma^5)\psi$, are singlets under the weak isospin group. For leptons and quarks the representations are:

$$\psi_L \;=\; \begin{pmatrix}\nu_\ell\\ \ell\end{pmatrix}_L\,,\quad \begin{pmatrix}u\\ d\end{pmatrix}_L\,,\quad \text{and}\quad \psi_R = \ell_R,\; u_R,\; d_R\,, \tag{2.15}$$

where a family index is implied. The gauge-invariant Lagrangian for leptonic fields is

$$\mathcal{L}_{\mathrm{EW}} \;=\; -\frac{1}{4}W^a_{\mu\nu}W^{a\,\mu\nu} - \frac{1}{2}B_{\mu\nu}B^{\mu\nu} + \overline{\psi^a_L}(i\gamma^\mu D^L_\mu)^{ab}\psi^b_L + \overline{\psi_R}(i\gamma^\mu D^R_\mu)\psi_R \tag{2.16}$$

with the field strengths defined by

$$W^I_{\mu\nu} \;\equiv\; \partial_\mu W^I_\nu - \partial_\nu W^I_\mu + ig\,\varepsilon^{IJK}W^J_\mu W^K_\nu \tag{2.17}$$

$$B_{\mu\nu} \;\equiv\; \partial_\mu B_\nu - \partial_\nu B_\mu \tag{2.18}$$

and covariant derivatives acting on LH and RH fields, respectively, are

$$(D^L_\mu)^{ab} \;\equiv\; \partial_\mu\delta^{ab} + ig\,\frac{1}{2}W^I_\mu(\sigma_I)^{ab} + ig'\frac{1}{2}YB_\mu\delta^{ab} \tag{2.19}$$

$$D^R_\mu \;\equiv\; \partial_\mu + ig'\frac{1}{2}YB_\mu\,. \tag{2.20}$$

Here $a,b = 1,2$ are weak isospin indices, W^I are the $\mathrm{SU}(2)_L$ gauge fields in the adjoint representation ($I = 1,2,3$), σ^I are the three Pauli matrices, ε^{IJK} is the totally antisymmetric structure constant and g is the weak coupling. B is the abelian gauge field associated with the hypercharge group and g' is the corresponding gauge coupling constant. Notice that since LH and RH fields have different representations under weak isospin, the fermion Dirac mass terms of the form $m\overline{\psi_L}\psi_R$ are strictly forbidden at this level in the theory.

In total, there are four gauge bosons that appear in the spectrum, two charged vectors $W^\pm_\mu$ given by the raising/lowering operators of $\mathrm{SU}(2)_L$, and two neutral vectors A_μ and Z_μ which are orthogonal admixtures of the diagonal weak isospin field W^3_μ and the hypercharge field B_μ:

$$W^\pm_\mu \;=\; \frac{1}{\sqrt{2}}(W^1_\mu \mp i\,W^2_\mu) \tag{2.21}$$

$$Z_\mu \;=\; \cos\theta_w\,W^3_\mu - \sin\theta_w\,B_\mu \tag{2.22}$$

$$A_\mu \;=\; \sin\theta_w\,W^3_\mu + \cos\theta_w\,B_\mu\,. \tag{2.23}$$

The massless photon is identified with A_μ, the massive W boson mediating β-decays and other charged weak interactions are identified with $W^\pm_\mu$, and the massive neutral boson Z^μ is responsible for new neutral interactions. The *weak mixing angle* θ_w is a new parameter of the SM measuring the relative strength between each gauge coupling $\tan\theta_w = g'/g$ and also relating them to the electromagnetic coupling via $e = g\sin\theta_w = g'\cos\theta_w$. This quantity is fixed by experiments to be $\sin^2\theta_w \approx 0.23$. The electroweak theory outlined above when taken to low energies gives rise to the Fermi theory for weak interactions and electromagnetism. It also predicts new phenomena, namely the presence of *neutral weak interactions*, where, e.g. neutrinos scatter off elastically from a target nucleus in a bubble chamber without changing the charge of the outgoing nucleus. This process was measured by the Gargamelle experiment at CERN and was the first indirect evidence of a new massive neutral Z boson, as well as the first measurement of the weak angle θ_w.

The Brout-Englert-Higgs mechanism

One direct consequence of gauge-invariance in non-abelian Yang-Mills theories like the electroweak theory is that all of the gauge bosons (2.21) are predicted to be massless. As discussed in the previous section, because of the different SU(2) representations of LH and RH fermion fields, explicit mass terms for fermions are also forbidden by gauge invariance. While elegant, the electroweak theory described up to now can't be a theory of nature because we know that fermions and vector bosons (aside from the photon) must be massive. One way to provide masses to the fields while keeping gauge invariance is via spontaneous symmetry breaking. Examples of such theories are common in condensed matter. For example, in a ferromagnetic system the Hamiltonian is rotationally symmetric, yet when the system is cooled down to the ground state, the spins of each atom collectively align into one randomly selected preferred direction. Hence, the ground state solution of a ferromagnet violates the rotational invariance of the system. In such a scenario, the symmetry of the system is said to be *spontaneously broken*. An important consequence is that the massless (gauge) modes in these models acquire an effective mass upon symmetry breaking. This idea was extended to relativistic quantum field theories by Higgs [43], Brout and Englert [44], Guralnik, Hagen and Kibble [45], for an abelian gauge theory with a complex scalar field. Besides providing a mass to the gauge boson via spontaneous symmetry breaking, the theory also predicts the existence of a new scalar propagating degree of freedom, *the Higgs boson*.

Along these lines, Salam [46] and Weinberg [47] incorporated into the electroweak theory of Glashow a new complex scalar field φ transforming as a doublet under weak isospin

$$\varphi \;=\; \begin{pmatrix} \varphi^+ \\ \varphi^0 \end{pmatrix} \tag{2.24}$$

and hypercharge $Y(\varphi) = 1/2$. The scalar sector Lagrangian is given by the usual kinetic term and the scalar potential V

$$\mathcal{L}_{\text{Higgs}} \;=\; (D_\mu\varphi)^\dagger D^\mu\varphi \;+\; V(\varphi) \tag{2.25}$$

given by

$$V(\varphi) \;\equiv\; -\mu^2\varphi^\dagger\varphi \;+\; \lambda(\varphi^\dagger\varphi)^2\,. \tag{2.26}$$

If $\mu^2 > 0$ then the quadratic term in the potential would be an ordinary scalar mass term with no symmetry breaking properties. On the other hand, if we chose instead $\mu^2 < 0$, then the surface $V(\varphi)$ has a non-trivial geometry near the origin with the well-known *mexican hat* shape (see Figure 2.2 for a visualization). The field configuration extremizing the Higgs potential is obtained from the equation $\varphi\,(\varphi^\dagger\varphi - v^2/2) = 0$, where $v \equiv \mu/\sqrt{\lambda}$. The ground-state vacuum solution is given by the vacuum expectation value (VEV) $\langle\,\varphi^\dagger\varphi\,\rangle = v^2/2$. In the electroweak theory, the VEV of the Higgs leaving electromagnetism unbroken is given by the vacuum configuration aligned with the lower component

$$\langle\,\varphi\,\rangle \;=\; \frac{1}{\sqrt{2}}\cdot\begin{pmatrix}0\\ v\end{pmatrix} \tag{2.27}$$

The scalar field excitations around this configuration correspond to a new physical scalar degree of freedom, the Higgs boson h. Notice that the symmetry breaking order parameter v is the only fundamental mass scale in the SM electroweak theory. The $W^\pm$ and Z electroweak gauge bosons acquire masses proportion to this quantity. These can be extracted by replacing the VEV inside the kinetic term of the Higgs doublet. The mass of the charged boson is found to be $M_W = g\,v/2$. From this and the value of the Fermi constant one can extract the value of the VEV to be at $v \approx 246$ GeV. The neutral gauge fields W^3 and B are described by the 2×2 mass matrix

$$M_0 \;=\; \frac{v^2}{8}\begin{pmatrix} g^2 & gg' \\ gg' & g'^2\end{pmatrix}\,. \tag{2.28}$$

This matrix can be diagonalized into the mass eigenstates Z and A in (2.21) by a 2×2 rotation matrix parametrized by the weak mixing angle θ_w. The photon here is the massless eigenvalue of (2.28) and the mass of the Z boson is predicted to be slightly higher than the $W^\pm$ boson mass, namely $M_Z = M_W/\cos\theta_w$.

Yukawa sector

The Higgs scalar doublet can also give mass to the fermions once Yukawa interactions are introduced. For both leptons and quarks these read:

$$-\;\mathcal{L}_{\text{yuk}} \;=\; y^d_{ij}\;(\overline{Q^i}_L\,\varphi\,d^j_R) \;+\; y^u_{ij}\,(\overline{Q^i_L}\,\tilde{\varphi}\,u^j_R) \;+\; y^e_{ij}\,(\overline{L^i}\varphi\,e^j) \;+\; \text{h.c.}\,, \tag{2.29}$$

where $\tilde{\varphi} = i\sigma_2\varphi^*$ and $y^{u,d,e}$ are 3×3 complex Yukawa matrices for the down-type quarks, up-type quarks and charged leptons. Notice that here we have included a *generation* index $i = 1,2,3$ accounting for the three lepton and quark generations in the SM, we describe this in more detail in the next section. Once spontaneous symmetry breaking occurs $\varphi^0 \to (v+h)/\sqrt{2}$, the Yukawa interactions in (2.29)

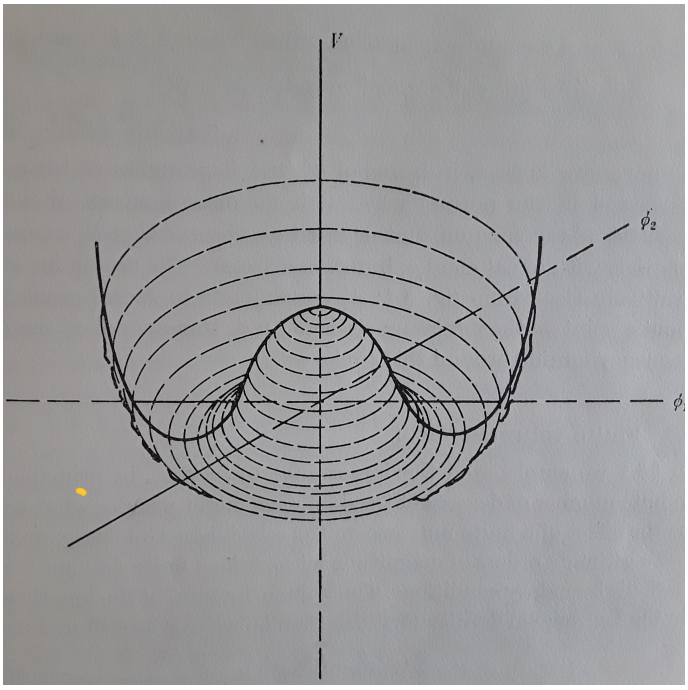


Figure 2.2: The Higgs *mexican hat* or *sobrero* potential.

give rise to mass terms for fermions. Neutrinos remain massless in the SM since no right-handed neutrino field is present in the matter content[3]:

$$-\mathcal{L}_{\text{mass}} = m^d_{ij}\,\bar{d}^i d^j + m^u_{ij}\,\bar{u}^i u^j + m^e_{ij}\,\bar{e}^i e^j \tag{2.30}$$

where m^d, m^u and m^e are 3×3 non-diagonal matrices for down quarks, up quarks and charged leptons, respectively, defined by

$$m^{d,u,e} = \frac{v}{\sqrt{2}}\, y^{d,u,e}\,. \tag{2.31}$$

These matrices have to be diagonalized using bi-unitary transformations for each type of field. The eigenvalues of these matrices correspond to the physical masses of the quarks and leptons, which are fundamental parameters of the SM that need to be extracted from experiments.

2.3 The Physics of Flavor

For being a fundamental theory of nature, the SM has a surprisingly rich particle content with a non-trivial structure. This can be attributed to the fact that quarks and lepton fields come in three identical replicas, or *generations*. First hints of this now well-known fact appeared in the '30s when a heavier "cousin" of the electron was discovered in cosmic rays. The muon is described by a fermionic field in exactly the same SM representation as the electron. The only difference is the value of its mass

[3]Of course, we now know that this prediction of the SM is wrong since neutrino oscillations from multiple sources (atmospheric, solar or artificial) have been carefully measured indicating that neutrinos are actually massive particles.

which is approximately ~ 200 times heavier than the electron, making it unstable. When the discovery of the muon was announced, a surprised I. Rabi famously asked: *who ordered that?* Similar discoveries where made in the following decades, like e.g. kaons, now known to be bound states of strange-quarks which are heavier cousins of the down-quark, or the top-quark a very heavy version of the up-quark. To this date, the reason why the fundamental building blocks of matter can be categorized into three identical replicas is still a mystery.

Flavor symmetries

We now know that three generations of fermions exist in nature:

- First Generation: the electron $e^{\pm}$, the electron-neutrino ν_e in the lepton sector, and the up-quark u and down-quark d in the quark sector.
- Second Generation: the muon $\mu^{\pm}$, the muon-neutrino ν_μ in the lepton sector, and the charm-quark c and strange-quark c, s in the quark sector.
- Third Generation: the tau-lepton $\tau^{\pm}$, the tau-neutrino τ_e in the lepton sector, and the top-quark t and bottom-quark b in the quark sector.

If one turns off all the Yukawa couplings ($y^{u,d,e} \to \mathbb{0}$) the SM Lagrangian exhibits a large global accidental symmetry $\mathcal{G}_\mathcal{F}$ known as *flavor symmetry*. The group $\mathcal{G}_\mathcal{F}$ is defined as

$$\mathcal{G}_\mathcal{F} \;=\; \mathcal{G}_\mathcal{F}^q \otimes \mathcal{G}_\mathcal{F}^\ell \,, \tag{2.32}$$

where each factor rotates quark and lepton fields of different chiralities independently in 3-dimensional flavor space:

$$\mathcal{G}_\mathcal{F}^q \;\equiv\; \mathrm{U}(3)_Q \otimes \mathrm{U}(3)_d \otimes \mathrm{U}(3)_u \,, \tag{2.33}$$
$$\mathcal{G}_\mathcal{F}^\ell \;\equiv\; \mathrm{U}(3)_L \otimes \mathrm{U}(3)_e \,. \tag{2.34}$$

These global flavor symmetries are not exact symmetries of nature. The non-zero hierarchical Yukawa interactions in (2.29) explicitly break these global symmetries into the subgroup $\mathcal{H}_\mathcal{F}$:

$$\mathcal{G}_\mathcal{F} \;\longrightarrow\; \mathcal{H}_\mathcal{F} \;=\; \mathrm{U}(1)_B \otimes \mathrm{U}(1)_e \otimes \mathrm{U}(1)_\mu \otimes \mathrm{U}(1)_\tau \otimes \mathrm{U}(1)_Y \tag{2.35}$$

where $\mathrm{U}(1)_B$ is the Baryon number, $\mathrm{U}(1)_{e,\mu,\tau}$ are the individual lepton numbers, and the last group is the gauged hypercharge. This breaking occurs because the Yukawa matrices $y^{u,d,e} \neq y\,\mathbb{1}$ where y would be some proportionality constant. Ultimately this is because we know that each fermion after spontaneous symmetry breaking has a different physical mass and are all far from being degenerate.

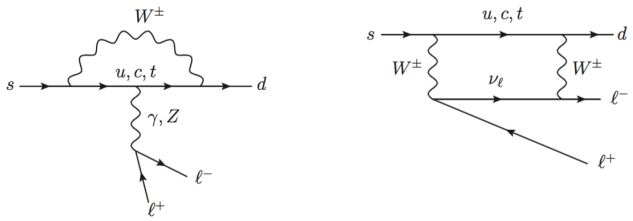


Figure 2.3: Two one-loop SM processes leading to FCNC.

The Cabibbo-Kobayashi-Maskawa matrix

In the SM, hadronic charged currents are where flavor physics has the largest phenomenological impact. These are described by the tree-level interactions between the $W^\pm$ boson and the quark currents:

$$\mathcal{L}_{\rm SM} \supset \mathcal{L}_{\rm c.c.} = \frac{g}{\sqrt{2}} V^{ij}_{\rm CKM} W^\pm_\mu \left(\bar{u}^i_L \gamma^\mu d^j_L\right) . \tag{2.36}$$

The fields u^i and d^j represent the physical mass eigenstates of the up and down quarks obtained once the mass matrices in (2.31) are diagonalized with the bi-unitary transformations $\Delta m^d \equiv \text{diag}(m_d, m_s, m_b) = V^d_L\, m^d\, (V^d_R)^\dagger$ and $\Delta m^u \equiv \text{diag}(m_u, m_c, m_t) = V^u_L\, m^u\, (V^u_R)^\dagger$. In this case the resulting 3×3 matrix $V_{\rm CKM}$ in (2.36), known as the Cabbibo-Kobayashi-Maskawa (CKM) mixing matrix [48, 49], reads $V_{\rm CKM} = (V^u_L)^\dagger V^d_L$. This matrix measures the mis-alignment in flavor space between each family. Notice that the CKM mixings, as well as the quark masses are not derived quantities in the SM. These mixing elements must be extracted directly or indirectly from experiments involving charged current transitions. Experiments have revealed that the structure of the CKM mixing matrix is given by a perturbation of the unit matrix. In fact, it can be conveniently parametrized via

$$V_{\rm CKM} = \begin{pmatrix} V_{ud} & V_{cd} & V_{td} \\ V_{us} & V_{cs} & V_{ts} \\ V_{ub} & V_{cb} & V_{tb} \end{pmatrix} \approx \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(1\rho - i\eta) \\ -\lambda & 1 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} \tag{2.37}$$

where we neglected contributions of order $\mathcal{O}(\lambda^4)$ or higher, and λ is the Cabibbo expansion parameter with experimental value of $\lambda \approx 0.2251(3)$. The Wolfenstein parametrization [50] given above has also three other parameters of order $\mathcal{O}(1)$ that must be fitted to data. The fit to flavor-changing processes yields [51]

$$A = 0.833(12)\,, \qquad \rho = 0.157(14)\,, \qquad \eta = 0.352(11)\,. \tag{2.38}$$

More interestingly, the imaginary parameter η, which is responsible for charge-parity (CP) violation in weak decays, can be traced back to the fact that the Yukawa couplings are in fact complex numbers. For two generation the complex phases in the 2×2 mixing matrices end up being unphysical because they can all be absorbed by redefining the quark fields, leading to CP conservation. For three generations the situation is different. All complex phases but one can be reabsorbed by field redefinition. The remaining physical complex phase, the Kobayashi-Maskawa phase $\delta_{\rm KM}$,

is the responsible quantity for CP violation in the SM [49]. In fact, the introduction of a third quark family (t and b quarks) in the '70s by Kobayashi and Maskawa came from the phenomenological necessity to explain the observation of CP violation in meson mixing. The discovery of the Υ meson (a quarkonia state of bottom quarks) and later the discovery of top-quarks at high-energy colliders confirmed the existence of the third quark family.

One interesting feature of the SM is the absence of flavor changing neutral currents (FCNC) in at tree-level. In the SM, the only source of flavor violation and CP violation come from the CKM mixing matrix and are mediated by the W boson. The Z boson, the photon and the gluon couplings to fermionic currents are flavor conserving because of the unitarity of the $V_{L,R}^{u,d}$ matrices. Nevertheless, FCNC does arise in the SM in processes generated at the one-loop level, see figure 2.3 for a couple of examples. These FCNC processes, being both loop and GIM suppressed [52], imply that these are very rare processes making them good places to search for deviations caused by New Physics (e.g. FCNC arising in tree-level exchanges of new mediators).

Lepton flavor univerality violation

One of the main predictions of the SM is that the coupling of leptons with gauge bosons are flavor-independent in the limit where the leptonic masses can be neglected. This is a direct consequence of the restoration of the global flavor symmetry $\mathcal{G}_{\mathcal{F}}^{\ell}$ in this limit. This is usually referred to as *lepton flavor universality* (LFU). Of course, since leptons have (small) masses, LFU in the SM is actually violated, but in a predictable way. LFU can be nicely tested in low energy precision experiments by measuring the same observable o in a (semi-)leptonic process and comparing the yields for the three different leptons $\ell_i = \{e, \mu, \tau\}$ by forming the ratios $R_{\ell_1\ell_2} = o_{\ell_1}/o_{\ell_2}$. These *LFU ratios* are theoretically clean observables and are predicted in the SM to satisfy $R_{\ell_1\ell_2} \to 1$ in the limit $m_{\ell_1}, m_{\ell_2} \to 0$.

LFU in weak interactions is one of the key predictions of the SM. Searching for the possibility of new sources of LFU violation in experiments is a promising way of looking for BSM physics. Experimentally it has been probed at the percent level precision both directly in W decays at LEP [53], but also indirectly via precision measurements of pion, kaon, D meson and tau lepton decays (see for example [54, 55, 56, 57]). Recently, evidence for LFU violation has accumulated in several experiments testing the SM hypothesis in semi-leptonic B decays. We describe some of these flavor *anomalies* in the context of BSM physics in top-quarks in more detail in chapter 6.

Chapter 3

Physics Beyond the Standard Model

3.1 Do we need to go beyond the SM?

For many decades, the consensus of the high energy physics community has been that the SM cannot be the ultimate theory of nature. The reasons for this can be divided into two categories: *theoretical* and *observational/experimental*. As we explain right below, the theoretical arguments usually arise from a top-down approach to BSM, while the observational/experimental arguments are more in accordance with a bottom-up approach to BSM. Interestingly, one of the main driving forces for building a multi-TeV hadron collider like the LHC at CERN was the almost certain expectation from the theoretical community that new physics states were going to be quickly discovered. Unfortunately, this "no-loose theorem" for the discovery of BSM physics at the LHC is currently in a tight spot. More than 100 fb^{-1} of data has been gathered during run-I and II and still no clear sign of NP has appeared in high energy experiments. Highly advertised extensions of the SM solving well-known problems, like SUSY, extra-dimensions or technicolor have been (and still are) meticulously confronted with data. Results show that large portions of parameter space have been excluded by direct searches, forcing some of these models to be modified into much less appealing forms.

On the other side, we now know that the SM cannot describe all the known phenomena in our Universe. The discovery of neutrino oscillations and the observation of dark matter are two very solid reasons that demand BSM physics. Both cases indicate the presence of at least one new particle[1]. Another interesting hint of NP is the flavor puzzle. The emerging patterns in the CKM mixing matrix and the regularities in the hierarchy of fermionic masses seem to indicate some underlying non-trivial flavor structure. In addition, the recent hints of LFU violation in B-decays, if true, may provide a new window into the flavor puzzle. These experimental and observational hints of BSM without any preferred theoretical framework to explain them is a good motivation for studying BSM physics in an agnostic way using a bottom-up approach. All of this tells us that the answer to the question posed in the title of this section is Yes!

To keep things clear, we now outline our main assumptions. Since no new physics has been directly discovered so far, we take a worst-case-scenario stance and assume that the leading effect of NP is non-resonant within the energy reach of the LHC.

[1] In the case of Dark Matter this is not necessarily true. Nonetheless, the dark particle hypothesis is by far the best explanation for Dark Matter phenomenology.

Since the first two quark generations are more constrained by low energy precision experiments, we assume that the leading effects of the new particles appear in the top-quark sector. In the following sections, we introduce the SM effective field theory and define top-philic NP within this framework. Before, we finish this section by briefly outlining the electroweak hierarchy problem, the most famous top-down argument for BSM, neutrino masses and the flavor puzzle.

The electroweak hierarchy problem

One of the leading theoretical arguments for physics beyond the SM is what is known as the *electroweak hierarchy problem*. This can be summarized in the following sentences: any fundamental scalar field, such as the Higgs field is quadratically sensitive to any mass scale beyond the electroweak scale. Quantum corrections to the bare Higgs mass m_h at the one-loop level will scale quadratically with the cut-off scale Λ of the theory:

$$\delta m_\varphi^2 = \Lambda^2 \frac{3}{16\pi^2}\big(- y_t^2 + \lambda + \frac{3}{8}g^2 + \frac{1}{8}g'^2\big). \tag{3.1}$$

The leading term corresponds to the top-quark contribution (we did not include here the contributions of the other fermions since their Yukawa couplings are much smaller) while the other sub-leading terms to the W, Z and Higgs boson contributions. Adding a bare mass term $m^2 = \frac{3\epsilon}{16\pi^2}\Lambda^2$ to (3.1), where ϵ is a dimensionless parameter, gives us the physical mass of the Higgs

$$m_\varphi^2 = \Lambda^2 \frac{3}{16\pi^2}\big(\epsilon - y_t^2 + \lambda + \frac{3}{8}g^2 + \frac{1}{8}g'^2\big) \tag{3.2}$$

Taking this expression at face value suggests that the mass of the Higgs should be of the same order of the cut-off $m_h \sim \Lambda$. Formally, this instability of the Higgs mass with quantum corrections is related to *technical naturalness* [58]. This definition of naturalness states that any dimensionless coupling c in a field theory can be arbitrarily small if in the limit $c \to 0$ the global symmetry of the theory is enhanced to a larger symmetry group. In this case, the parameter c is said to be technically natural and the radiative corrections δc remain small, i.e. proportional to the amount of symmetry breaking. A good example of technically natural quantities are the fermion masses in the SM. The mass term $m_f\,\bar{\psi}\psi$ has a global U(1) symmetry, which in the limit $m \to 0$ is enhanced to $\mathrm{U}(1) \otimes \mathrm{U}(1)_A$ where the second factor is the axial symmetry $\psi \to e^{i\theta\gamma_5}\psi$. In the limit of exact axial symmetry, fermions remain massless under radiative corrections. For small $m \neq 0$, axial symmetry is explicitly broken, yet, this approximate symmetry protects the fermion masses from arbitrarily large radiative corrections, enforcing $\delta m_f^2 \propto m_f^2$. Hence, fermion masses are technically natural and will remain small under radiative corrections. In contrast, the masses of scalar fields like the SM Higgs are in general not technically natural. The mass term $m^2\varphi^\dagger\varphi$ of the Higgs field does not enhance the global symmetry of the SM in the limit $m \to 0$, meaning that quantum corrections of its mass will generically be sensitive to the mass scales related to the cut-off regulator Λ. The EW hierarchy problem can now be summarized as follows: if the cut-off of the SM is very large $\Lambda >> m_\varphi^2$, e.g. at the Planck scale, then this means that ϵ in (3.2) would have to be uncomfortably fine-tuned in order for the expression inside the

bracket to almost completely cancel the large cut-off and get a Higgs mass at the electroweak scale. The untuned scenario, with $\epsilon \sim \mathcal{O}(1)$, would predict a cut-off at $\Lambda \lesssim 1$ TeV.

Many *top-down* solutions to the Hierarchy problem have been proposed during the years. Most of these theories predict new phenomena at the electroweak scale or slightly above. The so-far null results at the LHC indicates that it may be necessary to rethink the hierarchy problem.

Neutrino masses

The first real crack in the SM was discovered two decades ago when neutrino flavor oscillations were first confirmed [59]. In a strict sense, the experimental verification of massive neutrinos falsifies the SM. In the SM since no RH neutrino state is introduced in the matter content, neutrinos are predicted to be exactly massless states (the moment the SM was introduced, neutrinos were consistent with being massless and LH states). This means that the SM Lagrangian is not complete since a missing term providing (tiny) neutrinos masses is necessary. In particular, the fact that neutrinos are massive implies the existence of at least one new particle in nature interacting with at least two of the LH lepton doublet fields.

The Flavor Puzzle

A lot is known about the flavor structure of the SM. On one side, all masses of the quarks and leptons have been carefully measured with the exception of the absolute neutrino mass scale, which anyways must arise from physics beyond the SM. Low energy flavor experiments have also measured very precisely the CKM mixing angles. These fundamental parameters have a very particular pattern which has led many theoreticians to attempt to find a deeper explanation. The CKM mixing matrix is a hierarchical perturbation of the unit matrix with a very special pattern shown in figure 3.1 (left). When scanning the space of 3×3 unitary matrices, the probability of randomly picking a matrix with CKM features is extremely unlikely. This suggests some sort of "selection rule" coming from a (slightly broken) symmetry in an underlying theory. By contrast, neutrinos seem to follow a completely different pattern compared to quarks and charged leptons. The mixing patterns parametrized by the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [60, 61], the analog of the CKM matrix in the neutrino sector, have much large mixing angles. The neutrino oscillation data also favors a much more compressed mass spectrum for the neutrinos[2], with the ratio of massive neutrinos of order $\mathcal{O}(1)$. The known structure of the PMNS matrix is also shown in figure 3.1 (right). This matrix has apparently no special structure within the space of 3×3 unitary matrices. This suggests an *anarchical* scenario where the structure of the PMNS matrix may have been generated by accident (i.e. sampled randomly from the space of 3×3 complex unitary matrices with a uniform distribution) in the underlying theory without any fundamental symmetry involved.

[2]The absolute scale of neutrino masses is still unknown, upper bounds have set it below the sub-eV scale.

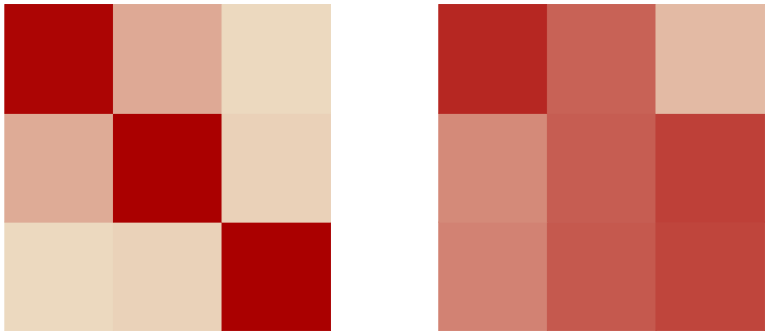


Figure 3.1: Visual representation of the mixing patterns in the CKM matrix (left) for quarks and the PMNS matrix (right) for neutrinos.

3.2 The SM as an effective theory

A very useful way of studying BSM physics is to proceed in a bottom-up approach and assume that the NP scale is much higher than the electroweak scale. Indeed, this scenario would point towards a mass gap separating the new heavy degrees of freedom from the heaviest particles in the SM, allowing us to focus on the low-energy dynamics below the cut-off without worrying in the details of the ultra-violet (UV) theory.

3.2.1 The effective SM Lagrangian

This particular setting allows us to dispose of any specific UV extension of the SM and parametrize new physics in a model-independent way by using an effective field theory (EFT). In general, EFTs describe the effects of unknown physics living at a heavy scale $\Lambda \gg v_{EW}$ through an *effective Lagrangian*, written as an expansion in powers of Λ^{-1} of local operators:

$$\mathcal{L}_{\text{eff}} \;=\; \mathcal{L}_{d\leq 4} \;+\; \sum_{d\,\geq\,5}^{\infty} \Big[\sum_{i=1}^{N_d} \frac{C_i^{(d)}(\mu)}{\Lambda^{d-4}}\, \mathcal{O}_i^{(d)}(\mu) \Big] \tag{3.3}$$

The first term in this effective Lagrangian is 4-dimensional or less and correspond to the renormalizable Lagrangian of the SM described in sec. 1. The low-energy effects of BSM interactions come from the second term in (3.3) and enter through a set of non-renormalizable d-dimensional operators $\mathcal{O}_i^{(d)}$ defined at a renormalization scale μ. The size N_d of the set of operators at a given order in d is finite. At a fixed dimension, each operator is formed by all possible Lorentz and gauge invariant products of the light degrees of freedom and derivatives. Schematically any d-dimensional operator can be written on very general grounds as

$$\mathcal{O}_i^{(d)} = \Lambda^d\, \hat{\mathcal{O}}^{(d)}\Big(\frac{D^\mu}{\Lambda}, \frac{H}{\Lambda}, \frac{\psi}{\Lambda^{3/2}}, \frac{\mathcal{F}^{\mu\nu}}{\Lambda^2} \Big) \tag{3.4}$$

where $\hat{\mathcal{O}}^{(d)}$ is a Lorentz and SM gauge invariant dimensionless monomial function, D_μ is the SM covariant derivative, H is the Higgs field, ψ are the SM fermions (we have dropped internal indices for the sake of clarity) and $\mathcal{F}$ is the SM field strength

tensor. The heavy degrees of freedom (the new heavy particles) have been integrated out and their information is encoded in an infinite number of dimensionless parameters, the *Wilson coefficients* $C_i^{(d)}$. These coefficients parametrize our ignorance about the underlying theory. Although the power expansion in d contains an infinite number of terms making the overall theory non-renormalizable, this is not an issue since at fixed order the low-energy theory is specified by a finite number $C_i^{(d)}$ $(i = 1, ..., N_d)$ of couplings allowing for an order-by-order renormalization via a finite number of counterterms.

When exploring the effects of NP in observables probed at experiments with a characteristic energy scale $E \ll \Lambda$, then it is enough to truncate (3.3) at fairly low orders in the expansion given that the effects of a d-dimensional operator will scale as $(E/\Lambda)^{d-4}$. In this case, the *irrelevant* operators with $d > 4$ will generally give small corrections to any observable when compared to the *marginal* $(d = 4)$ and *relevant* $(d < 4)$ operators. In practice, depending on the observable, it is enough to keep operators of order $d = 5$ and $d = 6$ in the expansion (3.4).

One important property of the SM effective Lagrangian 3.3 is that all the accidental symmetries of the renormalizable SM term $\mathcal{L}_{d\leq 4}$ are not necessarily preserved by the higher dimensional operators. Similarly, higher dimensional operators can also introduce a new source of symmetry violation for approximate symmetries in the SM. For example, lepton (L) and baryon (B) number can be explicitly broken by dimension $d = 5$ and $d = 6$ operators, respectively. Lepton flavor universality, which is explicitly violated by lepton masses in the SM, can receive new breaking contributions from $d = 6$ operators.

Operator mixing

In order to probe the Wilson coefficients at much lower energies than the cut-off, it is necessary to evolve $C^{(d)}(\mu = \Lambda)$ down to the electroweak scale $\lambda \sim m_t$ in accordance with the renormalization group (RG) equations

$$\frac{dC^i(\mu)}{d\log\mu} = \frac{1}{16\pi^2}\gamma^{ij}C^j, \tag{3.5}$$

where γ^{ij} is the $N_d \times N_d$ anomalous dimension matrix controlling the operator mixing. The solution of the RG equations are

$$C^i(m_W) = C^i(\Lambda) + \frac{1}{16\pi^2}\gamma^{ij}C^j \log\left(\frac{m_W}{\Lambda}\right) \tag{3.6}$$

At NLO in QCD, the renormalization of one operator typically involves other operators. This means that the anomalous dimension matrix γ^{ij} in (3.5) is usually non-diagonal, leading to the mixing of different operators at lower scales. Operator mixing could give rise to non-negligible effects when evolving down to lower mass scales. For example, an operator with vanishing Wilson coefficients at the cutoff $\Lambda = 1$ TeV could perfectly well pick up a non-zero Wilson coefficient of order $\mathcal{O}(1)$ at $\Lambda = m_t$ just by operator mixing effects. Fortunately, the anomalous dimensions for all dimension $d = 6$ operators have been computed at NLO in [62, 63, 64].

The $d = 5$ Weinberg operator

In the SMEFT, all possible operators of dimension $d = 5, 6$ have been completely classified. Remarkably, at dimension $d = 5$ there is only one operator [65]. For $d = 5$ there is only one possible effective operator preserving gauge and Lorentz invariance that can be built in the SM known as the Weinberg operator:

$$\mathcal{O}^{(5)} = (L_i^T H)\, i\sigma_2\, (H^\dagger L_j) + \text{h.c.} \tag{3.7}$$

This operator violates lepton number by two units $\Delta L = 2$. Once the Higgs acquires a VEV, this operator gives rise to a Majorana mass for neutrinos and the so-called see-saw mechanism $m_\nu = C^{(5)} v^2/\Lambda$. Interestingly, this mechanism for generating neutrino masses through lepton number violation has important phenomenological implications. On the downside, the natural scale of the cut-off is around $\Lambda \sim 10^{12}$ GeV for neutrinos with masses in the eV range, pretty much out of reach for any particle accelerator. Nonetheless, lepton number violation via (3.7) can in principle be accessed through precision measurements in nuclear reactions. The most promising $\Delta L = 2$ signature is the *neutrino-less double beta decay* $n + n \to e^- e^- pp$.

The $d = 6$ effective operators

For a complete list of the 59 dimension-6 effective operators, we refer the reader to Ref. [66]. There, the operators are expressed in the so-called *Warsaw basis*, which is the basis we will be using throughout this work. Here below we show a subset of such operators with at least two quark fields:

Four-quark operators:

$$\mathcal{O}_{qq}^{(1)\,[ijkl]} = (\overline{q}_i \gamma^\mu q_j)\,(\overline{q}_k \gamma_\mu q_l) \tag{3.8}$$
$$\mathcal{O}_{qq}^{(3)\,[ijkl]} = \left(\overline{q}_i \gamma^\mu \tau^I q_j\right)\left(\overline{q}_k \gamma_\mu \tau^I q_l\right) \tag{3.9}$$
$$\mathcal{O}_{qu}^{(1)\,[ijkl]} = (\overline{q}_i \gamma^\mu q_j)\,(\overline{u}_k \gamma_\mu u_l) \tag{3.10}$$
$$\mathcal{O}_{qu}^{(8)\,[ijkl]} = \left(\overline{q}_i \gamma^\mu T^A q_j\right)\left(\overline{u}_k \gamma_\mu T^A u_l\right) \tag{3.11}$$
$$\mathcal{O}_{qd}^{(1)\,[ijkl]} = (\overline{q}_i \gamma^\mu q_j)\left(\overline{d}_k \gamma_\mu d_l\right) \tag{3.12}$$
$$\mathcal{O}_{qd}^{(8)\,[ijkl]} = \left(\overline{q}_i \gamma^\mu T^A q_j\right)\left(\overline{d}_k \gamma_\mu T^A d_l\right) \tag{3.13}$$
$$\mathcal{O}_{uu}^{[ijkl]} = (\overline{u}_i \gamma^\mu u_j)\,(\overline{u}_k \gamma_\mu u_l) \tag{3.14}$$
$$\mathcal{O}_{ud}^{(1)\,[ijkl]} = (\overline{u}_i \gamma^\mu u_j)\left(\overline{d}_k \gamma_\mu d_l\right) \tag{3.15}$$
$$\mathcal{Q}_{quqd}^{(1)\,[ijkl]} = (\overline{q}_i u_j)\, i\tau^2\, (\overline{q}_k d_l) \tag{3.16}$$
$$\mathcal{Q}_{quqd}^{(8)\,[ijkl]} = \left(\overline{q}_i T^A u_j\right) i\tau^2 \left(\overline{q}_k T^A d_l\right) \tag{3.17}$$

Semi-leptonic operators:

$$\mathcal{O}_{lq}^{(1)\,[ijkl]} = \left(\bar{l}_i\gamma^\mu l_j\right)\left(\bar{q}_k\gamma^\mu q_l\right) \tag{3.18}$$
$$\mathcal{O}_{lq}^{(3)\,[ijkl]} = \left(\bar{l}_i\gamma^\mu \tau^I l_j\right)\left(\bar{q}_k\gamma^\mu \tau^I q_l\right) \tag{3.19}$$
$$\mathcal{O}_{lu}^{[ijkl]} = \left(\bar{l}_i\gamma^\mu l_j\right)\left(\bar{u}_k\gamma^\mu u_l\right) \tag{3.20}$$
$$\mathcal{O}_{eq}^{[ijkl]} = \left(\bar{e}_i\gamma^\mu e_j\right)\left(\bar{q}_k\gamma^\mu q_l\right) \tag{3.21}$$
$$\mathcal{O}_{eu}^{[ijkl]} = \left(\bar{e}_i\gamma^\mu e_j\right)\left(\bar{u}_k\gamma^\mu u_l\right) \tag{3.22}$$
$$\mathcal{Q}_{lequ}^{(1)\,[ijkl]} = \left(\bar{l}_i e_j\right)\varepsilon\left(\bar{q}_k u_l\right) \tag{3.23}$$
$$\mathcal{Q}_{lequ}^{(3)\,[ijkl]} = \left(\bar{l}_i\sigma^{\mu\nu} e_j\right)\varepsilon\left(\bar{q}_k\sigma_{\mu\nu} u_l\right) \tag{3.24}$$
$$\mathcal{Q}_{ledq}^{[ijkl]} = \left(\bar{l}_i e_j\right)\left(\bar{d}_k q_l\right) \tag{3.25}$$

Two-quark boson operators:

$$\mathcal{O}_{\varphi u}^{[ij]} = \left(\varphi^\dagger\varphi\right)\bar{q}_i u_j\tilde{\varphi} \tag{3.26}$$
$$\mathcal{O}_{\varphi q}^{(1)\,[ij]} = \left(\varphi^\dagger i\overleftrightarrow{D_\mu}\,\varphi\right)\left(\bar{q}_i\gamma^\mu q_j\right) \tag{3.27}$$
$$\mathcal{O}_{\varphi q}^{(3)\,[ij]} = \left(\varphi^\dagger i\overleftrightarrow{D_\mu^I}\,\varphi\right)\left(\bar{q}_i\gamma^\mu \tau^I q_j\right) \tag{3.28}$$
$$\mathcal{O}_{\varphi u}^{[ij]} = \left(\varphi^\dagger i\overleftrightarrow{D_\mu}\,\varphi\right)\left(\bar{u}_i\gamma^\mu u_j\right) \tag{3.29}$$
$$\mathcal{Q}_{\varphi ud}^{[ij]} = \left(\tilde{\varphi}^\dagger iD_\mu\varphi\right)\left(\bar{u}_i\gamma^\mu d_j\right) \tag{3.30}$$
$$\mathcal{Q}_{uW}^{[ij]} = \left(\bar{q}_i\sigma^{\mu\nu}\tau^I u_j\right)\tilde{\varphi}W_{\mu\nu}^I \tag{3.31}$$
$$\mathcal{Q}_{dW}^{[ij]} = \left(\bar{q}_i\sigma^{\mu\nu}\tau^I d_j\right)\varphi W_{\mu\nu}^I \tag{3.32}$$
$$\mathcal{Q}_{uB}^{[ij]} = \left(\bar{q}_i\sigma^{\mu\nu} u_j\right)\tilde{\varphi}B_{\mu\nu} \tag{3.33}$$
$$\mathcal{Q}_{uG}^{[ij]} = \left(\bar{q}_i\sigma^{\mu\nu}T^A u_j\right)\tilde{\varphi}G_{\mu\nu}^A \tag{3.34}$$

Baryon number violating operators:

$$\mathcal{Q}_{duq}^{[ijkl]} = \left(\overline{d^C}_{ia}u_{jb}\right)\left(\overline{q^C}_{kc}\varepsilon l_l\right)\epsilon^{abc} \tag{3.35}$$
$$\mathcal{Q}_{qqu}^{[ijkl]} = \left(\overline{q_{ia}^C}\varepsilon q_{jb}\right)\left(\overline{u_{kc}^C}e_l\right)\epsilon^{abc} \tag{3.36}$$
$$\mathcal{Q}_{qqq}^{(1)\,[ijkl]} = \left(\overline{q_{ia}^C}\varepsilon q_{jb}\right)\left(\overline{q^C}_{kc}\varepsilon l_l\right)\epsilon^{abc} \tag{3.37}$$
$$\mathcal{Q}_{qqq}^{(3)\,[ijkl]} = \left(\overline{q_{ia}^C}\tau^I\varepsilon q_{jb}\right)\left(\overline{q^C}_{kc}\tau^I\varepsilon l_l\right)\epsilon^{abc} \tag{3.38}$$
$$\mathcal{Q}_{duu}^{[ijkl]} = \left(\overline{d^C}_{ia}u_{jb}\right)\left(\overline{u^C}_{kc}e_l\right)\epsilon^{abc}\,. \tag{3.39}$$

When including flavor sub-indices Hermiticity of the $\mathcal{O}$ operators will imply $\mathcal{O}^{[ij]} = \mathcal{O}^{[ji]*}$ and $\mathcal{O}^{[ijkl]} = \mathcal{O}^{[jilk]*}$ as well as for the corresponding Wilson coefficients.

Loop vs tree induced operators

Not all Wilson coefficients in the operator product expansion (3.3) are expected to be of the same order. If the underlying theory is weakly coupled, some operators may

have a large suppression with respect to the others if, e.g. the former is generated at the one-loop level and the later appear from a tree-level exchange. It can be shown that some effective operators can never arise from integrating out one, or several, heavy mediator at tree-level in *any* possible underlying theory [67]. For instance the dipole operators in (3.31)–(3.34) of the form $\mathcal{Q}_{\psi V}^{[ij]} = (\bar{\psi}^i \sigma_{\mu\nu} \psi^j)\widetilde{\varphi} V^{\mu\nu}$, where V is any of the SM field strengths, is one example of such operators. Notice that since there are no fundamental antisymmetric 2-tensor mediators in Nature it is then not possible to contract the current $(\bar{\psi}^i \sigma_{\mu\nu} \psi^j)$ with the field strength at tree-level. These operators can therefore only be generated at higher orders in perturbation theory after integrating out heavy states running inside the loops. This suggests a natural partition of the full basis of $d = 6$ effective operators into *potential tree-generated* operators $\mathcal{O}^{\text{tree}}$ and *loop-generated* operators $\mathcal{O}^{\text{loop}}$, leading to the following expansion

$$\mathcal{L}_{\text{eff}} \;=\; \mathcal{L}_{d\leq 4} \;+\; \sum_i \; g_*^2 \; \frac{C_i^{\text{tree}}}{\Lambda^2} \mathcal{O}_i^{\text{tree}} \;+\; \sum_j \; \frac{g_*^2}{16\pi^2} \; \frac{C_j^{\text{loop}}}{\Lambda^2} \mathcal{O}_j^{\text{loop}} \;+\; \ldots \tag{3.40}$$

Here g_* represents a generic coupling, e.g. the coupling of each SM current to a heavy mediator in a simplified model. The coefficients C_i^{tree} are expected to be of $\mathcal{O}(1)$ while C_j^{loop} are proportional to other SM couplings (corresponding to vertices from external leg insertions of SM fields). Notice that in strongly coupled UV theories, the large g_* coupling could overcome the loop suppression factor for the loop-generated operators, potentially leading to Wilson coefficients of order $\mathcal{O}(1)$. For weakly coupled UV theories the coupling g_* is not large enough to overcome the loop suppression. In these theories, the Wilson coefficients of the loop-generated operators are always small.

3.3 Effective field theory in the top-quark sector

We now move to the central subject of this thesis: top-quarks. If the new physics degrees of freedom are very heavy $\Lambda \gg m_t$ and couple to top-quarks, we can then employ the SMEFT formalism described above to study the effects of the UV theory at the LHC or in low energy flavor experiments. In the SMEFT the leading effects come from $d = 6$ operators. Once the flavor indices are expanded in the three fermion families, the initial 59 $d = 6$ operators turn into a staggering $\sim \mathcal{O}(1000)$ flavorful operators (2499 to be precise [64]). Notice that most of these are two-fermion and four-fermion operators. In the following, we distinguish several flavor scenarios (at the cutoff scale) for the top-quark sector that are relevant for top-quark physics, especially for NP with large couplings to tops.

3.3.1 Baseline flavor symmetry $\mathrm{U}(2) \otimes \mathrm{U}(1)_{\text{top}}$

In order to keep our analysis tractable and single-out the top-quark sector from the plethora of flavorful $d = 6$ operators, it is necessary to introduce a flavor structure. One way is to assume an (approximate) global flavor symmetry of unspecified origin that distinguishes the third generation quarks from the light first and second generation quarks:

$$\mathcal{G}_{\mathcal{F}} = \mathcal{G}_{1+2} \otimes \mathcal{G}_3 \tag{3.41}$$

For the moment we do not impose any global symmetry on the lepton sector. For simplicity, we also restrict the Wilson coefficients of these operators to be real and leave out the possibility for CP-odd operators. In (3.41) it is understood that the third generation quark fields are singlets under $\mathcal{G}_{1+2}$ while the first and second generation quarks are singlets under $\mathcal{G}_3$. We adopt the following notation for quarks that realizes 3.41 and distinguishes the third generation (heavy) quarks from the first and second (light) generations:

$$\begin{aligned} q_1 &= (u,d)_L^T\,, \quad u_1 = u_R, \quad d_1 = d_R \qquad (\text{Light}) && (3.42)\\ q_2 &= (c,s)_L^T\,, \quad u_2 = c_R, \quad d_2 = s_R \qquad (\text{Light}) && (3.43)\\ Q &= (t,b)_L^T\,, \quad t = t_R, \quad b = b_R \qquad (\text{Heavy}). && (3.44) \end{aligned}$$

A very general choice aligning with what we are interested in, is to impose the following flavor groups on the quark fields

$$\begin{aligned} \mathcal{G}_{1+2} &= \mathrm{U}(2)_{\text{light}} \equiv \mathrm{U}(2)_{q_i+u_i+d_i}\,, && (3.45)\\ \mathcal{G}_3 &= \mathrm{U}(1)_{\text{top}} \equiv \mathrm{U}(1)_t\,. && (3.46) \end{aligned}$$

The group $\mathrm{U}(2)_{\text{light}}$ transforms all first and second generations identically. Imposing this symmetry indicates that NP is completely flavor blind when looking at the first two generations and does not distinguish between LH and RH fields, allowing for chirality mixing (scalar and tensors) operators of the form $(q_i u_j)$ and $(q_i d_j)$. The group $\mathrm{U}(2)_{\text{light}}$ in (3.45) corresponds to the less-restrictive baseline scenario adopted in Ref. [68] for the first two generations.

What makes our baseline flavor scenario particular is the symmetry imposed in the third generation. The group factor $\mathrm{U}(1)_{\text{top}}$ implies that NP distinguishes the RH top-quark from the rest of the remaining fields (Q and b). Notice that there is, in fact, an accidental (residual) group $\mathrm{U}(1)_{Q+b}$ acting on the remaining fields. The symmetry $\mathrm{U}(1)_{\text{top}}$ forbids bilinears with RH top-quark and other fields, i.e. (Qt) and (tb). The only possible bilinears are (tt), (QQ) (Qb) and (bb).

More-restrictive flavor scenario $\mathrm{U}(2)^3$ for light quarks

Another fairly general choice is to restrict the interactions between quarks of the first two generations with the following ansatz:

$$\mathcal{G}_{1+2} = \mathrm{U}(2)^3_{\text{light}} \equiv \mathrm{U}(2)_{q_i} \otimes \mathrm{U}(2)_{u_i} \otimes \mathrm{U}(2)_{d_i}\,. \tag{3.47}$$

An important consequence of $\mathrm{U}(2)^3_{\text{light}}$ is that chirality mixing scalar and tensor four-fermion operators in the first and second generation are strictly forbidden as well as charged current bilinears. Under this symmetry assumption the only allowed light quark bilinear operators are the flavor diagonal ones: $(q_i q_i) = (q_1 q_1 + q_2 q_2)$, $(u_i u_i) = (u_1 u_1 + u_2 u_2)$, $(d_i d_i) = (d_1 d_1 + d_2 d_2)$ for the first and second generation $i = 1, 2$. The group $\mathcal{G}_{1+2}$ in (3.45) corresponds to the baseline scenario adopted in Ref. [68] for the first two generations.

3.3.2 Top-philic flavor structure

Third family dominance

We are interested in NP coupling predominantly to the heaviest particles of the SM, namely, the third generation quarks. It is necessary to impose additional flavor assumptions besides the base-line symmetry $\mathcal{G}_\mathcal{F} = \mathrm{U}(2)_{\mathrm{light}} \otimes \mathrm{U}(1)_{\mathrm{top}}$ or the more restrictive symmetry $\mathcal{G}_\mathcal{F} = \mathrm{U}(2)^3_{\mathrm{light}} \otimes \mathrm{U}(1)_{\mathrm{top}}$. The symmetry $\mathcal{G}_\mathcal{F}$ distinguishes between the first two families and the third family where it singles out the RH top-quark. One can imitate what is observed in the Yukawa sector of the SM; the Higgs field couples predominantly via Yukawa interactions to the third generation fields. This assumption of *third family dominance* postulates that the heavy degrees of freedom in the UV theory only have couplings of order $\sim \mathcal{O}(1)$ with the bottom and top quark fields and negligible couplings with the first two generation quarks. A consequence is that any operator involving light quarks bilinear will have negligible Wilson coefficients.

Top-philia

We now define the main flavor scenario relevant for this work. Notice that third family dominance does not describe precisely the same situation in the SM Yukawa sector. In the SM, the bottom-quark Yukawa actually satisfies $y_b \ll y_t$ (the Yukawa terms with bilenar (Qb) is negligible with respect to (Qt)). To mimic this situation, we assume that NP couples predominantly to the RH top-quark t with order $\mathcal{O}(1)$ couplings and has a suppressed couplings with the third generation fields Q and b. This *top-philic* flavor scenario will be the focus of our attention in chapter 5 and 6.

We now give a more precise definition of the top-philic flavor structure in the context of the SMEFT framework. For the dimension-6 operators in the Warsaw we can impose top-philia by implementing the following prescription: for each quark field in the operators perform the replacements

(Top-philia definition SMEFT)

$$\begin{aligned} u_i &\to y_i^u\, u_i \\ d_i &\to y_i^d\, d_i\,, \\ q_i &\to y_i^d\, q_i\,. \end{aligned} \tag{3.48}$$

Here we have restored the usual notation with $i = 1, 2, 3$ and $y_i^q = \sqrt{2} m_i^q / v$ are the quark Yukawa couplings in the SM. Notice that our definition of top-philia differs with the one adopted in Ref. [68]. In that paper, they have chosen an alternative definition with the same prescription as in (3.48) but with $q_i \to y_i^u\, q_i$ for the left-handed fields. This implies that only operators with Q and t are relevant. This definition of top-philia also allows for operators with large LH bottom quarks, while our definition in (3.48) is genuinely top-philic.

When applying the top-philic flavor prescription (3.48) to the full set of 2499 dimension-6 operators, this singles out from (3.8)–(3.34) the following operators:

(Top-philic operators)

$$\mathcal{O}_{tt} \equiv \mathcal{O}_{uu}^{[3333]} = y_t^4 \,(\bar{t}\gamma^\mu t)(\bar{t}\gamma^\mu t) \tag{3.49}$$
$$\mathcal{O}_{lt}^{ij} \equiv \mathcal{O}_{lu}^{[ij33]} = y_t^2 \,\left(\bar{l}_i\gamma^\mu l_j\right)\left(\bar{t}\gamma^\mu t\right) \tag{3.50}$$
$$\mathcal{O}_{et}^{ij} \equiv \mathcal{O}_{eu}^{[ij33]} = y_t^2 \,\left(\bar{e}_i\gamma^\mu e_j\right)\left(\bar{t}\gamma^\mu t\right) \tag{3.51}$$
$$\mathcal{O}_{\varphi t} \equiv \mathcal{O}_{\varphi u}^{[33]} = y_t^2 \,(\varphi^\dagger i \overleftrightarrow{D_\mu}\, \varphi)\left(\bar{t}\gamma^\mu t\right) \tag{3.52}$$

All other operators are heavily suppressed by the small Yukawas and can be safely neglected in the limit $y_b \ll y_t$. An important remark is that a realistic UV theory may also generate other non-quark operators, e.g. in particular four-lepton operators of the form

$$\mathcal{O}_{ll}^{[ijkl]} = \left(\bar{l}_i\gamma^\mu l_j\right)\left(\bar{l}_k\gamma^\mu l_l\right) \tag{3.53}$$
$$\mathcal{O}_{uu}^{[ijkl]} = \left(\bar{e}_i\gamma^\mu e_j\right)\left(\bar{e}_k\gamma^\mu e_l\right) \tag{3.54}$$

Of course, the operators that can appear besides the ones in (3.49)–(3.52) will depend on the specific details of the UV scenario and will generally be constrained by processes not involving quarks. Additional flavor structures for leptons can be imposed in order to reduce the number of operators if necessary. Another important remark is that RGE mixing will unavoidably give rise to other operators besides (3.49)–(3.52) that were not generated at tree level at the matching scale. Nonetheless, the running of these coefficients from the few-TeV matching scale Λ_{NP} down to the LHC characteristic energies are expected to be negligible. Of course, this is not true when running to much below the EW scale where small mixing effects can produce very large effects compared to the level of precision measurements in flavor experiments. Interestingly, in sec. 5.5 and sec. 6.7 we will outline a class of UV complete models that can naturally give rise to the top-philic flavor assumption described here.

Chapter 4

Idealized Observables for tth production

4.1 Introduction

Characterizing in general terms the interactions between the top-quark and the Higgs boson – the heaviest elementary particles observed so far – is an important step in searching for clues of physics beyond the SM. At present, the direct experimental constraints on the top-quark Yukawa coupling are still affected by a large experimental uncertainty. The dominant source of information being the experimental searches for $\sigma(pp \to t\bar{t}h)$ from ATLAS [12, 69, 70, 71, 72] and CMS [13, 73, 74, 75]. In view of higher-statistics studies of this process, it is important to address the question of how to characterize this production cross-section in general extensions of the SM where the NP gives rise to non-resonant effects. For this, one is forced to introduce a new class of observables that is suitable for studying multi-resonance production like $pp \to t\bar{t}h$ at the LHC.

An interesting approach to characterize the Higgs boson interactions with minimum theoretical bias, in the vast class of models with heavy new particles, is that of the pseudo-observable formalism. These were first proposed in the context of electroweak observables at the Z pole [76] aimed for BSM explorations at the LEP experiments. This concept was later generalized to the recently discovered Higgs boson, [77, 78, 79]. The basic idea of Higgs pseudo-observables is to identify a set of quantities that are (i) theoretically motived from the QFT point of view, (ii) sensitive to different directions in the landscape of BSM physics in a way that is model independent, and (iii) accessible experimentally. More precisely, the Higgs pseudo-observable formalism developed in Ref. [77, 78], to describe electroweak production and decays modes of the Higgs boson, is based on a general decomposition of on-shell amplitudes based on analyticity, unitarity, crossing symmetry, and a momentum expansion based on the known singularity structure of the amplitudes in the low-energy limit. The Higgs pseudo-observables represent an optimized set of observables which can be computed in a wide class of explicit NP models, including Effective Field Theory approaches to physics beyond the SM.

Unfortunately, the task of classifying, extracting the pseudo-observables and matching these to specific NP models for $pp \to t\bar{t}h$ production, or any multi-resonance process in general, is not trivial. One of the main challenges is the large number of pseudo-observables (we provide some examples in the next section) that

typically appear in $2 \to N$ body processes for $N > 2$ and not enough experimental handles available to disentangle them.

The practical disadvantage of performing a general amplitude decomposition of $pp \to t\bar{t}h$ and the experimental challenges inherent to this process does not prevent us from identifying a set of idealized and more inclusive observables with similar advantages. The main objective of this chapter is to present a new set of observables that follow the general pseudo-observable criteria, namely the identification of a series of observables that are well defined both from the theoretical and the experimental point of view, and that capture all relevant (and practically accessible) non-resonant NP effects entering $t\bar{t}h$ production. We find that the different kinematical regimes of the top and Higgs resonances experimentally accessible using boosted substructure techniques naturally leads us to identify a series of idealized $pp \to t\bar{t}h$ cross-sections that fulfill this goal.

As explained later on, rather than being defined at the amplitude level, these idealized observables, are defined at the cross-sections level, following more closely the philosophy of the so-called *simplified template cross-sections* [79]. This fact implies some limitations. In particular, it will not be possible to unambiguously reconstruct the underlying amplitude using only the idealized observables. However, we argue that these observables represent the most useful set of independent information on the underlying amplitude that can be realistically extracted from $pp \to t\bar{t}h$, or from any multi-resonance process at the LHC.

This chapter is organized as follows. In sec. 4.2 we demonstrate the intractability of performing an amplitude decomposition for defining pseudo-observables for $pp \to t\bar{t}h$. In sec. 4.3 we define the idealized observables for a generic multi-resonance production at the LHC and discuss this for $pp \to t\bar{t}h$. In sec. 4.4 we calculate the idealized cross-section for $pp \to t\bar{t}h$ production for non-resonant BSM physics in the SMEFT framework and demonstrate that these observables disentangle different NP contributions. In the final section sec. 4.5 we provide a search strategy for measuring the idealized cross-sections at the HL-LHC.

4.2 Pseudo-observables and tth production

In principle, a set of pseudo-observables based on a general amplitude decomposition can be attempted for $pp \to t\bar{t}h$ production. However, in this case, most of the advantages of the formalism developed in Ref. [77, 78] get lost. The main problem being a large number of independent tensor structures following from a general decomposition of the amplitude and, at the same time, the impossibility to experimentally access all the kinematical variables that would characterize the amplitude in general terms. In this section, we demonstrate this with a couple of examples.

4.2.1 Amplitude decomposition

We now illustrate this by considering the amplitude of $t\bar{t}h$ production at parton level. We focus on the gluon initiated process $gg \to t\bar{t}h$ since this is the dominant production mode.

$$\mathcal{A}^{ab}(\ g(p_1)g(p_2)\to t(p_3)\bar{t}(p_4)h(k)\ ) = \epsilon_\mu(p_1)\epsilon_\nu(p_2)\,T^a_{ij}T^b_{kl}\ \left[\bar{u}^r_i(p_3)\,\Gamma^{\mu\nu}_{jk,rs}(q_1,q_2)\,v^s_j(p_4)\right] \tag{4.1}$$

where $p_{1,2}$ are the four-momentum of the colliding gluons with polarization vectors $\epsilon(p_{1,2})$, $p_{3,4}$ the outgoing momenta of the produced top-quarks with spinors u^r and v^s, k the four-momentum of the Higgs boson and T are SU(3) QCD generators defined via the Gell-Man matrices $T^a \equiv \lambda^a/2$ with normalization $\mathrm{Tr}(T^aT^b) = 1/2$. In (4.1), the quantity of interest characterizing the nature of the interaction is the momentum-dependent tensor $\Gamma^{\mu\nu}(q_1,q_2)$, which is a function of the transferred momenum $q_1 \equiv p_3 - p_1$ and $q_2 \equiv p_4 - p_2$. Being from different vector spaces, we factorize color and Lorentz structures of the tensor as $\Gamma^{\mu\nu}_{jk,rs} = \delta_{jk}\otimes\Gamma^{\mu\nu}_{rs}$. Next we expand Γ in the Clifford algebra basis $\{\mathbf{1},\gamma^\mu, i\sigma^{\mu\nu},\gamma^\mu\gamma^5,\gamma^5\}$ in full generality

$$\begin{aligned}\Gamma^{\mu\nu}_{rs}(q_1,q_2) &= S^{\mu\nu}\mathbb{1} + V^\mu_{(1)}\,\gamma^\nu + V^\nu_{(2)}\,\gamma^\mu + i\,T_{(0)}\,\sigma^{\mu\nu} + i\,T^\mu_{(1)\,\alpha}\,\sigma^{\nu\alpha} \\ &+ i\,T^\nu_{(2)\,\alpha}\,\sigma^{\mu\alpha} + i\,T^{\mu\nu}_{(3)\,\alpha\beta}\,\sigma^{\alpha\beta} + A^\mu_{(1)}\,\gamma^\nu\gamma^5 + A^\nu_{(2)}\,\gamma^\mu\gamma^5 + P^{\mu\nu}\gamma^5\end{aligned} \tag{4.2}$$

Here we can already see that there is a proliferation of Lorentz coefficients S, $V_{(i)}$, $T_{(i)}$,$A_{(i)}$, P, that complicate matters.

To keep things simple, we can only consider the main Feynman diagram contributing to the production mode, namely, the exchange $gg \to t\,(\bar{t}t)^*\,\bar{t} \to t\bar{t}h$ with t-channel topology[1]. The amplitude $\mathcal{M}$ at tree-level for this diagram simplifies to:

$$\mathcal{A}^{ab}(\ g(p_1)g(p_2)\to t(p_3)\bar{t}(p_4)h(k)\ ) = [\bar{u}^t_i(p_3)\,(\gamma^\mu)_{ts}\,T^a\epsilon_\mu(p_1)]\,[\epsilon_\mu(p_2)T^b\,(\gamma^\mu)_{rw}\,v^w_j(p_4)]\,\mathcal{V}^{ij}_{sr}(q_1,q_2) \tag{4.3}$$

(we omit the other u-channel diagram with $p_{1,2}\to p_{2,1}$). Here $\mathcal{V}^{ij}_{sr}(q_1,q_2)$ represents the off-shell three-point function tth depending on the momenta defined by $q_1 \equiv p_3 - p_1$ and $q_2 \equiv p_4 - p_2$. As before, we can factorize color from the rest as $\mathcal{V}^{ij} = \delta^{ij}\otimes\mathcal{V}$ and expand in the Clifford algebra basis $\{\mathbf{1},\gamma^\mu, i\sigma^{\mu\nu},\gamma^\mu\gamma^5,\gamma^5\}$:

$$\mathcal{V}(q_1,q_2) \;=\; S(q_1,q_2)\mathbb{1} + V^\mu(q_1,q_2)\,\gamma_\mu + T^{\mu\nu}(q_1,q_2)\,i\sigma_{\mu\nu} + A^\mu(q_1,q_2)\,\gamma_\mu\gamma^5 + P(q_1,q_2)\,\gamma^5 \tag{4.4}$$

Here S,V,T,A,P are generic scalar, vector, (anti-symmetric)tensor, axial-vector and pseudo-scalar functions. Each of these functions can be decomposed in full generality into Lorentz invariant structures:

$$S(q_1,q_2) \;=\; F^{(1)}_S(q_1^2,q_2^2) \;+\; F^{(2)}_S(q_1^2,q_2^2)\,\frac{q_1\cdot q_2}{m_t^2} \tag{4.5}$$

$$V^\mu(q_1,q_2) \;=\; F^{(1)}_V(q_1^2,q_2^2)\,\frac{q_1^\mu}{m_t} \;+\; F^{(2)}_V(q_1^2,q_2^2)\,\frac{q_2^\mu}{m_t} \tag{4.6}$$

$$T^{\mu\nu}(q_1,q_2) \;=\; F^{(1)}_T(q_1^2,q_2^2)\,\frac{q_1^{[\mu}q_2^{\nu]}}{m_t^2} \;+\; F^{(2)}_T(q_1^2,q_2^2)\,\frac{\epsilon^{\mu\nu\sigma\rho}\,q_{1\sigma}q_{2\rho}}{m_t^2} \tag{4.7}$$

$$A^\mu(q_1,q_2) \;=\; F^{(1)}_A(q_1^2,q_2^2)\,\frac{q_1^\mu}{m_t} \;+\; F^{(2)}_A(q_1^2,q_2^2)\,\frac{q_2^\mu}{m_t} \tag{4.8}$$

$$P(q_1,q_2) \;=\; F^{(1)}_P(q_1^2,q_2^2) \;+\; F^{(2)}_P(q_1^2,q_2^2)\,\frac{q_1\cdot q_2}{m_t^2} \tag{4.9}$$

[1] Other diagrams with exchanged $s-$*channel* off-shell top-quark are more suppressed because of the high virtuality.

The invariants $F^{(1,2)}_{S,V,T,A,P}$ are the *form factors* of the off-shell three-point function t^*t^*h. Even for this simplified scenario where the amplitude has the form (4.4), one finds a multitude of form factors that make any analysis using pseudo-observable highly non-trivial. Of course, some of the form factors are related to each other via Ward-Takahashi identities.

Notice that the SM is a very simplified case compared to the general expressions above. For instance, the off-shell t^*t^*h vertex in the SM simply reads

$$\mathcal{V}^{ij}_{rs}(q_1,q_2) \;= \frac{iy_t}{\sqrt{2}}\;\delta^{ij}\;S^F_{rt}(q_1)\;S^F_{ts}(q_2)\,, \tag{4.10}$$

where $S^F(p) \equiv (\not{p} - m_t)^{-1}$ is the fermion propagator. The form factors in the SM satisfy:

$$F_S^{(1,2)} = F_V^{(1,2)} = F_T^{(1)} \;=\; \frac{y_t\, m_t^2/\sqrt{2}}{(q_1^2-m_t^2)(q_2^2-m_t^2)} \tag{4.11}$$

$$F_T^{(2)} \;= F_A^{(1,2)} = F_P^{(1,2)} \;=\; 0. \tag{4.12}$$

with a double pole structure. Here, the second set of form factors (4.12) vanish because in the SM the Higgs is a CP-even boson.

4.3 Idealized observables for multi-resonance production

At present, the experimental searches for $pp \to t\bar{t}h$ are in the multi-lepton [69, 73, 75], $b\bar{b}$ [70, 74], or $\gamma\gamma$ [71] Higgs decay channel, and use advanced multivariate analyses in order to suppress the overwhelming backgrounds. In all these cases, the kinematics of the signal events is dominantly coming from threshold production of the $t\bar{t}h$ states where the decay products of the three resonances are expected to be individually resolved inside the detector.

Defining a set of idealized observables that can be matched to realistic observables, such as the cross-section in different decay channels, is not straightforward. We propose an alternative approach that bypasses the difficulty of using amplitude-decomposition pseudo-observables for multi-resonance processes like $pp \to t\bar{t}h$. In the following, we give a definition of these new observables for a generic multi-resonance process.

4.3.1 Idealized observables

The aim is to characterize at hadron colliders the NP entering in a $2 \to N$ body multi-resonance process with a generic form

$$pp \to\; \mathcal{X}^1\mathcal{X}^2 ... \mathcal{X}^N \;\to\; x^1 x^2 ... x^n\,, \tag{4.13}$$

where $\mathcal{X}^j$ represents any of the known SM heavy resonances $\mathcal{X} = \{W^\pm, Z, t, h\}$ and x^j represent any of the possible "massless" final state decay products $x = \{\gamma, \ell^\pm, \nu_\ell, u, d, s, c, b\}$. In what follows we assume that no other light resonance besides $\mathcal{X}$ are accessible at current LHC energies. The method we describe here is based on defining a set of observables o_i that we call *idealized observables*. Formally,

these observables $\{o_1, o_2, ..., o_k\}$ are constructed from a *base observable* o and a set of *categories* $\mathcal{K}$. These must satisfy the following defining criteria:

- (i) First, the base observable o has to be theoretically well-defined, hence, a function $o(\mathcal{A})$ of the amplitude $\mathcal{A}$ of the production process $pp \to \mathcal{X}^1\mathcal{X}^2...\mathcal{X}^N$. This amplitude can be calculated at any given order in perturbation theory using quantum field theory.

- (ii) $\mathcal{K} = \{\kappa_1, ..., \kappa_k\}$ is a (finite) set of categorical variables κ_i that label the partonic events based on a pre-defined criteria. The set must be *mutually exclusive* and defined independently of the possible types of decay modes of each resonance $\mathcal{X}^1\mathcal{X}^2...\mathcal{X}^N \to x^1x^2...x^n$. The idealized obervables $o_i(\mathcal{A})$ are defined as the restriction of o to the category $\kappa_i \in \mathcal{K}$, i.e. $o_i \equiv o|_{\kappa_i}$. The mutual exclusivity condition for $\mathcal{K}$ is then equivalent to the possibility of decomposing o into the sum

$$o \;=\; \sum_{i=1}^{k} o_i \,, \tag{4.14}$$

  over the whole set $\mathcal{K}$.

- (iii) The categories in $\mathcal{K}$ must be chosen in such a way that the resulting observables o_i have each different sensitivities to different directions in the parameter space of the theory used to compute $\mathcal{A}$. In particular, observables that can efficiently disentangle different BSM effects.

- (iv) These idealized observables must each be extracted from experimental analyses which allow for a model-independent unfolding procedure.

To make sense of these definitions let's visualize the parton-level[2] data from the multi-resonance production process (4.13) as it would come from a Monte Carlo (MC) event generator. Parton level events are generated by randomly sampling from the truth level probability distributions previously calculated from the matrix element computation (as a function of $\mathcal{A}$) and integrated over the phase space volume V of the $2 \to N$ process. The volume V lives in a multi-dimensional space with coordinates given by the usual four-momenta p^i_μ for each of the heavy resonance $\mathcal{X}^i$. Each event is defined by its kinematics and corresponds to a point in this space. The generated $pp \to \mathcal{X}^1\mathcal{X}^2...\mathcal{X}^N$ events populate the available phase space volume V. For hadron colliders one typically uses the transverse momenta p^i_T, the rapidity η^i and azimuthal angle ϕ^i. Other sets of kinematical variables can be used as well.

For the base observable o, one can choose any relevant quantity like for example the total production cross-section $o = \sigma(pp \to \mathcal{X}^1\mathcal{X}^2...\mathcal{X}^N)$. In this picture, one can now define a categorization of events by choosing a specific partition of the kinematical phase space into k regions, i.e. $V = V_1 \cup V_2 \cup ... \cup V_k$. Assuming that each region V_i is a smooth manifold, one can define their boundaries with a set of algebraic equations, or in the simplest scenario with a set of kinematical cuts. We

[2]Here by parton-level, we mean final states before QCD/QED showering and hadronization. We do include the PDF convolution for the initial parton in the proton.

can now classify the events by using the simple criteria: if a $pp \to \mathcal{X}^1\mathcal{X}^2...\mathcal{X}^N$ event falls inside the region V_i of phase space, we say that it belongs to the category κ_i. In this language, looking for different idealized observables translates into looking for different ways of splitting the phase space volume into a specific partition $V = V_1 \cup V_2 \cup ... \cup V_k$ such that conditions (iii) and (iv) are satisfied.

Much of what we have described above is not out of the ordinary in collider physics. In experiments, the kinematical space of the experimental data sets is usually partitioned into mutually exclusive signal, background, validation, and control regions. These phase space "bins" are typically defined via kinematic cuts that have been optimized after characterizing the signal. The criteria (i)–(iv) defining the idealized observables have many similarities with the simplified template cross-section framework [79] used by the Higgs Cross-section Working group at CERN. The main difference here is that we are applying an event categorization directly at the partonic level for one production process independent of the decays channels, while template cross-sections are defined at the experimental level and rely on categorizations specific to the several (Higgs) production channels.

As we show in the next section, it is possible to define idealized observables by splitting the phase space into regions using kinematical arguments that are pretty much model independent, but most importantly, directly motivated by a real experimental setup. This reversal of logic has the advantage that it can give rise to observables that match much better real experimental measurements, making the unfolding procedure less complicated. On the downside, predicting some of the idealized observables may be less direct and would need to be performed using MC simulation at parton level.

4.3.2 Kinematical categories for multi-resonances

The next step is to come up with a useful partition of phase space that defines a categorization of events to define the idealized observables o_i. What makes defining idealized observables non-trivial are both the need to satisfy criteria (iii) and (iv) described above. Events from $pp \to \mathcal{X}^1\mathcal{X}^2...\mathcal{X}^N$ in the SM populate the multi-dimensional phase space unevenly, with some regions more densely populated than others. This is, in particular, true for resonance production where the majority of events fall inside relatively small regions of phase space in the vicinity of the on-shell production threshold. Kinematically, these are the regions where the heavy particles are produced at rest in the laboratory frame or with very little transferred momentum. The regions of phase space far from these are on the other hand sparsely populated in comparison with the threshold regions.

BSM effects from very heavy degrees of freedom that are not kinematically accessible in direct searches will produce non-resonant deviations from the SM predictions. Naively, one expects these NP effects deforming the resonance production of SM particles to be more apparent outside the threshold dominated regions of phase space. This is a motivation to consider a purely kinematical categorization where one of the categories captures events consisting mostly of resonances produced at rest while the remaining categories capture events from resonances that are kinematically boosted. Intuitively, in the boosted categories non-resonant NP physics described by contact interactions would lead to an enhancement in the event rates that grow linearly in the partonic center-of-mass energy s.

We now propose a set of categories $\mathcal{K}$ for SM multi-resonance processes based on a binomial classification, $\kappa = \{B, R\}$, for each individual resonance $\mathcal{X}^j$. We assume that $\mathcal{X}^j$ can be found in either of the *two* kinematical states: *boosted* ($\kappa = B$) or *rest* ($\kappa = R$). To set notation, we label each resonance with the categorical variables κ_j as $\mathcal{X}^j_{\kappa_j}$ with $\kappa_i = \{B, R\}$ and denote the categories in $\mathcal{K}$ by concatenation $\mathcal{X}^1_{\kappa_1}\mathcal{X}^2_{\kappa_2}...\mathcal{X}^N_{\kappa_N}$. Notice that if all the resonances in $pp \to \mathcal{X}^1\mathcal{X}^2...\mathcal{X}^N$ are different (i.e. with different quantum numbers and masses) there will be $k = 2^N$ different categories in $\mathcal{K}$. The set of kinematical categories is then written as

$$\mathcal{K} \;=\; \left\{\mathcal{X}^1_R\mathcal{X}^2_R\mathcal{X}^3_R...\mathcal{X}^N_R\,,\;\mathcal{X}^1_B\mathcal{X}^2_R\mathcal{X}^3_R...\mathcal{X}^N_R\,,\;\mathcal{X}^1_R\mathcal{X}^2_B\mathcal{X}^3_R...\mathcal{X}^N_R\,,\;...\,,\;\mathcal{X}^1_B\mathcal{X}^2_B\mathcal{X}^3_B...\mathcal{X}^N_B\right\}\,. \tag{4.15}$$

$\mathcal{K}$ has only one *rest category*, $\mathcal{X}^1_R\mathcal{X}^2_R...\mathcal{X}^N_R$, where all the produced resonances are approximately at rest. We expect most of the events to fall into this category because of threshold production dominance. The remaining $2^N - 1$ *boosted categories* are expected to be less populated in the SM and more sensitive to NP.

Now that we have proposed a sensible way of partitioning phase space into kinematical categories, what remains is to provide a precise method on how to achieve this in practice. For this, we need to define precisely when a resonance $\mathcal{X}^j$ is classified as *boosted* and when it is classified as *rest*, in a way compatible with experiments.

4.3.3 Defining Boost/Rest categories

Fortunately, all SM resonances $\mathcal{X} = \{W^\pm, Z, t, h\}$ have decay modes $\mathcal{X} \to x^1x^2...x^n$ into leptons and quarks $x^j \in \{\ell, \nu, u, d, s, c, b\}$ with somewhat comparable rates. These have been extensively studied at the LHC (including recently the Higgs $h \to b\bar{b}$ and $h \to \ell\ell\ell\ell$ modes) and at previous colliders. This provides a rich arena to define experimentally-inspired kinematical categories for the idealized observables. Of particular interest to us are the hadronic decay modes into jets. The hadronic decays of boosted bosons $V = \{W^\pm, Z\}$, top-quarks, and Higgs into jets have been the subject of a lot of theoretical activity in the last decade. At large LHC energies, these heavy resonances will decay into several highly collimated partons producing one unresolved "fat" jet inside the detector. Fat jets that arise in this way will have distinctive QCD radiation patterns. Such patterns produce identifiable features inside the clustering history of the fat jet as well as non-trivial correlations between different energy deposits in the detector. Several techniques exploiting the substructure of fat jets have been proposed in order to infer the nature of the heavy resonance that produced the fat jet and discriminate it from the QCD multi-jet background. Many studies have shown that fat-jet tagging techniques can be used to identify hadronically decaying boosted Higgs bosons [80], boosted W and Z bosons [80, 81], and also boosted top-quarks [81, 82, 83, 84], with high efficiency and large QCD rejection rates. For example, the use of such boosted techniques for $pp \to t\bar{t}h$ has seen a fruitful development in recent years, and the feasibility of such analyses for the LHC has been proven, see e.g. [83, 85, 86].
The possibility of tagging V-jets, top-jets, and Higgs-jets experimentally can be exploited to define the kinematical categorization. For the top and Higgs tagging, the fat-jet clustering cones are usually fixed to satisfy $\Delta R \gtrsim 1$. These large cone radii are broad enough to capture the overlapping jets arising from each of the

hadronic daughter decays. Notice that the optimal cone radius ΔR for fat-jets can be fixed by estimating the average opening distance using the rule-of-thumb

$$\Delta R \;\approx\; \frac{2\, m_{\mathcal{X}}}{p_T^{\mathcal{X}}} \tag{4.16}$$

where $m_{\mathcal{X}}$ is the mass of the heavy hadronic SM resonance $\mathcal{X}$ and $p_T^{\mathcal{X}}$ its momentum. This can be used to define precisely at parton level the R and B categories for each resonance based on whether all the massless decay products of the partonic resonance fall inside a cone of fixed radius ΔR, or not, respectively. We classify parton level events using the following kinematical categorization

- *Cone-based classifier: A resonance $\mathcal{X}$ is classified as boosted ($\mathcal{X}_B$) if for all pair of daughters x^i and x^j in the decay $\mathcal{X} \to x^1 x^2 ... x^n$ the following is satisfied:*

$$max(R_{ij}) \;<\; \Delta R_{\mathcal{X}} \tag{4.17}$$

  else, it is classified at rest ($\mathcal{X}_R$). Here $R_{ij} \equiv R(x^i, x^j) = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}$ is the planar distance between the pair of daughter particles and $\Delta R_{\mathcal{X}}$ is a resonance-specific $\mathcal{O}(1)$ parameter.

The cone parameter $\Delta R_{\mathcal{X}}$ is a free parameter that can be optimized for each heavy resonance. This categorization is very much inspired by fat-jet tagging. The cone parameter $\Delta R_{\mathcal{X}}$ in (4.17) can be thought as a parton-level proxy for the fat-jet clustering parameter ΔR in (4.16) used in the real experiments. We then propose using $\Delta R_{\mathcal{X}} = \Delta R$ as a criterion for fixing the partonic cone parameter. For instance, for the top-quark, a good jet clustering choice for the cone is $\Delta R_t = 1.5$ while for Higgs decays a better choice is $\Delta R_h = 1.2$. Notice that the cone-based classifier is purely geometrical in origin. It is completely independent of the nature or the multiplicity of the decay daughters used in the procedure. It can indeed be computed for leptonic or hadronic $1 \to n$ decay processes.

The cone-based criteria for defining kinematical categories is not the only possibility. One, maybe simpler possibility is to classify resonances based on the transverse momentum that they carry without resorting to the decay product kinematics. We define:

- *p_T-based classifier: A resonance $\mathcal{X}$ is classified as boosted ($\mathcal{X}_B$) if $p_T^{\mathcal{X}} > q_T^{\mathcal{X}}$, else it is classified as rest ($\mathcal{X}_R$). Here $p_T^{\mathcal{X}}$ is the partonic transverse momentum of the resonance and $q_T^{\mathcal{X}}$ is a resonance-specific mass parameter of order $\mathcal{O}(10^2)$ GeV.*

Note that the approximate relation (4.16) relates the cone-based to the p_T-based classifiers. For the top-quark and the Higgs, a typical threshold value for the categorization parameter is $q_T^t = 250$ GeV and $q_T^h = 200$ GeV, respectively. Below we will use both categorizations to analyze $pp \to t\bar{t}h$ production both in the SM and in the SMEFT. While the cone-based observables are closer to the experimental analyses, in particular, to those using boosted techniques, defining the categories using the p_T of the mother particles makes them much easier to compute.

$\mu_{XYZ} = \sigma(t_X t_Y h_Z)/\sigma_{\text{tot}}^{\text{SM}}$	cone-categorization	p_T-categorization
$t_B t_B h_B$	7.7×10^{-3}	3.2×10^{-2}
$t_B t_R h_B$	3.4×10^{-2}	8.5×10^{-2}
$t_R t_R h_B$	5.3×10^{-2}	4.5×10^{-2}
$t_B t_B h_R$	2.4×10^{-2}	7.9×10^{-2}
$t_B t_R h_R$	0.111	0.11
$t_R t_R h_R$	0.765	0.65

Table 4.1: Ratio of the cross section in each category over the total $pp \to t\bar{t}h$ for the LHC at 13 TeV, in the SM. In the upper table we use the categories in ΔR while in the lower table we use the p_T (250 GeV for the top, 200 GeV for the Higgs).

4.3.4 Idealized cross-sections in $pp \to tth$ production

Now that we have defined precisely the kinematical categories $\mathcal{K}$ (one cone-based and the other p_T-based) for multi-resonances, we will use this to analyze $pp \to t\bar{t}h$ production at the LHC. For this multi-resonance process, we do not distinguish top-quarks and anti-top-quarks, meaning that there are six inequivalent kinematical categories $t_X t_Y h_Z$, with $X, Y, Z \in \{B, R\}$: one rest category $t_R t_R h_R$ and five boosted categories that partition the totality of phase space. These are

$$\mathcal{K}_{t\bar{t}h} \;=\; \{t_R t_R h_R,\; t_R t_R h_B,\; t_R t_B h_R,\; t_B t_R h_B,\; t_B t_B h_R, t_B t_B h_B\}\;. \tag{4.18}$$

For the base observable o we will simply use the total production cross-section $o = \sigma_{\text{tot}}(pp \to t\bar{t}h)$. This gives rise to six *idealized cross-sections* $o_i = \sigma(t_X t_Y h_Z)$. By definition these satisfy Eq.(4.14)

$$\sigma_{\text{tot}}(pp \to t\bar{t}h) \;=\; \sum_{X,Y,Z\in\{B,R\}} \sigma(t_X t_Y h_Z)\,. \tag{4.19}$$

It is also convenient to normalize the idealized cross-sections with respect to the total production cross-section:

$$\mu_{XYZ} \;\equiv\; \frac{\sigma(t_X \bar{t}_Y h_Z)}{\sigma_{\text{tot}}(pp \to t\bar{t}h)}\,. \tag{4.20}$$

Expression (4.19) now becomes a partition of unity $\sum \mu_{XYZ} = 1$. We now compute these idealized observables in the SM assuming no BSM physics for the moment. We will perform this for the cone-based observables as well as the p_T-based ones.

We use `MadGraph` to perform the leading-order calculation of the six idealized cross-sections $\sigma(t_X t_Y h_Z)$. For this we generate leading-order (LO) parton-level events of $pp \to t\bar{t}h$ at $\sqrt{s} = 13$ TeV followed by the hadronic decay chain ($h \to b\bar{b}$) ($t \to u\bar{d}b$) ($\bar{t} \to \bar{u}d\bar{b}$). For the cone-based classifier, the boosted label B (rest R) is assigned to the top-quark or to the Higgs boson if the decay products are found within the cone $\Delta R_t < 1.5$ ($\Delta R_t \geq 1.5$) or $\Delta R_h < 1.2$ ($\Delta R_h \geq 1.2$) respectively,

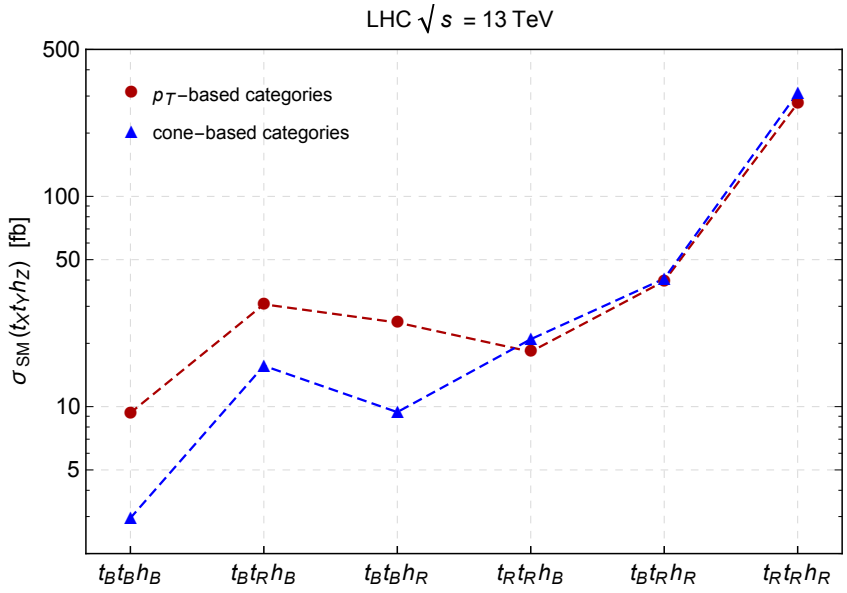


Figure 4.1: The idealized cross-sections for $pp \to t\bar{t}h$ production in the SM. Blue triangles correspond to the cone-based categorization, while red dots correspond to the p_T-based categorization.

which is consistent with the first step of clustering a stream of particles into a fat-jet, as implemented in the fat-jet reconstruction algorithms. For the p_T-based classifier, we label a top-quark or Higgs boson as boosted B if at parton level $p_T^t \geq 250$ GeV or $p_T^h \geq 200$ GeV, respectively, else these were label as rest R. We checked that the resonances satisfying the B label requirement in the cone-based classifier almost always implied a p_T of several hundred GeV for the mother particle. In our simulations, no cuts were applied at the generation level, and the event selection to these six categories was performed for both type of idealized cross-sections.

In Table 4.1 we show the normalized idealized cross sections for each category in the SM. As it can be seen, it drops by two orders of magnitude going from the completely resolved category $t_R t_R h_R$ to the completely boosted category $t_B t_B h_B$. As expected from threshold production dominance, the total cross-section of $pp \to t\bar{t}h$ is dominated by the rest category $t_R t_R h_R$ with approximately 75% (65%) of the events falling into this category for the cone-based (p_T-based) event classifiers. Next in line are the events with only one boosted resonance. Notice that for both categorizations, we obtain the same approximate relation $\mu_{BRR} \approx 2\mu_{RRB} \sim 10\%$. The factor of two is a simple combinatorial factor given that there are two top-quark resonances that contribute to $t_B t_R h_R$ while only one Higgs contributes to $t_R t_R h_B$. This suggests that in the SM the categories with exactly one boosted resonance is approximately blind to the type of resonance (whether it is a Higgs or a top-quark). And this holds true for the idealized cross-sections calculated from both categorizations.

The remaining multi-boosted categories $t_B t_R h_B$, $t_B t_B h_R$ and $t_B t_B h_B$ contain non-trivial information about $pp \to t\bar{t}h$ production in the SM. At this level, both categorizations give very different outcomes. In Fig. 4.1, we show the idealized cross-sections in the SM for these six categories for both the p_T based and cone-based observables. In the SM, the cone-base observables tend to have a much smaller cross-

section than the p_T based ones for categories with two or more boosted resonances in the final state.

4.4 Sensitivity of $\sigma(t_X t_Y h_Z)$ to non-resonant NP

In this section, we investigate the impact of non-resonant BSM physics on the idealized cross-sections in $pp \to t\bar{t}h$ production. We provide a *proof of principle* that demonstrates the capability of these observables to probe different directions in the New Physics parameter space. To this purpose, we compute the six kinematical categories using a well-defined subset of operators in the SM Effective Field Theory (SMEFT).

Assuming that new particles lie at an energy scale Λ much larger than the typical energy of the process under consideration, one can write an effective field theory as an expansion in powers of the inverse cut-off $1/\Lambda$, as discussed in sec. 3.2:

$$\mathcal{L}^{\text{EFT}} = \mathcal{L}^{\text{SM}} + \sum_i \frac{c_i^{(6)}}{\Lambda^2}\mathcal{O}_i^{(6)} + \sum_j \frac{c_j^{(8)}}{\Lambda^4}\mathcal{O}_i^{(8)} + \dots \tag{4.21}$$

In order to show how the six categories are sensitive to different directions in the EFT parameter space, we compute the $\sigma(t_X t_Y h_Z)$ using the following large subset of SMEFT dimension-6 operators:

$$\begin{aligned}
\mathcal{L}^{\text{EFT}} \supset & - \delta k_t \frac{y_t}{v^2}\bar{Q}t\tilde{\varphi}\left(\varphi^\dagger\varphi - \frac{v^2}{2}\right) + \text{h.c.} \\
& - c_{tg}\frac{g_s y_t}{4v^2}\bar{Q}\sigma^{\mu\nu}T^A t G^A_{\mu\nu}\tilde{\varphi} + \text{h.c.} \\
& + \frac{c_{4f}}{v^2}\sum_{i=1,2}\left[(\bar{Q}u_R^i)(\bar{u}_R^i Q) + (\bar{q}_L^i t)(\bar{t}q_L^i)\right],
\end{aligned} \tag{4.22}$$

where $Q = (t_L, b_R)^T$ is the third generation LH quark doublet, $q_L = (u_L^i, d_R^i)^T$ the first and second generation LH quark doublets, $t = t_R$ the RH top-quark, G the QCD field strength, y_t is the top-quark Yukawa coupling, g_s the strong coupling, $v = 256$ GeV the Higgs VEV and φ the Higgs doublet. For convenience we rescaled the coefficients so that $c_i^{(6)} \sim c\, v^2/\Lambda^2$ and where $\sigma^{\mu\nu} = \frac{1}{2}[\gamma^\mu, \gamma^\nu]$. The relevant Wilson coefficients are δk_t, c_{tg} and c_{4f}. Here we have chosen different top-philic operators that can arise from a large variety of UV complete models. The first term is an operator that could arise from integrating out a heavy vector-like fermion. The second term is a chromo-magnetic dipole operator that can only be generated by integrating out a heavy state at the one-loop level. The last operators are four-fermion contact interactions that are very common when integrating out a heavy scalar state. For a discussion on the contribution of these operators in $pp \to t\bar{t}h$ see e.g. Ref. [87, 88]. In principle, a large number of additional dimension-6 operators contribute at the tree-level to this process. However, our goal is not to perform a global EFT analysis, rather show the sensitivity of the proposed kinematical categories. This can be done more efficiently using the operators listed above that, as we will show, are sufficient to provide a large enough BSM parameter space.

As far as four-fermion operators are concerned, some of them can be tested in $pp \to t\bar{t}$ production [87, 88, 89], which however cannot provide enough independent observables to constrain all of them. For example, using the same categorization for $pp \to t\bar{t}$ production only leads to three kinematical categories $\mathcal{K}_{t\bar{t}} = \{t_R\bar{t}_R,\ t_R\bar{t}_B,\ t_B\bar{t}_B\}$, which is not enough to resolve the different non-resonant NP contributions in (4.22). Nonetheless, as we show later, these categories are useful in some directions of BSM parameter space.

In [88] it is argued that the four fermion operators contribute to $\sigma(pp \to t\bar{t}h)$ via a single combination called c_4, and the very similar direction also enters in $pp \to t\bar{t}$ production. While this is true when considering only the interference terms with the SM, this statement is no longer true for the quadratic terms, which can be dominant at high-p_T. This implies that a general EFT analysis of $pp \to t\bar{t}h$ (which is beyond the purpose of this thesis) should include independently all the four fermion operators. Other dimension-6 operators which can also contribute to $pp \to t\bar{t}h$ are the anomalous triple-gluon coupling GGG, the GGh operator, a deviation in the Higgs self-interaction hhh, as well as operators involving an initial-state bottom quarks. In principle, the GGG operator could give a large contribution to both $\sigma(pp \to t\bar{t}h)$ and $\sigma(pp \to t\bar{t})$ [89, 90]; however, it has been shown recently that this effect is negligible after the constraints from multi-jet production at the LHC are taken into account [91]. The GGh and hhh deformations turn out to have a negligible impact on $\sigma(pp \to t\bar{t}h)$, while the operators involving initial-state bottom quarks are suppressed by the small bottom PDF and can be probed much better in other processes like $pp \to t\bar{t}b\bar{b}$ or $pp \to t\bar{t}t\bar{t}$.[3]

4.4.1 EFT expansion: double-insertions and squared terms

Like in similar high-p_T processes at the LHC, such as VV, Vh or vector-boson fusion (VBF) Higgs production, it is expected that the squared dimension-6 terms interfering with the SM amplitude give the dominating contribution to the $pp \to t\bar{t}h$ cross-section. This is indeed the case for the chromo-magnetic dipole and the four-fermi operators in Eq. (4.22). While such terms are formally of order $\mathcal{O}(1/\Lambda^4)$ and are usually taken into account at the same time as the interference of dimension-8 operators with the SM, it is still consistent to include them since in various UV models, especially for strongly coupled theories, they can be enhanced with respect to the neglected dimension-8 ones [92]. For a recent example where this has been demonstrated see [93].

However, differently than in the electroweak Higgs-production processes mentioned above where the relevant dimension-6 operators affecting Higgs couplings have at most one insertion at the amplitude level, in $pp \to t\bar{t}h$ production there are also possible diagrams with two insertions of dimension-6 operators, in particular of the top chromo-dipole operator, see Fig. 4.2 for some examples. The general expression for the $pp \to t\bar{t}h$ amplitude is:

[3] The interference with the SM is given by

$$\sigma^{tth}/\sigma^{tth}_{\rm SM} \simeq 1 + 0.015 k_{gg} - 3.2 \cdot 10^{-4} k_\lambda,$$

where $k_\lambda = \lambda/\lambda_{\rm SM}$ and k_{gg} is defined from the interaction $\delta\mathcal{L} = -\frac{\epsilon^{\rm SM}_{gg}}{2v} k_{gg} G^A_{\mu\nu} G^{A\,\mu\nu} h$, with $\epsilon^{\rm SM}_{gg} = -6.5 \cdot 10^{-3}$, so that $k_{gg} = 1$ describes the SM contribution to $gg \to h$. The quadratic terms are even more suppressed, not being energy-enhanced.

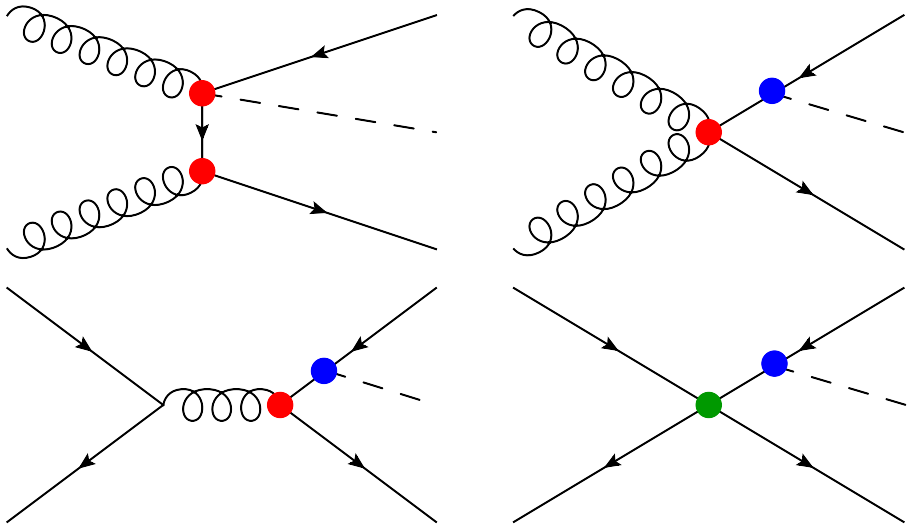


Figure 4.2: Few examples of $gg \to t\bar{t}h$ (top row) and $q\bar{q} \to t\bar{t}h$ (bottom row) diagrams with two insertions of dimension-6 operators. The red, blue, and green dot represent c_{tg}, δk_t, and c_{4f}, respectively.

$$\mathcal{A} = \mathcal{A}_{\text{SM}} + c_i^{(6)} \mathcal{A}^i_{\text{NP}_6=1} + c_i^{(6)} c_j^{(6)} \mathcal{A}^{ij}_{\text{NP}_6=2} + c_j^{(8)} \mathcal{A}^j_{\text{NP}_8=1} + \dots, \tag{4.23}$$

where we also added the neglected terms with one insertion of dimension-8 operators, with $c^{(8)} \sim c\, v^4/\Lambda^4$. The interference arising from the dim-6 double-insertion (NP_6=2) amplitude with the SM is expected to be in general of the same size as the dim-6 squared single-insertion (NP_6=1)^2 terms and thus have to be included for consistency if the latter are also included. Under certain assumptions on the new physics, in particular for $c^{(8)} \ll (c^{(6)})^2$, both terms can still be larger than the interference of (dim-8) operators with the SM. The presence of the (NP_6=2) terms at the amplitude level, however, complicates the situation when a cross-section is evaluated since in general, it will be a quartic function of the EFT coefficients. The cubic and quartic terms, necessarily present when performing, for example, a Monte Carlo evaluation of a cross section, should instead be neglected since their size is smaller than the neglected contributions from dimension-8 operators.

In order to address this, we construct the complete quartic function for any cross section computed starting with the amplitudes in eq. (4.23) (without the dim-8 terms), then neglect all cubic and quartic terms in the coefficients, and finally consider only regions in parameter space with $c^{(6)} \ll 1$, i.e. in the vicinity of the SM. The difference between what is obtained using this truncated quadratic expansion and the full quartic one can be considered as some estimate of the uncertainty due to missing higher-order operators in the EFT expansion.

When performing an EFT analysis of a high-p_T process such as $pp \to t\bar{t}h$, even more so when focussing specifically on boosted events, the issue of the validity of the EFT approach should be addressed. However, since our goal here is not to perform a global and consistent EFT fit, rather demonstrate the sensitivity of the proposed categories to different BSM scenarios, we do not address this issue further. In particular, we stress that a measurement of the idealized cross-sections proposed here could also be used to set bounds on new physics models with light new physics,

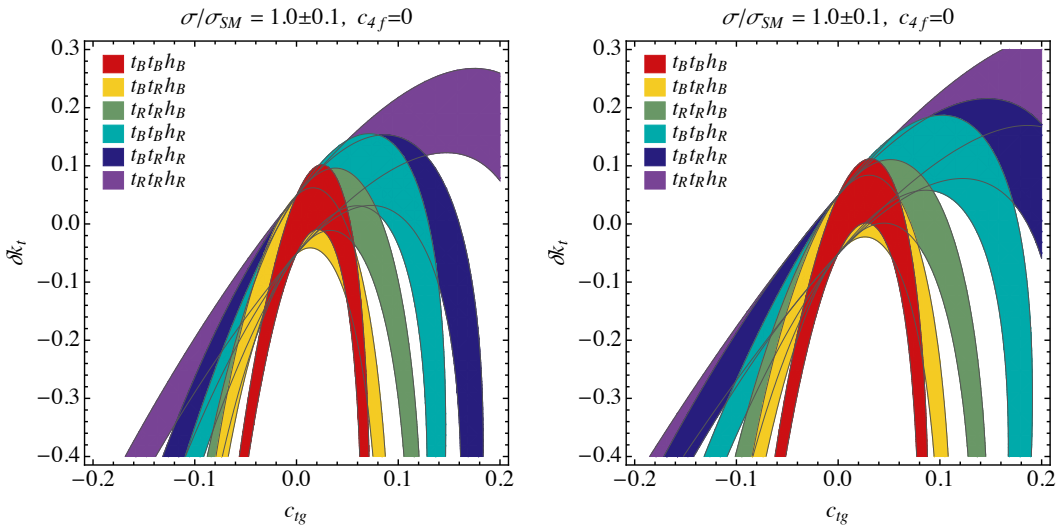


Figure 4.3: Sensitivity to the dim-6 EFT operators $(\delta\kappa_t, c_{tg})$ of each kinematical category $t_X t_Y h_Z$ for the cone-based categorization (left panel) and the p_T-based categorization (right panel). We have set $c_{4f} = 0$. The colored bands represent $\mu^{\text{EFT}}_{XYZ} = 1 \pm 0.1$.

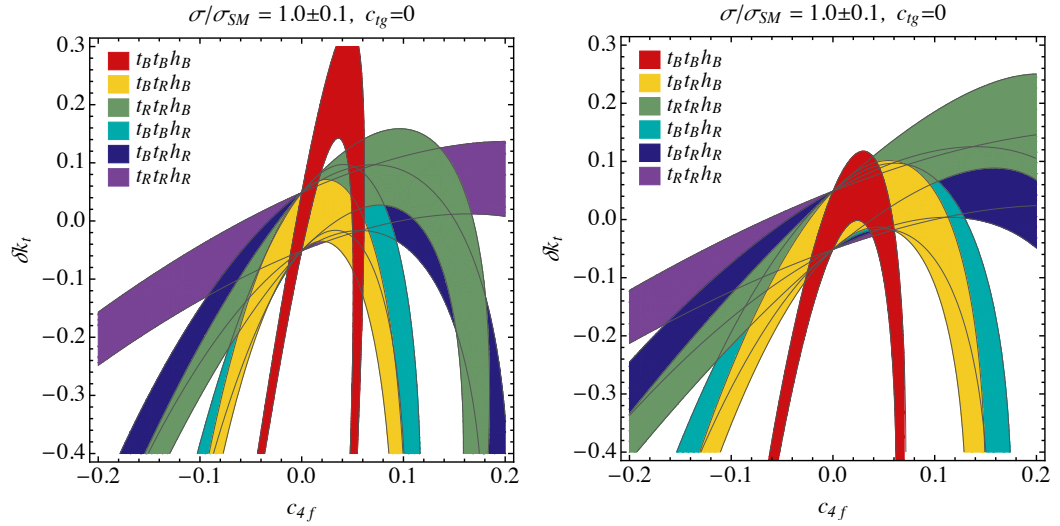


Figure 4.4: Sensitivity to the dim-6 EFT operators $(\delta\kappa_t, c_{4f})$ of each kinematical category $t_X t_Y h_Z$ for the cone-based categorization (left panel) and the p_T-based categorization (right panel). We have set $c_{tg} = 0$. The colored bands represent $\mu^{\text{EFT}}_{XYZ} = 1 \pm 0.1$.

which fall outside of the SMEFT range of applicability.

4.4.2 Sensitivity of the EFT coefficients

By implementing the operators in Eq. (4.22) in a `FeynRules` model, we generate the partonic process at LO in QCD:

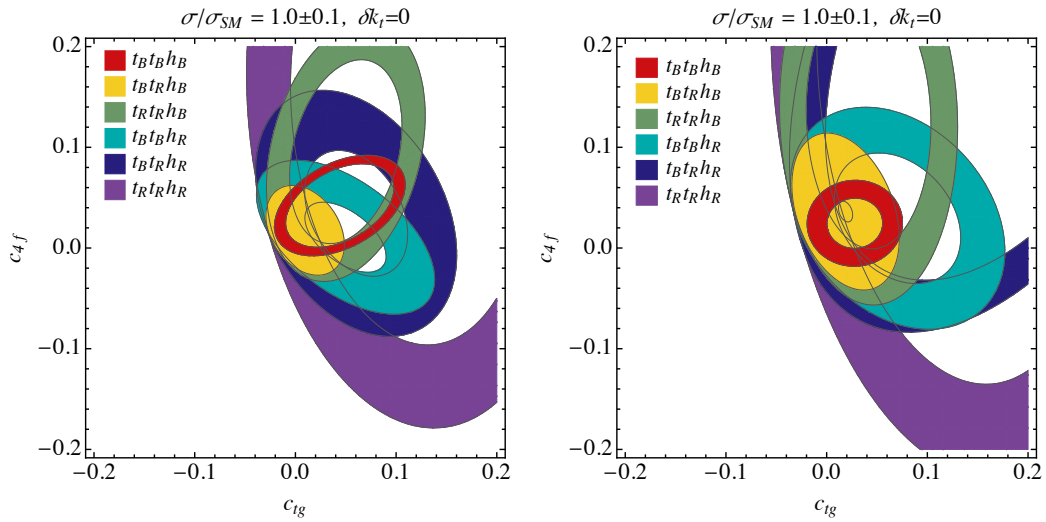


Figure 4.5: Sensitivity to the dim-6 EFT operators (c_{tG}, c_{4f}) of each kinematical category $t_X t_Y h_Z$ for the cone-based categorization (left panel) and the p_T-based categorization (right panel). We have set $\delta k_T = 0$. The colored bands represent $\mu^{\text{EFT}}_{XYZ} = 1 \pm 0.1$.

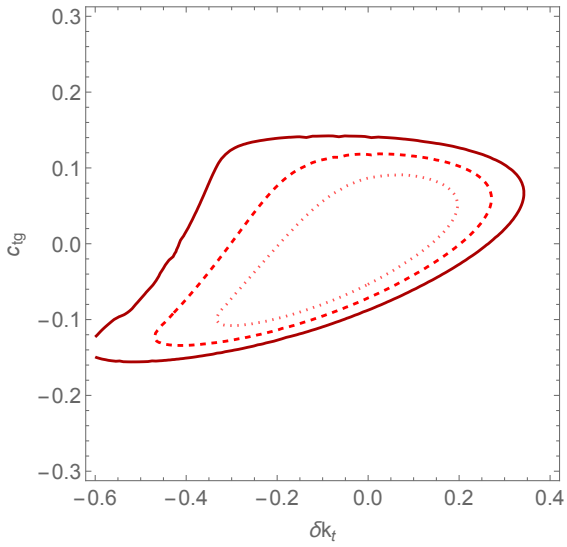


Figure 4.6: Fit to the Wilson coefficients δk_T and c_{tg} while marginalizing over the four-fermion operator c_{4f}. the contours corrspond to the 1σ (red dotted), 2σ (red dashed) and 3σ (red solid) regions.

$$pp \to t(\to u\bar{d}b)\ \bar{t}(\to d\bar{u}\bar{b})\ h(\to \bar{b}b) \tag{4.24}$$

events using `MadGraph5` for enough points in parameter space in order to be able to reconstruct the complete quartic polynomial for the cross-section. From the effective operators listed in Eq. (4.22), the generic quartic expression for the $pp \to t\bar{t}h$ cross-section with up to two EFT insertions at the amplitude level can be written as:

$$\begin{aligned}\sigma(pp \to t\bar{t}h) \;=\; & A^{\rm SM}\,(1+\delta k_t)^2 \;+\; A_1\,c_{tg}(1+\delta k_t) \;+\; A_2\,c_{tg}(1+\delta k_t)^2 \\ & +\; A_3\,c_{tg}^2 \;+\; A_4\,c_{tg}^2\delta k_t \;+\; A_5\,c_{tg}^2\delta k_t^2 \\ & +\; A_6\,c_{tg}^3 \;+\; A_7\,c_{tg}^3\delta k_t \;+\; A_8\,c_{tg}^4 \\ & +\; A_9\,c_{4f}(1+\delta k_t)^2 + A_{10}\,c_{4f}^2(1+\delta k_t)^2 \\ & +\; A_{11}\,c_{tg}c_{4f}(1+\delta k_t) \;+\; A_{12}\,c_{tg}c_{2u}(1+\delta k_t)^2\,,\end{aligned} \tag{4.25}$$

where $A^{\rm SM}$ describes the SM contribution and $A_{1,\ldots,12}$ correspond to higher order NP terms. In order to extract these 13 coefficients for each of the categories, we perform Monte Carlo simulations with `MadGraph` for 13 different points in parameter space, divide all the events into the six kinematical categories and then solve for the coefficients. We work with the normalized idealized cross-section defined by

$$\mu_{XYZ}^{\rm EFT} \;\equiv\; \frac{\sigma(t_X\bar{t}_Y h_Z)}{\sigma_{\rm SM}(t_X\bar{t}_Y h_Z)} \tag{4.26}$$

for each category as a quartic function of the three Wilson coefficients δk_t, c_{tg} and c_{4f}. Then, we truncate the polynomial leaving out any term beyond the quadratic degree in the EFT coefficients, as discussed above. The resulting quadratic polynomial for (4.26) for each kinematical category is given by

$$\mu_{XYZ}^{\rm EFT} \;=\; 1+\sum a_{ij}c_ic_j \;=\; 1 \;+\; a_1\,\delta k_t \;+\; a_2\,c_{tg} \;+\; a_3\,c_{4f} \tag{4.27}$$

$$+\; a_4\,(\delta k_t c_{tg}) \;+\; a_5\,(\delta k_t c_{c4}) \;+\; a_6\,(c_{tg}c_{c4}) \tag{4.28}$$

$$+\; a_7\,\delta k_t^2 \;+\; a_8\,c_{tg}^2 \;+\; a_9\,c_{4f}^2\,. \tag{4.29}$$

The extracted values of the nine coefficients $a_1, ..., a_9$ associated to each quadratic form c_ic_j can be found for each category in the matrix tables 4.2 by looking at the entry corresponding to the Wilson coefficients c_i and c_j at a given row and column, respectively. We show this only for the cone-based observables, similar coefficients can be extracted for the p_T-based ones by the same procedure. These results show that the values of a_i change drastically from category to category suggesting that the observables are sensitive to different NP contributions. Right next to each table we also display a visual representation that highlights the different patterns in constructive and destructive interference within each kinematical category. There, the dark red (dark blue) corresponds to the largest positive (negative) value for the coefficients in the corresponding polynomial.

In order to explicitly demonstrate the sensitivities of the categories $\mathcal{K}_{t\bar{t}h}$ to different combinations of the Wilson coefficients we provide in figure 4.3, figure 4.4 and figure 4.5 the regions for $\mu_{XYZ}^{\rm EFT} = 1 \pm 0.1$ for each category in the $(\delta k_t - c_{tg})$, $(\delta k_t - c_{4f})$ and $(c_{4f} - c_{tg})$ projection planes for which the third Wilson coefficients has been set to zero. In each figure we give results for the cone-based (left panel) and the p_T-based (right panel) idealized observables.

Remarkably, the set of idealized observables remove all flat directions when combined. All bands intersect at a closed region near the SM prediction (zero Wilson

coefficients). The most orthogonal categories are the rest category $t_R t_R h_R$ (purple regions) and completely boosted category $t_B t_B h_B$ (red regions) with all other categories in between. Of interest is the kinematical category $t_B t_R h_B$ (yellow regions) with an idealized cross-section much larger than $t_B t_B h_B$. This category when combined with the "rest" category (or even $t_B t_R h_R$) removes all the flat directions in the coupling planes. From the plots, one can also notice that the idealized cross-sections defined with the cone-based classifier resolve slightly better the different directions in EFT parameter space when compared to the observables defined with the p_T-based classifier.

Assuming that the idealized observables can be measured in the HL-LHC, one can perform a global fit to the set of kinematical categories and extract limits on the three Wilson coefficients. For example, let us assume that the LHC can measure three of the idealized cross-sections, $\sigma(t_B t_R h_B)$, $\sigma(t_B t_R h_R)$ and $\sigma(t_R t_R h_R)$ to be SM-like with a precision of 50 %, 20 % and 20 %, respectively[4]. In figure 4.6 we show the best fit contours at 1σ (red dotted), 2σ (red dashed) and 3σ (red solid) the $(\delta k_t, c_{tg})$–plane while marginalizing over the remaining Wilson coefficient c_{4f}.

4.5 Experimental probes for $t_X t_Y h_Z$

We now make contact with experiments. The idea behind the idealized observables is to maximize the sensitivity to NP while minimizing the (BSM) theory dependence with a set of experimentally-inspired observables that only depend on the production modes. On one hand, the idealized cross-sections can be easily calculated in an explicit new physics models using partonic Monte Carlo simulations, while, on the other hand, they can be extracted by an unfolding procedure from an experimental analysis. In this section, we outline a simple search strategy that can be used to measure the set of idealized observables for $pp \to t\bar{t}h$ production. Our aim is not to give a full experimental analysis including backgrounds and detector effects. We leave this task for the experimental collaborations. We will here demonstrate how to measure one of the idealized cross-section assuming SM physics, namely, $\sigma_{\rm SM}(t_B \bar{t}_R h_B)$.

In what follows we will only focus on the $pp \to t\bar{t}h$ signal process and completely ignore the backgrounds from other processes. Even in this idealized scenario, it is challenging to extract the individual cross-sections $\sigma(t_X t_Y h_Z)$. Because of the intrinsic inefficiencies when reconstructing via experiments a boosted or resolved top-quark or Higgs Boson, the different kinematical categories in $\mathcal{K}_{tth}$ will overlap at the experimental level. This has the effect of introducing sizeable systematic uncertainties in the measurements of the idealized cross-sections that need to be estimated before performing any real experimental analysis.

These observations imply that any dedicated search strategy aiming to measure just one of the idealized cross-section $\sigma(t_X t_Y h_Z)$, is really measuring a statistical mixtures of the cross-sections coming from all kinematical categories. We can write the measured idealized cross-section in the SM in the following way:

$$\sigma^{\rm exp}_{\rm SM}(t_X \bar{t}_Y h_Z) \;=\; \sum_{X' Y' Z'} \mathcal{E}^{\,X\,Y\,Z}_{X'Y'Z'} \cdot \sigma_{\rm SM}(t_{X'} \bar{t}_{Y'} h_{Z'}). \tag{4.30}$$

[4]These are just non-realistic benchmark values that are fixed to illustrate the quality of the fit.

μ^{EFT}_{RRR}	SM	δk_t	c_{tg}	c_{4f}
SM	1	2	−4.1	−1.5
δk_t		1	−6.3	−3.0
c_{tg}			16.6	2.5
c_{4f}				4.6

μ^{EFT}_{BRR}	SM	δk_t	c_{tg}	c_{4f}
SM	1	2	−4.3	−2.4
δk_t		1	−6.7	−4.9
c_{tg}			31	17
c_{4f}				20

μ^{EFT}_{RRB}	SM	δk_t	c_{tg}	c_{4f}
SM	1	2	−4.7	−3.8
δk_t		1	−6.1	−7.7
c_{tg}			70	−24
c_{4f}				26

μ^{EFT}_{BBR}	SM	δk_t	c_{tg}	c_{4f}
SM	1	2	−5.2	−4.2
δk_t		1	−8.8	−8.5
c_{tg}			46	59
c_{4f}				62

μ^{EFT}_{BRB}	1	δk_t	c_{tg}	c_{4f}
1	1	2	−3.1	−3.5
δk_t		1	−5.6	−6.9
c_{tg}			107	60
c_{4f}				82

μ^{EFT}_{BBB}	1	δk_t	c_{tg}	c_{4f}
1	1	2	−7.9	−17
δk_t		1	−13	−35
c_{tg}			184	−199
c_{4f}				311

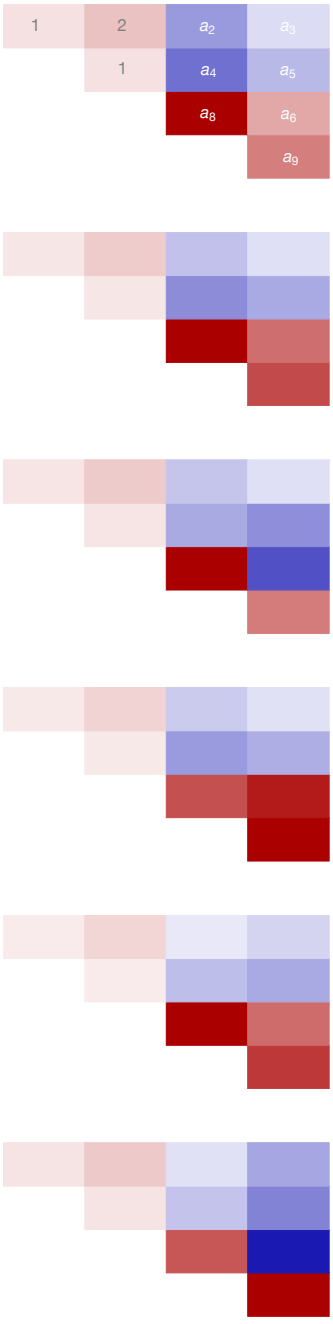


Table 4.2: Extracted values of the (normalized) idealized cross-sections with respect to the SM prediction $\mu^{\text{EFT}}_{XYZ} = \sigma(t_X t_Y h_Z)/\sigma_{\text{SM}}(t_X t_Y h_Z)$ for different EFT Wilson coefficients in $pp \to t\bar{t}h$ production. The numbers correspond to the a_i coefficients of the quadratic polynomial in Eq.(4.27). On the right-hand side of each table we provide a visual representation of the pattern in the coefficients a_i for the corresponding category using a color temperature map (red for positive and blue for negative values).

Each entry $\mathcal{E}^{X\,Y\,Z}_{X'Y'Z'}$ of the 6×6 migration matrix represents the fraction of events from the category $t_{X'}t_{Y'}h_{Z'}$ that satisfy the experimental cuts and selections of a dedicated signal region targeting the category $t_X t_Y h_Z$. Ideally, the experimentally measured idealized cross-section (4.30) would be dominated by only one of the categories, meaning that $\mathcal{E} \approx \mathbb{1}$. In practice, this is not what actually happens as shown below.

The efficiencies $\mathcal{E}^{X\,Y\,Z}_{X'Y'Z'}$ can be computed in the SM and are expected to be model-independent. Indeed, non-resonant BSM physics will in general modify each idealized cross-section differently (for example dimension-6 operators can easily enhance the categories with multi-boosted resonances as indicated in sec 4.4.2), but should not modify too much the weights $\mathcal{E}^{X\,Y\,Z}_{X'Y'Z'}$ since the kinematics inside each $t_X t_Y h_Z$ category should not vary considerably. The latter (very general) assumption allows performing the complicated unfolding procedure only once and for the SM. The $\mathcal{E}$ efficiencies extracted from the much simpler SM analysis could then be used for confronting the idealized observables to complicated BSM physics scenarios, leaving the matching to the explicit NP model only at the much simpler level of the theoretical calculation of the simplified cross sections $\sigma(t_X t_Y h_Z)$.

4.5.1 The search strategy

The idea is to devise six independent dedicated search strategies, one for each $t_X t_Y h_Z$ category. The task is to come up with cut-based searches such that the efficiency matrix in (4.30) is $\mathcal{E} \approx \mathbb{1}$, i.e. practically no event migration between categories.

From what was explained in sec. 4.3.4, the boosted resonances t_B and h_B can be measured in the experiments by binning events into fat-jet multiplicity N_J and simply counting top-tagged and Higgs-tagged fat jets using substructure techniques. Experimentally one can slice the phase space into six mutually exclusive regions defined by the output of tagging the reconstructed fat-jets:

- $N_J = 3$, with 2 top-jets and 1 fat Higgs-jet ($t_B t_B h_B$).
- $N_J = 2$, with 1 top-jet and 1 fat Higgs-jet ($t_B t_R h_B$).
- $N_J = 2$, with 2 top-jets and no Higgs-jets ($t_B t_B h_R$).
- $N_J = 1$, with 1 top-jet and no Higgs-jets ($t_B t_R h_R$).
- $N_J = 1$, with no top-jets and 1 Higgs-jet ($t_R t_R h_B$).
- $N_J = 0$, no reconstructed fat-jets ($t_R t_R h_R$).

In parenthesis we have included the kinematical category in $\mathcal{K}_{t\bar{t}h}$ that would in practice contribute the most to each signature. Of course, while the idealized cross-section is in principle independent of the decay channels of each resonance, in practice the experimental extraction of $\sigma(t_X t_Y h_Z)$ with boosted resonances relies on hadronically decaying boosted tops and Higgs.

The real challenge now lies in correctly identifying the resonances at rest t_R and h_R. The reason for this is two-fold: (i) if the resonance $\mathcal{X}_R$ decays hadronically into multiple jets, in practice it is easy to distinguish $\mathcal{X}_R$ from $\mathcal{X}_B$ (the former forms fat jets). The problem is that the large jet combinatorics forbids for an efficient

resonance reconstruction using mass window variables for categories with multiple $\mathcal{X}_R$. For instance, it would be hard to distinguish between events arising from $t_R t_R h_B$ and $t_B t_R h_R$ or events between $t_B t_R h_B$ and $t_B t_B h_R$. (ii) If the top-quark decays leptonically $t_R \to j\ell^{\pm}\nu$ the missing transverse energy from the neutrino also forbids for direct resonance reconstruction, making it difficult to infer if the underlying top-quark was either a t_B or a t_R. Notice that the misidentification of B with R resonances (hadronic or leptonic) are the main source of event migration between categories at the experimental level, unavoidably leading to $\mathcal{E} \neq \mathbb{1}$.

Anticipating a large QCD multi-jet background for the hadronically decaying modes of t_R, the least compromising scenario is to use at the experimental level leptonic top-decays to identify t_R. We supplement the six signal regions described above with a requirement for reconstructed leptons $\ell^{\pm}$ from the decaying top-quarks t_R. In addition, we also require extra narrow jets (possibly with b-tagging) accounting for the bottom jet from the resolved top-quark decay or $h \to b\bar{b}$. In table 4.3 we define the six dedicated signal regions `tX tY hZ` of our search strategy intended to measure the idealized cross-sections $\sigma(t_X t_Y h_Z)$. In order to discriminate between leptonic t_R and t_B we still need to impose selection cuts. For this we require that the visible part of the top to be separated by a distance cut $R_{\ell j} > \Delta R_0$ where $R_{\ell j}$ is the planar distance between any of the leptons in the event and the closest hard jet, and ΔR_0 is a distance parameter to be fixed at some optimized value. This selection requirement must accept most leptonic t_R with high efficiency while at the same time reject a large portion of leptonic boosted t_B resonances. This is a very simple and non-optimal requirement (since it ignores the neutrino) that aligns well with the cone-based categorization. Other requirements can be used in place (in the next section we use a modified version of what was described above).

4.5.2 Case study: the $t_B t_R h_B$ category

Apart from the completely boosted category $t_B t_B h_B$, which is challenging to measure because of its very small cross-section (e.g. around ≈ 1 fb in the SM), the category with the best BSM sensitivity is $t_B t_R h_B$. In order to reduce as much as possible the large QCD backgrounds, as explained in the previous subsection, we will focus

Signal Region	N_ℓ	N_j ($\Delta R_{\text{narrow}} = 0.4$)	N_J ($\Delta R_{\text{fat}} \sim 1$)	N_{top}	N_{Higgs}
`tB tB hB`	0	–	3	2	1
`tB tB hR`	0	≥ 2	2	2	0
`tB tR hB`	1	≥ 1	2	1	1
`tB tR hR`	1	≥ 3	1	1	0
`tR tR hB`	2	≥ 2	1	0	1
`tR tR hR`	2	≥ 4	0	0	0

Table 4.3: Event selections for the six experimental signal regions `tX tY hZ` optimized for extracting the idealized cross-sections $\sigma(t_X t_Y h_Z)$. N_ℓ, N_j and N_J corresponds to lepton, narrow-jet and fat-jet multiplicity. N_{top} and N_{Higgs} are the number of tagged fat top-jets and fat Higgs-jets using substructure techniques.

on the t_R decaying leptonically an apply the basic event selections in the `tB tR hB` signal region described in table 4.3. In our analysis, we use the most basic top-quark and Higgs taggers available in the market with out-of-the-box settings. Top-tagging was performed with the John-Hopkins tagger [82] implemented in `FastJet` with default values for the tagging parameters. The Higgs-tagger we performed with the Mass-Drop tagger supplemented with sub-jet filtering as described in [80], with the default parameter settings in `FastJet`.

The precise `tB tR hB` event selection is the following: we start by clustering jets using the Cambridge/Aachen algorithm with a large cone radius of $\Delta R = 1.5$. The resulting fat-jets are then run through a top-tagger. The tagged top-quarks are set aside. On the remaining particles, the same fat-jet clustering procedure is performed but this time with a cone size of $\Delta R = 1.2$. The resulting fat-jets are run through the Higgs-tagger and any tagged Higgs-jets is set aside. The event is selected if there is exactly one top-jet satisfying $p_T > 250$ GeV and exactly one Higgs-jet satisfying $p_T > 200$ GeV, else the event is rejected. For each selected event we remove the identified top and the Higgs fat-jets from the event and subject the remaining particles to narrow jet re-clustering using $\Delta R = 0.4$. The event is retained if at least one jet is present with minimum transverse momentum cut of

cone-based / `tB tR hB`	SM [%]	SMEFT ($c_{tg} = 1$) [%]	SMEFT ($c_{4f} = 1$) [%]
$\mathcal{E}_{BBB}$	5.24	5.52	3.54
$\mathcal{E}_{BBR}$	1.22	1.37	1.30
$\mathcal{E}_{BRB}$	3.90	3.45	3.72
$\mathcal{E}_{BRR}$	0.77	0.81	0.79
$\mathcal{E}_{RRB}$	1.11	1.24	1.00
$\mathcal{E}_{RRR}$	0.11	0.17	0.15

p_T-based / `tB tR hB`	SM [%]	SMEFT ($c_{tg} = 1$) [%]	SMEFT ($c_{4f} = 1$) [%]
$\mathcal{E}_{BBB}$	3.26	3.74	3.09
$\mathcal{E}_{BBR}$	0.79	1.00	0.93
$\mathcal{E}_{BRB}$	2.31	2.68	2.67
$\mathcal{E}_{BRR}$	0.60	0.62	0.54
$\mathcal{E}_{RRB}$	0.56	0.71	0.48
$\mathcal{E}_{RRR}$	0.06	0.09	0.05

Table 4.4: Efficiencies $\mathcal{E}^{BRB}_{XYZ}$ of selecting events from the pure kinematical categories $t_X t_Y h_Z$ when performing the search `tB tR hB` using the cone-based categorization (upper table) and the p_T-based categorization (lower plot). First column corresponds to the SM $pp \to t\bar{t}h$ while the second and third have non-resonant NP contributions from the chromo-magnetic oparator ($c_{tg} = 1$) and and the four-fermion operator ($c_{4f} = 1$).

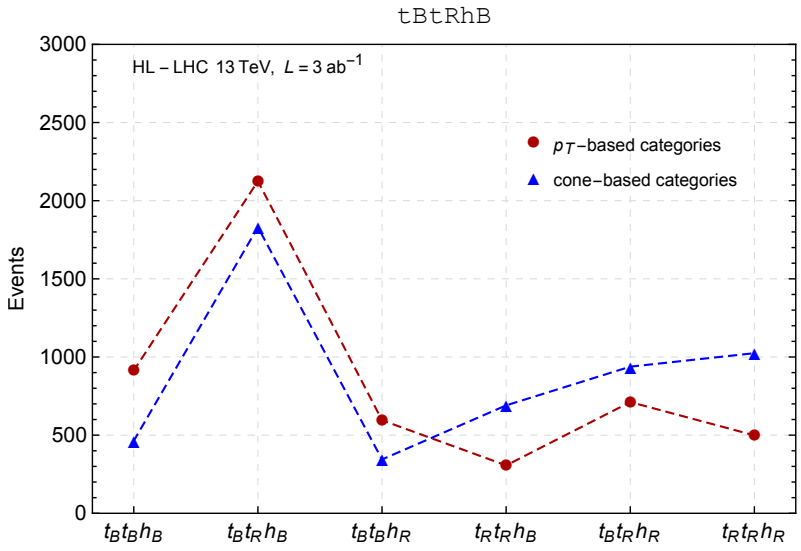


Figure 4.7: Expected number of SM $pp \to t\bar{t}h$ events at the LHC (at 3 ab^{-1}) from each kinematic category when implementing the event selections in `tBtRhB`. The blue triangles (red dots) correspond to the cone-based (p_T-based) categorization.

$p_T^j > 20$ GeV. Finally, we ask for the event to have exactly one isolated lepton, $e^\pm$ or $\mu^\pm$, with minimum transverse momentum of $p_T^\ell > 20$ GeV. For the leptonic isolation requirement, we use the mini-isolation cone [94] that depends on the momentum of the candidate lepton. We use

$$R_{\rm iso} \;=\; \max\left(0.2\,,\,\frac{20\,\text{GeV}}{p_T^\ell}\right)\,. \tag{4.31}$$

The lepton is considered isolated if $p_T^{\rm tot}/p_T^\ell$ is smaller than 5 % where $p_T^{\rm tot}$ is the sum of the transverse momentum of all particles (excluding the lepton) inside the isolation cone $R_{\rm iso}$. Compared to the leptonic isolation with fixed-cone radius, mini-isolation improves signal efficiency considerably for environments with large hadronic activity. In order to guarantee that the leptonic top daughters are consistent with t_R, we demand the isolated lepton and the leading jet in the event to be separated from each other at distance larger than $\Delta R_{j\ell\nu} > 1.5$.

In order to extract the efficiencies $\mathcal{E}^{BRB}_{XYZ}$, we first simulated in `MadGraph` a sample of 1M $pp \to t\bar{t}h$ events at $\sqrt{s} = 13$ TeV decaying into $t \to b\ell\nu$, $\bar{t} \to \bar{b}jj$ and $h \to b\bar{b}$ in the SM and in the SMEFT for two representative cases: $\{\delta k_T, c_{tg}, c_{4f}\} = \{0, 1, 0\}$ and $\{\delta k_T, c_{tg}, c_{4f}\} = \{0, 0, 1\}$. These partonic events were then partitioned into the six categories in $\mathcal{K}_{t\bar{t}h}$ based on their kinematics. This was done for both the cone-based and p_T-based categorizations. Each of the partonic samples was then showered and hadronized using `Pythia8` and jet clustering was performed with `FastJet`. Next, the resulting events for each kinematical category were subject to the selection cuts in `tBtRhB` described above. Results for the search efficiencies $\mathcal{E}^{BRB}_{XYZ}$ for events in each of the six categories $t_X t_Y h_X$ are displayed in table 4.4 for the cone-based categorization (upper table) and for the p_T-based categorization (lower table).

A few comments enter into place: for both categorizations the efficiency $\mathcal{E}^{BRB}_{XYZ}$

peaks for the corresponding target category $t_B t_R h_B$ and is suppressed for all other categories except for the fully boosted one $t_B t_B h_B$, where the efficiency is even a bit larger than for $t_B t_R h_B$. The reason why this occurs can be traced back to the difficulty of discriminating a boosted leptonic top-quark from a resolved one. Another important remark is that the efficiencies do not vary much when computed in the SM or in any of the SMEFT scenarios. The largest deviations are not larger than 50% for the rest category. This shows that our procedure is indeed model-independent.

We finalize this chapter by giving the expected number of SM signal events from the `tB tR hB` search strategy for each kinematical category at the High-Luminosity LHC. For an integrated luminosity of $\mathcal{L}_{int} = 3000$ fb^{-1} , the results for each kinematical categorization is shown in figure 4.7. There we can see that the number of expected events indeed peaks at the intended category $t_B t_R h_B$. In conclusion, if the $t\bar{t}$ plus jet backgrounds can be kept under experimental control, then these kinematical categories could potentially be measured at the HL-LHC phase.

Chapter 5

New Physics in Four Top-quarks

5.1 Introduction

In our experimental quest towards discovering NP at the LHC, it is important to scrutinize all possible final states, including those produced in rare SM processes with small cross-sections, and usually considered to be beyond the LHC Run-II sensitivity reach. Most current studies of this kind are geared towards final states and processes involving the heaviest known particles, the Higgs boson, and the top quark. In particular, many recent proposals exist to study the Higgs boson in final states with extra radiation [95], differential distributions in Higgs production [96, 97, 98, 99], Higgs pair production [100, 101, 102] and Higgs production in association with other massive particles [103] including top-quarks [72, 73, 104]. On the other hand, the top quark has been studied extensively in LHC mainly through top-quark pair production [105, 106, 107, 108], as well as through single-top production [109, 110], both within the SM and beyond. More recently, associate $t\bar{t}H$ and $t\bar{t}Z$ production have also become objects of intense study [111, 112]. Finally, one of the long-standing challenges in top physics is to measure four-top quark production. This process is interesting given that its cross-section in the SM can be significantly enhanced in many NP scenarios, see Ref. [113, 114, 115, 116, 117, 118, 119, 120].

In this chapter, we propose a dedicated search strategy for the SM production of four-top quarks at the 13 TeV LHC in the multi-lepton decay channels. We probe the reach of our search and show that it is possible for the LHC to find evidence for four-top production before the end of Run-II. A measurement or an upper bound on the SM four-top process using our proposed search strategy can be directly used to constrain interesting (non-resonant) NP scenarios coupling dominantly to top-quarks. These *top-philic* NP scenarios with SM-like kinematics would easily avoid bounds from top-pair production and other searches and would most naturally manifest themselves as deviations in the total four-top production. To illustrate this, we present current and projected constraints on two top-philic simplified models (a neutral vector and a scalar mediator) entering four-top production and show that our search strategy can cover important regions of parameter space. In order to further motivate our search, we present a UV complete model that can give rise to top-philic vector boson and scalars. This particular model can evade all standard searches at low and high energies except for $pp \to t\bar{t}t\bar{t}$ production where large portions of parameter space can be probed at the LHC. As a final note, the search proposed here can also be used to probe very heavy top-philic NP that gives rise

to effective operators of the form $(\bar{t}_R\gamma^\mu t_R)(\bar{t}_R\gamma_\mu t_R)$. This opens the possibility of including measurements of four-top production in global fits for effective operators, see [93] for a recent analysis.

Four-top production in the SM is challenging due to its small cross-section at the LHC, of about 10 fb at 13 TeV [121]. There are only few existing proposals for beyond SM four-top searches [113, 115, 116, 117, 118, 120, 122, 123, 124, 125, 126, 127] and also some preliminary experimental reports [128, 129, 130, 131, 132, 133, 134] in this direction. One of the main issues with the existing phenomenological analyses is that these do not include all of the relevant background processes for the multi-lepton decay channels, making the predictions difficult to trust. Here, we present for the first time a detailed classification of all relevant backgrounds for the four-top multi-lepton channels (same-sign dilepton and trileptons). Special emphasis is given to the reducible backgrounds comprised of fake leptons and charge-flip. These backgrounds, which are always difficult to model in a phenomenological analysis, must be correctly estimated since they turn out to be important for giving reliable predictions.

This chapter is structured as follows. In Section 5.2 we describe the expected features of the SM four-top process at the LHC relevant for the multi-lepton channels. In Section 5.3 we design dedicated searches for the SS dilepton and trilepton channels separately and then combine them statistically to present discovery/evidence luminosities as well as exclusion limits for the SM four-top signal strength. To be as general as possible, we also consider the impact on signal sensitivity when using different estimations of the reducible backgrounds (fake leptons and charge-flip). In Section 5.4 we use our SM four-top strategy to give bounds on two top-philic non-resonant NP models that contribute to $pp \to t\bar{t}t\bar{t}$ at the LHC. Finally, in Section 5.5 we provide a description of a top-philic UV complete model that gives rise to the Z' simplified model probed in Section 5.4.

5.2 Signal features

The SM four-top process produces a rich set of final states all giving rise to interesting signatures at the LHC. A useful way of illustrating all four-top decay channels is depicted in figure 5.1, where each side of the unit *square* is partitioned into the leptonic, semi-leptonic and hadronic branching ratio of $t\bar{t}$. Regions with the same shade represent a particular $t\bar{t}t\bar{t}$ decay mode and the total area of the shades gives the corresponding branching ratio. The dominant decay mode is the mono-leptonic channel with a branching ratio of $\sim 40\%$, followed by the fully hadronic and opposite-sign (OS) dilepton modes with $\sim 20\%$ each, the same-sign (SS) dilepton (represented by the dotted blue contour) and the trilepton modes with $\sim 10\%$ each and finally the fully leptonic mode with $\sim 1\%$. Because of large backgrounds in the mono-leptonic and hadronic channels, the SS dilepton channel is usually considered to be the most promising search channel for the SM four-top process at the LHC. Bellow, we argue that, for higher luminosities within Run-II, a dedicated search based on trileptons can achieve comparable or even better sensitivities to the four-top signal than the SS dilepton channel.

In the SM, the production of $t\bar{t}t\bar{t}$ is predominantly a QCD process of order $\mathcal{O}(\alpha_S^4)$ that requires a partonic center-of-mass energy of at least $4m_t \sim 692$ GeV, resulting in a very small cross-section at the LHC. Besides QCD, there is also a sub-leading

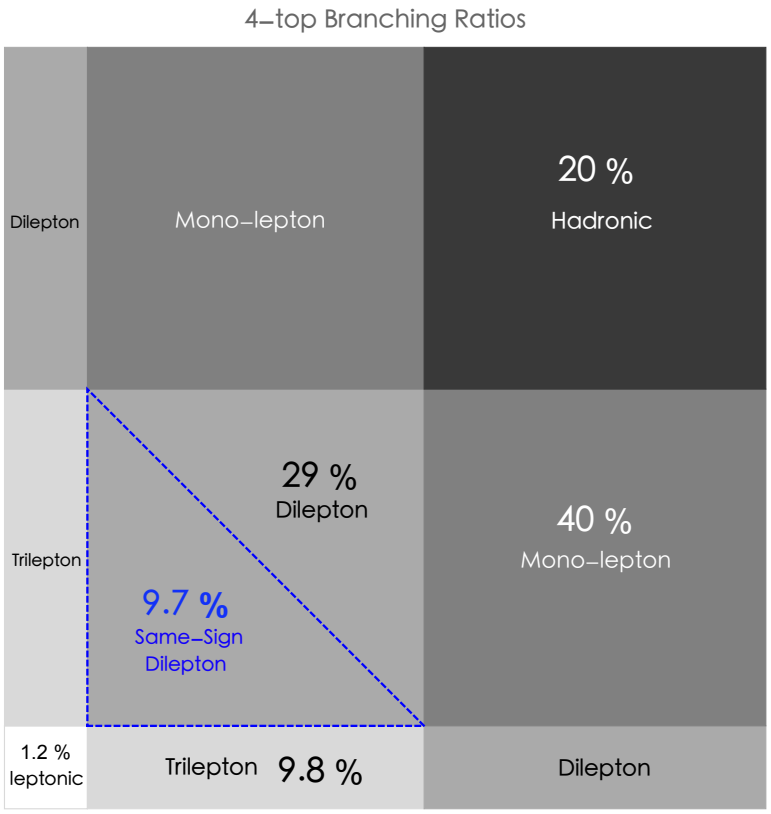


Figure 5.1: Branching ratios for the possible $t\bar{t}t\bar{t}$ decay channels. The total square has unit area and each side represents either the leptonic, semi-leptonic or hadronic branching ratio of $t\bar{t}$ (undecayed τ-lepton is included). For the sake of clarity we have rendered all areas corresponding to a particular final state with the same color (even disconnected in the graphic), whereas the quoted percentage corresponds to the sum of all areas associated to a particular final state.

Higgs boson exchange contribution of order $\mathcal{O}(\alpha_S^2 y_t^4)$ and an EW contribution of order $\mathcal{O}(\alpha_S^2 \alpha^2)$, both accounting for ~10% of the total cross-section. The QCD driven four-top process is given at leading order (LO) by 72 and 12 topologically inequivalent diagrams from initial gg and $q\bar{q}$ scattering, respectively. At the LHC, the gluon-initiated process accounts for approximately 95 % of the total QCD cross-section. Using `MadGraph5` [35] we calculated at $\sqrt{s} = 13$ TeV the four-top production cross-section and found $\sigma^{LO}(t\bar{t}t\bar{t}) = 9.7$ fb at LO and $\sigma^{NLO}(t\bar{t}t\bar{t}) = 12.32$ fb at next-to-leading order (NLO). For a complete analysis of $pp \to t\bar{t}t\bar{t}$ at NLO in QCD, the reader is referred to Ref. [121]. In figure 5.2 we show three representative Feynman diagrams contributing to the $gg \to t\bar{t}t\bar{t}$ process.

Since $t\bar{t}t\bar{t}$ decays into a large number of final states (of order $\mathcal{O}(10)$ particles) predominantly into the central direction, it is important to give a rough estimate of the amount of accidental object overlaps expected in the detector. For this, we considered the minimum angular separation ΔR_{2t}^{min} between any pair of top-quarks in each event (irrespective of their charge). In figure 5.3 we have plotted the truth-level distribution of this variable for the four-top signal (black), compared to top-pair production $pp \to t\bar{t}$ (red). The plot suggests that we should expect a considerable amount of object overlaps between the decay products of the four-top signal. This

accidental overlap will manifest itself as a drop in isolation efficiency at the object reconstruction level, in particular, it will show up as an increase in the fraction of leptons from top decays that end up accidentally close to jets.

The main signature of a four-top event is the large number of b-jets coming from the weak decay of each top-quark. All of the dominant $t\bar{t}$ backgrounds are expected to have less b-jets per event, making the b-jet multiplicity N_b the most important signal to background discriminant. In order to exploit this fact, we will use in our analysis a high-efficiency operating point for the b-jet tagging algorithm and put a cut on the number of b-jets. This should be enough to raise the signal-to-background ratio considerably.

When focusing on the multi-lepton channels, besides having many b-jets, events are also expected to have a fair amount of hard light-quark and charm jets coming from the hadronic decay of top-quarks plus the expected additional soft jets from QCD radiation. For this reason, the total jet multiplicity N_j will also be a relevant variable in our search. In figure 5.4 (left) we plot N_j for the four-top sample in the SS dilepton channel. Here, the signal is characterized by a mean of $N_j = 7$ hard jets, while the two leading backgrounds $t\bar{t}W$ and $t\bar{t}Z$ have jet multiplicity distributions peaking at lower values around $N_j = 4$. Similar arguments hold for the trilepton channel, as can be seen in figure 5.4 (right), the signal jet multiplicity peaks around $N_j = 6$ also above the two leading backgrounds.

The remaining features characterizing the four-top signal in the multi-lepton channel are in part related to the total transverse energy of the process. Highly energetic events will have more boosted top-quarks giving rise to a very different signature when compared to events close to $t\bar{t}t\bar{t}$ production threshold. In order to get an idea of the distribution of events according to their transverse energy, we show the truth level p_T-distribution for top-quarks in figure 5.5. We see that a significant fraction of events with boosted top-quarks (with $p_T \gtrsim 300$ GeV) can be expected. In fact, we find that 51% of the events contain at least one top-quark with $p_T > 300$ GeV, and 28% (6%) of the events contain at least two (three) top quarks with $p_T > 300$ GeV.

Hadronic boosted top-quarks can be tagged using jet substructure techniques, see Ref. [135]. Unfortunately, at expected LHC luminosities and current top-tagging efficiencies, we find that even these sizeable boosted-top fractions are not sufficient

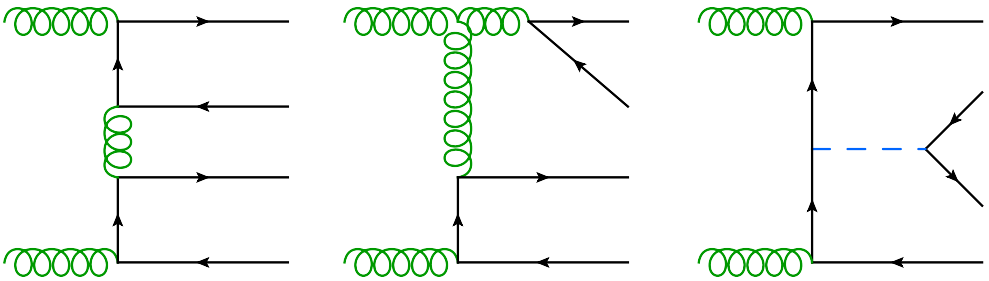


Figure 5.2: Gluon initiated representative LO diagrams contributing to $pp \to t\bar{t}t\bar{t}$ in the SM. The first two diagrams give leading contributions of order $\mathcal{O}(\alpha_S^4)$ while the last diagram includes a sub-leading contribution of order $\mathcal{O}(\alpha_S^2 y_t^4)$ with a Higgs boson exchange (blue dashed propagator).

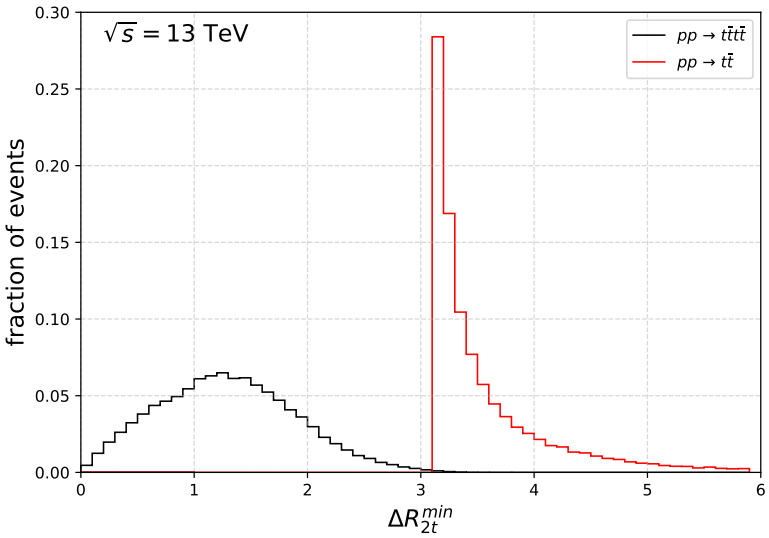


Figure 5.3: Minimum truth-level distance between any pair of top-quarks for both four-top production (black) and top-pair production (red).

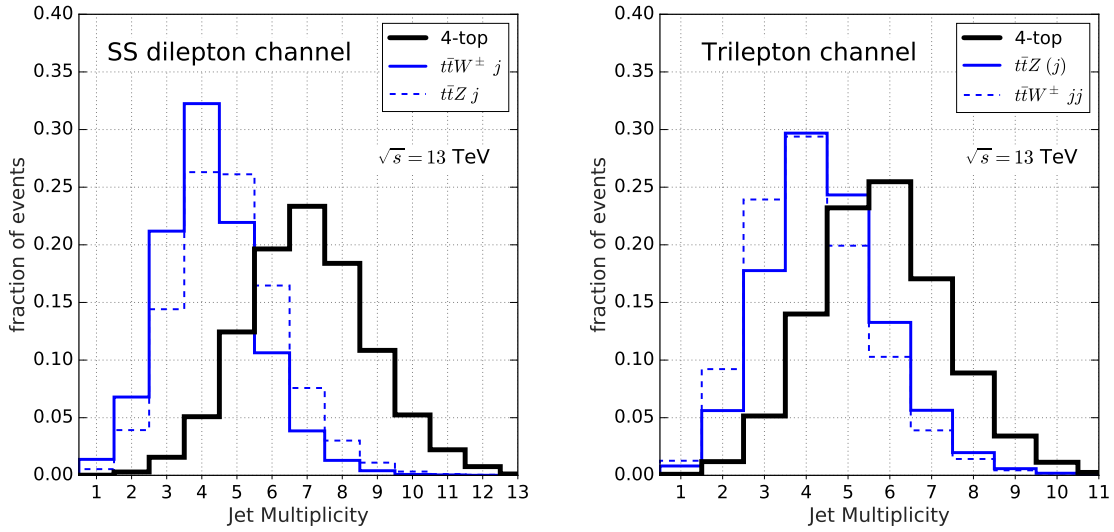


Figure 5.4: Distribution of the number of jets for the SS dilepton (left) and trilepton (right) channels for signal and expected backgrounds.

and we do not include top-tagging in a competitive four-top search based on the multi-lepton signature. Such techniques might, however, be of more relevance in the single-lepton or fully hadronic $t\bar{t}t\bar{t}$ signatures and/or for larger signal event samples expected from the high-luminosity LHC phase.

Another important consequence of boosted top-quarks is, however, the reduction of the signal efficiency in the multi-lepton channels. The sought lepton coming from a boosted top-quark decay will usually be collimated with the b-quark and end up close to the b-jet axis, eventually overlapping with the b-jet induced hadronic activity in the calorimeters. Such signal leptons will fail the standard lepton isolation

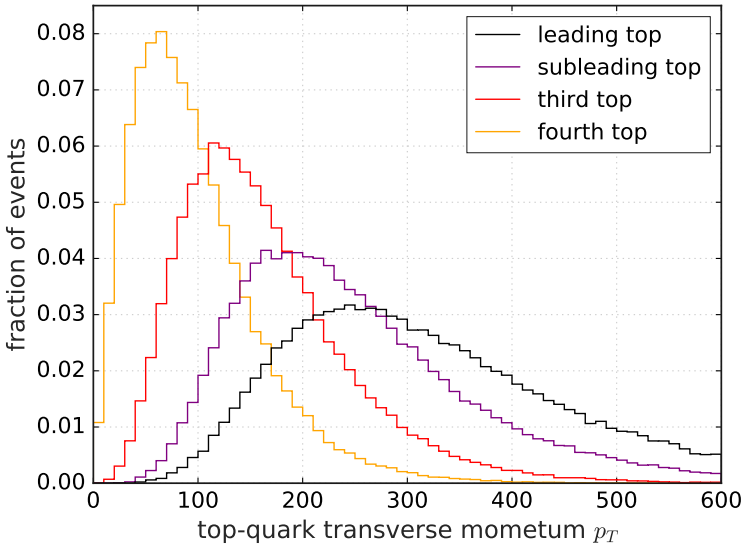


Figure 5.5: p_T-distributions of p_T-ordered top quarks in $t\bar{t}t\bar{t}$ production at 13TeV LHC.

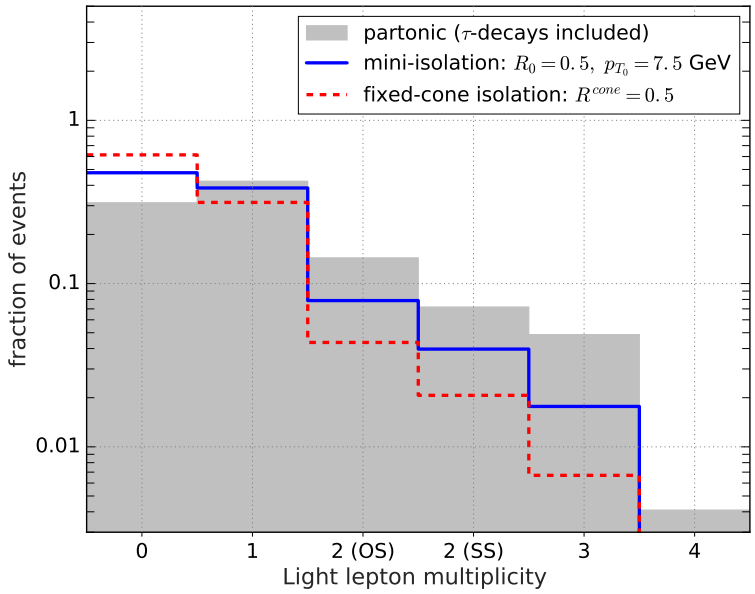


Figure 5.6: Four-top lepton multiplicity after taking into account detector effects and leptonic isolation criteria: mini-isolation (solid blue), traditional isolation criteria with cone radius of 0.5 (dashed red).

criteria. To bypass this issue as well as accidental lepton overlaps with other jets in the event, we use the mini-isolation technique, useful for identifying signal leptons close or even inside hadronic jets. The details of mini-isolation can be found in Ref. [94], however for the purposes of the following paragraphs we briefly describe it as follows. In contrast to the standard lepton isolation requirement based on a fixed isolation cone radius $R^{\text{cone}} \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2}$, the mini-isolation criterium adopted in

this work is based on defining a variable isolation cone radius

$$R^{\text{cone}}\left(p_{T\ell}\right) = \min\Big(R_0,\ \frac{p_{T_0}}{p_{T_\ell}}\Big)\ , \tag{5.1}$$

where R_0 and p_{T_0} are fixed parameters and p_{T_ℓ} is the transverse momentum of the candidate lepton.

In order to illustrate the impact of lepton isolation on the four-top multi-lepton channels, we show in figure 5.6 the lepton multiplicity distributions before and after imposing isolation requirements in absence of kinematic cuts (apart from lepton isolation). Parton shower and hadronization effects are included via `Pythia8` [36] and leptonic isolation cuts plus detector smearing are implemented with `Delphes3` [38]. As we can see in the figure, the fraction of events with more than one lepton in the final state at truth-level (grey histogram) is significantly reduced once the standard isolation requirement is imposed (dashed red line). As stated above, many signal leptons fail to pass the standard isolation because at least one of the top-quarks is typically boosted or because it accidentally ends up close to an unrelated jet. We have also plotted in figure 5.6 the performance of the mini-isolation requirement (solid blue line) for the parameters in Eq. (5.1) set to $R_0 = 0.5$ and $p_{T_0} = 7.5$ GeV. This shows that the signal efficiencies for both the SS dilepton and trilepton channels can improve by a factor of ~ 2 when compared to the standard isolation criteria.

As a last signal feature related to the presence of boosted top-quarks, we consider the truth-level distance between leptons and b-quarks $\Delta R_{\ell b}$ as a function of the top-quark boost. As explained above, the lepton coming from a boosted top-quark decay is expected to be close to the b-quark. This is not the case for an energetic lepton arising from the leptonic decay of a W boson in, e.g. the $t\bar{t}W^{\pm}$ background. For this reason functions of $\Delta R_{\ell b}$ could potentially work as good discriminants in a multi-lepton four-top search. In order to test this idea, we plot at parton level in figure 5.7 the maximum separation $\Delta R_{\ell b}^{\text{max}}$ between each lepton and the set of b-quarks against the lepton transverse momentum for a sample of $pp \to t\bar{t}t\bar{t}$ decaying in the SS dilepton channel (blue dots) and compare it to a $t\bar{t}W^{\pm}$ background sample (red dots). Notice the mild separation between signal and background in the high-p_T region, where signal events tend to cover the low $\Delta R_{\ell b}^{\text{max}}$ region while background events tend to cover larger $\Delta R_{\ell b}^{\text{max}}$ values. Unfortunately, the signal and background separation is not good enough and the expected number of events at LHC is too low, leading to a few-percent improvement of the signal significance for optimal choices of the cut on this variable. Nevertheless, at higher luminosities, boost-sensitive angular discriminants such as $\Delta R_{\ell b}$ could eventually be exploited to improve sensitivity.

5.3 Standard model four-top search strategy

The experimental program targeting the four-top signal at the LHC is currently in a preliminary stage. Both CMS and ATLAS collaborations have released a hand-full of dedicated searches for the SM four-top signal at the LHC. These are mainly based on the hadronic and mono-leptonic channels [128, 129, 130, 131]. Despite having much larger branching ratios than the multi-lepton decay modes, searches in these channels suffer from a tiny signal-to-background ratio due to overwhelming QCD and $t\bar{t}$ backgrounds. Better exclusion limits on the $t\bar{t}t\bar{t}$ cross-section have been obtained

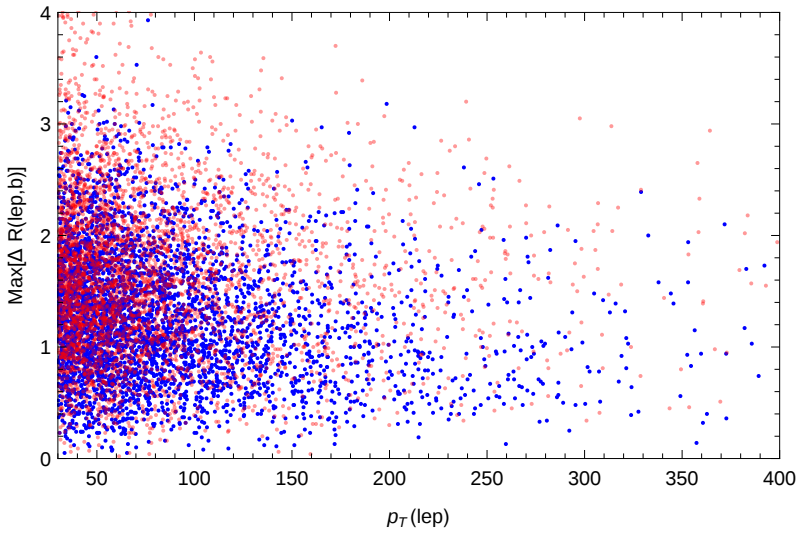


Figure 5.7: Distribution of signal (blue) versus background ($ttW^{\pm}b$ + jets, red) events for the maximum distance between a lepton and a b-quark as a function of the transverse momentum of the lepton.

in a set of SUSY searches based on the SS dilepton signature [1, 132, 133, 136]. To our knowledge, the best upper bound to this date has been recently presented in a 13 TeV CMS search [1] and reads $\sigma^{\rm SM}_{t\bar{t}t\bar{t}} < 57\,\text{fb}$. The signal regions constructed for these SUSY searches aim for NP signatures and as a consequence make use of strong cuts on kinematic variables such as the effective mass $m_{\rm eff}$ and missing transverse energy $\not{E}_T$. The same is true for the recently proposed (resonant) $t\bar{t}t\bar{t}$ search strategy in the SS dilepton channel [125] which otherwise shares several features with our proposal. Unfortunately, these cuts are not optimal for the SM four-top signature at lower luminosities where signal events are far from abundant. Another problem with existing searches is that the SS dilepton and trilepton channels are usually combined during event selection in the same signal regions (e.g. see Ref. [132]) and the differentiating features of the two signals with respect to their dominant backgrounds cannot be fully exploited.

In the following subsections, we describe our SM four-top search strategy in the multi-lepton channels. In order to maximize signal sensitivity, we consider separately the SS dilepton and trilepton channels and show that each is optimized in a different signal region. Interestingly, for integrated luminosities ranging approximately between $50 - 75$ fb^{-1}, the trilepton channel surpasses in sensitivity the SS dilepton channel. The two searches described below avoid hard cuts on kinematic variables given that the four-top signal is statistically limited and also given the lack of a clear signal-to-background separation in the differential distributions (evidenced in, e.g. figure 5.7). For this reason, both searches essentially consist of an optimization of jet and b-jet multiplicity selections. The main goal of this section is to demonstrate that by combining both dedicated multi-lepton searches, the LHC should achieve sensitivity to the SM four-top signal earlier than expected and possibly claim evidence of $t\bar{t}t\bar{t}$ production during Run-II.

5.3.1 Same-sign dilepton channel

At hadron colliders, events with SS dileptons are rarely produced in the SM. On the other hand, a SS dilepton signal is predicted in a variety of beyond the SM scenarios such as universal extra-dimension models, supersymmetry and left-right symmetric models, making this signature a promising place to discover new physics at the LHC. Among the rare SM processes giving rise to SS dileptons, $pp \to t\bar{t}t\bar{t}$ is one of the sub-dominant contributions. Consequently, searches for SM $t\bar{t}t\bar{t}$ production in the SS dilepton channel need to discriminate the signal from much larger backgrounds. We performed a detailed classification of all irreducible and reducible backgrounds in Appendix A.1. These fall into the following main categories: $t\bar{t}W$, $t\bar{t}Z$, $t\bar{t}H$ and "Others" for the irreducible backgrounds; in addition, jets faking leptons and lepton charge mis-measurement (Q-flip) constitute the dominant reducible backgrounds.

We generated Monte-Carlo samples for both the signal and backgrounds using `MadGraph` interfaced with `Pythia8` to account for hadronization and showering effects. Background samples with large jet multiplicities were generated with `AlpGen` [137], see Appendix A.1. The detector response was simulated with `Delphes3` in order to parametrically reproduce the LHC ATLAS detector.

Our search strategy then proceeds as follows. For b-jet reconstruction, we emulate a high operating-point b-tagging algorithm with a reconstruction efficiency around 75% for b-jets and rejection rates for charm and light jets of 8 and 400 respectively, see Ref. [138]. We isolate leptons using the mini-isolation requirement in Eq. (5.1) with cone parameters $R_0 = 0.2$ and $p_{T_0} = 8$ GeV for electrons, and $R_0 = 0.3$ and $p_{T_0} = 10$ GeV for muons. In order to deal with isolated non-prompt leptons arising from heavy meson decays inside jets, we reject any isolated leptons within a distance of $\Delta R_{\ell j} < 0.4$ from a reconstructed jet if the following condition is satisfied:

$$p_{T_\ell} < \alpha \, (\Delta R_{\ell j})^{\,\beta} \; p_{T_{\text{jet}}} \,, \tag{5.2}$$

where p_{T_ℓ} and $p_{T_{\text{jet}}}$ are the transverse momenta of the lepton and the jet, respectively. The parameters in Eq. (5.2) are fixed at $\alpha = 0.19$ for electrons, $\alpha = 0.17$ for muons and $\beta = -1$ for the angular exponent. A recent search by CMS [1] used a similar condition to reduce isolated non-prompt backgrounds. For more details see Appendix A.4. In addition, the following selection criteria are imposed on the reconstructed objects: only electrons with $p_T > 15$ GeV and $|\eta| < 2$[1], muons with $p_T > 10$ GeV and $|\eta| < 2.5$, and jets of any flavor with $p_T > 25$ GeV and $|\eta| < 2.5$ are retained.

Events are then required to contain:

- Exactly one SS dilepton (events with additional leptons are vetoed).
- Jet multiplicity (of any flavor) satisfying $N_j \geq 6$.
- b-jet multiplicity satisfying $N_b \geq 3$.

Finally, we bin the selected events into signal regions, denoted by $\mathbf{SR}\boldsymbol{n}\mathbf{j}$, and defined by different threshold values $n = 6, 7, 8, ...$ of the jet multiplicity $N_j \geq n$.

The results of such an analysis are shown in Table 5.1. The expected number of total events N_{exp} in each signal region as well as the expected number of events for the signal and each background category correspond to an integrated LHC luminosity of

[1] This specific cut in η removes large $e^{\pm}$ Q-flip background near the end-caps.

$\mathcal{L}$=300 fb^{-1}	**SR6j**	**SR7j**	**SR8j**
N_{exp}	139 (171)	85 (101)	43 (51)
$t\bar{t}t\bar{t}$	**16.7**	**13.5**	**8.9**
$t\bar{t}W$	60.7	35.0	17.1
$t\bar{t}Z$	32.1	20.3	10.7
$t\bar{t}h$	5.5	3.1	1.3
Fakes	12.5 (17.3)	7.1 (9.8)	3.3 (4.6)
Q-flip	7.6 (34.4)	3.7 (16.6)	1.6 (7.4)
Other	4.4	2.4	1.0
S/B	**0.14 (0.11)**	**0.19 (0.15)**	**0.26 (0.21)**
$\mathbf{S/\sqrt{B}}$	**1.51 (1.34)**	**1.60 (1.44)**	**1.53 (1.37)**

Table 5.1: Four-top SS dilepton channel event yields assuming 300 fb^{-1} at 13TeV LHC. Results are calculated using the benchmark $\epsilon_{\text{fake}} = 7.2 \times 10^{-5}$ and $\epsilon_{\text{Qflip}} = 2.2 \times 10^{-4}$, estimated in Appendix A.3. In the parenthesis results are calculated using the conservative benchmark $\epsilon_{\text{fake}} = 10^{-4}$ and $\epsilon_{\text{Qflip}} = 10^{-3}$. Here **SRnj** denotes the signal region with at least n tagged jets. The expected number of events is given by $N_{\text{exp}} = \text{Round}(S + B)$.

$\mathcal{L} = 300$ fb^{-1}. The reducible backgrounds have been estimated using two benchmark values for the fake lepton ($j \to \ell^{\pm}$) and Q-flip ($e^{\mp} \to e^{\pm}$) probabilities. The first benchmark defined by $\epsilon_{\text{fake}} = 7.2 \times 10^{-5}$ and $\epsilon_{\text{Qflip}} = 2.2 \times 10^{-4}$, is obtained by fitting MC simulations to existing 13 TeV results based on data-driven methods. All results are then calculated using these mis-identification probabilities. In addition, for illustration purposes, results inside the parenthesis are calculated using a set of more conservative benchmark values, $\epsilon_{\text{fake}} = 10^{-4}$ and $\epsilon_{\text{Qflip}} = 10^{-3}$. For a detailed discussion on both benchmark choices see Appendix A.3.

For the SS dilepton channel, the dominant background is the irreducible $t\bar{t}W$ background followed by $t\bar{t}Z$ and the Q-flip background. As expected, the signal-to-background ratio increases with the jet multiplicity and the signal sensitivity is maximized for **SR7j** but drops for **SR8j** due to limited statistics. Consequently, we find that using the projected LHC luminosity of 300 fb^{-1}, a four-top search in the SS dilepton channel is expected to yield a signal significance around $S/\sqrt{B} = 1.60$ with $S/B = 0.19$ if at least 7 jets are selected in the final state. We note that dedicated experimental analyses are expected to outperform our chosen fake lepton and Q-flip probability benchmarks. In the limit, where the associated backgrounds can be completely neglected, the projected signal significance at 300 fb^{-1} luminosity improves to $S/\sqrt{B} = 1.73$ with $S/B = 0.22$ in the most sensitive **SR7j** region.

5.3.2 Trilepton channel

The trilepton signature has been used in the past at the Tevatron [139, 140] and more recently at the LHC [111, 112] to search for SUSY and other beyond SM scenarios that predict much larger rates than the SM. Because it is more difficult to isolate three leptons, the four-top trilepton channel suffers from lower acceptance and efficiency when compared to the SS dilepton channel. This can be seen when comparing the bins $N_\ell = 3$ and $N_\ell = 2$ (SS) in figure 5.6, where the ratio of events between the trilepton and the SS dilepton channels is predicted to be approximately $1 : 2$. This implies that the four-top trilepton channel may achieve a comparable sensitivity to the SS dilepton channel if their backgrounds in a given signal region satisfy a hierarchy of at least $1 : 4$.

The backgrounds for the trilepton channel have been classified in Appendix A.2. In contrast to the SS dilepton signature, trileptons have the advantage of a lower instrumental background. The reason for this is (i) Q-flip is no longer a background, (ii) the SM processes producing fake trileptons have a lower rate when compared to those processes producing fake SS dileptons. This last point is evident if one keeps in mind that fake trileptons arise from mis-reconstructed $j \to \ell^\pm$ in $\ell^\pm\ell^\mp j$ while fake SS dileptons arise from mis-reconstruction of the much more abundant $\ell^\pm j$ final states.

We apply the same lepton mini-isolation and b-tagging algorithm as in the SS dilepton case, as well as the condition in Eq. (5.2) for rejecting non-prompt lepton backgrounds. We also use the same kinematic cuts for physical objects and select events according to the following criteria:

- Exactly three charged leptons (events with additional leptons are vetoed).
- Jet multiplicity (of any flavor) satisfying $N_j \geq 4$.
- b-jet multiplicity satisfying $N_b \geq 3$.
- A Z-mass window veto: the invariant mass $m_{\ell\ell}$ of all possible same-flavor OS dileptons $\ell^+\ell^-$ must fall outside the mass window 70 GeV $< m_{\ell\ell} <$ 105 GeV.

Selected events are then binned into signal regions **SR*n*j** with $n = 4, 5, 6, ...$, defined in the same way as in the previous section.

We present the results for the trilepton analysis in Table 5.2 for a projected LHC luminosity of 300 fb^{-1}. The Z-mass veto used in this analysis has a signal acceptance of approximately 90% and a very large rejection rate of the otherwise overwhelming $t\bar{t}Z$ background. This cut is thus instrumental in making the trilepton channel competitive with the SS dilepton channel despite smaller signal event rates. For the signal region **SR5j** we obtain a maximum significance of $S/\sqrt{B} = 1.87$ with a high signal-to-background ratio of $S/B = 0.45$. In case the jet to lepton fake-rate can be significantly reduced compared to our conservative benchmark value $\epsilon_{\text{fake}} = 10^{-4}$ (see Appendix A.3), the significance can be improved up to $S/\sqrt{B} = 1.97$ with a high signal-to-background ratio of $S/B = 0.50$.

5.3.3 Results

We are now in position to give results for the SM four-top search in the combined SS dilepton and trilepton channels. We implemented a statistical analysis based on

$\mathcal{L} = 300$ fb^{-1}	**SR4j**	**SR5j**	**SR6j**
N_{exp}	31 (32)	25 (26)	17 (17)
$\mathbf{t\bar{t}t\bar{t}}$	**8.6**	**7.8**	**6.0**
$t\bar{t}Z$	9.9	8.0	5.1
$t\bar{t}W$	6.7	4.9	2.9
$t\bar{t}h$	2.3	1.8	1.2
Fakes	2.5 (3.5)	1.7 (2.4)	0.9 (1.3)
Other	1.4	1.0	0.5
S/B	**0.38 (0.36)**	**0.45 (0.43)**	**0.57 (0.54)**
$\mathbf{S/\sqrt{B}}$	**1.80 (1.76)**	**1.87 (1.84)**	**1.84 (1.80)**

Table 5.2: Four-top trilepton channel event yields assuming 300 fb^{-1} at 13TeV LHC. Results are calculated using the benchmark value $\epsilon_{\text{fake}} = 7.2 \times 10^{-5}$ estimated in Appendix A.3. In the parenthesis results are calculated using the conservative benchmark value $\epsilon_{\text{fake}} = 10^{-4}$.

a log-likelihood test [141] using the most sensitive signal regions for each channel, i.e. **SR7j** for the SS dilepton channel and **SR5j** for the trilepton channel. Our analysis is based on the following main assumptions: (i) For the total number of observed events we use the sum of the expected signal and total background events. (ii) We only consider uncertainties for the dominant backgrounds, i.e. $\sim 12\%$ and $\sim 13\%$ of theoretical uncertainties (on the NLO cross-sections) for $t\bar{t}Z$ and $t\bar{t}W$ respectively [35] and a $\sim 50\%$ uncertainty fixed at $\mathcal{L}_{\text{int}} = 13.2$ fb^{-1} for the Fakes and Q-flip backgrounds extracted from Ref. [6]. For each uncertainty we assign an independent nuisance parameter following a gaussian prior. (iii) In order to project our results to arbitrary luminosities, the number of signal and background events are scaled with $\mathcal{L}_{\text{int}}$, while the statistical uncertainties for the Fakes and Q-flip backgrounds are scaled with $\sqrt{\mathcal{L}_{\text{int}}}$. (iv) Estimated backgrounds for the Fake leptons and Q-flip are based on the mis-identification probabilities $\epsilon_{\text{fake}} = 7.2 \times 10^{-5}$ and $\epsilon_{\text{Qflip}} = 2.2 \times 10^{-4}$ estimated in Appendix A.3.

We now present *evidence* and *discovery* luminosities for the SM four-top signal based on our search strategy, i.e. predict when the LHC should exclude the background-only hypothesis at 3σ and 5σ respectively. We find that the LHC experiments should be able to establish evidence (p-value of 3×10^{-3}) for the SM four-top process at a combined integrated luminosity of $\mathcal{L}_{\text{int}} \simeq 215$ fb^{-1}, and claim discovery (p-value of 3×10^{-7}) at a higher integrated luminosity of $\mathcal{L}_{\text{int}} \simeq 1060$ fb^{-1}. These results show that with our search strategy, the LHC starts becoming sensitive to four-top production in the SM at luminosities achievable within Run-II.

As already alluded to above, these projections are very sensitive to the Fake and Q-flip background estimations, which are based on a particular set of values for the lepton-faking and Q-flipping probabilities. To illustrate this point, we plot in figure 5.8 (top) iso-contours of both 3σ evidence (shaded contours) and 5σ dis-

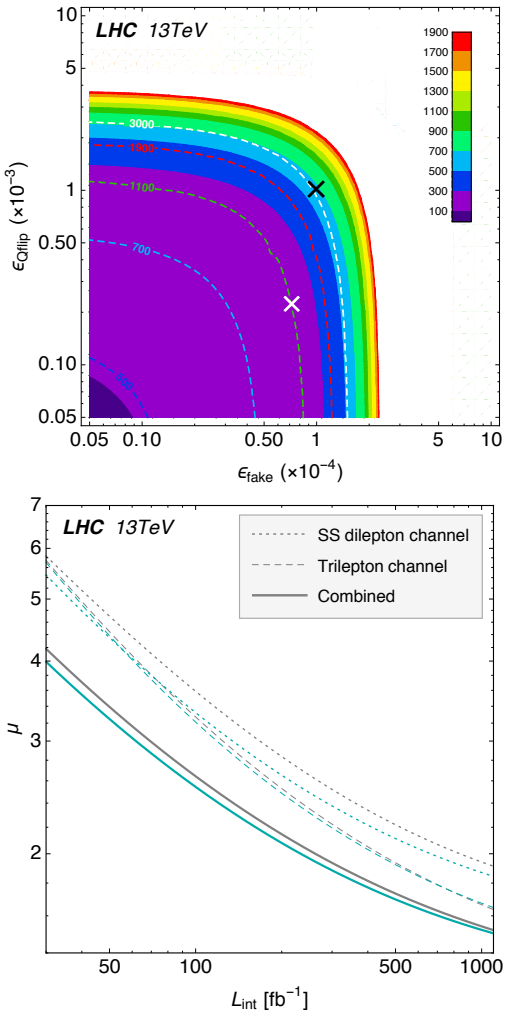


Figure 5.8: (top panel) Iso-contours of 3σ evidence luminosities (fb^{-1}) (shaded contours) and 5σ discovery luminosities (fb^{-1}) (dashed lines) in the $(\epsilon_{\text{fake}}, \epsilon_{\text{Qflip}})$-plane of lepton-faking and Q-flip probabilities. The white cross corresponds to the estimated values $(7.2 \times 10^{-5}, 2.2 \times 10^{-4})$ from Appendix A.3, while the black cross corresponds to the conservative benchmark values $(10^{-4}, 10^{-3})$. (Bottom panel) Projected ATLAS 95% CL exclusion limits for the SM four-top signal strength as a function of the integrated luminosity for the SS dilepton channel (blue dotted line), the trilepton channel (blue dashed line) and the combination (blue solid line). We also include the results obtained from reducible backgrounds estimated using the conservative benchmark (gray lines).

covery (dashed lines) luminosities in the $(\epsilon_{\text{fake}}, \epsilon_{\text{Qflip}})$-plane. The misidentification probabilities used for our results is marked with a white cross, while the much more conservative benchmark $\epsilon_{\text{fake}} = 10^{-4}$ and $\epsilon_{\text{Qflip}} = 10^{-3}$ is marked with a black cross. For example, improving upon our estimates for $\epsilon_{\text{fake,Qflip}}$ by an order of magnitude would make the associated backgrounds completely negligible. In that limit we find that the LHC experiments should be able to establish 3σ evidence for SM $t\bar{t}t\bar{t}$ production already at an integrated luminosity of $\mathcal{L}_{\text{int}} \simeq 98$ fb^{-1}, and claim discovery at $\mathcal{L}_{\text{int}} \simeq 420$ fb^{-1}.

The extracted upper limit from our search for the SM four-top signal strength is

$$\mu_{t\bar{t}t\bar{t}}^{\mathrm{SM}} \;\leq\; 1.87 \quad \text{at } 95\% \text{ CL}, \tag{5.3}$$

for an integrated luminosity of $\mathcal{L}_{\mathrm{int}} = 300$ fb^{-1}. In figure 5.8 (bottom) we give the 95% CL exclusion limits on the signal strength $\mu_{t\bar{t}t\bar{t}}^{\mathrm{SM}}$ for each individual multi-lepton channel (blue dotted and dashed lines) and for both combined (blue solid line) as a function of the integrated luminosity. We also present results (gray lines) obtained using the benchmark $\epsilon_{\mathrm{fake}} = 10^{-4}$ and $\epsilon_{\mathrm{Qflip}} = 10^{-3}$ to estimate the reducible backgrounds. Notice that the extracted upper bound for the signal strength in the combined analysis does not change considerably when using this more conservative benchmark. As expected, for higher luminosities the trilepton channel achieves better sensitivity than the SS dilepton channel and drives the search.

5.4 Application to New Physics

Many NP models addressing the SM hierarchy or flavor puzzles predict new TeV scale dynamics coupling most strongly to the third generation, in particular, the top quark (see e.g. Refs. [114, 142]). Such interactions can most naturally be searched for through multi-top-quark production. Several existing proposals target pair- or $t\bar{t}$-associated production of heavy resonances decaying to top quark pairs [116, 117, 118, 122, 123, 124, 125, 126]. In this case, the dominant signature is the appearance of resonances in $t\bar{t}$ invariant mass spectra. In all cases, for multi-TeV resonance masses, boosted top searches can be effective in these scenarios [127].

On the other hand, color-neutral particles coupling predominantly to the third generation with masses below the $t\bar{t}$ threshold are at present only weakly constrained [143]. Such states appear in models addressing recent B-meson decay anomalies [144, 145] or in scenarios of cosmological thermal relic dark matter (see e.g. Ref. [146, 147]). The exchange of such particles mediating four-top production would generically result in kinematics, not strikingly different from the dominant QCD contributions. Their dominant effect is thus expected to be a modification of the inclusive four-top production cross-section. We study this possibility in more detail using two representative toy model examples in which we extend the SM with respectively a new vector or scalar boson affecting the four-top production at the tree-level: (1) a top-philic neutral Z' vector boson and (2) a neutral scalar ϕ with Yukawa couplings to the top. In the first model we assume for simplicity that the Z' with mass $m_{Z'}$ only couples significantly to right-handed top quarks[2]. The relevant interaction Lagrangian then reads

$$\mathcal{L}_{Z'} \;=\; -\, g_{tZ'}\, \bar{t}_R \not{Z}' t_R \,. \tag{5.4}$$

We note that the chiral top-current, to which Z' is coupled is broken explicitly by the top quark mass (and by anomalies) and the $m_{Z'} \to 0$ limit cannot be approached trivially in this model. Nonetheless, well-defined UV completions exist in the literature where these issues are properly addressed with no immediate consequences for $t\bar{t}t\bar{t}$ phenomenology (see e.g. Refs. [148, 149, 150] for an explicit example as well as

[2]This same interaction was studied in Ref. [125], but in a different parameter region.

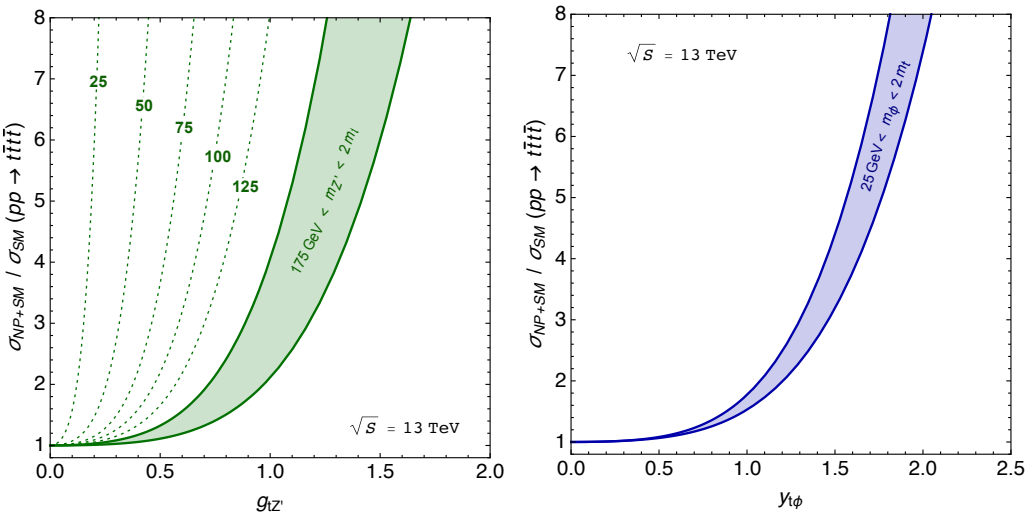


Figure 5.9: Predictions for the deviation $\sigma_{NP+SM}/\sigma_{SM}$ in the $pp \to t\bar{t}t\bar{t}$ cross-section at $\sqrt{s} = 13$ TeV within the simplified NP Z' (left-hand side) and ϕ (right-hand side) models as a function of the couplings $g_{tZ'}$ and $y_{t\phi}$, for different Z' and ϕ masses, respectively.

Ref. [151] for a more general discussion). In the second model, the relevant $\phi - t$ interactions are on the other hand described by

$$\mathcal{L}_\phi \;=\; -\,y_{t\phi}\,\bar{t}_L \phi t_R + \text{h.c.}\,. \tag{5.5}$$

Depending on the phase of $y_{t\phi}$, this interaction is in general CP violating. While the form of interactions above is not manifestly invariant under the SM EW gauge symmetry, suitable UV completions in terms of multiple Higgs doublet or singlet SM extensions can be easily constructed where the dominant effects in $t\bar{t}t\bar{t}$ phenomenology are captured by the effective Lagrangian in Eq. (5.5) (see e.g. Ref. [120]). By choosing $m_\phi = m_h = 125$ GeV, this second example also covers the interesting case of a modified top Yukawa coupling of the SM Higgs boson.

In both models, since the new degrees of freedom are never produced on-shell, their effects in four-top production are largely independent of their possible other interactions. [3] In our case, contributions can be parametrized in terms of the mediator mass and the relevant coupling to top quarks, $(m_{Z'}, g_{tZ'})$ for model (1) and $(m_\phi, y_{t\phi})$ for model (2), respectively, and in particular do not depend on the mediator decay width. This is in contrast to direct $t\bar{t}$ resonance searches, where the resonance width can play an important role (see e.g. [126, 152, 153, 154]).

In figure 5.9 we show the predicted four-top production cross-section including NP contributions $\sigma_{NP+SM}(pp \to t\bar{t}t\bar{t})$ in both models, normalized to the SM cross-section prediction $\sigma_{SM}(pp \to t\bar{t}t\bar{t})$, all computed at LO in QCD. We find that for both NP models, the off-shell mediator contributions to $pp \to t\bar{t}t\bar{t}$ can considerably enhance the four-top production at the LHC. For the top-philic Z' model (left panel), the enhancement with respect to the SM cross-section becomes almost independent of the Z' mass in the range $m_t \lesssim m_{Z'} \lesssim 2m_t$ being roughly a factor of

[3] This should be compared to Ref. [125], where the new particle is assumed to be heavier than $2m_t$, appearing as an on-shell resonance that decays to $t\bar{t}$, and consequently motivating and resulting in a somewhat different search strategy.

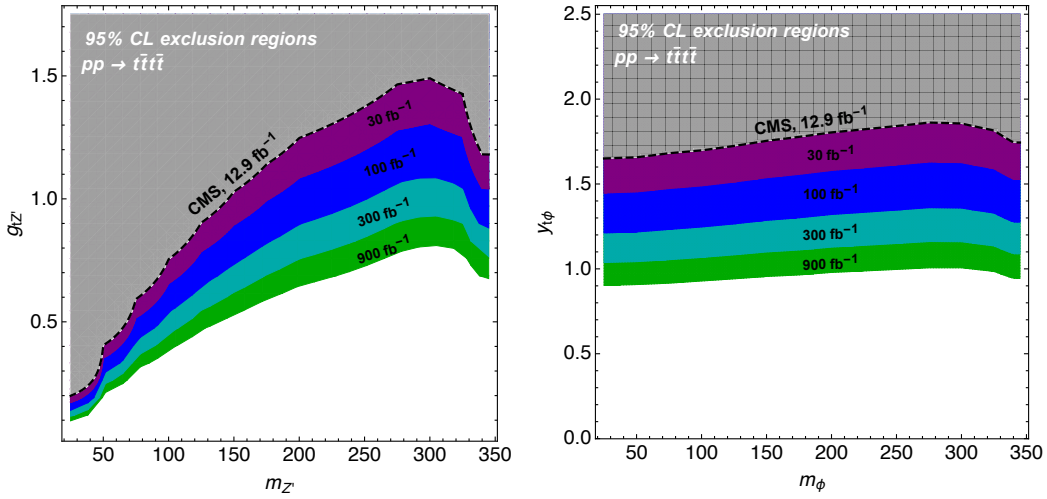


Figure 5.10: NP exclusion regions for LHC luminosities of 30fb^{-1} (Purple), 100 fb^{-1} (Blue), 300 fb^{-1} (Dark Cyan) and 900 fb^{-1} (Green) respectively, for the Z' model (Left) and the ϕ scalar model (Right). The existing bound extracted from the recent CMS search [1] is shown in gray shade and bounded by a dashed contour.

two for couplings satisfying $0.5 \lesssim g_{tZ'} \lesssim 1$. On the other hand, the strong Z' mass dependence for $m_{Z'} \ll m_t$ can be easily understood since the Z' couples to an unconserved current and the scattering amplitudes receive contributions proportional to its breaking due to m_t, and thus grow with $m_t/m_{Z'}$. For the neutral scalar model (right panel) the enhancement in the cross-section is practically mass-independent in the whole considered ϕ mass range. Here, one can expect an enhancement of order $\sigma_{NP+SM}/\sigma_{SM} \sim 2$ for couplings of order $y_{t\phi} \sim 1$. In both models, for mediator masses close to $t\bar{t}$ threshold, interference effects make the mass dependence non-monotonous. We note that the sign of these interference effects is fixed irrespective of the sign (or phase) of $g_{tZ'}(y_{t\phi})$, since the NP-mediated amplitudes are proportional to $g^2_{tZ'}(|y_{t\phi}|^2)$ in both models, respectively. Finally, resonance width effects (not included in our analysis) are expected to become relevant in the region a few GeV below threshold.

In the mass region of interest, the main effect of the considered NP contributions is to rescale the four-top cross-section while leaving the main kinematic distributions SM-like. Indeed, we have verified that the acceptances and efficiencies of the NP contributions for both proposed multi-lepton four-top search strategies are approximately constant in all of the NP parameter space and comparable to the SM signal. We are therefore able to apply the upper limits on the SM four-top production extracted from our search in the previous section directly. The resulting 95% CL exclusion regions in the NP parameter space (mass vs. coupling) for both simplified models are shown in figure 5.10. For comparison, we also show the corresponding bounds extracted from a recent CMS search [1]. Finally we note that, while we do not project our constraints to larger mediator masses, where the NP-mediated $t\bar{t}t\bar{t}$ production is resonantly enhanced, we expect our search strategy to retain good sensitivity also in this region of the NP models' parameter space, as long as the resonant $t\bar{t}$ pairs are not significantly boosted ($m_{Z',\phi} \not\gg 2m_t$).

5.5 UV complete model: top-philic U(1)′

Now that we have demonstrated the sensitivity of our $t\bar{t}t\bar{t}$ search strategy to new (colorless) top-philic degrees of freedom, we provide in this last section a sketch of one possible UV complete model realizing the simplified top-philic models discussed above. These models happen to be transparent to most LHC and low-energy observables except for four-top production.

One of the simplest extensions of the SM is to add a new abelian group U(1)′:

$$\mathcal{G} \;=\; \mathrm{SU}(3)_c \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y \otimes \mathrm{U}(1)' \tag{5.6}$$

The new group gives rise to a Z' gauge boson that couples to (charged) fermionic states with a gauge coupling that we denote as g_T. To give mass to the Z' boson we introduce a complex scalar field Φ that acquires a VEV v' and breaks spontaneously the U(1)′ group, i.e. $\mathcal{G} \to \mathcal{G}_{\rm SM}$. The scalar is taken as a SM gauge singlet with charge assignment $\Phi \sim (\mathbf{1}, \mathbf{1}, 0, Q'_\phi)$ under (5.6) with primed charge Q'_ϕ to be fixed later on. In order to obtain top-philic interactions, we could charge the RH top quarks under the U(1)′ group and leave the rest of the SM matter content uncharged. The problem with assigning flavor non-universal U(1)′ charges to the SM fermions is that it automatically forbids a renormalizable Yukawa interaction between t_R, the Higgs doublet and the quark doublet $Q^3 = (t_L, b_L)^T$, thus affecting the well established flavor structure of masses and mixings in the SM. One solution [148, 149, 155, 156, 157] is to leave all SM fields, including t_R, uncharged and introduce a new *hidden* matter sector with one vector-like fermion with identical representation as t_R with respect to $\mathcal{G}_{\rm SM}$ and a non-vanishing U(1)′. This scenario leaves unspoiled the SM Yukawa sector but forbids a direct coupling of the Z' boson to the t_R state. Interestingly, an effective coupling between the Z' and the physical RH top-quark can nonetheless be induced in the broken phase $\mathrm{SU}(3)_c \otimes \mathrm{U}(1)_{\rm em}$. This is possible because of an inherent misalignment between the interaction and the mass basis that leads to fermion mixing between the U(1)′–charged vector-like fermion and the uncharged top-quark state. As a result, when rotating back to the physical fields two eigenmass states arise, one light quark identified with the SM RH top-quark and one heavy quark -a heavy *top-partner*-, both with non-vanishing couplings to the Z'.

5.5.1 The model

To illustrate this we introduce a top-partner $\mathcal{T}$ with both chiralites [4] with charge assignment $\mathcal{T}_{L,R} \sim (\mathbf{3}, \mathbf{1}, +\frac{2}{3}, Q'_{\mathcal{T}})$ under (5.6) and $Q'_{\mathcal{T}} \neq 0$. The rest of matter content (prior to symmetry breaking) remains uncharged under the new abelian gauge group. The fermionic Lagrangian for the quark sector is given by three parts:

$$\mathcal{L} \;\supset\; \mathcal{L}_{\rm light} + \mathcal{L}_{\rm heavy} + \mathcal{L}_{\rm mix} \tag{5.7}$$

where

$$\mathcal{L}_{\rm light} \;=\; \overline{Q}^i (i\not{D}) Q^i + \overline{u}^i (i\not{D}) u^i + \overline{d}^i (i\not{D}) d^i - \Big(y^d_{ij}\, \overline{Q}^i \varphi d^j + y^u_{ij}\, \overline{Q}^i \tilde{\varphi} u^j + \text{h.c.} \Big) \tag{5.8}$$

$$\mathcal{L}_{\rm heavy} \;=\; \overline{\mathcal{T}} (i\not{D}) \mathcal{T} \;-\; m_{\mathcal{T}} \overline{\mathcal{T}} \mathcal{T} \tag{5.9}$$

$$\mathcal{L}_{\rm mix} \;=\; -\,\lambda^i_{\mathcal{T}}\, \overline{\mathcal{T}}_L \Phi\, u^i_R \;+\; \text{h.c.} \tag{5.10}$$

[4] We introduce both chiralites in order not to make the theory anomalous

fields	$SU(3)_c$	$SU(2)_L$	$U(1)_Y$	$U(1)'$
Q_L^i	**3**	**2**	1/6	0
L_L^i	**1**	**2**	−1/2	0
u_R^i	**3**	**1**	2/3	0
d_R^i	**3**	**1**	−1/3	0
e_R^i	**1**	**1**	−1	0
φ	**1**	**2**	1/2	0
$\mathcal{T}_{L,R}$	**3**	**1**	2/3	+1
Φ	**1**	**1**	0	+1

Table 5.3: Matter and scalar field content of the U(1)′ model.

Here $Q_L^i = (u_L^i, d_L^i)^T$ and u_R^i are the LH quark doublets and RH up-quark singlets of the SM with family index $i = 1, 2, 3$, and φ is the SM scalar doublet and $\tilde{\varphi} \equiv i\tau_2\varphi$. y^u is the up-type Yukawa matrix and $\lambda_{\mathcal{T}}^i$ the Yukawa coupling bvetween the SM up-type quarks and the vector-like fermion $\mathcal{T}$, and $m_{\mathcal{T}}$ is the Dirac mass. In order to mix heavy and light degrees of freedom via (5.10) it is necessary to fix the U(1)′ charges of the vector-like fermion and the scalar singlet to satisfy $Q'_{\mathcal{T}} = Q'_{\Phi} \neq 0$. In Table. 5.3 we summarize the gauge representations and charge assignments of the fields in the U(1)′ model where we have fixed the "free" charge assignment of the scalar singlet to $Q'_{\Phi} = +1$.

In order to complete the UV model we need to write down the scalar Lagrangian and specify the most general renormalizable scalar potential consistent with gauge symmetry:

$$\mathcal{L}_{\Phi} \;=\; |D_{\mu}\Phi|^2 \;-\; V(\varphi, \Phi) \tag{5.11}$$

where $V(\varphi, \Phi) = V_{\varphi} + V_{\Phi} + V_{\text{portal}}$ such that

$$V_{\varphi} \;=\; \mu_{\varphi}^2(\varphi^{\dagger}\varphi) \;-\; \lambda_{\varphi}(\varphi^{\dagger}\varphi)^2 \tag{5.12}$$

$$V_{\Phi} \;=\; \mu_{\Phi}^2|\Phi|^2 \;-\; \lambda_{\phi}|\Phi|^4 \tag{5.13}$$

$$V_{\text{portal}} \;=\; \lambda_{\varphi\phi}\,(\varphi^{\dagger}\varphi)|\Phi|^2 \tag{5.14}$$

It can be shown that the parameters of the potential, assumed all to be real, can be chosen such that the potential V is bounded from below and that the structure of the vacuum has the following form:

$$\langle\,\varphi\,\rangle \;=\; \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix}, \qquad \langle\,\Phi\,\rangle \;=\; \frac{v'}{\sqrt{2}}\,. \tag{5.15}$$

where v' is identified with the SM VEV $v \approx 246$ GeV. From the kinetic term of the scalar Lagrangian $\mathcal{L}_{\Phi}$, the new gauge boson Z' acquires a mass via spontaneous symmetry breaking given by $M_{Z'} = v' g_T$. The new scalar will have the representation $\Phi = (v' + \phi + ia)/\sqrt{2}$, where the radial mode ϕ is a propagating degree of freedom and a is the would-be Goldstone boson absorbed by the longitudinal component of Z'. Notice that the Higgs h and the complex scalar ϕ will actually mix for non-zero portal coupling $\lambda_{\varphi\phi} \neq 0$. For more details see [158, 159]. We will assume here that the mixings between both states can be safely neglected.

5.5.2 Fermion mixing

Once the spontaneous symmetry breaking chain $\mathcal{G}(1) \to \mathcal{G}_{\rm SM} \to {\rm SU}(3)_c \otimes {\rm U}(1)_{\rm em}$ takes place, the resulting fermion mass matrix for up-quarks will have the structure:

$$-\mathcal{L}_{\rm mass} \;=\; \left(\bar{u}'_L\,,\bar{\mathcal{T}}'_L\right)\begin{pmatrix} m_u & 0 \\ \mu^\dagger & m_{\mathcal{T}} \end{pmatrix}\begin{pmatrix} u'_R \\ \mathcal{T}'_R \end{pmatrix} \;+\; \text{h.c.}\,, \tag{5.16}$$

where $m_u^{ij} \equiv \frac{v}{\sqrt{2}}\, y_u^{ij}$ and $\mu^i \equiv \frac{v'}{\sqrt{2}}\,\lambda_{\mathcal{T}}^i$ are the entries of 3×3 and 3×1 mass matrices, respectively, and the "primed" fields indicate the interaction basis. The full mass matrix in (5.16) can be diagonalized via 4×4 bi-unitary transformation matrices $\mathcal{V}_{L,R}$. As shown in Appendix C, these mixing matrices can be conveniently parametrized on general grounds using the block form given in Eq. C.18:

$$\mathcal{V}_X \;=\; \left(\begin{array}{c|c} \left[\mathbb{1}_{3\times 3} \;-\; \frac{\Theta_X(\Theta_X)^\dagger}{1+\cos\theta_X}\right] V_X & \Theta_X \\ \hline -(\Theta_X)^\dagger\, V_X & \cos\theta_X \end{array}\right) \tag{5.17}$$

where $X = \{L, R\}$ is a chiral index, V_X are 3×3 unitary matrices and Θ_X are light–heavy 3×1 mixing matrices. The angle θ_X is a function of the parameters Θ_X^i, defined in Eq. (C.10) of Appendix C. One important feature of the model is that the LH light–heavy mixing parameters Θ_L are of order $\mathcal{O}(m_t/M_{\mathcal{T}})$. These are naturally suppressed for vector-like fermion masses satisfying $M_{\mathcal{T}} \gg v, v'$. We derive this in the appendix (sec. C.2).

5.5.3 Induced Z' interactions

We now take a look at the interactions arising between the gauge boson Z' and fermions. These interactions turn out to be chiral in the mass eigenbasis because of the large inequality between LH mixings (suppressed) and RH mixings (unsuppressed). The Lagrangian for the Z' boson in the mass eigenbasis reads

$$\mathcal{L}_{Z'} \;=\; g_T\, Z'_\mu \begin{pmatrix} \bar{u}_X & \bar{\mathcal{T}}_X \end{pmatrix} \gamma^\mu \cdot K_X \cdot \begin{pmatrix} u_X \\ \mathcal{T}_X \end{pmatrix}. \tag{5.18}$$

where $K_{L,R}$ are 4×4 Hermitian mixing matrices defined by

$$K_X \;=\; \left(\begin{array}{c|c} K_X^{(u)} & K_X^{(u\mathcal{T})} \\ \hline K_X^{(u\mathcal{T})\dagger} & K_X^{(\mathcal{T})} \end{array}\right) \;=\; \mathcal{V}_X \begin{pmatrix} \mathbb{0} & \mathbb{0} \\ \mathbb{0} & 1 \end{pmatrix} \mathcal{V}_X^\dagger\,. \tag{5.19}$$

The blocks $K_X^{(u)}$, $K_X^{(u\mathcal{T})}$ and $K_X^{(\mathcal{T})}$ are a 3×3, 3×1, 1×1 matrices describing the interaction of the Z' boson with light–light, light–heavy and heavy-heavy chiral currents, respectively. A direct computation using Eq. (5.17) yields the general expression for the Z' couplings to RH fields

$$K_R^{(u)} \;=\; \Theta_R(\Theta_R)^\dagger \tag{5.20}$$

$$K_R^{(u\mathcal{T})} \;=\; -(\cos\theta_R)\,\Theta_R\,, \tag{5.21}$$

$$K_R^{(\mathcal{T})} \;=\; \cos^2\!\theta_R\,. \tag{5.22}$$

A similar expression for the couplings to LH fields can be obtained by replacing $R \leftrightarrow L$ in the equations above. Notice that the Z' couplings to the SM RH up-quarks, via $K_R^{(u)}$, are actually flavor-changing for generic mixings, potentially leading to flavor violating transitions of the form $Z' \to u^i_{L,R}\bar{u}^j_{L,R}$.

We now use the solutions (C.25)–(C.24) to write down the general expressions for the Z' coupling to LH fields:

$$K_L^{(u)} = \left(\frac{m_t}{M_{\mathcal{T}}\cos\theta_R}\right)^2 \left[\frac{\Delta\hat{m}\,\Theta_R(\Theta_R)^\dagger\Delta\hat{m}}{1+(\frac{m_t}{M_{\mathcal{T}}\cos\theta_R})^2\,(\Theta_R)^\dagger\Delta\hat{m}^2\Theta_R}\right] \tag{5.23}$$

$$K_L^{(u\mathcal{T})} = -\left(\frac{m_t}{M_{\mathcal{T}}\cos\theta_R}\right) \left[\frac{\Delta\hat{m}\,\Theta_R}{1+(\frac{m_t}{M_{\mathcal{T}}\cos\theta_R})^2\,(\Theta_R)^\dagger\Delta\hat{m}^2\Theta_R}\right] \tag{5.24}$$

$$K_L^{(\mathcal{T})} = \frac{1}{1+(\frac{m_t}{M_{\mathcal{T}}\cos\theta_R})^2\,(\Theta_R)^\dagger\Delta\hat{m}^2\Theta_R}\,. \tag{5.25}$$

where $\Delta\hat{m} \equiv \Delta m/m_t = \mathrm{diag}(m_u/m_t, m_c/m_t, 1)$ is a diagonal mass hierarchy matrix. As for the RH case, the Z' coupling to LH up-quark fields are in general not flavor diagonal. Notice that these couplings are nonetheless highly suppressed for $M_{\mathcal{T}} \gg m_t$ and generic $\mathcal{O}(1)$ RH mixings, where $M_{\mathcal{T}}$ is the physical mass of the heavy top-partner.

Similar expressions can be derived for the SM electroweak gauge bosons and the Higgs boson. These are provided in appendix D. One important peculiarity of the U(1)′ model is the appearance of flavor violating transitions at tree level in the Z and in Higgs interactions with up-quarks as well as violation of CKM unitarity in W boson interactions (see e.g. Ref. [160]). One nice feature is that all of the deviations patterns from the SM can be traced back to one common origin. Indeed, as shown in appendix D, the structures of the Z boson interactions with LH up-quarks parametrized by the matrix $X_L^{(u)}$, the CKM mixing matrix V_{CKM} and the Higgs boson couplings to up-quarks parametrized by $Y^{(u)}$ are predicted to satisfy the relation:

$$X_L^{(u)} \;=\; V_{\mathrm{CKM}}(V_{\mathrm{CKM}})^\dagger \;=\; \frac{v}{\sqrt{2}}\,Y^{(u)}\,(\Delta m_u)^{-1} \;=\; \mathbb{1} \;-\; K_L^{(u)} \tag{5.26}$$

Notice that in the SM each of the 3×3 matrices in the first three equalities in (5.26) are all equal to $\mathbb{1}$. In the last equality one can then interpret the departures from SM behaviour as originating from the appearance (via mixing) of new interactions of order $\mathcal{O}(m_t/M_{\mathcal{T}})^2$ between the (LH) up-quarks and the Z' boson.

5.5.4 Top-philia: t–$\mathcal{T}$ mixing

Up to now we have described a general scenario where all three up-quarks mix with the vector-like fermion $\mathcal{T}$. Interestingly, it turns out that the LH interactions in the U(1)′ inherit the highly hierarchical from the SM. Indeed, the flavor structure governing the deviations from the SM interactions, encoded in the matrix $K_L^{(u)}$ in

(5.23), is determined by the flavor pattern in the 3×3 matrix

$$\Delta\hat{m}\,\Theta_R(\Theta_R)^\dagger\Delta\hat{m} \;=\; \begin{pmatrix} \frac{m_u^2}{m_t^2}\,|\Theta_1|^2 & \frac{m_u m_c}{m_t^2}\,\Theta_1\Theta_2^* & \frac{m_u}{m_t}\,\Theta_1\Theta_3^* \\ \frac{m_c m_u}{m_t^2}\,\Theta_2\Theta_1^* & \frac{m_c^2}{m_t^2}\,|\Theta_2|^2 & \frac{m_c}{m_t}\,\Theta_2\Theta_3^* \\ \frac{m_u}{m_t}\,\Theta_3\Theta_1^* & \frac{m_c}{m_t}\,\Theta_3\Theta_2^* & |\Theta_3|^2 \end{pmatrix}, \tag{5.27}$$

where $\Theta_R = (\Theta_1, \Theta_2, \Theta_3)^T$. The large masses ratios in the up-quarks sector makes this texture highly hierarchical. The largest potential deviations are in the (33) entry and are at most of order $(m_t/M_{\mathcal{T}})^2$ with sub-leading effects in the (23) and (32) entries with an extra suppression factor of order Cabibbo cube $\lambda_c^3 \sim m_c/m_t$. From this we see that the NP effects in the LH sector, besides being highly suppressed, are pretty much top-philic.

On the other hand, the NP flavor structure in the RH currents is quite different. The couplings of light fermions to the $W^\pm$ and Z bosons are unaffected by NP and are as in the SM. The only new interactions for RH up-quarks arise via mixing with the Z' boson. These are governed by the light–light matrix $K_R^{(u)} = \Theta_R(\Theta_R)^\dagger$ which in principle can be large and anarchichal giving rise to sizeable flavor changing transitions between different RH up-quark types. In order to obtain a top-philic model in the RH sector, additional flavor assumptions must be introduced. For this we mimic the flavor structure of the SM in the quark sector and assume that the Yukawa interactions (5.10) are hierarchical and exhibit third generation dominance, i.e. $\lambda_{\mathcal{T}}^3 \gg \lambda_{\mathcal{T}}^1,\ \lambda_{\mathcal{T}}^2$. As a consequence, light–heavy RH mixings between fermions will only be relevant between the top-quark and $\mathcal{T}$. We assume that the Yukawa column matrix has the following minimal structure

$$\lambda_{\mathcal{T}} = \begin{pmatrix} 0 \\ 0 \\ \lambda_t \end{pmatrix}, \tag{5.28}$$

where $\lambda_t \sim \mathcal{O}(1)$. In this case the two 4×4 mixing matrices $\mathcal{V}_{L,R}$ each collapse to 2×2 unitary matrices parametrized by angles $\theta_{L,R}$ describing two-fermion mixing t–$\mathcal{T}$. With this in mind, one can make the replacements $\Theta_R \to (0, 0, \sin\theta_R)^T$ and $V_R^u \to \mathbb{1}$ in the general formulas. This yields for the non-vanishing entries

$$K_R^{(t)} \;=\; \sin^2\theta_R\,, \tag{5.29}$$

$$K_R^{(t\mathcal{T})} \;=\; -\frac{1}{2}\sin 2\theta_R\,, \tag{5.30}$$

$$K_R^{(\mathcal{T})} \;=\; \cos^2\theta_R\,, \tag{5.31}$$

for expressions (5.20)–(5.22), and

$$K_L^{(t)} = \left(\frac{m_t}{M_{\mathcal{T}}}\right)^2 \left[\frac{\tan^2\theta_R}{1+\left(\frac{m_t}{M_{\mathcal{T}}}\right)^2\tan^2\theta_R}\right] \tag{5.32}$$

$$K_L^{(t\mathcal{T})} = -\left(\frac{m_t}{M_{\mathcal{T}}}\right) \left[\frac{\tan\theta_R}{1+\left(\frac{m_t}{M_{\mathcal{T}}}\right)^2\tan^2\theta_R}\right] \tag{5.33}$$

$$K_L^{(\mathcal{T})} = \frac{1}{1+\left(\frac{m_t}{M_{\mathcal{T}}}\right)^2\tan^2\theta_R} \tag{5.34}$$

for expressions (5.23)–(5.25). These exact two-fermion mixing results coincide with those calculated in Ref. [157]. We can now make contact with the simplified Z' model with Lagrangian Eq (5.4) studied in sec. 5.4. We find

$$|g_{tZ'}| = g_T \sin^2\theta_R\,. \tag{5.35}$$

In the top-philic U(1)$'$ model, the scalar boson ϕ (the radial mode) will also have top-philic couplings, giving rise to the simplified Lagrangian (5.5). The coupling on the other hand will tend to be small because of the LH mixings Θ_L. Indeed one obtains the following matching:

$$|y_{t\phi}| = \frac{1}{\sqrt{2}}\left(\frac{m_t}{M_{\mathcal{T}}}\right)\lambda_t\,\sin\theta_R \tag{5.36}$$

This implies that the impact of the scalar ϕ for the LHC phenomenology is negligible compared to the Z' in the limit $M_{\mathcal{T}} \gg v, v'$ for small Yukawa coupling λ_t. On the other hand, if the value of λ_t is stretched to larger values right at the perturbativity limit, then the top-philic couplings of ϕ can be comparable to the Z'.

The effects of NP in the electroweak sector can be neglected for heavy vector-like $\mathcal{T}$. One can assess the NP deviations using the relation (5.26). For instance, no tree-level FCNC transitions are present, but we do get a small violation of flavor universality in the Z boson interactions with LH tops as well as a small modification of the top-quark Yukawa interaction with the Higgs. For example the Higgs couplings in this model reads

$$Y^{(u)} = \sqrt{2}\,\frac{\Delta m_u}{v} - \left(\frac{m_t}{M_{\mathcal{T}}}\right)^2 \left[\frac{\tan^2\theta_R}{1+\left(\frac{m_t}{M_{\mathcal{T}}}\right)^2\tan^2\theta_R}\right] \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{5.37}$$

The violation of CKM unitarity only occurs in the third row of the CKM matrix and affects each of the elements universally:

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ \alpha\, V_{td} & \alpha\, V_{ts} & \alpha\, V_{tb} \end{pmatrix} \tag{5.38}$$

such that

$$\alpha = \frac{1}{\sqrt{1 + \left(\frac{m_t}{M_{\mathcal{T}}}\right)^2 \tan^2 \theta_R}} \,. \tag{5.39}$$

These results show that the U(1)′ top-philic model presented in this section will mostly have a relevant phenomenological impact in the quarks sector at tree-level in top-quark processes involving the Z' boson, and to a lesser extent the scalar singlet ϕ. Therefore, the main channel to test this model is $pp \to t\bar{t}t\bar{t}$ production.

Chapter 6

Top-philic Forces: from B–decays to Colliders

6.1 Introduction

As discussed in the introduction of this work, there is no compelling theoretical requirement for discovering NP in resonant processes at particle colliders in the high-energy frontier. One plausible scenario is for physics beyond the SM to manifest itself for the first time in small deviations from the SM outside the realm of colliders in precision measurements of rare processes at low energies. Indirect searches are known to be very sensitive probes of NP scales above the current and the projected reach of the LHC or future colliders. A dramatic example are the limits coming from the proton lifetime measurements. Large underground precision experiments are able to probe Baryon number violating effective interactions predicted in Grand Unified Theories (GUT) [161, 162, 163]. For instance, the current bound on the lifetime of the proton decay process $p \to e^+\pi^0$ is around $\tau \sim 1.5 \times 10^{34}$ years given by super-Kamiokande [164]. These experimental results have ruled out the minimal SU(5) model and are currently probing other GUT theories like SUSY SU(5).

One important class of low energy experiments are flavor precision measurements at heavy meson factories. In the past, the search for rare or forbidden flavor processes has been used extensively to probe indirectly NP scales in the second and third generation at energies much higher than the electroweak scale. For instance, searches for flavor violating decays such as $K \to \mu^\pm e^\mp$, essentially forbidden in the SM, set model-independent limits on NP scales around 10^3 TeV. More recently, large amounts of data from B-factories like BaBar and Belle and data from the first runs of the LHCb experiment at CERN have been used for the first time to probe the less constrained third generation sector, in particular, precision observables in B-meson decays.

If deviations from the SM are discovered at low energies in B-factories or LHCb, it would then be natural to search for potential NP effects in top-quark processes at the high-p_T LHC. In this chapter we demonstrate with one particular example how NP in tops can produce observable effects in rare B-decays as well as rare top-quark processes at hadron colliders. As a concrete example we focus on $b \to s$ flavor chaging transitions which have been measured recently by the LHCb collaboration. Remarkably, a pattern of deviations from the Standard Model (SM) in $b \to s\ell^+\ell^-$ transitions [14, 15, 16, 17, 18] has emerged over the last decade, suggesting the pres-

ence of new physics (NP) that violates lepton flavour universality. Global analyses of the experimental measurements within the low energy effective theory find that only a few effective operators are needed to consistently explain the observed deviations from the SM [165, 166, 167, 168, 169, 170]. The experimentalists of LHCb measured the LFU ratio

$$R_{K^{(*)}}^{[q_1^2,q_2^2]} = \frac{\mathrm{Br}(B \to K^{(*)}\mu^+\mu^-)}{\mathrm{Br}(B \to K^{(*)}e^+e^-)}\,, \tag{6.1}$$

where Br stands for the partial branching fraction comprising $q^2 = (p_{l^+} + p_{l^-})^2$ between q_1^2 and q_2^2 (in units of GeV^2). They reported [14, 15]:

$$R_K^{[1,6]} = 0.745\,^{+0.090}_{-0.074} \pm 0.036\,, \tag{6.2}$$

$$R_{K^*}^{[1.1,6]} = 0.660\,^{+0.110}_{-0.070} \pm 0.024\,, \tag{6.3}$$

$$R_{K^*}^{[0.045,1.1]} = 0.68\,^{+0.113}_{-0.069} \pm 0.47, \tag{6.4}$$

which are $\approx 2.5\,\sigma$ *smaller* than the values predicted in the SM [171] which read $R_K^{[1,6]} = 1.00 \pm 0.01$, $R_K^{[1.1,6]} = 0.92 \pm 0.02$ and $R_K^{[0.045,1.1]} = 1.00 \pm 0.01$, respectively. Although the experimental confirmation of these results is still lacking and further improvement is needed to increase the significance of the observed deviations with respect to the SM, these indications of LFU violation have stimulated considerable activity in the flavor physics community. The observations that $R_{K^{(*)}}^{\mathrm{exp}} < R_{K^{(*)}}^{\mathrm{SM}}$ and the deviation in the angular observable P_5' in the muonic channel $b \to s\mu^+\mu^-$ [18] will be referred to in the remainder of this work as the "*B-physics anomalies*". Note, however, that these neutral current anomalies in $b \to s$ transitions are not the only B-anomalies in the data. Other significant departures from LFU in the third generation have been measured at B-factories. Another set of puzzling measurements have appeared in charged currents, namely, semileptonic $b \to c$ transitions. In the last decade, several tests of LFU violation in tree-level processes $B \to D^{(*)}\ell\nu$ by BaBar, Belle and LHCb, have consistently shown large deviations from the SM predictions in the LFU ratio $R_{D^{(*)}}^{\mathrm{exp}} > R_{D^{(*)}}^{\mathrm{SM}}$. Since these large discrepancies appear in tree level processes, they can't be directly explained using top-philic models. For this reason we do not explore these flavor anomalies in this work.

A NP scenario that has been explored recently in this context is that of a mediator that couples predominantly to right-handed up-type quarks and to muons. The required contributions to explain the $b \to s\ell^+\ell^-$ anomalies arise at the one-loop level in this case. This idea has been presented in terms of the Standard Model Effective Field Theory (SMEFT) [172], as well as with specific models: a scalar leptoquark [5, 173], and a top-philic Z' boson [156, 157]. Given the large value of the top-quark mass and the structure of the Cabibbo–Kobayashi–Maskawa (CKM) matrix, the largest effects in $b \to s$ transitions will be generated when the mediator couples to top-quarks.

We, therefore, focus on this case and demonstrate that LHC searches of rare top-quark processes can be used as a complementary probe for these models. We analyze this scenario within the EFT framework as well as with particular models, considering all possible mediators described with a simplified dynamical model that gives rise to tree-level matching contributions to the SMEFT operators in which we are interested. Two top-philic NP scenarios are identified, one with a neutral vector boson, and the other with a vector $\widetilde{U}_1 \sim (\mathbf{3}, \mathbf{1}, 5/3)$ and scalar leptoquark

state $R_2 \sim (\mathbf{3}, \mathbf{2}, 7/6)$. We also consider constraints from LEP-I on the modifications of the Z properties, including the necessary one-loop matching corrections at the electroweak scale.

The last stage in the bottom-up approach to BSM physics is to write down a UV theory that gives rise to the necessary degrees of freedom responsible for the low-energy and LHC phenomenology. This is even more pressing if the NP mediators are non-standard vector leptoquark bosons such as $\widetilde{U}_1$. In the last section of this chapter, we provide a new class of UV complete models that can give rise to *any* TeV scale $\mathrm{SU}(2)_L$-singlet vector leptoquark as a gauge boson of SU(4). We use this to build a consistent UV complete model with a reasonable flavor structure that gives rise to top-philic leptoquark states $\widetilde{U}_1$ and R_2 responsible for the deviations in $b \to s$ transitions. This last model can be viewed as a non-abelian generalization of the top-philic $\mathrm{U}(1)'$ models provided in [156, 157] and described in sec. 5.5.

6.2 Effective Field Theory description

Effective Hamiltonian for $d_i \to d_j \ell\ell$ transitions

The first step is to describe in a model independent way the possible interactions that enter into $d_i \to d_j \ell_k \ell_l$ transitions with $i \neq j$ and $\ell = \{e, \mu\}$ using the effective field theory approach. Since the processes we focus on are B-meson decays occurring at a characteristic energy scale set by the b-quark mass $m_b \approx 4.8$ MeV $\ll M_W$, we can write down the most general semi-leptonic $\Delta F = 1$ effective weak Hamiltonian consistent with the SM in the broken phase $\mathrm{SU}(3) \otimes \mathrm{U}(1)_{\mathrm{em}}$:

$$\mathcal{H}_{\mathrm{eff}}^{d_i \to d_j \ell^+ \ell^-} = -\frac{4G_F}{\sqrt{2}} \frac{\alpha}{4\pi} V_{td_i} V_{td_j}^* \sum_I \left(\mathcal{C}_I^{[ij\ell\ell]} \mathcal{O}_I^{[ij\ell\ell]} + \mathcal{C}_I'^{[ij\ell\ell]} \mathcal{O}_I'^{[ij\ell\ell]} \right), \tag{6.5}$$

where V is the CKM matrix, $\mathcal{C}_I^{(\prime)}$ are the Wilson coefficients associated with the effective operators $\mathcal{O}_I^{(\prime)}$ and $I = \{7, 8, 9, 10, P, S, T, T_5\}$. For a definition of the other operators see Ref. [174]. Once specifying the Hamiltonian (6.5) for $\Delta B = \Delta S = 1$ observables, the most relevant operators for our analysis are the well known $\mathcal{O}_{9,10}^{(\prime)}$ four-fermion semi-leptonic operators, defined at the scale $\mu \sim m_b$ as

$$\mathcal{O}_9^{[ij\ell\ell]} = (\bar{d}^j \gamma_\mu P_L d^i)(\bar{\ell}\gamma^\mu \ell) \tag{6.6}$$
$$\mathcal{O}_9'^{[ij\ell\ell]} = (\bar{d}^j \gamma_\mu P_R d^i)(\bar{\ell}\gamma^\mu \ell) \tag{6.7}$$
$$\mathcal{O}_{10}^{[ij\ell\ell]} = (\bar{d}^j \gamma_\mu P_L d^i)(\bar{\ell}\gamma^\mu \gamma_5 \ell) \tag{6.8}$$
$$\mathcal{O}_{10}'^{[ij\ell\ell]} = (\bar{d}^j \gamma_\mu P_R d^i)(\bar{\ell}\gamma^\mu \gamma_5 \ell) \tag{6.9}$$

Where $P_{L,R} = (\mathbb{1} \mp \gamma^5)/2$ are the chirality projectors. Crossing symmetry also implies that the Hamiltonian described above will give rise to $X \to \ell^+ \ell$ decays, where $X = (d_i d_j)$ is a bound state meson. One can compute the decay rates for the $B_s \to \ell_1 \ell_2$ and $B \to K^{(*)} \ell_1 \ell_2$ processes, cf. Ref. [175].
By assuming that the NP couplings to electrons are negligible, it has been established that R_K and R_{K^*} can be very well explained by a purely left-handed combination

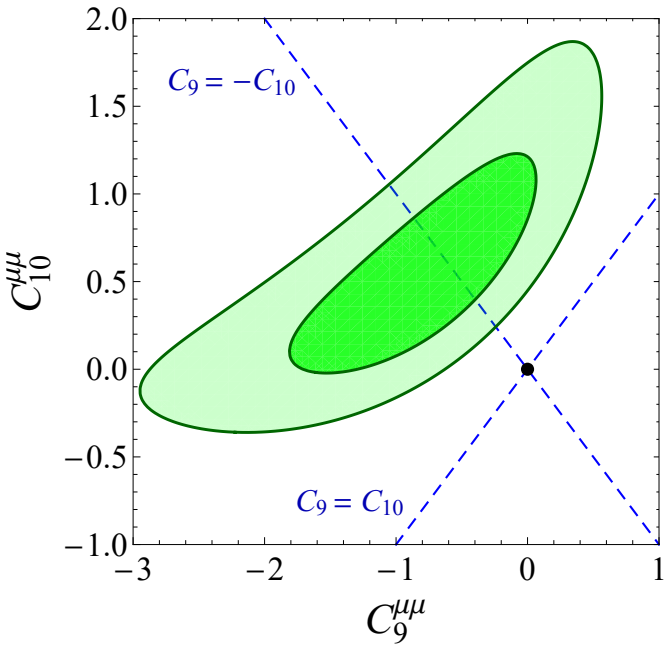


Figure 6.1: Low-energy fit to R_K, R_{K^*} and $\mathcal{B}(B_s \to \mu\mu)$ in the plane $C_9^{\mu\mu}$ vs. $C_{10}^{\mu\mu}$. Darker (lighter) region is allowed to $1\,\sigma$ ($2\,\sigma$) accuracy. Blue dashed lines correspond to scenarios with $C_9^{\mu\mu} = \pm C_{10}^{\mu\mu}$, while the black dot denotes the SM point. Taken from Ref. [2].

$C_9^{\mu\mu} = -C_{10}^{\mu\mu} < 0$. [1] The result of the fit for the $C_9^{\mu\mu} = -C_{10}^{\mu\mu}$ solution is illustrated in figure 6.1 by the dashed line crossing the 1σ region (dark green) given by [2]

$$C_9^{\mu\mu} = -C_{10}^{\mu\mu} \in (-0.85, -0.50)\,, \tag{6.10}$$

which deviates from the SM (black dot) by almost 4σ. In this fit we used $R_{K^{(*)}}^{\rm exp}$ [14, 15], and the theoretically clean $\mathcal{B}(B_s \to \mu\mu)^{\rm exp} = \left(3.0 \pm 0.6^{+0.3}_{-0.2}\right) \times 10^{-9}$ [176] [2]. On the other hand, the solution $C_9^{\mu\mu} = C_{10}^{\mu\mu}$ is completely ruled out. Notice that the vectorial solution $C_9^{\mu\mu} < 0$ can explain the anomaly within the 1σ range, but an additional contribution from a non-vanishing Wilson coefficient $C_{10}^{\mu\mu} > 0$ is preferable.

6.2.1 SM effective field theory

As discussed in in sec. 3.2, when the new degrees of freedom are much heavier than the electroweak scale, a useful approach is to parametrize their effects at low energies via the SMEFT. We assume that the NP above the cutoff scale Λ couples predominantly to the RH top-quark $t \equiv t_R$ in the quark sector, in accordance with the top-philic flavor structure described in sec. 3.3.

[1] An explanation of $R_{K^{(*)}}$ by NP couplings to electrons is disfavored by global analysis of the $b \to s\mu\mu$ observables, cf. Ref. [165, 166, 169].

[2] The measured $\mathcal{B}(B_s \to \mu\mu)^{\rm exp}$, agrees with the SM value $\mathcal{B}(B_s \to \mu\mu)^{\rm SM} = (3.65 \pm 0.23) \times 10^{-9}$ [177].

Top-philia

The full set of top-philic dimension-6 effective operators were provided in (3.49)-(3.52). Here, we consider a more minimalistic setup where at the cutoff scale only a subset of operators are generated with non-vanishing Wilson coefficients. These are:

$$\mathcal{O}_{lt}^{ij} = (\bar{l}_i\gamma^\mu l_j)(\bar{t}\gamma_\mu t)\,, \tag{6.11}$$
$$\mathcal{O}_{et}^{ij} = (\bar{e}_i\gamma^\mu e_j)(\bar{t}\gamma_\mu t)\,, \tag{6.12}$$

where $l_i = (\nu_L^i, \ell_L^i)^T$ are the LH lepton doublets and $e_i = \ell_R^i$ the RH lepton singlets. Here we work in the weak basis where the down-type quark and charged lepton mass matrices are diagonal. As for the flavor structure of the leptonic currents in (6.11)–(6.12), we assume second family dominance. The only real motivation for this choice is phenomenological. All deviations in the B-anomalies (i.e. $R_{K^{(*)}}$ and angular $b \to s\mu^+\mu^-$) suggest NP coupling dominantly to muons. We then impose from now on the muonic flavor ansatz

$$\mathcal{C}_{lt,et}^{ij} = \delta^{i2}\delta^{j2}\,\mathcal{C}_{lt,et}\,. \tag{6.13}$$

Of course, when considering a specific UV complete model, other top-philic operators besides (6.11)–(6.12) may arise at the Λ scale as well. In what follows we consider their effect negligible below the electroweak scale. Above the electroweak scale, some of these other operators would have to be considered for a reliable LHC analysis, like for example the 4-top operator $(\bar{t}\gamma_\mu\bar{t})(t\gamma^\mu t)$ which is indeed a relevant byproduct in UV models with top-philic vector bosons.

The top-philic operators (6.11)–(6.12) mix under RG evolution (3.6) when running from the high scale Λ down to the electroweak scale $\mu_{\rm EW}$. The one-loop anomalous dimension matrix for the SMEFT dimension six operators have been calculated in [62, 64]. Some operators receive a leading contribution from RG mixing that is enhanced by the large top-quark Yukawa coupling and is therefore particularly interesting phenomenologically. The resulting dimension six operators are

$$\begin{aligned}
\mathcal{O}_{\varphi l}^{(1)} &= (\varphi^\dagger i\overleftrightarrow{D}_\alpha\varphi)(\bar{l}_2\gamma^\alpha l_2)\,,\\
\mathcal{O}_{\varphi e} &= (\varphi^\dagger i\overleftrightarrow{D}_\alpha\varphi)(\bar{\mu}_R\gamma^\alpha\mu_R)\,,\\
\mathcal{O}_{lq}^{(1)} &= (\bar{l}_2\gamma^\alpha l_2)(\bar{q}_i\gamma_\alpha q_j)\,,\\
\mathcal{O}_{qe} &= (\bar{q}_i\gamma^\alpha q_j)(\bar{\mu}_R\gamma_\alpha\mu_R)\,.
\end{aligned} \tag{6.14}$$

Here $q_i = (u_L^i, d_L^i)^T$ are the quark doublets, φ is the SM Higgs doublet and i, j are flavour indices. In the first leading logarithm approximation we have

$$\begin{aligned}
\mathcal{C}_{\varphi l}^{(1)}(\mu_{\rm EW}) &= \mathcal{C}_{\varphi e}(\mu_{\rm EW}) \;=\; 6y_t^2 L\,\mathcal{C}_{lt}(\Lambda)\,,\\
\mathcal{C}_{lq}^{(1)}(\mu_{\rm EW}) &= y_t^2\lambda_t^{ij} L\,\mathcal{C}_{lt}(\Lambda)\,,\\
\mathcal{C}_{qe}(\mu_{\rm EW}) &= y_t^2\lambda_t^{ij} L\,\mathcal{C}_{eu}(\Lambda)\,,
\end{aligned} \tag{6.15}$$

with

$$L \equiv \frac{1}{16\pi^2}\,\log\left(\frac{\Lambda}{\mu_{\rm EW}}\right)\,, \tag{6.16}$$

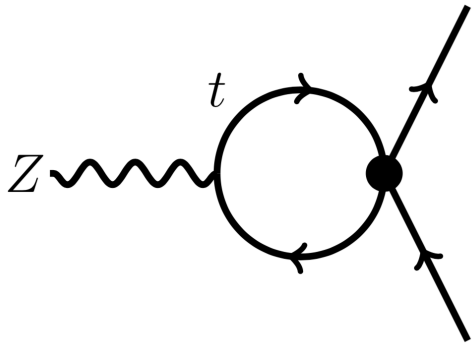


Figure 6.2: One loop correction to $Z \to \mu^+\mu^-$.

$y_t = \sqrt{2}m_t/v \sim 1$ being the top-quark Yukawa and $\lambda_t^{ij} = V_{ti}^* V_{tj}$, with V the Cabibbo-Kobayashi-Maskawa (CKM) matrix. Other operators are generated at the electroweak scale in a similar way, but the corresponding RG mixing contributions are suppressed by the $U(1)_Y$ gauge coupling or by the lepton Yukawa.

The first two operators in (6.14) give rise to modified Z couplings to muons and neutrinos after electroweak symmetry breaking. The last two operators in (6.14) encode the effects we are interested, as they will contribute to $\Delta F = 1$ semileptonic meson decays. Note that the flavour structure of these operators ($\propto \lambda_t^{ij}$) implies a very strong hierarchy of these contributions for different flavor transitions: $|V_{ts}^* V_{td}| \sim \lambda^5 \ll |V_{td}^* V_{tb}| \sim \lambda^3 \ll |V_{ts}^* V_{tb}| \sim \lambda^2$.

Modification to Z boson couplings

The operators in (6.11)–(6.12) modify the Z boson couplings to the muons at the quantum level (see figure 6.2) via RG mixing with $(\varphi^\dagger i \overleftrightarrow{D}_\mu \varphi)(\bar{\ell}\gamma^\mu \ell) \supset \frac{g_Z}{2} v^2 Z_\mu (\bar{\ell}\gamma^\mu \ell)$ and $(\varphi^\dagger i \overleftrightarrow{D}_\mu \varphi)(\bar{e}\gamma^\mu e) \supset \frac{g_Z}{2} v^2 Z_\mu (\bar{e}\gamma^\mu e)$ after electroweak symmetry breaking, where φ is the SM Higgs doublet.

We can parametrize these anomalous couplings of the Z boson to muons on general grounds as:

$$\mathcal{L} = \frac{g_Z}{2} \bar{\mu}\gamma_\alpha \left(\delta g_L P_L + \delta g_R P_R\right) \mu Z^\alpha \,, \tag{6.17}$$

with $g_Z = g/c_W$. Taking into account the first leading logarithm from renormalization group evolution together with the finite parts of the one-loop correction we obtain

$$\begin{aligned} \delta g_L &= -\frac{3y_t^2}{8\pi^2}\frac{v^2}{\Lambda^2}\left[\log\left(\frac{m_t}{\Lambda}\right) - \frac{4s_{\theta_W}^2}{9}(F+1) + \frac{F}{2}\right]\mathcal{C}_{lt} \,, \\ \delta g_R &= -\frac{3y_t^2}{8\pi^2}\frac{v^2}{\Lambda^2}\left[\log\left(\frac{m_t}{\Lambda}\right) - \frac{4s_{\theta_W}^2}{9}(F+1) + \frac{F}{2}\right]\mathcal{C}_{eu} \,, \end{aligned} \tag{6.18}$$

with $y_t = \sqrt{2}m_t/v$. We have retained terms that are enhanced by the top-quark Yukawa. We use the notation $s_\alpha \equiv \sin\alpha$, θ_W is the weak angle. The Wilson

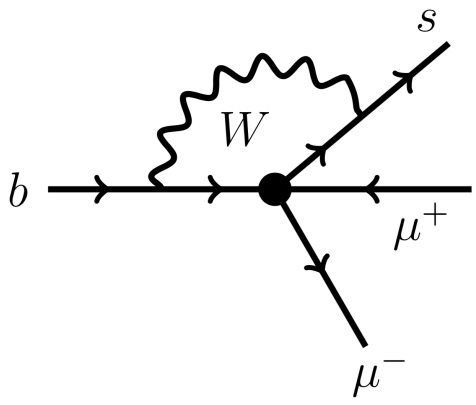


Figure 6.3: One loop contribution in unitary gauge to $b \to s\mu^+\mu^-$ from a four-fermion operator involving top quarks and muons.

coefficients $\mathcal{C}_{lt}$ and $\mathcal{C}_{et}$ are evaluated at the scale Λ and the loop function F is given by ($\tau_t = 4m_t^2/M_Z^2$)

$$F = -2 + 2\sqrt{\tau_t - 1}\arctan\left(\frac{1}{\sqrt{\tau_t - 1}}\right). \tag{6.19}$$

The inclusion of the one-loop matching corrections cancels the scale dependence of the leading renormalization group contribution. We also verified the corresponding entry of the anomalous dimension matrix calculated in Ref. [63]. Note that one-loop finite corrections could also originate from a UV completion in our framework. We assume that these model dependent finite corrections are subdominant, similar assumptions have been made for instance in [178].

Weak Effective Theory for $b \to s\ell\ell$

Below the electroweak scale, the top-quark is integrated out together with the $W^\pm$, Z and the Higgs boson. The top-philic operators $\mathcal{O}_{lt}$ and $\mathcal{O}_{et}$ give a one-loop matching contribution via the diagram shown in figure 6.3 to the semileptonic operators $O_9^{ij,\ell}$ $O_{10}^{ij,\ell}$. The leading contribution due to RG evolution can be obtained using the one-loop anomalous dimension matrix obtained in [63]. The finite parts from the one-loop correction were calculated in [179]. Keeping top-Yukawa enhanced contributions, the final results read [172]

$$\begin{aligned}
\mathcal{C}_9^{ij,\mu} &\simeq \frac{x_t v^2}{8 s_{\theta_W}^2 \Lambda^2}\left[\log\left(\frac{\Lambda}{M_W}\right) + I_0(x_t)\right](\mathcal{C}_{lt} + \mathcal{C}_{et}), \\
\mathcal{C}_{10}^{ij,\mu} &\simeq \frac{-x_t v^2}{8 s_{\theta_W}^2 \Lambda^2}\left[\log\left(\frac{\Lambda}{M_W}\right) + I_0(x_t)\right](\mathcal{C}_{lt} - \mathcal{C}_{et}),
\end{aligned} \tag{6.20}$$

with $x_t = m_t^2/M_W^2$ and $I_0(x_t) \simeq -0.74$, where the loop function is given by

$$I_0(x) = \frac{x-7}{4(1-x)} - \frac{x^2 - 2x + 4}{2(1-x)^2}\log(x) \tag{6.21}$$

and $\mathcal{C}_{lt}$ and $\mathcal{C}_{et}$ are evaluated at the scale Λ. The Wilson coefficients $\mathcal{C}_{9,10}^{ij,\mu}$ are proportional to m_t^2 due to the required chirality flip in each top-quark propagator. We verified that one-loop matching corrections cancel the scale dependence of the leading renormalization group contribution in (6.20).

6.3 Low energy phenomenology

We now confront the effective operators described above with experimental data from precision measurements and flavor experiments.

Z-pole measurements

Modifications of the $Z \to \mu^+\mu^-$ decay rate are constrained by lepton flavor universality tests in Z decays performed at LEP-I. We use the following measurement (see Sec. 7.2.1 in [180])

$$\frac{\Gamma_{\mu\mu}}{\Gamma_{ee}} = 1.0009 \pm 0.0028\,, \qquad \frac{\Gamma_{\tau\tau}}{\Gamma_{ee}} = 1.0019 \pm 0.0032\,, \tag{6.22}$$

with a correlation $\rho = 0.63$, where the notation $\Gamma_{ff} \equiv \Gamma(Z \to f^+f^-)$ has been used. The partial decay width for $Z \to \mu^+\mu^-$ taking into account (6.17) is given to linear order in the NP contributions by

$$\frac{\Gamma(Z \to \mu^+\mu^-)}{\Gamma(Z \to e^+e^-)} = 1 + \frac{c_{2\theta_W}\delta g_L - 2s^2_{\theta_W}\delta g_R}{A}\,, \tag{6.23}$$

where $A \equiv 1 + c_{4\theta_W}/2 - c_{2\theta_W}$ and $c_\alpha \equiv \cos\alpha$.
Modifications of the Z coupling to leptons are also constrained by the leptonic asymmetry parameter determined by the Aleph experiment at LEP-I. We use the measurement (see Table 7.4 in [180])

$$\mathcal{A}_\mu = 0.1456 \pm 0.0091\,. \tag{6.24}$$

The leptonic asymmetry parameter is defined by

$$\begin{aligned} \mathcal{A}_\mu &= \frac{\Gamma(Z \to \mu_L^+\mu_L^-) \;-\; \Gamma(Z \to \mu_R^+\mu_R^-)}{\Gamma(Z \to \mu^+\mu^-)} \\ &= \frac{B}{A} \;+\; \frac{c_{2\theta_W}\delta g_L \;+\; 2s^2_{\theta_W}\delta g_R}{A} \;-\; \frac{B(c_{2\theta_W}\delta g_L \;-\; 2s^2_{\theta_W}\delta g_R)}{A^2}\,, \end{aligned} \tag{6.25}$$

with $B = -1/2 + c_{2\theta_W}$ and A defined below (6.23).

B-anomalies

The semileptonic operators (6.6) and (6.8) can in principle accommodate the anomalies observed in $b \to s$ transitions. To analyze this, we reconstruct the likelihood for $b \to s\mu^+\mu^-$ observables from the 1σ and 2σ contours in the $\mathcal{C}_9 - \mathcal{C}_{10}$ plane

χ^2	$b \to s\mu^+\mu^-$	$R_{K^{(*)}}$	$Z \to \ell^+\ell^-$
SM	25.8	22.5	0.5
$\Lambda = 1$ TeV	2.5	5	7.9
$\Lambda = 1.5$ TeV	2.5	5	7.8
$\Lambda = 1.8$ TeV	2.4	5	7.8

Table 6.1: Contribution to the χ^2 from each sector at the minimum of the global χ^2 and in the SM.

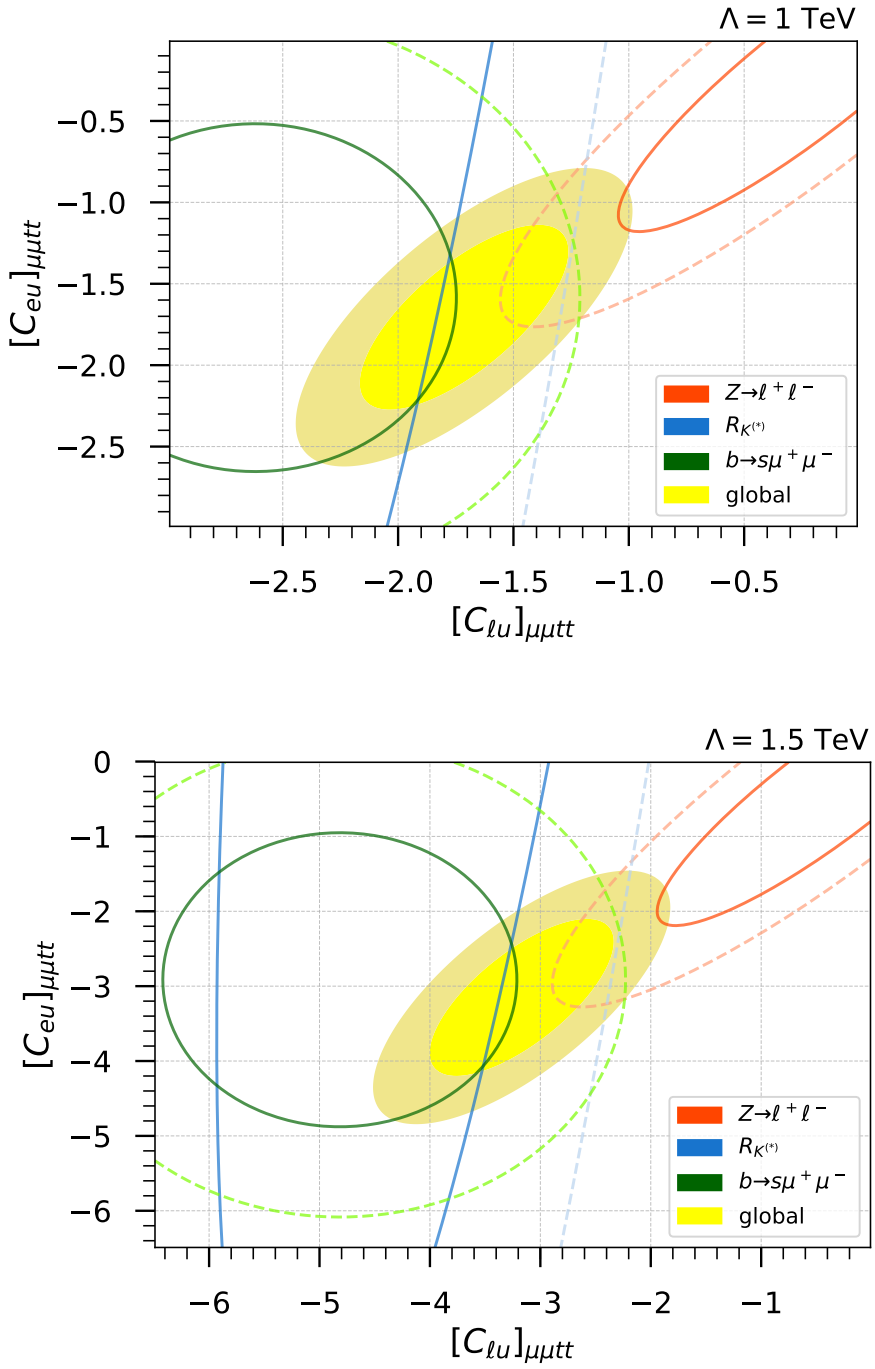
provided in [181], assuming a bivariate normal distribution. We obtain $(\mathcal{C}_9, \mathcal{C}_{10}) = (-1.11, 0.273)$ for the mean values, $\sigma_{C_9} = \sigma_{C_{10}} = 0.24$ for the standard deviation, and a correlation $\rho = 0.20$. We also include in our analysis the ratios R_K and R_{K^*}, using the general formulas derived in [172] and the experimental values reported in [14, 15]. Contributions to $b \to s\nu\bar{\nu}$ and $s \to d\nu\bar{\nu}$ are related in this framework to those in $b \to s\mu^+\mu^-$ due to the $\mathrm{SU}(2)_L$ gauge symmetry and the predictive flavour structure [156, 182]. Current bounds from B and K meson decays into final states with neutrinos do not set any relevant constraint in our framework.

A global χ^2 function is built with all these observables. The results of the fit are summarized in Table 6.1 and in figure 6.4. Table 6.1 shows the contributions to χ^2 from each sector within the SM and at the minimum of the global χ^2 for three benchmark values of Λ. Figure 6.4 shows the isocontours of $\Delta\chi^2 \equiv \chi^2 - \chi^2_{\text{min}} = \{2.3, 5.99\}$ in the plane $\{\mathcal{C}_{lt}, \mathcal{C}_{et}\}$ for the same benchmarks. The preferred region by the global fit (shown in figure 6.4 as a yellow ellipse) lies is the third quadrant along the direction $\mathcal{C}_{et} \sim \mathcal{C}_{lt}$. In this region, the NP contribution to the effective Hamiltonian for $b \to s\ell^+\ell^-$ transitions enters mainly in the Wilson coefficient $\mathcal{C}_9$, implying that the fit has a peference for vectorial muon couplings.

One important observation is that the NP effects considered cancel accidentally for $\mathcal{C}_{et} \sim \mathcal{C}_{lt}$ in the decay width for $Z \to \mu^+\mu^-$, see Eq. (6.23) (notice that $c_{2\theta_W} \simeq 2s^2_{\theta_W}$). The leptonic asymmetry parameter $\mathcal{A}_\mu$ breaks this blind direction of the LEP-I χ^2 to some degree, but a very strong correlation between these two variables remains. We have compared the LEP-I bounds we obtain with those derived using the results of [4] and found good agreement. For this comparison we use the following values reported in Ref. [4]: $\delta g_L/2 = (0.1 \pm 1.1) \times 10^{-3}$, $\delta g_R/2 = (0.0 \pm 1.3) \times 10^{-3}$, with a correlation $\rho = 0.90$. Another observation is that current data for $b \to s\ell^+\ell^-$ and LEP-I show some slight tension within the framework analyzed here, which is reflected in the contribution of LEP observables to the χ^2 in Table 6.1. The combined fit would be better if the deviations from the SM observed in $b \to s\ell^+\ell^-$ transitions decrease slightly with future measurements.

6.4 Simplified dynamical models

Now that we have identified the effective operators capable of generating semi-leptonic $b \to s$ transitions at the one-loop level, the next step is to identify the possible new degrees of freedom in the UV that give rise to the EFT at the elec-


Λ = 1 TeV
−0.5
−1.0
−1.5
−2.0
−2.5
$[C_{eu}]_{\mu\mu tt}$
−2.5
−2.0
−1.5
−1.0
−0.5
$[C_{\ell u}]_{\mu\mu tt}$
$Z \to \ell^+\ell^-$
$R_{K^{(*)}}$
$b \to s\mu^+\mu^-$
global
Λ = 1.5 TeV
0
−1
−2
−3
−4
−5
−6
$[C_{eu}]_{\mu\mu tt}$
−6
−5
−4
−3
−2
−1
$[C_{\ell u}]_{\mu\mu tt}$
$Z \to \ell^+\ell^-$
$R_{K^{(*)}}$
$b \to s\mu^+\mu^-$
global

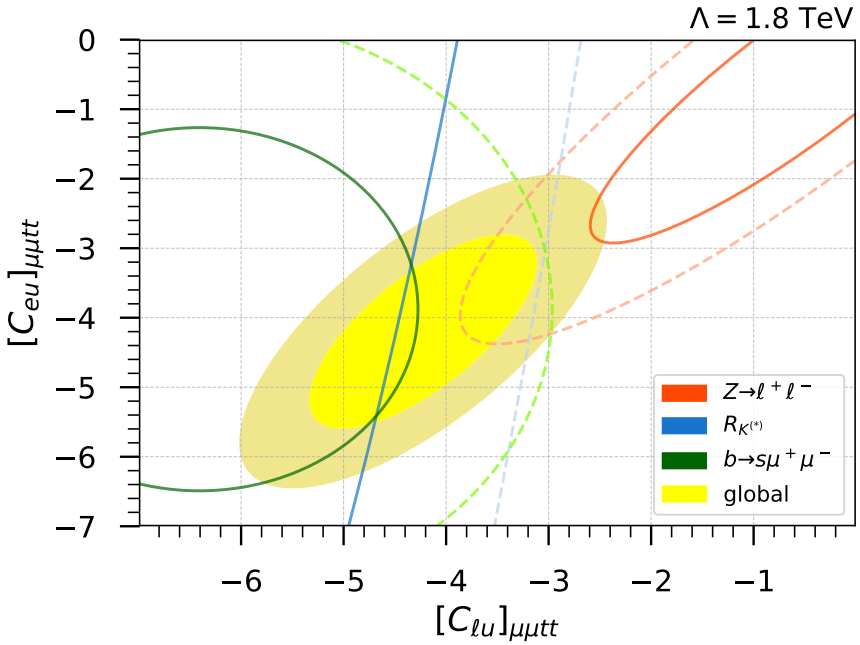


Figure 6.4: Preferred regions at 68% and 95% confidence level (CL) in the $(\mathcal{C}_{lt}, \mathcal{C}_{et})_{\mu\mu tt}$ plane from the global χ^2 (yellow-filled), $b \to s\mu^+\mu^-$ observables (green), $R_{K^{(*)}}$ (blue) and LEP-I measurements (red). Three benchmark values of the high scale $\Lambda = 1, 1.5, 1.8$ TeV have been chosen.

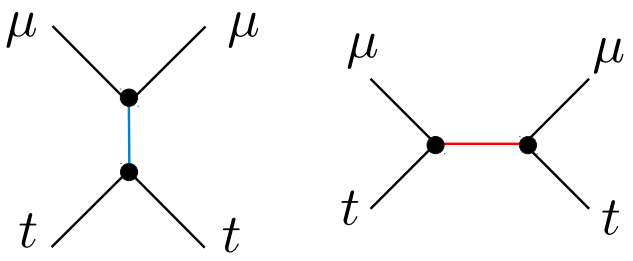


Figure 6.5: color-singlet (left) and colorful (right) mediators giving rise to the top-philic operators.

troweak scale and below. Since the top-philic operators are four-fermion operators, it is a simple exercise to classify all possible heavy mediators that couple to the relevant vector currents and generate at tree level the top-philic operators $\mathcal{O}_{lt}$ and $\mathcal{O}_{et}$. Taking into account the different irreducible representations of the Lorentz and SM gauge symmetry groups, one finds that there are only five different states that can generate these operators at tree-level. These are shown in Table 6.2. We distinguish two categories of mediators shown by the two possible Feynman diagrams in figure 6.5: color-singlet and color-triplet.

The only color-singlet mediator giving rise to four-fermion current × current

mediator	$Z' \sim (\mathbf{1},\mathbf{1},0)$	$S_1 \sim (\bar{\mathbf{3}},\mathbf{1},\frac{1}{3})$	$R_2 \sim (\mathbf{3},\mathbf{2},\frac{7}{6})$	$\widetilde{U}_1 \sim (\mathbf{3},\mathbf{1},\frac{5}{3})$	$\widetilde{V}_2 \sim (\bar{\mathbf{3}},\mathbf{2},-\frac{1}{6})$
$\mathcal{O}_{lt}$	✓	✗	✓	✗	✓
$\mathcal{O}_{et}$	✓	✓	✗	✓	✗
$\mathcal{C}_{lt} < 0$	✓	✗	✓	✓	✗
$\mathcal{C}_{et} < 0$	✓	✗	✓	✓	✗

Table 6.2: Possible mediators generating at tree-level the two relevant operators. The Z' represents a vector boson in the singlet representation of the SM gauge group while the nomenclature used for the leptoquarks corresponds to that in [7]. The last row shows those for which the Wilson coefficients are negative, as required by the low-energy fit.

interaction is a vector boson with SM representation $Z' \sim (\mathbf{1},\mathbf{1},0)$. Once this heavy state is integrated out it can give rise to the two semi-leptonic top-philic operators $\mathcal{O}_{lt}$ and $\mathcal{O}_{et}$ both with negative Wilson coefficients necessary for the low-energy fit.

Heavy bosons carrying color are also a possibility, in particular, if these are SU(3) color triplets. These states, which can be either scalar or vector bosons, are referred to in the literature as it leptoquarks. Leptoquarks have also non-vanishing hypercharges that are fixed in order for it to couple simultaneously to quark-lepton bilinears. These states first appeared in the Pati-Salam unification model [161] based on the gauge group $\mathrm{SU}(4) \otimes \mathrm{SU}(2)_L \otimes \mathrm{SU}(2)_R$ where quarks and leptons are unified in the same representation. The full classification of leptoquarks based on their quantum numbers can be found in Ref. [7].

6.4.1 Top-philic Z' model

The most minimal Z' model giving rise to a top-philic theory without any initial constraints on the leptonic sector corresponds to the simplified dynamical Lagrangian:

$$\mathcal{L}_{Z'} = -\frac{1}{4} Z'_{\mu\nu} Z'^{\mu\mu} + \frac{1}{2} M^2_{Z'} Z'_\mu Z'^\mu + Z'_\mu \left(g_t \, \mathcal{J}^\mu_t + g_\ell \, \mathcal{J}^\mu_\ell \right) \tag{6.26}$$

The fermionic currents are defined in terms of the electroweak unbroken fields

$$\mathcal{J}^\mu_t = c^t_R (\bar{t} \gamma^\mu t) \tag{6.27}$$

$$\mathcal{J}^\mu_\ell = \sum_{\ell = e,\mu,\tau} c^\ell_L (\bar{\ell}_L \gamma^\mu \ell_L) + c^\ell_R (\bar{e}_R \gamma^\mu e_R) \tag{6.28}$$

The Z' coupling is decomposed into a flavor universal component g_f and (chiral) flavor non-universal components given by the column vector $c^f_{L,R}$. It is convenient for the upcoming analysis to define the Z' coupling constant with the RH top-quark current as $\epsilon^{tt}_R \equiv g_t \, c^t_R$ and with each muonic current as $\epsilon^{\mu\mu}_{L,R} \equiv g_\mu \, c^\mu_{L,R}$. Notice that we have explicitly forbidden lepton flavor violating (LFV) couplings of the Z'. This simplified Lagrangian is very useful for phenomenological studies at the LHC and low-energy physics below the EW scale.

Once integrating out the heavy Z' it is straightforward to obtain at the matching scale $\Lambda = M_{Z'}$ the EFT Lagrangian with the expected current $\times$ current interactions

$$\mathcal{L}_{\text{eff}} = \frac{1}{2\,M_{Z'}} \left(g_t^2 \left[\mathcal{J}_t^\mu \mathcal{J}_t^\mu\right] + 2 g_t g_\ell \left[\mathcal{J}_t^\mu \mathcal{J}_\ell^\mu\right] + g_\ell^2 \left[\mathcal{J}_\ell^\mu \mathcal{J}_\ell^\mu\right] \right) \tag{6.29}$$

One can identify three tree-generated top-philic operators $\mathcal{O}_{tt}$, $\mathcal{O}_{lt}$ and $\mathcal{O}_{et}$ and the Wilson coefficients $\mathcal{C}_{tt}$, $\mathcal{C}_{lt}$ and $\mathcal{C}_{et}$ matched to the simplified model parameters at the scale $M_{Z'}$:

$$\begin{aligned}
\mathcal{O}_{tt} &= (\bar{t}\gamma_\alpha t)\,(\bar{t}\gamma^\alpha t)\,, & \mathcal{C}_{tt} &= -(\epsilon_R^{tt})^2 && (6.30)\\
\mathcal{O}_{lt} &= \left(\bar{l}_2\gamma_\alpha l_2\right)(\bar{t}\gamma^\alpha t)\,, & \mathcal{C}_{lt} &= -\epsilon_R^{tt}\,\epsilon_L^{\mu\mu} && (6.31)\\
\mathcal{O}_{et} &= (\bar{\mu}_R\gamma_\alpha \mu_R)\,(\bar{t}\gamma^\alpha t)\,, & \mathcal{C}_{et} &= -\epsilon_R^{tt}\,\epsilon_R^{\mu\mu} && (6.32)
\end{aligned}$$

Where we have imposed a flavor structure in order for the Z' to couple exclusively to second generation leptons, i.e. $c_{L,R} = (0, \star, 0)^T$ in Eq. (6.26) with the star indicating $\mathcal{O}(1)$ coupling. In addition to these top-philic operators, additional four-fermion operators with other chiral structures will be generated as well. We find purely leptonic operators of the form:

$$\begin{aligned}
\mathcal{O}_{ll} &= (\bar{l}_2\gamma^\alpha l_2)(\bar{l}_2\gamma_\alpha l_2), & C_{ll} &= -(\epsilon_L^{\mu\mu})^2 && (6.33)\\
\mathcal{O}_{le} &= (\bar{l}_2\gamma^\alpha l_2)(\bar{e}\gamma_\alpha e), & C_{le} &= -\epsilon_L^{\mu\mu}\epsilon_R^{\mu\mu} && (6.34)\\
\mathcal{O}_{ee} &= (\bar{e}\gamma^\alpha e)(\bar{e}\gamma_\alpha e), & C_{ee} &= -(\epsilon_R^{\mu\mu})^2 && (6.35)
\end{aligned}$$

In the second column of these equations, we have explicitly written the expressions for the matched Wilson coefficients, respectively. Interestingly, the four-top operator $\mathcal{O}_{tt}$ and the purely leptonic operators $\mathcal{O}_{ll,el}$, which are irrelevant for the B-anomalies, will have important phenomenological consequences for the simplified Z' model. In particular, we will see that $pp \to t\bar{t}t\bar{t}$ production at the LHC and neutrino trident production at beam dump experiments turn out to be important probes for the model that give complementary limits.

The fact that the preferred region of the low-energy fit provided in figure 6.4 lies around the direction $\mathcal{C}_{et} \approx \mathcal{C}_{lt}$ implies that the couplings of the heavy vector boson to muons has approximately a vectorial Lorentz structure $\epsilon_R^{\mu\mu} \approx \epsilon_L^{\mu\mu}$. Remarkably, of all the possible mediators in Table 6.2, the Z' boson is the only single state capable of simultaneously generating both operators $\mathcal{O}_{lt}$ and $\mathcal{O}_{et}$ with the correct negative sign.

6.4.2 Top-philic Leptoquarks

Of all possible leptoquark models, just a few can generate $\mathcal{O}_{lu,eu}$ at tree level. Only those that give rise in the SM broken phase to leptoquark states with electromagnetic charge $|Q| = 5/3$ and $|Q| = 1/3$ can potentially couple to bilinears of the form $(\bar{t}_R \ell_{L,R})$ and $(\bar{t}_R^c \ell_{L,R})$, respectively. The only top-philic leptoquarks capable of this are the two scalar and the two vector leptoquarks shown in table 6.2. The SM gauge representations and the necessary interaction Lagrangian between these leptoquarks and the RH top-quark and muons are:

$$S_1 \sim (\bar{\mathbf{3}}, \mathbf{1}, 1/3)\,, \qquad \mathcal{L}_{S_1} \supset (y_1^{RR})_{t\mu}\, S_1\, (\bar{t}^c \mu_R) \tag{6.36}$$

$$R_2 \sim (\mathbf{3}, \mathbf{2}, 7/6)\,, \qquad \mathcal{L}_{R_2} \supset -(y_2^{RL})_{t\mu}\, R_2^{(5/3)}\, (\bar{t}\mu_L) \tag{6.37}$$

$$\widetilde{U}_1 \sim (\mathbf{3}, \mathbf{1}, 5/3)\,, \qquad \mathcal{L}_{\widetilde{U}_1} \supset -(x_1^{RR})_{t\mu}\, \widetilde{U}_1^{\alpha}\, (\bar{t}\gamma_\alpha \mu_R) \tag{6.38}$$

$$\widetilde{V}_2 \sim (\bar{\mathbf{3}}, \mathbf{2}, -1/6)\,, \qquad \mathcal{L}_{\widetilde{V}_2} \supset -(x_2^{RL})_{t\mu}\, \widetilde{V}_2^{(1/3)\,\alpha}\, (\bar{t}^c \gamma_\alpha \mu_L) \tag{6.39}$$

Here we have adopted the same notation as in Ref. [7] for the symbols of the leptoquarks and for the 3×3 coupling matrices $y_{1,2}^{RR,RL}$ and $x_{1,2}^{RR,RL}$. The interactions shown above are the only ones for top and muon couplings. We have ignored here many other interaction terms that are perfectly consistent with gauge and Lorentz invariance. For instance, for the $\mathrm{SU}(2)_L$ doublets R_2 (scalar) and $\widetilde{V}_2$ (vector), we have only displayed in the equations above the field component with charge $Q = 5/3$ and left out for the sake of simplicity the $Q = -2/3$ leptoquark interactions. Even more pressing are the neglected diquark interactions for the S_1 and $\widetilde{V}_2$ leptoquarks which. If allowed, these would mediate proton decay and easily exclude TeV mass mediators [7].

One can scrutinize each of these top-philic leptoquarks by integrating out the heavy fields and matching to the SMEFT in the Warsaw basis. One important remark is that, in contrast to the Z' model, none of these single leptoquark mediators can simultaneously generate the two top-philic operators $\mathcal{O}_{lt}$ and $\mathcal{O}_{et}$ in one shot. In fact, because of the chiral structures of each of the four leptoquarks in (6.36)–(6.39), only one of the operators can be generated at a time. A direct inspection shows that S_1 and $\widetilde{U}_1$, being weak isospin singlets, can only give rise to $\mathcal{O}_{et}$ while R_2 and $\widetilde{V}_2$, being doublets can only give rise to $\mathcal{O}_{lt}$. This means that the B-anomaly can only be explained with a two-mediator leptoquark model. The possible models are the following four singlet–doublet combinations: $S_1 - R_2$, $S_1 - \widetilde{V}_2$, $\widetilde{U}_1 - R_2$ and $\widetilde{U}_1 - \widetilde{V}_2$. Remarkably, the required sign of the Wilson coefficients, i.e. the negative signs of $\mathcal{C}_{lt}$ and $\mathcal{C}_{et}$, provides important information about the possible leptoquark models that can accommodate the anomalies. Matching the Wilson coefficients at the cutoff scale with the parameters of the simplified models one finds that the leptoquarks S_1 and $\widetilde{V}_2$ give rise to operators with positive Wilson coefficients, ruling out these two possible mediators as explanations for the B-anomalies. All of this relevant information is displayed in the last four columns in table. 6.2. From this, we conclude that if the top-philic NP are leptoquarks, then only a combination of the vector singlet $\widetilde{U}_1$ and scalar doublet R_2 can potentially explain the B-anomalies in $b \to s\ell\ell$ at the one-loop level.

The $\widetilde{U}_1 - R_2$ model

The most minimal interaction Lagrangian for the $\widetilde{U}_1 - R_2$ model with the RH tops and second generation leptons is described by

$$\mathcal{L}_{\mathrm{LQ}} \;=\; \mathcal{L}_{\mathrm{kin}} - \frac{1}{2}\, M_{\widetilde{U}}^2 \widetilde{U}^{\mu\,\dagger} \widetilde{U}_\mu - \frac{1}{2}\, M_R^2 R_2^\dagger R_2 + \kappa_S\, \bar{t}\, R_2^T (i\tau_2) l_2 + \kappa_V (\bar{t}_R \gamma^\alpha \mu_R) \widetilde{U}_{1\alpha} + \mathrm{h.c.} \tag{6.40}$$

where τ_2 is the Pauli matrix, the first term is the kinetic term, M_R and κ_S are the mass and Yukawa coupling of the scalar leptoquark R_2, and $M_{\widetilde{U}}$ and κ_V are the

mass and coupling of the vector leptoquark $\widetilde{U}_1$. The kinetic Lagrangian for both leptoquarks read

$$\mathcal{L}_{\text{kin}} \;=\; (D_\mu R_2)^\dagger (D^\mu R_2) \;-\; \frac{1}{2}\,\widetilde{U}^\dagger_{\alpha\beta}\widetilde{U}^{\alpha\beta} \tag{6.41}$$

where D is the SM covariant derivative and $\widetilde{U}^{\alpha\beta} \equiv D^\alpha \widetilde{U}_1^\beta \;- D^\beta \widetilde{U}_1^\alpha$ is the field strength for the vector leptoquark. The scalar leptoquark, being an weak isospin doublet, is decomposed in the SM broken phase into $R_2 = (R_2^{(5/3)}, R_2^{(2/3)})^T$ where the superindex Q in $R_2^{(Q)}$ corresponds to the elctromagentic charge of each component.

In (6.40) we have written the simplest Lagrangian with the smallest number of couplings necessary for the B-anomalies and left out other allowed interactions. For instance, R_2 in principle has also interactions that are not exclusively with up-quarks. Gauge and Lorentz symmetry allow for a term of the form $\bar{q}_i R_2 \ell_R$ which generates couplings to LH down-type quarks like, e.g. b_L, in the broken phase. Removing these interactions in down-quarks would require some sort of symmetry in the UV completions. We assume this scenario, which is indeed realized in the UV model presented in the last section of this chapter. The kinetic term in (6.41) is also not the most general one. In sec. 6.5.3 we will take into account additional model dependent interaction terms for the vector leptoquark with the SM gauge bosons, namely, the gluons.

Integrating out these heavy leptoquark states at tree-level gives the matching conditions [183, 184, 185]

$$\mathcal{C}_{lt} \;=\; -\frac{|\kappa_S|^2}{2}\,, \qquad \mathcal{C}_{eu} \;=\; -|\kappa_V|^2\,. \tag{6.42}$$

If the leptoquarks have similar mass, the preferred region from the fit in figure 6.4 implies that $|\kappa_S| \sim \sqrt{2}|\kappa_V|$. Notice that the Wilson coefficients in (6.42) have the negative sign necessary to accommodate the low-energy fit and that no other top-philic operators besides $\mathcal{O}_{lt}$ and $\mathcal{O}_{et}$ are generated at tree-level from integrating out these heavy states. From this point of view, one can argue that the $\widetilde{U}_1 - R_2$ model is simpler than the Z' model at low energies. Of course, none of these models with massive vector bosons are UV complete theories and must be interpreted as non-renormalizable intermediate vector models that parametrize and characterize in a very convenient and economical way the effective operators below the cutoff scale.

To finalize, notice that the two remaining leptoquarks, $S_1 \sim (\bar{\mathbf{3}}, \mathbf{1}, 1/3)$ and $\widetilde{V}_2^\mu \sim (\bar{\mathbf{3}}, \mathbf{2}, -1/6)$, in table 6.2 generate these same operators but with a Wilson coefficient matching with opposite sign as those in Eq. (6.42) and are therefore not suitable for the B-anomalies.

6.5 High-p_T phenomenology

The top-philic Z' and $\widetilde{U}_1 - R_2$ models have a relatively small imprint in low-energy observables. As shown before, below the electroweak scale the only important contribution arises from virtual exchanges in $b \to s\ell\ell(\nu\nu)$ transitions and $Z \to \mu^+\mu^-$ decays. On the other hand, at energies above the electroweak breaking scale, we should expect new (non-resonant) phenomena to appear in the top-quark sector. The TeV scale mediators in one-loop solutions to the B-anomalies imply Wilson coefficients of $\mathcal{O}(1)$ for the top-philic operators $\mathcal{O}_{et}$ and $\mathcal{O}_{lt}$. This automatically

gives rise to a sizeable amplitude for $t\bar{t} \to \mu^+\mu^-$ transitions in tree-level processes at high-energy colliders. Since the top-quark content of the proton PDF is completely negligible at the 13 TeV LHC, the only relevant process that can directly probe these operators are $pp \to t\bar{t}\mu^+\mu^-$ from (mostly) gluon fusion. Given that the cutoff scale $\Lambda \sim M_{Z'}$ (or $\Lambda \sim M_{\rm LQ}$) could be within LHC reach, we abandon the SMEFT approach for our collider studies and instead focus on the simplified models. Indeed, a considerable amount of LHC collisions with large momentum exchange (i.e. large parton momentum fraction x) will be able to resolve the propagator leading to a deviation from the EFT expectation and possibly to a breakdown of the effective theory. This breakdown happens when the collider energy satisfies $E \sim \Lambda$ which is the case for LHC energies and mediator masses near the TeV scale.

Abandoning the EFT description implies that the process $pp \to t\bar{t}\mu^+\mu^-$ gives rise to different phenomenological outcomes depending on the quantum numbers of the exchanged intermediate state. For instance, if the Z' color-singlet is exchanged in the s-channel via $pp \to t\bar{t}Z'$, we should expect a resonance in the invariant mass spectrum $m_{\mu\mu}$ of the di-muon pair and no resonant effects in the spectra of the associated $t\bar{t}$ final states. For leptoquarks, the main process contributing to $t\bar{t}\mu^+\mu^-$ come from pair production of the leptoquarks via QCD interactions, i.e. $gg \to {\rm LQ}^\dagger{\rm LQ}$, that subsequently decay via ${\rm LQ} \to t\mu^+$. Here we use the notation ${\rm LQ} = \{R_2, \widetilde{U}_1^\mu\}$ to collectively denote either leptoquark. We should then expect a resonance in the invariant mass spectrum $m_{t\mu}$ of the top-muon pairs. Unfortunately, none of these processes have been searched in a dedicated analysis by the LHC collaborations. Luckily, many different searches in high-multiplicity jets and leptons have been performed by ATLAS and CMS. Some of these searches look for NP in final states approximating a $t\bar{t}\mu^+\mu^-$ signal after the decay of the tops. Of these, the ones that have generic signal regions (where no sophisticated final state reconstruction has been performed) can be recast and used to perform a detailed collider study for the top-philic models. In the following, we demonstrate that the LHC can indeed probe the interesting portions of parameter space relevant for the B-anomalies.

6.5.1 Searching for NP in $pp \to tt\mu\mu$ at the LHC

The main collider signature present in both top-philic models is a pair of muons produced in association with a $t\bar{t}$ pair. Several available searches can be used to set limits on the two simplified models using this process. We focus on three different searches each targeting one of the three possible decay modes for $t\bar{t}$ (fully hadronic, semi-leptonic and fully leptonic):

- *Inclusive di-muon tails for $pp \to t\bar{t}\mu^+\mu^-$ with hadronic tops:* searches for resonant effects in the invariant mass spectrum of di-lepton pairs are one of the most important NP probes at the LHC. Both ATLAS and CMS have performed a variety of Z' searches targeting several BSM models like the Sequential SM (SSM), E_6-motivated GUT models, $B - L$ Z' models and also contact interaction of the form $\ell\ell qq$. We revisit the 13 TeV search strategy by ATLAS [3] at an integrated luminosity of 36.1 fb^{-1} . The main signal category we are interested in is the *inclusive category* targeting the process $pp \to Z' \to \mu^+\mu^- + X$ where X can be any final state spectator, like for example hadronic activity (jets), lower p_T additional leptons or both. In our recast, X will play the role of the spectator $\bar{t}t$ decaying hadronically to jets.

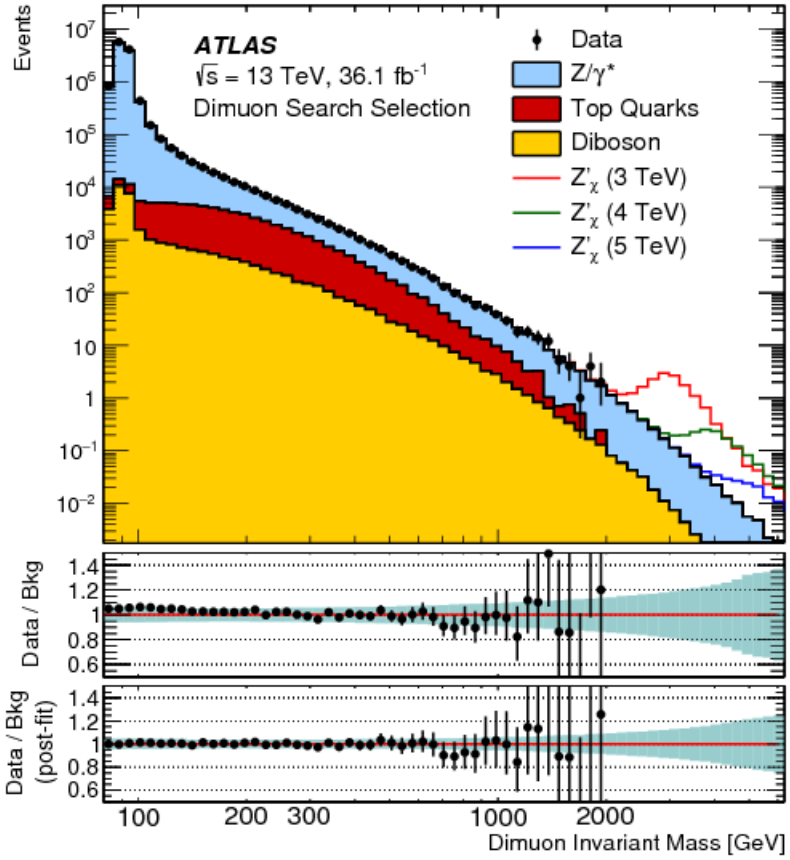


Figure 6.6: Di-muon spectrum extracted from LHC data at $\sqrt{s} = 13$ TeV and 36.1 fb^{-1} from Ref. [3].

In figure. 6.6 taken from Ref. [3] we show the measured data and expected backgrounds in the tail of the distribution used in our analysis. In the recast we used the same event selection as in [3]: we select from the MC data events that have at least one opposite sign di-muon pair with transverse momentum satisfying $p_T^\mu > 30$ GeV and rapidity $|\eta^\mu| < 2.5$. Following [3] we employed the p_T-dependent mini-isolation cone $R_{\rm cone} = 10$ GeV$/p_T^\mu$ requirements for muons.

- *SUSY tri-lepton searches for $pp \to t\bar{t}\mu^+\mu^-$ with semi-leptonic tops:* in SUSY models, the decay of gluinos and squarks produced in pairs are expected to produce final states with several leptons, jets and missing energy. Some of the signal regions defined in these searches, namely, those with tri-leptons, can be recasted to search for NP in $t\bar{t}\mu^+\mu^-$ where the $t\bar{t}$ decays into the semi-leptonic channel. For this we use the 13 TeV SUSY search by ATLAS [8] at a luminosity of 36.1 fb^{-1} . Before applying the basic selection cuts, leptons are reconstructed using the mini-isolation technique C.18 with dynamical cone of $R_{\rm cone} = \min(0.4, 0.1 + 9.6 \text{ GeV}/p_T^\ell)$ where p_T^ℓ is the transverse momentum of the candidate lepton. Isolated leptons are then rejected if $p_T^\ell < 10$ GeV and $|\eta^\ell| > 2.5$. Jets are clustered using the anti-k_T algorithm and cone $\Delta R = 0.4$ and kept if $p_T^j > 20$ GeV and $|\eta^j| < 2.8$. The efficiency for $b-$tagging a jet

with $|\eta^b| < 2.5$ was fixed at an operating point defined by the tagging efficiency $\epsilon_b \approx 70\%$ for a mis-tag rate of 1/380 for light quark/gluon jets and 1/12 for charm jets. Events in this search are categorized depending on the number of signal leptons N_ℓ, number of tagged b-jets N_b, number of jets N_j as well as cuts in the transverse momentum of the jet p_T^j, the transverse missing energy $E_T^{\rm miss}$, the effective mass $m_{\rm eff} = E_T^{\rm miss} + \sum_{i\in\{j,\ell\}} p_T^i$ and the ratio $E_T^{\rm miss}/m_{\rm eff}$. For our recast we focused on the R-parity conserving trilepton signal regions `Rpc3L0bS`, `Rpc3L0bH`, `Rpc3L1bS` and `Rpc3L1bH`, each differing in their b-jet content and $m_{\rm eff}$ cuts. The precise definitions of each of these signal region are displayed in table 6.3.

Signal Regions	N_ℓ	N_b	N_j	p_T^j [GeV]	$E_T^{\rm miss}$ [GeV]	$m_{\rm eff}$ [GeV]
`Rpc3L0bS`	≥ 3	$= 0$	≥ 3	≥ 40	≥ 200	≥ 600
`Rpc3L0bH`	≥ 3	$= 0$	≥ 3	≥ 40	≥ 200	≥ 1200
`Rpc3L1bS`	≥ 3	≥ 1	≥ 3	≥ 40	≥ 200	≥ 600
`Rpc3L1bH`	≥ 3	≥ 1	≥ 3	≥ 40	≥ 200	≥ 1200

Table 6.3: Selection cuts defining the tri-lepton signal regions used in our work from recasting Ref. [8] for $t\bar{t}\mu^+\mu^-$ search.

- *SUSY four-lepton searches for $pp \to t\bar{t}\mu^+\mu^-$ with leptonic tops:* since final states with four leptons are rare in the SM, searches for NP in the four-lepton channels can be quite efficient. These final states have been used by ATLAS in searches for R-parity violating SUSY in Ref. [9] at 13 TeV at a luminosity of 36.1 fb^{-1} . The search identifies light leptons (e and μ) as well as hadronic τ-jets as final states. The preselected muons (electrons) must have $p_T^\mu > 5$ GeV ($p_T^e > 7$ GeV) and $|\eta^\mu| < 2.7$ ($|\eta^\mu| < 2.47$), clustered jets with the anti-k_T algorithm with cone radius $\Delta R = 0.4$ must satisfy $p_T^j > 20$ GeV and $|\eta^j| < 2.8$. Of the selected jets, the visible portion of the hadronic $\tau_{\rm vis}$ are selected by tagging the preselected jets satisfying $|\eta^j| < 2.47$ with a tau-tagger. For di-lepton pairs with opposite sign, we remove the leptons for which the invariant mass falls inside the $\Upsilon(1S)$–$\Upsilon(3S)$ mass range and if $m_{\ell^+\ell^-} < 4$ GeV. Leptons with invariant masses falling inside the Z mass peak range 81.2–101.2 GeV are also removed from the events. Events with at least four leptons, defined by e, μ and $\tau_{\rm vis}$ are classified into four signal regions `SR0A`, `SR0B`, `SR1` and `SR2`, defined by cuts on $m_{\rm eff}$ and the multiplicity of hadroinic tau jets $\tau_{\rm vis}$. The signal regions are defined in table 6.4.

We now use these searches to set high-p_T limits on the two top-philic models. For the Z' mediator, we use the inclusive di-muon resonance search for $Z' \to \mu^+\mu^-$ since this is optimized for this scenario. Results will be presented in sec. 6.5.2 along with other limits from other processes. For the $\widetilde{U}_1 - R_2$ leptoquark model it will be necessary to carefully evaluate which search gives the best bound. Results for this scenario are given in sec. 6.5.3.

Signal Regions	N_ℓ	N_τ	p_T^τ [GeV]	m_{eff} [GeV]
SR0A	≥ 4	$= 0$	≥ 20	≥ 600
SR0A	≥ 4	$= 0$	≥ 20	≥ 1100
SR1	$= 3$	≥ 1	≥ 30	≥ 700
SR2	$= 2$	≥ 2	≥ 30	≥ 650

Table 6.4: Selection cuts defining the four-lepton signal regions used in our work from recasting Ref. [9] for $t\bar{t}\mu^+\mu^-$ search.

6.5.2 Limits on the Z' model

We now turn to the phenomenological implications of the Z' mediator assuming it has a mass around the TeV scale and a vectorial coupling to muons defined by $\epsilon_V^{\mu\mu} \equiv \epsilon_L^{\mu\mu} = \epsilon_R^{\mu\mu}$. When extracting limits from the LHC we will focus on tree-level Z' exchanges and omit from our analysis loop-induced processes such as $gg \to gZ'$. The latter are sensitive to details of the ultraviolet completion such as effects from heavy fermionic top-partners. These exotic fermions are not uncommon when trying to build an ultraviolet completion for the Z' model at hand and, while being too heavy to be directly produced on-shell they still may give non-negligible contributions to the production of the lighter Z' through loop-level non-decoupling effects, see Ref. [156, 157] for more details.

At tree level, the most important constraints come from Z' production in association with $t\bar{t}$ at the LHC. Once produced, the Z' boson can decay into muons, muon-neutrinos, and top-quarks. After neglecting small lepton masses, the partial decay widths for these channels are given by

$$\begin{aligned}
\Gamma(Z' \to \mu\bar{\mu}) &\simeq \frac{M_{Z'}}{24\pi}(|\epsilon_L^{\mu\mu}|^2 + |\epsilon_R^{\mu\mu}|^2)\,, \\
\Gamma(Z' \to t\bar{t}) &\simeq \frac{\lambda^{1/2}(1, z_t, z_t) N_C M_{Z'}}{24\pi}(1 - z_t)|\epsilon_R^{tt}|^2\,, \\
\Gamma(Z' \to \nu_\mu\bar{\nu}_\mu) &\simeq \frac{M_{Z'}}{24\pi}|\epsilon_L^{\mu\mu}|^2\,,
\end{aligned} \tag{6.43}$$

where $z_t = m_t^2/M_{Z'}^2$, $N_C = 3$, and λ represents the Källén function:

$$\lambda(x, y, z) \;\equiv\; x^2 + y^2 + z^2 - 2xy - 2xz - 2zy \tag{6.44}$$

Each of these decay channels give rise to three complementary LHC signatures: $t\bar{t}\mu^+\mu^-$, $t\bar{t}t\bar{t}$ and $t\bar{t} + \not{E}_T$. In order to set limits in the coupling plane $\{\epsilon_V^{\mu\mu}, \epsilon_R^{tt}\}$ of the model, we have recast a set of existing LHC searches for two benchmark masses of $M_{Z'} = 0.7$ TeV and $M_{Z'} = 1$ TeV. For each benchmark mass, the $1\,\sigma$ favored region fitting the $b \to s\ell^+\ell^-$ anomalies and the LEP-I measurements is given by the black dashed contours in figure 6.7. Limits for the process $pp \to t\bar{t}Z' \to t\bar{t}\mu^+\mu^-$ were extracted from the generic Z' di-muon resonance search by ATLAS [3] (Sec.10.3) at $36.1\,\text{fb}^{-1}$, assuming a detector acceptance of 40% and a decay width dominated by the three channels in Eq. (6.43). The 95% CL exclusion limits from this search are shown in the purple region in figure 6.7. Projections to a higher luminosity

of $300\,\text{fb}^{-1}$ are also given by the dotted purple contour in the same figure. For the process $pp \to t\bar{t}Z' \to t\bar{t}t\bar{t}$ we used the current best upper limit on the SM four-top cross-section by CMS [186] at $35.9\,\text{fb}^{-1}$ of data. The 2σ exclusion bound is given by the red region in figure 6.7. Projections to $300\,\text{fb}^{-1}$, given by the dotted red contour, were estimated using the multi-leptonic analysis performed in Ref. [187], where the 95% CL upper limit on the SM cross-section was found to be approximately $\sigma^{\text{SM}}_{t\bar{t}t\bar{t}} < 23$ fb. Notice that a dedicated resonance search for this channel can considerably improve this bound (especially at higher luminosities) if a high-mass cut is applied on the top-quark decay products or a top-tagger is used in order to improve sensitivity to the boosted tops from the decaying resonance, see also Ref. [127].

Another relevant probe of the Z' boson is the neutrino trident production [188]. The process $\nu_\mu\gamma^* \to \nu_\mu\mu\bar{\mu}$ occurring in a fixed target from a highly energetic neutrino beam gives important limits on the Z' boson coupling to muonic currents for a wide range of Z' masses. These constraints will be complementary to those from the LHC. The cross-section for this process normalized by the SM prediction is given by [188]

$$\frac{\sigma^{\text{NP}}_{\nu_\mu\mu\bar{\mu}}}{\sigma^{\text{SM}}_{\nu_\mu\mu\bar{\mu}}} = \frac{1 + \left(1 + 4s^2_{\theta_W} + \dfrac{2v^2\,(\epsilon^{\mu\mu}_V)^2}{M^2_{Z'}}\right)^2}{1 + (1 + 4s^2_{\theta_W})^2}. \tag{6.45}$$

This quantity has been measured at CCFR to be $\sigma^{\text{NP}}_{\nu_\mu\mu\bar{\mu}}/\sigma^{\text{SM}}_{\nu_\mu\mu\bar{\mu}} = 0.82 \pm 0.28$ [189], giving a strong constraint on the Z' muonic couplings. The $2\,\sigma$ upper limit on $\epsilon^{\mu\mu}_V$ is represented by the vertical blue region in figure 6.7.

We observe that all the constraints are complementary and exclude different regions of the available parameter space in figure 6.7. For $M_{Z'} = 0.7$ TeV, the preferred 1σ region from the global fit of flavour and LEP observables is already excluded, with each of the different constraints considered playing an important role. For $M_{Z'} = 1$ TeV, an allowed region remains centered around the point $\{|\epsilon^{\mu\mu}_V|, |\epsilon^{tt}_R|\} = \{1.2, 1.2\}$. Future searches from the LHC with 300 fb^{-1} will be sensitive to this region.

6.5.3 Limits on top-philic $\tilde{U}_1 - R_2$ leptoquark model

For this model the most important LHC bound comes from leptoquark (LQ) pair production $gg\,(q\bar{q}) \to \widetilde{U}^\dagger_1\widetilde{U}_1,\ R^\dagger_2 R_2$. The consequences of having the interactions in Eq. (6.40) plus a negligible top-quark PDF for the proton are: (i) LQ pair production is independent of the size of the couplings $\kappa_{S,V}$, hence driven completely by QCD interactions (see figure 6.8 (left) for a representative Feynman diagram), (ii) the absence of all $2 \to 2$ single LQ production channels of the form $qg \to \text{LQ}\,\ell$ at the LHC and (iii) the absence of $q\bar{q} \to \ell\bar{\ell}$ production via LQ exchange in the t-channel. The only relevant process at the LHC at leading order besides QCD pair production is the $2 \to 3$ single LQ production mode $gg \to \widetilde{U}_1(R_2)\,t\mu$ shown in figure 6.8 (right). This last process only becomes competitive with LQ pair production if the couplings $|\kappa_{S,V}| \gtrsim 1$ are large enough to overcome the $2 \to 3$ body phase space suppression.

The scalar leptoquark doublet R_2 (see Sec. 6.4.2) when decomposed into its $\text{SU}(2)_L$ components gives rise to the following interactions

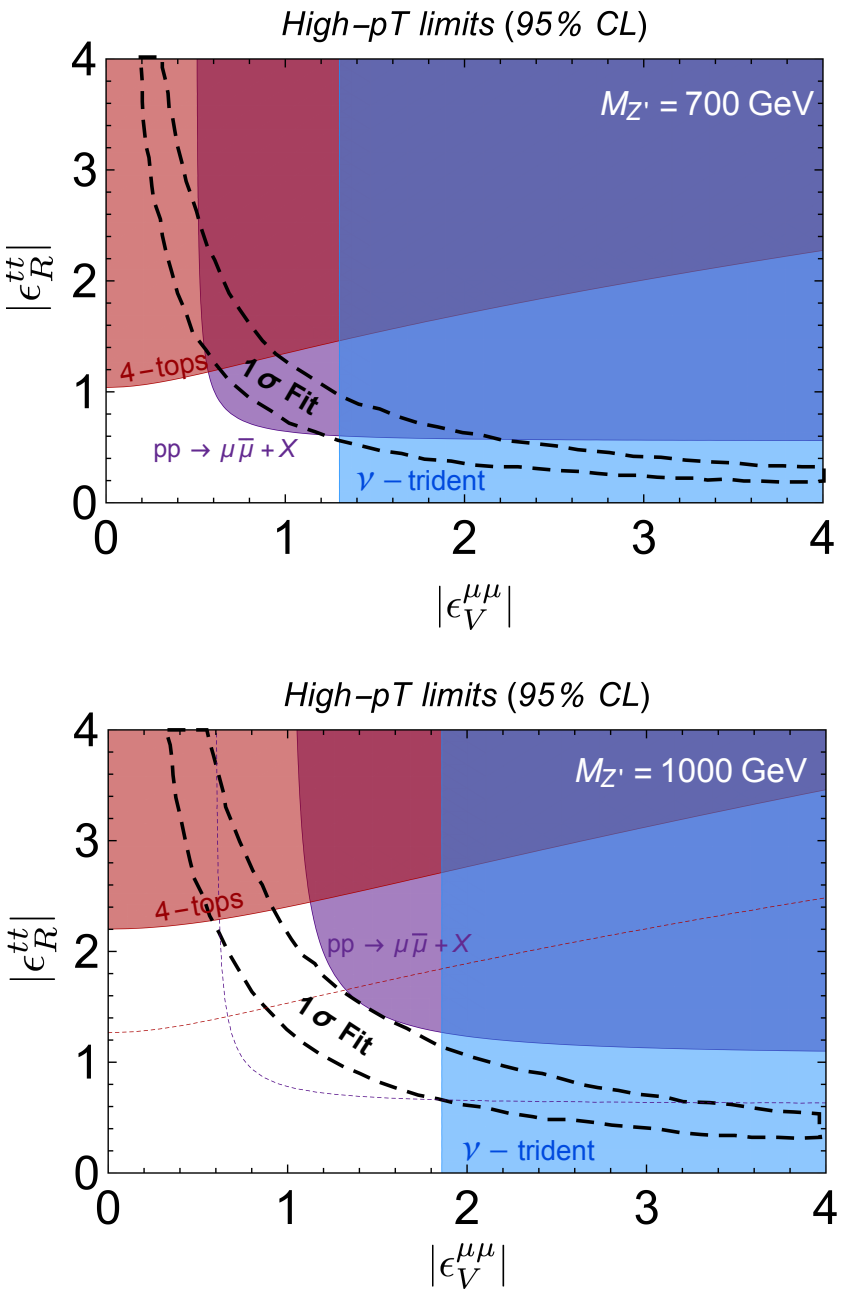


Figure 6.7: Summary of high-p_T bounds for the Z' model. The red, purple and blue 95% CL exclusion regions correspond to the LHC 4-top search, LHC di-muon tail search and the CCFR neutrino trident experiment, respectively. Dotted colored contours represent LHC bounds at a future luminosity of $300\,\text{fb}^{-1}$. The black dashed region corresponds to the $1\,\sigma$ global fit to $R_{K^{(*)}}$, LEP-I data and the $b \to s\mu\mu$ observables.

$$\mathcal{L} \supset \kappa_S \left[\bar{t}_R \, \nu_\mu R_2^{(2/3)} - \bar{t}_R \, \mu_L R_2^{(5/3)} \right] + \text{h.c.}\,. \tag{6.46}$$

In this case, the branching ratio for each state reads

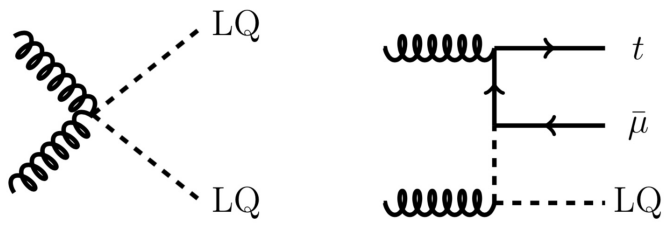


Figure 6.8: Representative diagrams for the QCD LQ pair production (left) and for the single LQ production mode $gg \to \tilde{U}_1(R_2)\, t\mu$ (right).

$$\beta(R_2^{(2/3)} \to t\nu_\mu) = 1\,, \qquad \beta(R_2^{(5/3)} \to t\mu) = 1\,, \tag{6.47}$$

where $\beta(X \to Y) \equiv \Gamma(X \to Y)/\Gamma_{\rm tot}$. The vector leptoquark singlet in Eq. (6.40) has a branching fraction of $\beta(\widetilde{U}_{1\alpha} \to t\mu) = 1$.

We derive constraints on the LQs in the $t\bar{t}\mu^+\mu^-$ channel by a recast of recent SUSY searches by ATLAS in the four-lepton [9] and same-sign di-lepton + tri-lepton channels [8]. We also derive bounds by a recast of an inclusive di-muon resonance search [3]. In order to estimate the number of signal events in each signal region, we first wrote UFO model files for R_2 and $\widetilde{U}_1^\alpha$ using `FeynRules` [33] and generated large LQ pair production samples in `MadGraph5` [35]. The decays of the tops into all channels were performed directly in `Pythia8` [36] as well as parton showering and hadronization. Finally, for each search, detector effects where simulated with `Delphes3` [38]. Selection cuts for the signal regions for each search were applied to the samples in order to extract the signal efficiencies.

For the scalar LQ pair production cross-section we used the NLO parametric representation given in [190]. For the vector leptoquark, the calculation of the pair production cross-section requires additional model-dependent assumptions about the underlying theory generating such state. The vector LQ-gluon interactions are parametrized by the following general terms in the Lagrangian [191]

$$\mathcal{L}_{\widetilde{U}-G} \;=\; -ig_s \left[\omega_G\, \widetilde{U}^\dagger_{1\alpha} G^{\alpha\beta} \widetilde{U}_{1\beta} \;+\; \frac{\lambda_G}{M^2_{\widetilde{U}}} \widetilde{U}^\dagger_{\sigma\mu} G^{\nu\sigma}\, \widetilde{U}^\mu_\nu \right], \tag{6.48}$$

where $G^{\alpha\beta}$ represents the gluonic field strength tensor and D_α is the SM gauge covariant derivative. The parameters ω_G and λ_G depend on the nature of the vector LQ. In our analysis we assume $\widetilde{U}_1^\alpha$ to be a fundamental gauge boson arising from an extended gauge group. This choice fixes $\omega_G = 1$ and $\lambda_G = 0$. For this benchmark, the production cross-section for the vector LQ was calculated with `MadGraph5` at leading order in QCD. We cross-checked our results with Ref. [192]. Note that the production cross-section for the vector LQ is a factor of $\sim \mathcal{O}(10)$ larger with respect to that of a scalar LQ with the same mass.

For the SUSY searches, we used the 95% CL limits provided by ATLAS on the number of allowed NP events in each signal region. Of all the SUSY searches, we found that the signal region `Rpc3L1bH` of the tri-lepton search [8] gives the best SUSY limits on the LQ masses: $M_R \gtrsim 1180$ GeV and $M_{\widetilde{U}} \gtrsim 1720$ GeV.

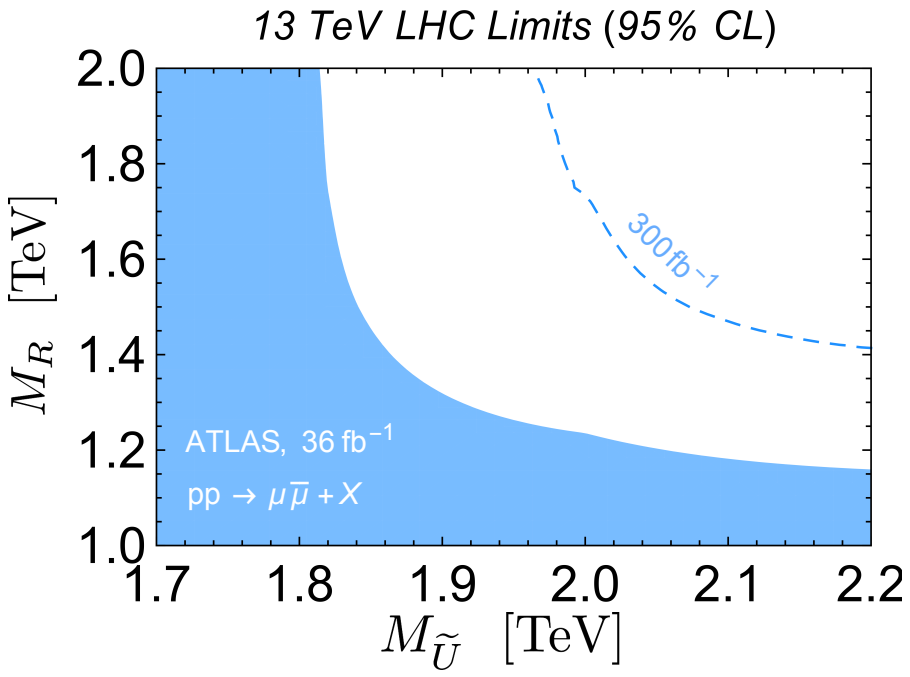


Figure 6.9: Excluded mass region for the R_2 plus $\tilde{U}_1$ model from the LHC di-muon tail search at 36.1 fb^{-1} . The dashed contour shows the projected limit to 300 fb^{-1} of LHC data.

Finally, we turn to the inclusive di-muon tail search [3]. The effect of the LQ resonant decay into $t\mu$ pairs is to modify the high-p_T tails of the di-muon invariant mass spectrum. We compare signal and background events above an invariant mass cut of $m_{\mu\mu} > 1200$ GeV (we find this value of the cut to be optimal for LQs above 1 TeV) and perform a statistical analysis by log-likelihood minimization to extract the 95% CL limits. In figure 6.9 we show the excluded region in the $\{M_{\tilde{U}}, M_R\}$ plane from this search at 36.1 fb^{-1} in red and a high luminosity projection with 300 fb^{-1} of data is given by the dashed red contour. Notice that having $M_R \sim M_{\tilde{U}}$, for example, is allowed for masses above 1.9 TeV. Below we give general bounds on scalar and vector LQs decaying to $t\mu$ as a function of the branching fraction β for one LQ at a time. These results from the SUSY tri-lepton search and the $pp \to \mu^+\mu^- + X$ tails give the most stringent bounds up to date for this particular decay channel. For the pair production of the scalar LQ component $R_2^{(2/3)}$, we use a dedicated search by CMS [193] in the $t\bar{t}\nu\bar{\nu}$ channel at 35.9 fb^{-1} . This search, however, sets a weaker limit on the R_2 mass, $M_R > 1020$ GeV.

General limits on leptoquark pair production in the $t\bar{t}\mu^+\mu^-$ channel

Above we provided limits for both leptoquarks in the special case where each branching ratio is equal to 1. Here we present for the sake of completion general limits for scalar and vector leptoquarks from pair production $pp \to \mathrm{LQ}^\dagger\mathrm{LQ} \to t\bar{t}\mu^+\mu^-$ and generic branching ratios $\beta \leq 1$. The 95% CL exclusion limits can be found in figure 6.10 for both LQs with mass M_{LQ} and branching ratio $\beta(\mathrm{LQ} \to t\mu)$. The solid lines represent the exclusion bounds from the searches with current luminosity

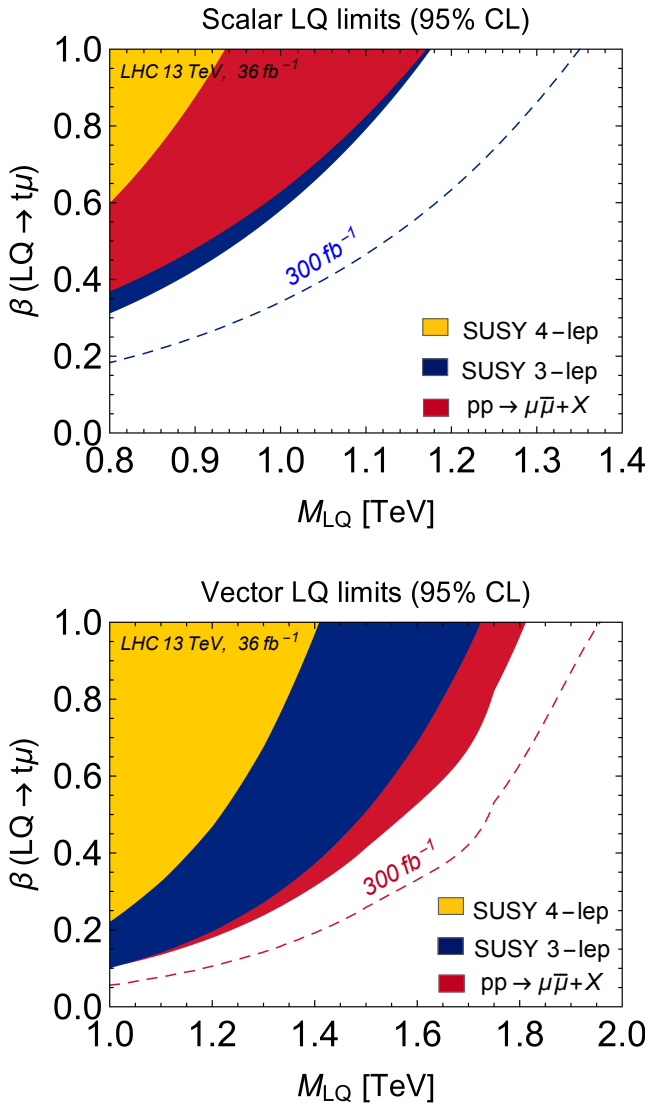


Figure 6.10: LHC bounds for a pair produced LQ decaying into the $t\bar{t}\mu^+\mu^-$ channel.

while the dashed lines are for a projected LHC luminosity of 300 fb^{-1} . The di-muon tail search produces the most stringent bounds for the vector leptoquark, while for the scalar leptoquark these limits are comparable with those coming from the SUSY tri-lepton search. These figures give an indication of how our results get modified when one allows for additional decay channels for the leptoquarks.

6.6 Discussions

Let us summarize and expand on what we have achieved so far.

• In Sec. 6.4 we proposed two possible top-philic scenarios that generate the pattern of NP for the B-anomalies at the one-loop level. The first model is based on the color-singlet Z' boson and the second on a combination of two (color-triplets) leptoquarks $\widetilde{U}_1 - R_2$. In principle, one could also consider other possible models mixing these two. For instance, a Z' with right-handed coupling to muons could be com-

bined with the scalar leptoquark R_2, in order to generate the two Wilson coefficients $\mathcal{C}_{lt}$ and $\mathcal{C}_{et}$. Alternatively, one could consider a Z' boson with left-handed coupling to muons combined with the vector leptoquark $\widetilde{U}_1$. One can even envision a model giving rise to the three states simultaneously.

• Ref. [172] performed a model independent analysis based on the SMEFT. It was advocated that having $\mathcal{C}_{lt} \sim -\mathcal{O}(1)$ for $\Lambda \sim 1$ TeV can provide a viable explanation of the $b \to s\ell^+\ell^-$ anomalies. In this work we have performed a similar analysis, including a more careful treatment of the LEP-I constraints. We have included the required one-loop matching corrections at the electroweak scale that are relevant to estimate $Z \to \mu^+\mu^-$ in this framework. From our analysis, we find that $\mathcal{C}_{lt} \sim -\mathcal{O}(1)$ for $\Lambda \sim 1$ TeV has some important tension with LEP-I measurements and a better solution is to have $\mathcal{C}_{lt} \sim \mathcal{C}_{et} \sim -\mathcal{O}(1)$ for $\Lambda \sim 1$ TeV.

We have verified that the finite corrections to $Z \to \mu^+\mu^-$ not included in Ref. [172] are small and cannot explain the discrepancy. We find that the reason of the discrepancy was the large correlation ($\rho = 0.9$) between $(\delta g_L, \delta g_R)$ from Eq. (6.17), which was not taken into account in [172] when using the bounds from [4]. To illustrate this, we show in figure 6.11 the values of $\Delta\chi^2 = \chi^2 - \chi^2_{\text{min}}$ as a function of the Wilson coefficient $\mathcal{C}_{lt}$ assuming $\mathcal{C}_{et} = 0$. Each sector included in the fit as well as the global χ^2 are shown. On the upper plot we show our results including the LEP-I measurements in (6.22) and (6.24). In the lower plot we show what happens when one uses instead the bounds from Ref. [4] for $(\delta g_L, \delta g_R)$, without taking into account the correlation. In Table 6.5 we show the values of $\mathcal{C}_{lt}$ at the minimum of the χ^2 for each sector, taking $\Lambda = 1$ TeV. When using the results from Ref. [4], missing the correlation between $(\delta g_L, \delta g_R)$ has the effect of reducing considerably the tension between LEP-I and $b \to s\ell^+\ell^-$. As remarked in Sec. 6.3, using the bounds from (6.22) and (6.24) leads to very similar results to taking the bounds on $(\delta g_L, \delta g_R)$ derived in Ref. [4] if the correlation is included.

		$b \to s\mu^+\mu^-$	$R_{K^{(*)}}$	$Z \to \ell^+\ell^-$	global
this work	$\mathcal{C}_{lt}$	-2.6	-2.4	-0.03	-0.75
	χ^2	5	2.1	0.5	36
Ref. [172]	$\mathcal{C}_{lt}$	-2.6	-2.4	0	-1.6
	χ^2	5	2.1	0	21.6

Table 6.5: Values of $\mathcal{C}_{lt}$ at the minimum of the χ^2 for each sector fixing $\Lambda = 1$ TeV.

• It is important to stress that using the results derived in Sec. 6.3 within the EFT framework to infer possible ultraviolet completions should be done carefully. We can illustrate the possible subtleties with an extension of the SM with a $\mathrm{U}(1)'$ gauge symmetry. After spontaneous symmetry breaking of the $\mathrm{U}(1)'$ symmetry, the Z' dynamics is described by the Lagrangian in

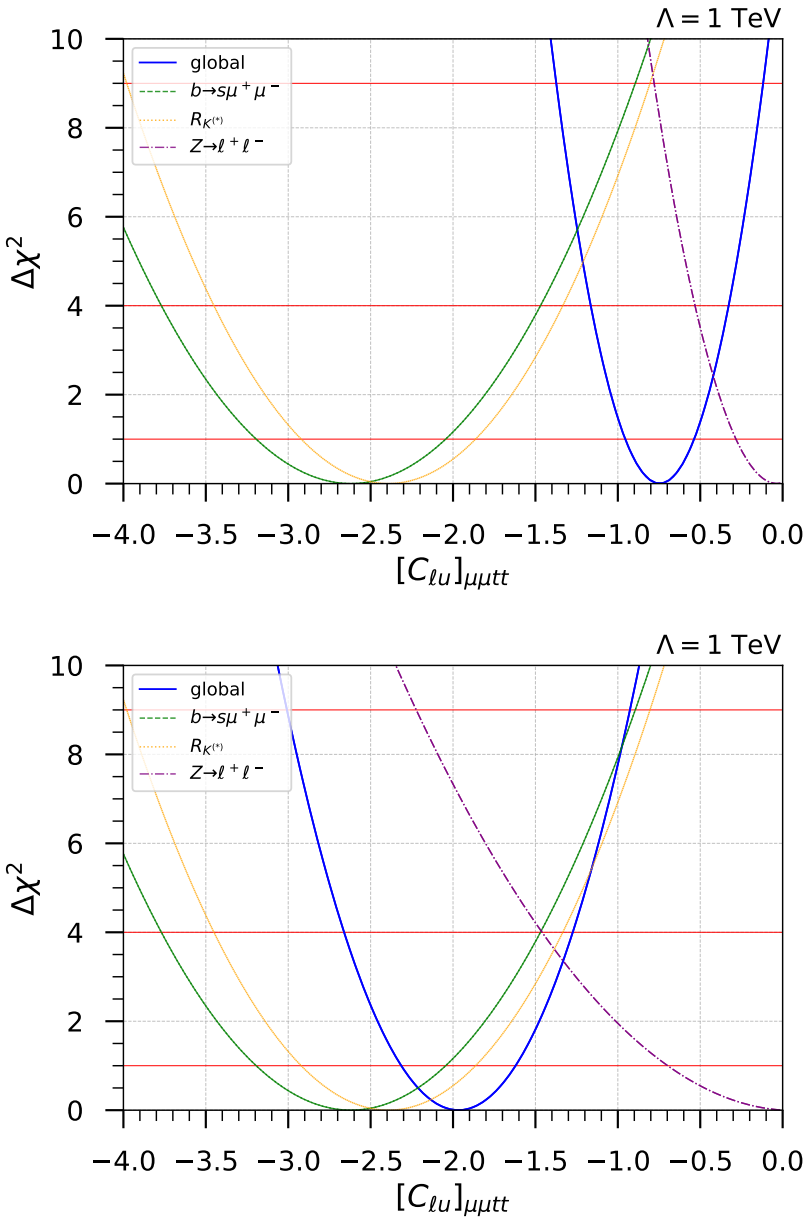


Figure 6.11: Values of $\Delta\chi^2 = \chi^2 - \chi^2_{\text{min}}$ against the Wilson coefficient $[\mathcal{C}_{lt}]_{\mu\mu tt}$, taking $[\mathcal{C}_{et}]_{\mu\mu tt} = 0$. Horizontal lines show the values of $\Delta\chi^2 = 1, 4, 9$. The high scale Λ has been fixed to 1 TeV. Upper plot: Results obtained in this work including the LEP-I measurements in (6.22) and (6.24). Lower plot: Results obtained using the bounds derived in Ref. [4] for $(\delta g_L, \delta g_R)$, but without taking into account the correlation.

$$\mathcal{L} = -\frac{1}{4} Z'^{\mu\nu} Z'_{\mu\ \nu} + \frac{1}{2} M^2_{Z'} Z'^{\mu} Z'_{\mu} - \frac{\kappa}{2} B^{\mu\nu} Z'_{\mu\nu} + g_{Z'} Z'_{\mu} J^{\mu}_{f}\,, \tag{6.49}$$

with $Z'_{\mu\nu} = \partial_\mu Z'_\nu - \partial_\nu Z'_\mu$, $g_{Z'}$ the U(1)$'$ gauge coupling, and J^μ_f representing the associated fermion current. The term proportional to κ is the kinetic mixing between the two abelian factors of the gauge group. Integrating out the Z' field at tree-level gives rise to the dimension six effective Lagrangian (see for instance [194])

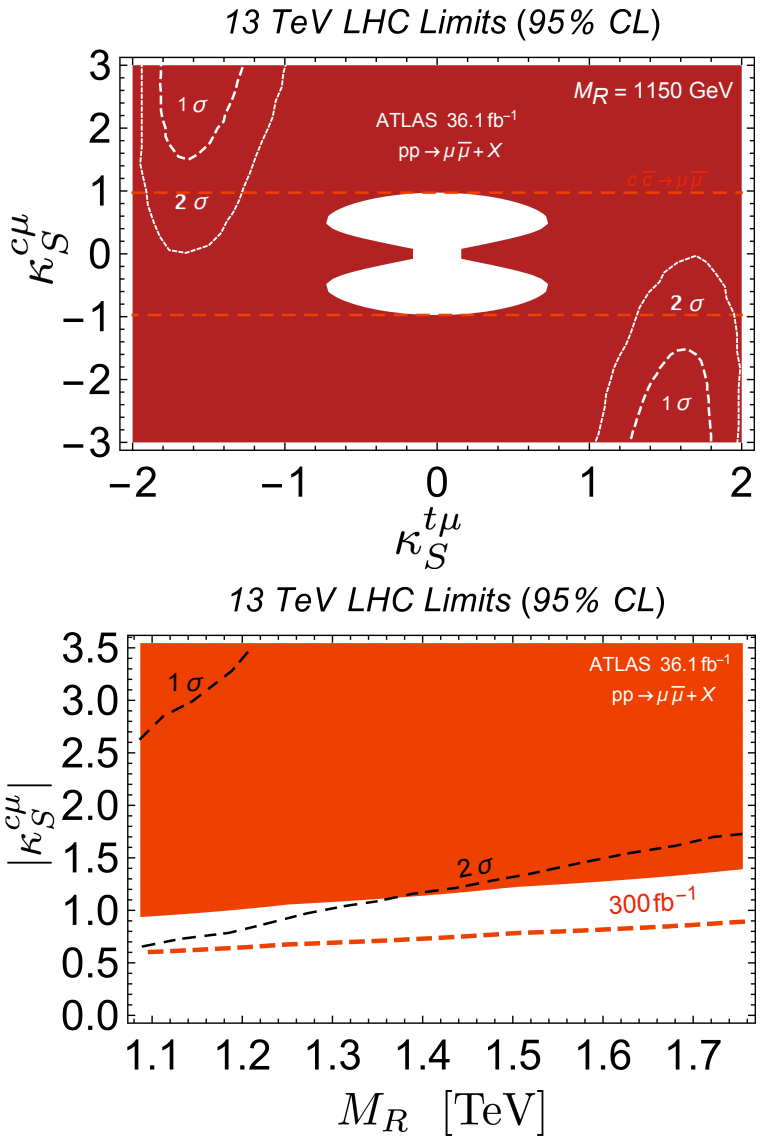


Figure 6.12: Scenario of R_2 coupling to charm and top motivated by [5]. Excluded regions by the LHC at 95% CL from a recast of a dimuon search at 36.1 fb^{-1}, including projections to 300 fb^{-1}. The preferred region by a global fit of $b \to s\ell^+\ell^-$ and LEP observables at 68% CL and 95% CL is shown by dashed contours.

$$\mathcal{L}_{\text{eff}} = \frac{g_{Z'}\kappa}{M_{Z'}^2}(\partial_\nu B^{\mu\nu})J_{f\mu} - \frac{g_{Z'}^2}{2M_{Z'}^2}(J_f^\mu)^2 - \frac{\kappa^2}{2M_{Z'}^2}(\partial_\nu B_{\mu\nu})^2\,, \tag{6.50}$$

which can be brought to the Warsaw operator basis using the SM equations of motion [66]. After doing this, one obtains matching contributions to the operators $(\varphi^\dagger i\overleftrightarrow{D}_\mu\varphi)(\bar{\ell}\gamma^\mu\ell)$ and $(\varphi^\dagger i\overleftrightarrow{D}_\mu\varphi)(\bar{e}\gamma^\mu e)$ which depend on the kinetic mixing parameter κ. This scenario lies outside of the framework assumed in this work, as these operators will contribute to Z-decay observables and compete with the loop-induced effects considered here.

However, if the kinetic gauge mixing parameter vanishes at the matching scale Λ, we can conclude from our analysis that a viable scenario would be a Z' boson with

vectorial coupling to muons. Having an explicit ultraviolet completion would allow us to calculate one-loop finite corrections to the matching at the high energy scale and test our assumption that the low-energy processes considered are dominated by logarithmic renormalization group evolution induced terms, such task is however beyond the scope of this work.

• It was originally proposed in [5] that the scalar leptoquark R_2 can accommodate $b \to s\ell^+\ell^-$ anomalies at the one-loop level. This scenario was also analyzed later in [195]. As we saw, the LQ R_2 only generates the operator $\mathcal{O}_{lt}$. It is worth noting that, as evidenced in figure 6.11, we can conclude that this scenario has an important tension with LEP-I measurements.

The model presented in [5] reduces this tension slightly by including a coupling of the leptoquark to the charm quark, besides the coupling to the top-quark. In this case, there is another relevant contribution to the effective Hamiltonian for $b \to s\ell^+\ell^-$ transitions from an operator involving both the charm and the top quark $\mathcal{C}_{lct} = (\bar{l}_2\gamma_\mu l_2)(\bar{t}_R\gamma^\mu c_R)$. This contribution can alleviate the tensions between $b \to s\ell^+\ell^-$ anomalies and LEP-I, however, since this new contribution is suppressed relative to the top-philic one $\mathcal{C}_{lt}$ by a factor $m_c/(m_t V_{tb}V_{ts}^*) \sim 1/6$, the required charm coupling of R_2 is larger than the top coupling in this model.

In the following, we show that high-p_T searches at the LHC set stringent constraints on this model excluding most of the preferred region by a global fit of $b \to s\ell^+\ell^-$ and LEP observables. Because of the large R_2 coupling to charm, the model predicts a large deviation in the high-p_T di-muon tails at the LHC [196]. For this we recast once again the inclusive $pp \to \mu^+\mu^- + X$ search by ATLAS [3] with the NP signal given by the combination of the t-channel exchange of $R_2^{(5/3)}$ in $c\bar{c} \to \mu^+\mu^-$ via the charm-muon Yukawa coupling and pair production of LQs decaying into $R_2^\dagger R_2 \to t\bar{t}\mu^+\mu^-$, $c\bar{t}\,(t\bar{c})\mu^+\mu^-$. We refer the reader to appendix B, where we have performed a full phenomenological analysis of all leptoquark models at the LHC for these two type of production processes.

We find that the R_2 model as an explanation of the $R_{K^{(*)}}$ and $b \to s\ell^+\ell^-$ anomalies is excluded for LQ masses below 1.15 TeV.[3] In figure 6.12 (left) we illustrate this with the dark red exclusion region at 95% CL for the benchmark $M_R = 1.15$ TeV in the Yukawa coupling plane $\{\kappa_S^{t\mu}, \kappa_S^{c\mu}\}$, following a notation analogous to (6.46) for the LQ couplings. The allowed region at 68% CL and 95% CL from a global fit to $b \to s\ell^+\ell^-$ and LEP-I observables are shown in figure 6.12 as dashed contours. The horizontal red dashed contours represent the limit extracted if we had only considered the t-channel $c\bar{c} \to \mu^+\mu^-$ in our analysis. Notice that including the final states $t\bar{t}\mu^+\mu^-$ and $c\bar{t}(t\bar{c})\mu^+\mu^-$ from pair production in the di-muon recast removes this flat direction in $\kappa_S^{t\mu}$. For a LQ mass above ~ 1.2 TeV, LQ pair production becomes negligible leaving only the t-channel mediated process $c\bar{c} \to \mu^+\mu^-$ as the only contribution to the di-muon tails. In figure 6.12 (right) we give the 95% exclusion regions in the $\{M_R, |\kappa_S^{c\mu}|\}$ plane for this scenario in orange. These bounds only rely on the size of the charm-muon coupling of R_2, so they apply to the model in [5]. Between $1.15 < M_R < 1.35$ TeV the allowed region at 95% CL from the low energy fit is not completely excluded by this LHC search. Our projections of the di-muon

[3]Here we assume all tau-lepton and down-type Yukawa couplings of R_2 to be zero while [5] does not make this assumption. Additional decay channels of R_2 into tau-leptons would reduce the branching fractions for the muonic decay channels making this bound weaker.

bound to $300\,\text{fb}^{-1}$ of LHC data, given by the dashed orange contour, cover this last piece of parameter space.

• As shown in Sec. 6.5.3, these strong tensions of R_2 with current LHC data can be avoided if one trades the dangerous couplings of R_2 to charm by a new vector LQ state $\widetilde{U}_1$ coupling to top. LQ pair production searches, shown in figure 6.9, put a current lower bound on both masses at about $M_{\widetilde{U}} \sim 1.9$ TeV and $M_R \sim 1.2$ TeV. For these masses, and for couplings of moderate size, the $\widetilde{U}_1 - R_2$ model can successfully explain the $b \to s\ell^+\ell^-$ anomalies without large tensions with high-p_T and low-energy observables. While the inclusive di-muon searches for LQ pair production is already giving relevant limits on this model, a dedicated search by the LHC for $t\mu$ resonances in $t\bar{t}\mu^+\mu^-$ final states will considerably improve them. In particular, the necessity for large couplings $|\kappa_{S,V}| \gtrsim 2$ to explain the B-anomalies singles out the single LQ production mode $pp \to \text{LQ}\,\mu t$ as an additional probe for this model.

6.7 Towards UV complete models

For each top-philic model identified in sec. 6.4, the low-energy phenomenology has forced us to introduce one massive vector boson (Z' or $\widetilde{U}_1$) with an explicit mass term leading to a non-renormalizable theory [4]. This implies that these models cannot be viewed as complete theories for the B-anomalies. Since we are proceeding in a data-driven bottom-up approach to BSM physics, this is not too worrying at this stage. In fact, the SM was historically conceived in the same way. The W boson was initially introduced as a massive vector boson with an explicit mass term in order to make sense of Fermi's four-fermion interactions describing nuclear β decay. Later on, this intermediate vector model was upgraded to a gauge boson in a UV theory with electroweak unification and spontaneous symmetry breaking via the Higgs mechanism, leading to renormalizable interactions and (tree-level) unitarity conservation at arbitrary energies.

Introducing a UV origin for the top-philic simplified models is the next logical step in our analysis. This is also necessary to perform a more detailed phenomenological analysis where one would need to calculate one-loop finite corrections to the matching at the high energy scale or include possible non-decoupling effects in order to confront theory with experiment.

In this final section, we present one class of UV completions for the top-philic models explaining the B-anomalies. We will focus primarily on the top-philic $\widetilde{U}_1 - R_2$ leptoquark model since UV theories for the top-philic Z' as a $\text{U}(1)'$ gauge boson have been constructed in sec. 5.5 and [156]. To our knowledge, no UV model for the vector leptoquark $\widetilde{U}_1$ has been proposed in the leptoquark literature until now. Interestingly, the UV model described here can be viewed as a non-abelian generalization of the top-philic $\text{U}(1)'$ model. The complete phenomenology of these theories will not be discussed here since it falls outside the main goal of this thesis (this is left for future work). Instead, our goal will be somewhat more modest and show that any $\text{SU}(2)_L$ – singlet vector leptoquark with non-universal interactions can arise as a gauge boson of a simple UV extension of the SM. We then show this explic-

[4] Contrary to the scalar leptoquarks.

fields	SU(N)	$SU(3)_{c'}$	$SU(2)_{L'}$	$U(1)_{Y'}$
Q^i	**1**	**3**	**2**	1/6
L^i	**1**	**1**	**2**	−1/2
u^i	**1**	**3**	**1**	2/3
d^i	**1**	**3**	**1**	−1/3
e^i	**1**	**1**	**1**	−1
ν_R^i	**1**	**1**	**1**	0
φ	**1**	**1**	**2**	1/2
$\Psi_{L,R}$	□	**1**	**a**	ψ
Ω	□	**b**	**c**	ω

Table 6.6: Matter content of the $\mathcal{G}(N)$ model. The new fields of the model are displayed in yellow shaded rows. The representations **a**, **b** and **c** and the primed hypercharges $\psi \equiv Y'(\Psi)$ and $\omega \equiv Y'(\Omega)$ of the fermion Ψ and scalar Ω are model building parameters to be fixed later on.

itly for the vector leptoquark $\widetilde{U}_\mu \sim (\mathbf{3}, \mathbf{1}, 5/3)$. In order to make contact with the B-anomalies, we also demonstrate that imposing a minimal flavor structure forces this vector leptoquark to couple predominantly to top-quarks and muons.

6.7.1 UV setup

We now describe the UV completion on general grounds. It is based on the semi-simple gauge group $\mathcal{G}(N)$ defined by:

$$\mathcal{G}(N) = \mathrm{SU}(N) \otimes \mathcal{G}_{\mathrm{SM}'} \tag{6.51}$$

$$\mathcal{G}_{\mathrm{SM}'} \equiv \mathrm{SU}(3)_{c'} \otimes \mathrm{SU}(2)_{L'} \otimes \mathrm{U}(1)_{Y'}\,. \tag{6.52}$$

This gauge group was recently discussed in [197, 198] in a very different context. Notice that the "primed SM" group $\mathcal{G}_{\mathrm{SM}'}$ in (6.52) does not necessarily correspond to the usual color, weak isospin and hypercharge groups of the SM. The SM group $\mathcal{G}_{\mathrm{SM}}$ is in fact embedded in $\mathcal{G}(N)$ and appear after spontaneous symmetry breaking above the electroweak scale. The usual SM quarks, leptons[5] and Higgs field φ are taken as singlets under the SU(N) group but are charged under $\mathcal{G}_{\mathrm{SM}'}$ with the conventional charges and representations. As a consequence, none of the gauge bosons related to the SU(N) group couple (directly) to the SM fields.

These couplings only emerge in the broken symmetry phase through fermionic mixing. For this, we introduce an additional heavy vector-like fermion Ψ in the fundamental representation □ of SU(N) and transforming as a singlet under $SU(3)_{c'}$. The dimension of the $SU(2)_{L'}$ representation, which we denote as **a**, and the primed hypercharge denoted as $\psi \equiv Y'(\Psi)$ are at this level model-building quantities that will be fixed later on. To break SU(N) we introduce at least one new scalar field Ω transforming in the fundamental representation of SU(N). The representation of this field under $\mathcal{G}_{\mathrm{SM}'}$ is taken as $(\mathbf{b}, \mathbf{c}, \omega)$. At this stage, the dimension of the representations, **b**, **c**, and the charges $\omega \equiv Y'(\Psi)$ are also model-building parameters to be fixed further on.

The matter content of our UV setup is given in table 6.6. Notice that more than one vector-like fermion Ψ and more than one scalar Ω can in principle be introduced if necessary. We can distinguish the following cases of interest:

[5] In our setup we also introduce RH neutrinos ν_R^i responsible for neutrino masses.

- If $N = 1$, we can identify $\mathcal{G}_{\rm SM'} = \mathcal{G}_{\rm SM}$ and $\mathrm{SU}(1) = \mathrm{U}(1)'$ is the abelian group. The fields Ψ and Ω must have non-vanishing charges under $\mathrm{U}(1)'$. This scenario gives rise to a Z' and no vector leptoquarks. We briefly discuss this model in sec. 6.7.4.

- If $N = 4$, then the weak isospin group can be trivially embedded in the full $\mathcal{G}(N)$ group as $\mathrm{SU}(2)_L = \mathrm{SU}(2)_{L'}$, while the SM color and hypercharge are non-trivially embedded inside $\mathrm{SU}(4) \otimes \mathrm{SU}(3)_{c'} \otimes \mathrm{U}(1)_{Y'}$. This scenario gives rise to gauge boson leptoquarks that are singlets of $\mathrm{SU}(2)_L$. We present in detail the model building for this case in sec. 6.7.2.

- If $N \geq 5$, then the whole SM can be non-trivially embedded in the full gauge group such that $\mathrm{SU}(2)_L \neq \mathrm{SU}(2)_{L'}$. In such a case the resulting gauge leptoquarks can in principle have higher dimensional $\mathrm{SU}(2)_L$ representations. For instance, if $N = 5$ then vector-leptoquark doublets such as $V_2 \sim (\mathbf{3}, \mathbf{2}, 5/6)$ and $\widetilde{V}_2 \sim (\bar{\mathbf{3}}, \mathbf{2}, -1/6)$ can potentially arise as gauge bosons. We will not discuss this case here since it is not relevant for the B-anomalies.

6.7.2 SU(4) models for vector leptoquarks

We now describe the simplest non-abelian scenario giving rise to a (TeV scale) vector leptoquark as a gauge boson with non-universal couplings to SM fermions. Since the gauge boson has to be colored this can only be achieved if the SM color group is non-trivially embedded in the subgroup $\mathrm{SU}(N) \otimes \mathrm{SU}(3)_{c'}$. We focus on the case $N = 4$ in our setup, hence

$$\mathcal{G}(4) \;=\; \mathrm{SU}(4) \otimes \mathrm{SU}(3)_{c'} \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1)_{Y'}\,, \tag{6.53}$$

As a consequence, vector leptoquarks arising from this case are necessarily weak singlets. Before we continue we set some notation. We denote the gauge couplings of each group factor in $\mathcal{G}(4)$ by g_4, g_3, g_2, g_1. The group SU(4) has associated 15 gauge fields that we denote H^A_μ with index $A = 1, ..., 15$ and its corresponding group generators Γ^A (see appendix). The remaining primed SM group $\mathrm{SU}(3)_{c'} \otimes \mathrm{SU}(2)_{L'} \otimes \mathrm{U}(1)_{Y'}$ have the more familiar gauge fields, G'^a_μ, W^i_μ and B'_μ and group generators by T^a (the Gell-man matrices), τ^i (the Pauli matrices), Y' with indices running as $a = 1, ..., 8$ and $i = 1, 2, 3$, respectively.

Generalities

After the spontaneous breaking of $\mathcal{G}(4) \to \mathcal{G}_{\rm SM}$ a rich spectrum of new massive gauge bosons arise [197]. In particular a vector leptoquark boson that we denote as $U \sim (\mathbf{3}, \mathbf{1}, p)$ with hypercharge p to be derived later on, a coloron $g' \sim (\mathbf{8}, \mathbf{1}, 0)$ and a vector singlet $Z' \sim (\mathbf{1}, \mathbf{1}, 0)$. The massless bosons of the SM also appear at this stage of the symmetry breaking, namely, the gluons and the neutral hypercharge boson B. To achieve the symmetry breaking of $\mathcal{G}(4)$ down to the SM we introduce two scalar field:

$$\Omega_3 \;\sim\; (\overline{\mathbf{4}}, \mathbf{3}, \mathbf{1}, \omega_3)\,, \tag{6.54}$$

$$\Omega_1 \;\sim\; (\overline{\mathbf{4}}, \mathbf{1}, \mathbf{1}, \omega_1)\,, \tag{6.55}$$

with primed hypercharges $\omega_{1,3}$, respectively. These charges will eventually be fixed to different values, each leading to different vector leptoquark models in the IR. The symmetry breaking pattern is given by

$$\mathrm{SU}(4) \otimes \mathrm{SU}(3)_{c'} \otimes \mathrm{U}(1)_{Y'} \quad \longrightarrow \quad \mathrm{SU}(3)_c \otimes \mathrm{U}(1)_Y \,, \tag{6.56}$$

and is accomplished through the vacuum expectation values [197, 199]

$$\langle \Omega_3^T \rangle = \frac{v_3}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \langle \Omega_1^T \rangle = \frac{v_1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} . \tag{6.57}$$

The scalar potential can be found in appendix E. The scalar Ω_3 is actually enough to produce the pattern (6.56). The second scalar Ω_1 is necessary in order to achieve the effective coupling between the vector LQ and the SM quark–lepton current. The breaking of the SM group $\mathcal{G}_{\rm SM} \to \mathrm{SU}(3)_c \otimes \mathrm{U}(1)_{\rm em}$ is then achieved with the Higgs field φ singlet under SU(4) and charged under $\mathcal{G}_{\rm SM'}$ in the ususal way $\varphi \sim (\mathbf{1}, \mathbf{1}, \mathbf{2}, 1/2)$ with an electroweak breaking vev at $\langle \varphi \rangle = v \approx 246$ GeV.

We now demand that the SM hypercharge Y remains unbroken after the first stage of symmetry breaking via (6.57), giving rise to the familiar massless boson B_μ of the SM. We find that this condition[6] is fulfilled only if the primed hypercharges of Ω_3 and Ω_1 satisfy the relation

$$\omega \;\equiv\; \omega_3 \;=\; -\frac{\omega_1}{3} \tag{6.58}$$

By annihilating the vacuum with a linear combination of the diagonal SU(4) generator Γ^{15} with the Y' generator one can express the SM hypercharge as function of ω:

$$Y \;=\; Y' \;+\; 2\sqrt{6}\,\omega\,\Gamma^{15}. \tag{6.59}$$

Notice that for the SM quarks and leptons, being SU(4) singlets, this formula simply reduces to $Y = Y'$. The SM gauge couplings for color (g_s) and hypercharge (g_Y) can be extracted by matching the couplings in the covariant derivative acting on an arbitrary SM fermion. This gives:

$$g_s \;=\; \frac{g_3 g_4}{\sqrt{g_3^2 + g_4^2}} \qquad \text{and} \qquad g_Y \;=\; \frac{g_1 g_4}{\sqrt{24\,\omega^2\, g_1^2 + g_4^2}}\,. \tag{6.60}$$

What we have presented so far has exactly the same structure as the gauge sector in the models described in Ref. [199, 200] for the charge assignment $\omega = 1/6$ for the primed hypercharge of Ω_3. Notice that the value of the couplings satisfy $g_4 \gg g_3 \sim g_s \gg g_1 \sim g_Y$. A typical benchmark is $g_4 \approx 3$, $g_3 \approx 1$ and $g_1 \approx 0.37$ [199]. For more details on the generalities of the SU(4), like e.g. the spectrum of the gauge boson masses, etc, see appendix E.

[6] which is equivalent to a vanishing determinant of the neutral boson mass matrix (appendix E).

Gauge leptoquarks in the $\mathrm{SU}(4)_R$ models

In order to generate non-vanishing couplings between the gauge leptoquark and the SM matter via fermionic mixing, we introduce only one vector-like fermion Ψ and fix the $\mathrm{SU}(2)_{L'}$ representation to be a singlet ($\mathbf{a} = \mathbf{1}$):

$$\Psi_{L,R} \;\sim\; (\mathbf{4}, \mathbf{1}, \mathbf{1}, \psi). \tag{6.61}$$

This fermion, being vectorial, can have a Dirac mass term $\mathcal{L} \supset -m_\Psi \bar{\Psi}\Psi$ in the Lagrangian, where the mass m_Ψ is a free parameter of the theory. Since Ψ is a singlet of $\mathrm{SU}(2)_L$, it will mostly mix with RH fields in the SM (we discuss mixing below). For this reason we refer to these models as $\mathrm{SU}(4)_R$ *models* and denote by U_R^μ the associated gauge leptoquark. Our model differs from [199, 200] in the weak isospin representation of the vector-like fermion (6.61) which is a doublet in their case ($\mathbf{a} = \mathbf{2}$). The fermionic Lagrangian is given by the three terms:

$$\mathcal{L} \;\supset\; \mathcal{L}_{\text{light}} \;+\; \mathcal{L}_{\text{heavy}} \;+\; \mathcal{L}_{\text{mix}} \tag{6.62}$$

where the first two Lagrangians, in the interaction basis, are defined by

$$\begin{aligned} \mathcal{L}_{\text{light}} \;=\;& \overline{Q^i} i \not{D} Q^i + \overline{u^i} i \not{D} u^i + \overline{d^i} i \not{D} d^i - (y^d_{ij} \overline{Q^i} \varphi d^j + y^u_{ij} \overline{Q^i} \tilde{\varphi} u^j + \text{h.c.}) \\ & \overline{L^i} i \not{D} L^i + \overline{e^i} i \not{D} e^i + \overline{\nu^i_R} i \not{D} \nu^i_R - (y^e_{ij} \overline{L^i} \varphi e^j + y^\nu_{ij} \overline{L^i} \tilde{\varphi} \nu^j_R + \text{h.c.}) \end{aligned} \tag{6.63}$$

$$\mathcal{L}_{\text{heavy}} \;=\; \overline{\Psi} \not{D} \Psi \;-\; m_\Psi (\overline{\Psi}_L \Psi_R + \overline{\Psi}_R \Psi_L)\,. \tag{6.64}$$

Here Q^i and L^i are the LH quark and lepton fields, u^i, d^i the RH up-quark and down-quark and e^i and ν^i_R the RH lepton and RH neutrino fields with family index $i = 1, 2, 3$, and H is the SM scalar doublet. The mixed light–heavy Yukawa Lagrangian in the same basis can be written on general grounds as

$$\mathcal{L}_{\text{mix}} = -\,\lambda^i_{\mathsf{q}} \,(\overline{\Psi}_L \,\Omega_3^\dagger \,\mathsf{q}^i_R) \;-\; \lambda^i_{\mathsf{e}} \,(\overline{\Psi}_L \,\Omega_1^\dagger \,\mathsf{e}^i_R) \;+\; \text{h.c.} \tag{6.65}$$

where $\lambda^i_{\mathsf{q},\mathsf{e}}$ are the Yukawa coupling of the RH quark fields $\mathsf{q}_R = \{u, d\}$ and RH lepton fields $\mathsf{e}_R = \{e, \nu_R\}$.

Not all Yukawa terms in $\mathcal{L}_{\text{mix}}$ are allowed because of $\mathrm{SU}(3)_{c'} \otimes \mathrm{U}(1)_{Y'}$ gauge invariance. The possibilities depend on the choices of the primed hypercharges ω and ψ of the scalar and the vector-like fermion. These have to satisfy simultaneously the equations

$$\omega \;=\; \frac{1}{4}\left(Y'_{\mathsf{q}} \,-\, Y'_{\mathsf{e}}\right) \tag{6.66}$$

$$\psi \;=\; \frac{1}{4}\left(3\,Y'_{\mathsf{q}} \,+\, Y'_{\mathsf{e}}\right) \tag{6.67}$$

where Y'_{q} and Y'_{e} are the $\mathrm{U}(1)_{Y'}$ charges of the quark q and lepton e fields, respectively. Each solution of (6.66)–(6.67) leads to a different Lagrangian $\mathcal{L}_{\text{mix}}$. We distinguish only four different model-building possibilities:

$$\omega = \frac{5}{12}\ ,\ \psi = \frac{1}{4} \quad \Longrightarrow \quad \mathcal{L}_{\rm mix} \supset -\lambda_u^i(\overline{\Psi}_L\Omega_3^\dagger u^i) - \lambda_e^i(\overline{\Psi}_L\Omega_1^\dagger e^i) + \text{h.c.} \tag{6.68}$$

$$\omega = \frac{1}{6}\ ,\ \psi = \frac{1}{2} \quad \Longrightarrow \quad \mathcal{L}_{\rm mix} \supset -\lambda_u^i(\overline{\Psi}_L\Omega_3^\dagger u^i) - \lambda_\nu^i(\overline{\Psi}_L\Omega_1^\dagger \nu_R^i) + \text{h.c.} \tag{6.69}$$

$$\omega = \frac{1}{6}\ ,\ \psi = -\frac{1}{2} \quad \Longrightarrow \quad \mathcal{L}_{\rm mix} \supset -\lambda_d^i(\overline{\Psi}_L\Omega_3^\dagger d^i) - \lambda_e^i(\overline{\Psi}_L\Omega_1^\dagger e^i) + \text{h.c.} \tag{6.70}$$

$$\omega = -\frac{1}{12}\ ,\ \psi = -\frac{1}{4} \quad \Longrightarrow \quad \mathcal{L}_{\rm mix} \supset -\lambda_d^i(\overline{\Psi}_L\Omega_3^\dagger d^i) - \lambda_\nu^i(\overline{\Psi}_L\Omega_1^\dagger \nu_R^i) + \text{h.c.} \tag{6.71}$$

Under the SM gauge group $\mathcal{G}_{\rm SM}$, the vector-like fermion Ψ can be decomposed into a color triplet component Ψ^{q} and a colorless component Ψ^{e} such that

$$\Psi_{L,R} \ = \ \begin{pmatrix} \Psi^{\mathsf{q}} \\ \\ \Psi^{\mathsf{e}} \end{pmatrix}_{L,R} \ . \tag{6.72}$$

The values of the hypercharges for each fermionic component can be extracted from (6.59), (6.66) and (6.67). This yields:

$$Y_{\Psi^{\mathsf{q}}} \ = \ \psi + \omega \ = \ Y'_{\mathsf{q}} \ = \ Y_{\mathsf{q}} \tag{6.73}$$

$$Y_{\Psi^{\mathsf{e}}} \ = \ \psi - 3\omega \ = \ Y'_{\mathsf{e}} \ = \ Y_{\mathsf{e}}\,. \tag{6.74}$$

We see that $\Psi^{\mathsf{q}} \sim (\mathbf{3},\mathbf{1},Y_{\mathsf{q}})^T$ is either a vector-like up-quark partner or vector-like down-quark partner that can mix with the RH light quarks u^i or d^i, depending on the values of $\{\psi,\omega\}$. In a similar fashion, the color singlet $\Psi^{\mathsf{e}} \sim (\mathbf{1},\mathbf{1},Y_{\mathsf{e}})^T$ is either a vector-like charged lepton or vector-like neutrino mixing with the RH leptons e^i or the RH neutrinos ν_R^i. From the Dirac Lagrangian in (6.64) we can extract the relevant leptoquark Lagrangian

$$i\,\overline{\Psi}\not{D}\Psi \ \supset \ \mathcal{L}_{LQ} \ = \frac{g_4}{\sqrt{2}}\,U_R^\mu \ \left(\overline{\Psi}_R^{\mathsf{q}}\gamma^\mu\Psi_R^{\mathsf{e}} \ + \ \overline{\Psi}_L^{\mathsf{q}}\gamma^\mu\Psi_L^{\mathsf{e}}\right) \ + \ \text{h.c.} \tag{6.75}$$

From this expression we see that the hypercharge $p \equiv Y_U$ of the gauge leptoquark is given exactly by $p = Y'_{\mathsf{q}} - Y'_{\mathsf{e}} = Y_{\mathsf{q}} - Y_{\mathsf{e}}$, hence related to the primed hypercharge ω via equation (6.66). We conclude that the representation of the gauge leptoquark in the SU(4) model has the general form

$$U_R^\mu \ \sim \ (\mathbf{3},\mathbf{1},4\,\omega) \tag{6.76}$$

under the SM gauge group. Notice that for each of the three possible primed hypercharge assignments for Ω_3 in equations (6.68)–(6.71) we obtain the three well-known (RH) vector singlet leptoquark models [7, 201]:

- $U_R = \widetilde{U}_1 \sim (\mathbf{3},\mathbf{1},5/3)$ if the primed hypercharge of Ω_3 satisfies $\omega = 5/12$. In such a model, the vector leptoquark couples to the heavy fermion partners of the u^i and e^i fields.

fields	$SU(4)_R$	$SU(3)_{c'}$	$SU(2)_L$	$U(1)_{Y'}$	$U_R = \widetilde{U}_1$	$U_R = U_1$	$U_R = \overline{U}_1$
$\Psi_{L,R}$	**4**	**1**	**1**	ψ	1/4	$\pm 1/2$	$-1/4$
Ω_3	$\bar{\mathbf{4}}$	**3**	**1**	ω	5/12	1/6	$-1/12$
Ω_1	$\bar{\mathbf{4}}$	**1**	**1**	$-3\,\omega$	$-5/4$	$-1/2$	1/4

Table 6.7: New (BSM) field content of the $SU(4)_R$ models. In each of the last three columns we specify the values in each row of the primed hypercharges ψ, $\omega_3 = \omega$ and $\omega_1 = -3\omega$ for each field that gives rise to a different gauge leptoquark representation: $\widetilde{U}_1 \sim (\mathbf{3}, \mathbf{1}, 5/3)$, $U_1 \sim (\mathbf{3}, \mathbf{1}, 2/3)$ or $\overline{U}_1 \sim (\mathbf{3}, \mathbf{1}, -1/3)$. See the text for more details.

- $U_R = U_1 \sim (\mathbf{3}, \mathbf{1}, 2/3)$ if the primed hypercharge of Ω_3 satisfies $\omega = 1/6$. In this model, the vector leptoquark either couples to the fermion partners of u^i and ν_R^i for $\psi = 1/2$, or couples to the fermion partners d^i and e^i for $\psi = -1/2$.

- $U_R = \overline{U}_1 \sim (\mathbf{3}, \mathbf{1}, -1/3)$ if the primed hypercharge of Ω_3 satisfies $\omega = -1/2$. In this model the vector leptoquark couples to the heavy partners of d^i and ν_R^i.

These three bosons $\widetilde{U}_1^\mu$, U_1^μ and $\overline{U}_1^\mu$ (for the notation see Ref. [7]) are the only three possible vector leptoquarks that are singlets under $SU(2)_L$. We summarize in table 6.7 the possible charge assignments leading to the four RH leptoquark in the $SU(2)_R$ model. Once full spontaneous symmetry breaking takes place, the heavy fermion partners mix with the (would-be) SM fields and induces the effective leptoquark couplings to the SM fields.

Gauge leptoquarks in $SU(4)_L$ models

Here we briefly comment on an alternative scenario. Above we purposefully chose the vector-like quark to be a singlet of $SU(2)_L$ in order to generate RH vector leptoquarks. One can in place change the weak isospin representation of the heavy fermion from a singlet ($\mathbf{b} = \mathbf{1}$) to be a doublet ($\mathbf{b} = \mathbf{2}$), namely,

$$\Psi_{L,R} \;\sim\; (\mathbf{4}, \mathbf{1}, \mathbf{2}, \psi). \tag{6.77}$$

This scenario gives rise to LHC chiral models, or $SU(4)_L$ *models*. In this case one can write light–heavy Yukawa interactions between Ψ_R and the left-handed would-be SM quark and lepton doublets, Q^i and L^i. In the interaction basis, this reads

$$\mathcal{L}_{\rm mix} = -\,\lambda_q^i\,(\overline{Q}^{\,i}\,\Omega_3^\dagger\,\Psi_R) \;-\; \lambda_\ell^i\,(\overline{L}^{\,i}\,\Omega_1^\dagger\Psi_R) \;+\; \text{h.c.} \tag{6.78}$$

Now there is only one model-building possibility. One finds that the primed hypercharges must necessarily satisfy $\{\omega, \psi\} = \{1/6, 0\}$. The resulting vector leptoquark in the SM group representation is $U_1 \sim (\mathbf{3}, \mathbf{1}, 2/3)$, this time with dominant couplings to LH SM quarks and leptons. The couplings arise from the fermion mixing of the light would-be SM states with heavy vector-like fermions $\Psi = (\Psi^q, \Psi^l)^T$. This particular model (dubbed "4321" model in the literature) equipped with a special flavor structure has been recently studied in [200, 202, 203] as a potential unified solution to the B-anomalies in both charged and neutral currents.

Fermionic mixing in $\mathrm{SU}(4)_R$

We now address fermionic mixing in the $\mathrm{SU}(4)_R$ models. After the final stage of symmetry breaking down to $\mathrm{SU}(3)_c \otimes \mathrm{U}(1)_{\rm em}$, the mass terms for the quark and lepton sectors read

$$\mathcal{L}_{\rm mass} \;=\; \mathcal{L}^{\mathsf{q}}_{\rm mass} \;+\; \mathcal{L}^{\mathsf{e}}_{\rm mass} \tag{6.79}$$

defined in the interaction basis (denoted below by primed fields) by

$$\mathcal{L}^{\mathsf{q}}_{\rm mass} \;=\; (\bar{\mathsf{q}}'_L, \bar{\Psi}'^{\mathsf{q}}_L) \begin{pmatrix} m_{\mathsf{q}} & 0 \\ \mu^\dagger_{\mathsf{q}} & m_\Psi \end{pmatrix} \begin{pmatrix} \mathsf{q}'_R \\ \Psi'^{\mathsf{q}}_R \end{pmatrix} \tag{6.80}$$

$$\mathcal{L}^{\mathsf{e}}_{\rm mass} \;=\; (\bar{\mathsf{e}}'_L, \bar{\Psi}'^{\mathsf{e}}_L) \begin{pmatrix} m_{\mathsf{e}} & 0 \\ \mu^\dagger_{\mathsf{e}} & m_\Psi \end{pmatrix} \begin{pmatrix} \mathsf{e}'_R \\ \Psi'^{\mathsf{e}}_R \end{pmatrix} . \tag{6.81}$$

We have defined the 3×3 Dirac mass blocks $m_q \equiv v\, y_{\mathsf{q}}/\sqrt{2}$ and $m_{\mathsf{e}} \equiv v\, y_{\mathsf{e}}/\sqrt{2}$, and the 3×1 mass blocks $\mu_{\mathsf{q}} = v_3 \lambda_{\mathsf{q}}/\sqrt{2}$ and $\mu_{\mathsf{e}} = v_1 \lambda_{\mathsf{e}}/\sqrt{2}$. Each of the mass matrices in (6.80) and (6.81) can be diagonalized by their own bi-unitary transformations $\mathcal{V}^{\mathsf{q}}_X$ and $\mathcal{V}^{\mathsf{e}}_X$, with $X = \{L, R\}$. These 4×4 mixing matrices can be parametrized in a very general way as a function of 3×3 (light–light) unitary mixing matrices V^{q}_X, V^{e}_X and 3×1 (light–heavy) mixing column matrices Θ^{q}_X, Θ^{e}_X into the the block decomposition:

$$\mathcal{V}^{\mathsf{q}}_X \;=\; \left(\begin{array}{c|c} \left[\mathbb{1}_{3\times 3} \;-\; \frac{\Theta^{\mathsf{q}}_X (\Theta^{\mathsf{q}}_X)^\dagger}{1+\cos\theta^{\mathsf{q}}_X}\right] V^{\mathsf{q}}_X & \Theta^{\mathsf{q}}_X \\ \hline -(\Theta^{\mathsf{q}}_X)^\dagger\, V^{\mathsf{q}}_X & \cos\theta^{\mathsf{q}}_X \end{array}\right) \tag{6.82}$$

with mixing angles θ^{q}_X defined via: $\sin(\theta^{\mathsf{q}}_X) \equiv \sqrt{(\Theta^{\mathsf{q}}_X)^\dagger \Theta^{\mathsf{q}}_X}$. See appendix C for a detailed derivation and discussion of this parametrization. The leptonic mixing matrices $\mathcal{V}^{\mathsf{e}}_X$ is defined by (6.82) by replacing $\mathsf{q} \leftrightarrow \mathsf{e}$. One of the main features of the $\mathrm{SU}(4)_R$ model is that the LH mixings $\Theta^{\mathsf{q},\mathsf{e}}_L$ turn out to be heavily suppressed, of order $\mathcal{O}(m_{\mathsf{q},\mathsf{e}}/m_\Psi)^2$, in the heavy vector-like fermion limit $m_\Psi \gg m^i_{\mathsf{q}}, m^i_{\mathsf{e}}$. We assume from now on this limit. For a detailed proof see appendix C. This implies that the gauge leptoquarks couples dominantly to RH fields.

We can write down the leptoquark interactions with fermions in the mass basis. We find that

$$\mathcal{L}_{\rm LQ} \;=\; \frac{g_4}{\sqrt{2}}\, U^\mu_R\, (\bar{\mathsf{q}}_X, \bar{\Psi}^{\mathsf{q}}_X)\, \gamma_\mu \cdot W_X \cdot \begin{pmatrix} \mathsf{e}_X \\ \Psi^{\mathsf{e}}_X \end{pmatrix} \;+\; \text{h.c.} \tag{6.83}$$

where $W_{L,R}$ is a 4×4 mixing matrix that measures the misalignment between quarks and leptons within the same multiplet. The matrix $W_{L,R}$ is analogous to the CKM matrix in the SM (which measures the misalignment between LH up and down quarks in the doublets) and the vector leptoquarks U_R plays the same role as the $W^\pm$ boson. Explicitly the W matrix is given by

$$W_X \;=\; \left(\begin{array}{c|c} W_X^{(\mathsf{qe})} & W_X^{(\mathsf{q}\Psi)} \\ \hline W_X^{(\mathsf{e}\Psi)^\dagger} & W_X^{(\Psi)} \end{array}\right) \;=\; \mathcal{V}_X^{\mathsf{q}} \left(\begin{array}{cc} \mathbb{0} & \mathbb{0} \\ \mathbb{0} & 1 \end{array}\right) (\mathcal{V}_X^{\mathsf{e}})^\dagger \,. \tag{6.84}$$

Here $W_{L,R}^{(\mathsf{qe})}$ are the 3×3 submatrices describing the induced interactions between the gauge leptoquark and the light–light SM currents $\bar{\mathsf{q}}_{L,R}\gamma^\mu \mathsf{e}_{LR}$. $W_X^{(\mathsf{q}\Psi)}$ and $W_X^{(\mathsf{e}\Psi)}$ correspond to 3×1 blocks describing the light–heavy couplings of the leptoquark while W_X^{Ψ} describe the coupling of to the vector-like quark and lepton partners.

The general expressions for the U_R couplings to SM fields of both chiralities are given as a function of the RH mixings $\Theta_R^{\mathsf{q},\mathsf{e}}$ and masses. A direct computation yields the expressions

$$W_R^{(\mathsf{qe})} \;=\; \Theta_R^{\mathsf{q}} \Theta_R^{\mathsf{e}\,\dagger} \tag{6.85}$$

$$W_L^{(\mathsf{qe})} \;=\; \frac{1}{\cos\theta_R^{\mathsf{q}} \cos\theta_R^{\mathsf{e}}} \left(\frac{m_t}{M_\Psi}\right)^2 (\Delta\hat{m}_{\mathsf{q}} \Theta_R^{\mathsf{q}} \Theta_R^{\mathsf{e}\,\dagger} \Delta\hat{m}_{\mathsf{e}}) \;+\; \mathcal{O}\left(\frac{m_t}{M_\Psi}\right)^3 \tag{6.86}$$

where M_Ψ is the physical mass of the heavy fermion and $\Delta\hat{m}_{\mathsf{q},\mathsf{e}} = \mathrm{diag}(m^1_{\mathsf{q},\mathsf{e}}, m^2_{\mathsf{q},\mathsf{e}}, m^3_{\mathsf{q},\mathsf{e}})/m_t$ are 3×3 dimensionless diagonal matrices that measure the mass hierarchy between SM fermions with respect to the heaviest fermion in the SM – the top-quark. Notice that the couplings of the gauge leptoquark to LH currents are heavily suppressed by the ratio $(m_t/M_\Psi)^2$ in the limit of large vector-like fermion mass, while the coupling to RH currents are unconstrained and can have large mixings of order $\mathcal{O}(1)$. Putting together expression (6.83) and (6.85) yields the effective leptoquark couplings in the mass eigenbasis to RH fields

$$\mathcal{L}_{U_R} \;=\; \frac{g_4}{\sqrt{2}} (\Theta_R^{\mathsf{q}} \Theta_R^{\mathsf{e}\,\dagger})_{ij}\, U_R^\mu\, (\bar{\mathsf{q}}_R^i \gamma_\mu \mathsf{e}_R^j) \;+\; \text{h.c.} \tag{6.87}$$

The Lagrangian (6.87) is general and describes the coupling of the RH leptoquark $U_R^\mu \sim (\mathbf{3},\mathbf{1},p)$ with $p = \{\frac{2}{3},\frac{5}{3},-\frac{1}{3}\}$ independent of $\{\psi,\omega\}$. When specifying the charges $\{\psi,\omega\}$ for each model-building scneario in (6.68)–(6.71), the expression (6.87) reduces to one of the following vector leptoquark model familiar from the literature:

$$\mathcal{L}_{\widetilde{U}} \;=\; \frac{g_4}{\sqrt{2}} (\Theta_R^{u} \Theta_R^{e\,\dagger})_{ij}\, \widetilde{U}_1^\mu\, (\overline{u}_R^i \gamma_\mu e_R^j) \;+\; \text{h.c.} \tag{6.88}$$

$$\mathcal{L}_{U} \;=\; \frac{g_4}{\sqrt{2}} (\Theta_R^{u} \Theta_R^{\nu\,\dagger})_{ij}\, U_1^\mu\, (\overline{u}_R^i \gamma_\mu \nu_R^j) \;+\; \text{h.c.} \tag{6.89}$$

$$\mathcal{L}_{U} \;=\; \frac{g_4}{\sqrt{2}} (\Theta_R^{d} \Theta_R^{e\,\dagger})_{ij}\, U_1^\mu\, (\overline{d}_R^i \gamma_\mu e_R^j) \;+\; \text{h.c.} \tag{6.90}$$

$$\mathcal{L}_{\overline{U}} \;=\; \frac{g_4}{\sqrt{2}} (\Theta_R^{d} \Theta_R^{\nu\,\dagger})_{ij}\, \overline{U}_1^\mu\, (\overline{d}_R^i \gamma_\mu \nu_R^j) \;+\; \text{h.c.} \tag{6.91}$$

Up to now, we haven't discussed the flavor structure of the model. For the moment the Yukawa couplings in (6.65) have been left completely generic meaning that the mixing pattern in (6.87) is still unspecified. Nevertheless, from these results we can expect the leptoquark couplings to be non-universal for generic mixing patterns.

6.7.3 Top-philic $\mathrm{SU}(4)_R$ model for the B-anomalies

We now focus on one specific $\mathrm{SU}(4)_R$ model that can give rise to the necessary degrees of freedom and pecularities of the $\widetilde{U}_1 - R_2$ model for the B-anomalies. Interestingly, this UV model turns out to be a non-abelian generalization of the top-philic $\mathrm{U}(1)'$ model from [156] discussed in sec. 5.5.

Of the four different $\mathrm{SU}(4)_R$ models described above we are primarily interested in the one giving rise to the gauge leptoquark $\widetilde{U}_1$ with hypercharge $p = 5/3$ coupling dominantly to RH up-quarks ($\mathtt{q} = u$) and charged leptons ($\mathtt{e} = e$). The heavy fermion decomposes as (6.72)

$$\Psi_{L,R} \;=\; \begin{pmatrix} \mathcal{T} \\ \mathcal{E} \end{pmatrix}_{L,R} \tag{6.92}$$

where in the SM group representation $\mathcal{T} \sim (\mathbf{3}, \mathbf{1}, 2/3)$ is a RH top-partner and $\mathcal{E} = (\mathbf{1}, \mathbf{1}, -1)$ a RH lepton partner. Each of these states mix with the (would-be) SM RH top-quark and RH muon fields, respectively.

We now specify the flavor structure of the Yukawa couplings in (6.65). We assume third generation dominance for quarks and second-generation dominance for leptons

$$\lambda_u \;=\; \begin{pmatrix} 0 \\ 0 \\ \star \end{pmatrix}, \qquad \lambda_e \;=\; \begin{pmatrix} 0 \\ \star \\ 0 \end{pmatrix}, \tag{6.93}$$

with the stars meaning entries of order $\mathcal{O}(1)$. This flavor structure reduces the general 4×4 fermion mixings down to a two-fermion mixing t–$\mathcal{T}$ (μ–$\mathcal{E}$) in the quark (lepton) sector, governed by one mixing angle θ^t_R (θ^μ_R). With this particular choice, we can make contact with the top-philic simplified $\widetilde{U}_1$ model that was used to explain the $R_{K^{(*)}}$ anomalies at the one-loop level in sec. 6.4. Once we include this particular flavor structure, and notice from Eq. (6.82) that $\Theta^u_R = (0, 0, \sin\theta^t_R)^T$ and $\Theta^e_R = (0, \sin\theta^\mu_R, 0)^T$, we find that the effective couplings of the leptoquark to the RH tops and muons is simply given by

$$\mathcal{L}_{\widetilde{U}} \;\supset\; \frac{g_4}{\sqrt{2}} \left(\sin\theta^t_R \sin\theta^\mu_R\right) \widetilde{U}_1^\mu \left(\bar{t}_R \gamma_\mu \mu_R\right) \;+\; \text{h.c.} \tag{6.94}$$

Here we have recovered the Lagrangian of the simplified model presented in (6.40) where we can make the identification $\kappa_V = g_4 \sin\theta^t_R \sin\theta^\mu_R/\sqrt{2}$. Integrating out the gauge leptoquark in (6.94) gives rise to the dimension-6 effective operator $\mathcal{O}_{et}$ valid below the cutoff scale set by $\Lambda = M_{\widetilde{U}}$.

We can also examine the vector leptoquark interactions with LH fields. A direct computation yields the highly suppressed interactions

$$\mathcal{L}_{\widetilde{U}} \;\supset\; \frac{g_4}{\sqrt{2}} \left(\tan\theta^t_R \tan\theta^\mu_R\right) \left(\frac{m_t}{M_\Psi}\right)^2 \widetilde{U}_1^\mu \left(\bar{t}_L \gamma_\mu \mu_L\right) \;+\; \text{h.c.} \tag{6.95}$$

that can be safely neglected for all practical purposes in the limit of heavy vector-like fermion.

Including the scalar leptoquark R_2

The model presented so far does not include the scalar leptoquark doublet $R_2 \sim (\mathbf{3}, \mathbf{2}, 7/6)$, necessary for generating the operator $\mathcal{O}_{lt}$ at lower energies. This can be achieved by introducing a new scalar field that we denote as Ω_2. This new scalar transforms as a bi-fundamental in $\mathrm{SU}(4)_R$ and $\mathrm{SU}(2)_L$:

$$\Omega_2 \;\sim\; (\overline{\mathbf{4}}, \mathbf{1}, \mathbf{2}, \omega_2)\,. \tag{6.96}$$

Depending on the value of the primed hypercharge ω_2, the new state can potentially couple to the SM LH doublets, either quarks $Q^i = (u^i, d^i)^T_L$ or leptons $L^i = (\nu^i, e^i)^T_L$. In order not to break the electroweak sector and potentially destabilize the vacuum structure of the scalar potential, this state is taken to have a trivial vacuum $\langle\, \Omega_2 \,\rangle = 0$, leaving the breaking of electroweak symmetry to the SM Higgs field (see appendix E). The mass M_{Ω_2} of the scalar is a new parameter of the model with a value to be chosen from phenomenological considerations.

After fixing the primed hypercharge to $\omega_2 = -3/4$ we obtain Yukawa interactions between the RH vector-like fermion $\Psi = (\mathcal{T}, \mathcal{E})^T$ and the LH lepton doublet

$$\mathcal{L}_{\text{mix}} \;\supset\; -\lambda^i_L \left(\overline{\Psi}_R\, \Omega_2\, (i\tau_2)\, L_i\right) \;+\; \text{h.c.}\,, \tag{6.97}$$

where λ_L is a new 3×1 Yukawa column matrix. The decomposition of the new scalar under the SM gauge group is $\Omega_2 \sim R_2 \oplus H_2$, where R_2 is a color triplet that we identify with the scalar leptoquark and H_2 a new color singlet. Both are $\mathrm{SU}(2)_L$ doublets with the following assignments $R_2 \sim (\mathbf{3}, \mathbf{2}, 7/6)$ and $H_2 \sim (\mathbf{1}, \mathbf{2}, -1/2)$ under the unbroken SM gauge group. After electroweak symmetry breaking the leptoquark itself decomposes into electric charge eigenstates as $R_2 = (R_2^{(5/3)}, R_2^{(2/3)})^T$. The interactions between the $R_2^{(5/3)}$ leptoquark and the fermionic currents in the mass eigenbasis are

$$\mathcal{L}_{R_2} \;\supset\; -R_2^{(5/3)}\, (\bar{u}_R, \bar{\mathcal{T}}_R) \cdot C \cdot \begin{pmatrix} e_L \\ \mathcal{E}_L \end{pmatrix} \;+\; \text{h.c.} \tag{6.98}$$

where the 4×4 matrix C is defined by

$$C \;=\; \left(\begin{array}{c|c} C^{(ue)} & C^{(u\Psi)} \\ \hline C^{(e\Psi)\dagger} & C^{(\Psi)} \end{array}\right) \;=\; \mathcal{V}^u_R \begin{pmatrix} \mathbb{0} & \mathbb{0} \\ \lambda_L & \mathbb{0} \end{pmatrix} (\mathcal{V}^e_L)^\dagger\,. \tag{6.99}$$

The 3×3 interaction submatrix of the leptoquark with SM currents is recovered from a straightforward calculation using Eq. (6.82):

$$C^{(q\ell)} \;=\; \Theta^u_R \cdot \lambda_L \cdot (V^e_L)^\dagger \left(\mathbb{1} \;-\; \frac{\Theta^e_L (\Theta^e_L)^\dagger}{1 + \cos\theta^e_L}\right) \tag{6.100}$$

$$\approx\; \Theta^u_R \cdot \lambda_L \cdot (V^e_L)^\dagger \tag{6.101}$$

where in the second line we neglected the small LH leptonic mixings. In order to induce couplings to LH muons alone, it is then necessary to assume a second-generation dominance flavor structure $\lambda_L = (0, \lambda_\mu, 0)^T$. With this flavor ansatz we recover the following top-philic leptoquark interactions:

fields	SU(4)	$SU(3)_{c'}$	$SU(2)_L$	$U(1)_{Y'}$
Q_L^i	**1**	**3**	**2**	1/6
L_L^i	**1**	**1**	**2**	−1/2
u_R^i	**1**	**3**	**1**	2/3
d_R^i	**1**	**3**	**1**	−1/3
e_R^i	**1**	**1**	**1**	−1
φ	**1**	**1**	**2**	1/2
$\Psi_{L,R}$	**4**	**1**	**1**	1/4
Ω_3	$\bar{\mathbf{4}}$	**3**	**1**	5/12
Ω_2	$\bar{\mathbf{4}}$	**1**	**2**	−3/4
Ω_1	$\bar{\mathbf{4}}$	**1**	**1**	−5/4

Table 6.8: Matter and scalar field content of the top-philic SU(4) model. The new fields are displayed in the yellow rows.

$$\mathcal{L}_{R_2} \supset -(\lambda_\mu \sin\theta_R^u)\, R_2^{(5/3)}(\bar{t}_R\, \mu_L) + \text{h.c.} \tag{6.102}$$

This is the same interaction Lagrangian in (6.40) with $\kappa_S = \lambda_\mu \sin\theta_R^u$. This model gives rise to the operator $\mathcal{O}_{lt}$ for the B-anomaly below the cutoff scale $\Lambda = M_R$. Similar interactions arise between the $R_2^{(2/3)}$ component and the scalar bilinear $(\bar{t}_R \nu_L)$. As for the color neutral second Higgs doublet H_2, this heavy state only interacts with second generation leptons via $H_2^-(\bar{\mu}_R\, \nu_\mu) + \text{h.c.}$ and $H_2^0(\bar{\mu}_R\, \mu_L)$ with couplings proportional to $\lambda_\mu \sin\theta_R^e$. One peculiarity of this UV model is that the R_2 leptoquark does not couple at all to LH quark fields because of gauge invariance. Notice that the interactions of the form $\bar{q}_i R_2 \ell_R$ (in principle allowed by SM gauge group) are forbidden because Ω_2 is a $SU(3)_{c'}$ singlet. In table. 6.8 we summarize the matter content of the top-philic $SU4_R$ model.

We finalize with some comments. In this UV complete model, additional massive gauge bosons appear at the same mass scale as the vector leptoquark. In particular, the gauge boson spectrum includes a top-philic coloron g' and a top-philic Z'. It would be important to revisit the low-energy fits taking into account these new degrees of freedom. For instance, the Z' couples with $\mathcal{O}(1)$ couplings to RH tops and RH muons giving an additional contribution to $\mathcal{O}_{et}$ besides the leptoquark. The vector-like fermion states also run inside the quantum loop in $b \to s\ell\ell$ giving additional contributions as well. LHC bounds must also be reanalyzed in more detail. This was done carefully in Ref. [200] for the $SU(4)_L$ model. There, the authors have shown that resonance searches in dijet and $t\bar{t}$ final states, as well as dilepton bounds from $pp \to Z' \to \ell^+\ell^-$, exclude masses for these states below ~ 3 TeV. In our model heavier leptoquark masses are necessary to evade these bounds. This would push up the values of the gauge couplings (for instance g_4) to higher values closer to the non-perturbative limit. We leave for future work a complete phenomenological analysis of the top-philic $SU(4)_R$ model and explorations of larger groups such as SU(5) and SU(6) leading to gauge vector leptoquarks with larger $SU(2)_L$ representation.

6.7.4 Top-philic U(1)′ model

An alternative UV model for the B anomalies is the top-philic Z' model. For a UV completion, the simplest scenario is to fix $N = 1$ in (6.51) and identify $\mathcal{G}_{\text{SM}'} = \mathcal{G}_{\text{SM}}$.

fields	$\mathrm{SU}(3)_c$	$\mathrm{SU}(2)_L$	$\mathrm{U}(1)_Y$	$\mathrm{U}(1)'$
Q_L^i	**3**	**2**	$1/6$	0
L_L^i	**1**	**2**	$-1/2$	0
u_R^i	**3**	**1**	$2/3$	0
d_R^i	**3**	**1**	$-1/3$	0
e_R^i	**1**	**1**	-1	0
φ	**1**	**2**	$1/2$	0
$\mathcal{T}_{L,R}$	**3**	**1**	$2/3$	$+1$
$\mathcal{E}_{L,R}$	**1**	**1**	-1	$+1$
Ω_1	**1**	**1**	0	$+1$
Ω_2	**1**	**2**	$1/2$	$+1$

Table 6.9: Matter and scalar field content of the top-philic $\mathrm{U}(1)'$ model. The new fields are displayed in the yellow rows.

This corresponds to the well known abelian extension of the SM,

$$\mathcal{G}(1) \;=\; \mathrm{SU}(3)_c \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y \otimes \mathrm{U}(1)' \,, \tag{6.103}$$

giving rise to an effective Z' gauge boson. In this model, the Z' boson couples predominantly to top-quarks through the fermion mixing mechanism described in sec. 5.5, where the heavy vector-like fermion is a top-partner with representation $\mathcal{T} \sim (\mathbf{3}, \mathbf{1}, 2/3, 1)$ under (6.103) and the scalar is the SM singlet $\Omega_1 \sim (\mathbf{1}, \mathbf{1}, 0, 1)$ (denoted as Φ in sec. 5.5). There we did not induce effective coulings of Z' to leptons. In order generate these to RH charged leptons it is necessary to introduce besides the top-partner, one additional heavy vector-like fermion $\mathcal{E} \sim (\mathbf{1}, \mathbf{1}, -1, Q'_{\mathcal{E}})$, i.e. a lepton-partner. The leptonic Lagrangian then reads

$$\mathcal{L}^{\ell}_{\text{light}} \;=\; \overline{L^i}(i\not{D}L^i) + \overline{e^i}(i\not{D})e^i - \left(y^e_{ij}\overline{L^i}\varphi e^j + \text{h.c.}\right) \tag{6.104}$$

$$\mathcal{L}^{\ell}_{\text{heavy}} \;=\; \overline{\mathcal{E}}(i\not{D})\mathcal{E} \;-\; M_{\mathcal{E}}\,\overline{\mathcal{E}}\mathcal{E} \tag{6.105}$$

$$\mathcal{L}^{\ell}_{\text{mix}} \;=\; -\,\lambda^i_{\mathcal{E}}\,\overline{\mathcal{E}_L}\Phi\, e^i_R + \text{h.c.} \tag{6.106}$$

where we have fixed the $\mathrm{U}(1)'$ of the new fields to satisfy $Q'_{\mathcal{E}} = Q'_{\Phi} = 1$ and $\lambda_{\mathcal{E}}$ are Yukawa couplings. This model gives rise at low energies to a scenario with $\mathcal{O}_{et}$ only, which is not favored by the low energy fit to the B-anomalies. Notice that to generate Z' couplings to LH charged leptons, and generate $\mathcal{O}_{lt}$, one would need to introduce a second scalar doublet with charges $\Omega_2 \sim (\mathbf{1}, \mathbf{2}, 1/2, 1)$, leading to the additional Yukawa interaction term $-\tilde{\lambda}^i_{\mathcal{E}}\,\overline{L_i}H\,\mathcal{E}^i_R$. We will not expand on the abelian scenario further on and refer the reader to the original material in Ref. [156]. In table. 6.9 we summarize the matter content of the top-philic $\mathrm{U}(1)'$ model.

Chapter 7

Concluding Remarks

For many decades, high-energy colliders have been the main tool for discovering New Physics. After operating for almost a decade, the LHC has produced a huge amount of data that up to now is compatible with the SM. Highly advertised theories from the past predicting new phenomena beyond the SM at the TeV scale are now almost excluded. The field of particle physics has entered into what seems like a vicious cycle: experiments lack concrete theoretical guidance to search for new phenomena, while theorists lack new data to guide the building of models beyond the SM necessary to give robust predictions for experimentalists.

The current situation seems to indicate that if BSM physics is at the reach of experiments then its effects might be hidden in non-standard signatures giving rise to non-resonant behavior at colliders or low-energy precision experiments. Since the top-quark is the least constrained colored fermion in the SM, in this thesis we have focused on analyzing NP scenarios with dominant couplings to top-quarks that give rise mostly to non-resonant phenomena.

The general strategy that we have adopted throughout this work is based on the bottom-up approach to BSM physics guided by data. The intended path we have taken towards discovery is summarized in the following: (i) propose new kinds of observables that characterize potential (non-resonant) NP in top-quark processes. (ii) Propose new experimental searches for measuring rare SM top-quark processes at hadron colliders with the aim of detecting deviations from the SM predictions. (iii) Use (i) and (ii) to analyze (non-resonant) possible BSM effects in real data (e.g. set exclusion limits or try to explain existing experimental anomalies not accounted by the SM) in a model-independent approach using effective field theories. (iv) Next, identify the quantum numbers and mass scale of the new (top-philic) degrees of freedom necessary for giving rise to the low-energy effective description explaining the data. (v) Confront these NP models with complementary data from other high-p_T colliders and low energy precision experiments. (vi) If the simplified dynamical models for the mediators are not excluded, build a UV complete model accounting for the new phenomena that are consistent with all of the data.

We now summarize the main results presented in this thesis:

- We have proposed a new class of observables designed to characterize multi-resonance process at hadron colliders. These *idealized observables* have most of the advantages present in the pseudo-observables framework used to analyze Higgs production at the LHC but are better suited to study more compli-

cated signatures like $pp \to t\bar{t}h$ production. Inspired by fat-jet taggers based on substructure techniques we partitioned the different kinematical regimes of $pp \to t\bar{t}h$ production and used this categorization to define a set of idealized cross-sections $\sigma(t_X t_Y h_Z)$ where $X, Y, Z \in \{B, R\}$ labels a resonance as boosted (B) or rest (R). As a proof-of-principle, we demonstrated how these idealized cross-sections are very sensitive to different directions in BSM parameter space. For this, we proved how they can be used for disentangling different contributions of the dimension-6 operators (the modified Yukawa operator, the chromo-magnetic top operator and a four-fermion contact interaction $qqtt$) entering $pp \to t\bar{t}h$ production. We then provided a realistic search strategy based for measuring the idealized cross-sections $\sigma(t_X t_Y h_Z)$ based on binning events by lepton multiplicity, narrow jet multiplicity and the number of top-tagged and Higgs-tagged fat-jets.

- We explored the possibility of measuring the rare SM process $pp \to t\bar{t}t\bar{t}$ at the LHC. For this, we proposed a simple search strategy combining both the same-sign dilepton and trilepton decay channels in order to enhance the significance of this signal. We made an exhaustive analysis of all possible reducible and irreducible backgrounds and provided reliable bounds for the SM production cross-section at current and future LHC luminosities. Our results showed that 3σ evidence of $\sigma_{\rm SM}(pp \to t\bar{t}t\bar{t})$ could be measured at run-II or run-III and a 5σ discovery of this landmark process at the LHC is possible. Indeed, very recently the CMS collaboration has announced the first evidence of four-top production based on the same search proposed in this thesis [186, 204]. In [204], CMS has managed to measure with a luminosity of 137 fb^{-1} the four-top cross-section

$$\sigma_{\rm SM}(pp \to t\bar{t}t\bar{t}) \;=\; 9.4^{+6.2}_{-5.6}\ \text{fb}\,. \tag{7.1}$$

  This result is in agreement with the SM prediction. We used our search strategy to set limits on non-resonant BSM physics, in particular on the couplings of a top-philic mediator with masses below the top-pair production threshold. We analyzed this for two simplified dynamical models: a top-philic vector boson (Z') and a top-philic scalar (Φ). We then went further up in the "bottom-up" ladder and proposed a very simple top-philic U(1)$'$ model giving rise to the Z' mediator as a gauge boson and Φ as a symmetry breaking scalar field, both coupling to the top-quark only via fermion mixing with a heavy top-partner $\mathcal{T}$. As it was demonstrated, this UV completion only gave rise to top-philic forces affecting RH top-quarks with very little effects in all other fermions. This implies that $pp \to t\bar{t}t\bar{t}$ is the best way to probe such theories.

- In the last chapter we proposed a viable explanation of the current hints of violation of lepton flavor universality in $b \to s\ell^+\ell^-$ transitions measured recently by the LHCb collaboration. We showed that two top-philic operators $\mathcal{O}_{lt} = (\bar{l}_2\gamma^\mu l_2)(\bar{t}_R\gamma_\mu t_R)$ and $\mathcal{O}_{et} = (\bar{\mu}\gamma^\mu\mu)(\bar{t}_R\gamma_\mu t_R)$ give rise to $\mathcal{O}_{9,10}$ necessary for explaining the B-meson anomalies at the one-loop level. Fits to $R_K^{(*)}$ and Z-pole observables require the muonic currents in the top-philic operators to

be vectorial. Next, we classified all possible NP mediators that generate the desired operators below the cutoff scale. We found that two top-philic scenarios are compatible with the data: one with a single vector mediator (the Z' model) and another with a combination of two leptoquarks (the $\widetilde{U}_1$–R_2 model); a vector singlet $\widetilde{U}_1 \sim (\mathbf{3}, \mathbf{1}, 5/3)$ and a scalar doublet $R_2 \sim (\mathbf{3}, \mathbf{2}, 7/6)$. We then studied the signatures of these mediators at the LHC in $pp \to t\bar{t}\mu^+\mu^-$, $pp \to t\bar{t}t\bar{t}$ showing that the LHC can probe the regions of parameter space relevant for fitting the B-anomaly. Since vector bosons like Z' and $\widetilde{U}_1$ need to be embedded in a renormalizable theory, we proposed a class of UV models giving rise to these vector states as gauge bosons of an extended gauge group. For this, we first constructed a model with gauge group $\mathrm{SU}(4) \otimes \mathrm{SU}(3) \otimes \mathrm{SU}(2) \otimes \mathrm{U}(1)$ with an extended fermion and scalar sector. By carefully selecting different U(1) charges for the new matter we showed that the resulting UV theories can give rise as a gauge boson any of the well-known vector leptoquark singlets (U_1, $\widetilde{U}_1$ $\overline{U}_1$) in the literature. These gauge leptoquarks were shown to couple predominantly to RH SM fields through fermion mixing in the broken group phase. We then used this general construction supplemented with a very simple flavor structure to write down a UV theory for the top-philic $\widetilde{U}_1$–R_2 simplified model explaining the flavor anomalies in $R_{K^{(*)}}$.

The work presented here is hopefully just the tip of the iceberg. Many interesting directions can be pursued.

For instance, we could perform a more complete analysis of the idealized cross-sections for $pp \to t\bar{t}h$ production by including the $t\bar{t}$ plus jet backgrounds. We could also use the idealized observables to characterize other interesting multi-resonance signals, like e.g. $t\bar{t}V$ ($V = W^{\pm}, Z$), etc. Another interesting possibility we wish to explore is to use these new observables as input to train supervised deep neural networks with the aim of extracting more precisely the Wilson coefficients of the EFT operators from several multi-resonance processes.

Given that four-top production will be measured at the LHC, it would be interesting to analyze this same process at future colliders like the HE-LHC (27 TeV) or at the 100 TeV super-colliders. At much higher energies, top-quarks become effectively massless states and the hadronic decay channels could be accessed using boosted techniques to select events with several top-tagged fat-jets. A dedicated search strategy for four-top production at future colliders would improve considerably the bounds on elusive NP mediator in UV theories like the top-philic U(1)′ model.

If in the next years the B-anomaly is confirmed then more efforts could be invested in probing the one-loop top-philic models. We provided in this work the first attempt to build UV complete models for any vector leptoquark that is a $\mathrm{SU}(2)_L$ singlet. It would be interesting to investigate in more detail the $\mathcal{G}(N) = \mathrm{SU}(N) \otimes \mathrm{SU}(3) \otimes \mathrm{SU}(2) \otimes \mathrm{U}(1)$ gauge theories for $N > 4$. For $N = 5$ ($N = 6$), these models can give rise to light gauge leptoquarks that are doublets (triplet) under $\mathrm{SU}(2)_L$. In particular, vector leptoquark doublets are usually considered very dangerous at the TeV scale because of their couplings to both leptoquark and diquark currents that typically lead to the decay of the proton at experimentally excluded rates. However, since the $\mathcal{G}(N)$ has a larger gauge structure than GUT models, one could in principle forbid the diquark couplings via gauge invariance by carefully

selecting the abelian charges of the heavy fermions and scalars. This interesting possibility remains to be explored in future work.

Bibliography

[1] CMS collaboration, C. Collaboration, *Search for SUSY in same-sign dilepton events at 13 TeV*, CMS-PAS-SUS-16-020.

[2] A. Angelescu, D. Becirevic, D. A. Faroughy and O. Sumensari, *Closing the window on single leptoquark solutions to the B-physics anomalies*, *JHEP* **10** (2018) 183, [1808.08179].

[3] ATLAS collaboration, M. Aaboud et al., *Search for new high-mass phenomena in the dilepton final state using 36 fb^1 of proton-proton collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **10** (2017) 182, [1707.02424].

[4] A. Efrati, A. Falkowski and Y. Soreq, *Electroweak constraints on flavorful effective theories*, *JHEP* **07** (2015) 018, [1503.07872].

[5] D. Becirevic and O. Sumensari, *A leptoquark model to accommodate $R_K^{\rm exp} < R_K^{\rm SM}$ and $R_{K^*}^{\rm exp} < R_{K^*}^{\rm SM}$*, *JHEP* **08** (2017) 104, [1704.05835].

[6] ATLAS collaboration, T. A. collaboration, *Search for supersymmetry with two same-sign leptons or three leptons using 13.2 fb^1 of $\sqrt{s} = 13$ TeV pp collision data collected by the ATLAS detector*, ATLAS-CONF-2016-037.

[7] I. Dorsner, S. Fajfer, A. Greljo, J. F. Kamenik and N. Kosnik, *Physics of leptoquarks in precision experiments and at particle colliders*, *Phys. Rept.* **641** (2016) 1–68, [1603.04993].

[8] ATLAS collaboration, M. Aaboud et al., *Search for supersymmetry in final states with two same-sign or three leptons and jets using 36 fb^{-1} of $\sqrt{s} = 13$ TeV pp collision data with the ATLAS detector*, *JHEP* **09** (2017) 084, [1706.03731].

[9] ATLAS collaboration, M. Aaboud et al., *Search for supersymmetry in events with four or more leptons in $\sqrt{s} = 13$ TeV pp collisions with ATLAS*, *Phys. Rev.* **D98** (2018) 032009, [1804.03602].

[10] ATLAS collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett.* **B716** (2012) 1–29, [1207.7214].

[11] CMS collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett.* **B716** (2012) 30–61, [1207.7235].

[12] ATLAS collaboration, M. Aaboud et al., *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, *Phys. Lett.* **B784** (2018) 173–191, [1806.00425].

[13] CMS collaboration, A. M. Sirunyan et al., *Observation of $t\bar{t}H$ production*, *Phys. Rev. Lett.* **120** (2018) 231801, [1804.02610].

[14] LHCb collaboration, R. Aaij et al., *Test of lepton universality using $B^+ \to K^+\ell^+\ell^-$ decays*, *Phys. Rev. Lett.* **113** (2014) 151601, [1406.6482].

[15] LHCb collaboration, R. Aaij et al., *Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays*, *JHEP* **08** (2017) 055, [1705.05802].

[16] LHCb collaboration, R. Aaij et al., *Measurement of Form-Factor-Independent Observables in the Decay $B^0 \to K^{*0}\mu^+\mu^-$*, *Phys. Rev. Lett.* **111** (2013) 191801, [1308.1707].

[17] LHCb collaboration, R. Aaij et al., *Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using 3 fb^{-1} of integrated luminosity*, *JHEP* **02** (2016) 104, [1512.04442].

[18] Belle collaboration, S. Wehle et al., *Lepton-Flavor-Dependent Angular Analysis of $B \to K^*\ell^+\ell^-$*, *Phys. Rev. Lett.* **118** (2017) 111801, [1612.05014].

[19] T. D. Lee and C.-N. Yang, *Question of Parity Conservation in Weak Interactions*, *Phys. Rev.* **104** (1956) 254–258.

[20] E. C. G. Sudarshan and R. e. Marshak, *Chirality invariance and the universal Fermi interaction*, *Phys. Rev.* **109** (1958) 1860–1860.

[21] H. Fritzsch, M. Gell-Mann and H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, *Phys. Lett.* **47B** (1973) 365–368.

[22] C.-N. Yang and R. L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, *Phys. Rev.* **96** (1954) 191–195.

[23] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, *Phys. Rev. Lett.* **30** (1973) 1343–1346.

[24] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, *Phys. Rev. Lett.* **30** (1973) 1346–1349.

[25] J. D. Bjorken and E. A. Paschos, *Inelastic Electron Proton and gamma Proton Scattering, and the Structure of the Nucleon*, *Phys. Rev.* **185** (1969) 1975–1982.

[26] J. C. Collins, D. E. Soper and G. F. Sterman, *Factorization of Hard Processes in QCD*, *Adv. Ser. Direct. High Energy Phys.* **5** (1989) 1–91, [hep-ph/0409313].

[27] G. Sterman and S. Weinberg, *Jets from quantum chromodynamics*, *Phys. Rev. Lett.* **39** (Dec, 1977) 1436–1439.

[28] G. P. Salam, *Towards Jetography*, *Eur. Phys. J.* **C67** (2010) 637–686, [0906.1833].

[29] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, *Longitudinally invariant K_t clustering algorithms for hadron hadron collisions*, *Nucl. Phys.* **B406** (1993) 187–224.

[30] M. Cacciari, G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, *JHEP* **04** (2008) 063, [0802.1189].

[31] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001, [hep-ph/9707323].

[32] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross-sections in deep inelastic scattering*, in *Monte Carlo generators for HERA physics. Proceedings, Workshop, Hamburg, Germany, 1998-1999*, pp. 270–279, 1998. hep-ph/9907280.

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

[34] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer and T. Reiter, *UFO - The Universal FeynRules Output*, *Comput. Phys. Commun.* **183** (2012) 1201–1214, [1108.2040].

[35] 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,"* *JHEP* **1407** (2014) 079, [1405.0301].

[36] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten et al., *An Introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159–177, [1410.3012].

[37] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, *Eur. Phys. J.* **C72** (2012) 1896, [1111.6097].

[38] DELPHES 3 collaboration, J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens et al., *DELPHES 3, A modular framework for fast simulation of a generic collider experiment*, *JHEP* **02** (2014) 057, [1307.6346].

[39] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, *Nucl. Instrum. Meth.* **A389** (1997) 81–86.

[40] E. Conte, B. Fuks and G. Serret, *MadAnalysis 5, A User-Friendly Framework for Collider Phenomenology*, *Comput. Phys. Commun.* **184** (2013) 222–256, [1206.1599].

[41] E. Fermi, *An attempt of a theory of beta radiation. 1.*, *Z. Phys.* **88** (1934) 161–177.

[42] S. L. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl. Phys.* **22** (1961) 579–588.

[43] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, *Phys. Lett.* **12** (1964) 132–133.

[44] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321–323.

[45] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, *Phys. Rev. Lett.* **13** (1964) 585–587.

[46] A. Salam, *Weak and Electromagnetic Interactions*, *Conf. Proc.* **C680519** (1968) 367–377.

[47] S. Weinberg, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (1967) 1264–1266.

[48] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, *Phys. Rev. Lett.* **10** (1963) 531–533.

[49] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652–657.

[50] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, *Phys. Rev. Lett.* **51** (1983) 1945.

[51] UTfit collaboration, M. Bona et al., *The Unitarity Triangle Fit in the Standard Model and Hadronic Parameters from Lattice QCD: A Reappraisal after the Measurements of Delta m(s) and BR(B —> tau nu(tau))*, *JHEP* **10** (2006) 081, [hep-ph/0606167].

[52] S. L. Glashow, J. Iliopoulos and L. Maiani, *Weak interactions with lepton-hadron symmetry*, *Phys. Rev. D* **2** (Oct, 1970) 1285–1292.

[53] Particle Data Group collaboration, K. A. Olive et al., *Review of Particle Physics*, *Chin. Phys.* **C38** (2014) 090001.

[54] Heavy Flavor Averaging Group (HFAG) collaboration, Y. Amhis et al., *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2014*, 1412.7515.

[55] A. Filipuzzi, J. Portoles and M. Gonzalez-Alonso, *$U(2)^5$ flavor symmetry and lepton universality violation in $W \to \tau\nu_\tau$*, *Phys. Rev.* **D85** (2012) 116010, [1203.2092].

[56] V. Cirigliano, M. Gonzalez-Alonso and M. L. Graesser, *Non-standard Charged Current Interactions: beta decays versus the LHC*, *JHEP* **02** (2013) 046, [1210.4553].

[57] M. Gonzalez-Alonso and J. Martin Camalich, *Global Effective-Field-Theory analysis of New-Physics effects in (semi)leptonic kaon decays*, *JHEP* **12** (2016) 052, [1605.07114].

[58] G. 't Hooft, *Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking*, *NATO Sci. Ser. B* **59** (1980) 135–157.

[59] Super-Kamiokande collaboration, Y. Fukuda et al., *Evidence for oscillation of atmospheric neutrinos*, *Phys. Rev. Lett.* **81** (1998) 1562–1567, [hep-ex/9807003].

[60] B. Pontecorvo, *Inverse beta processes and nonconservation of lepton charge*, *Sov. Phys. JETP* **7** (1958) 172–173.

[61] Z. Maki, M. Nakagawa and S. Sakata, *Remarks on the unified model of elementary particles*, *Prog. Theor. Phys.* **28** (1962) 870–880.

[62] E. E. Jenkins, A. V. Manohar and M. Trott, *Renormalization Group Evolution of the Standard Model Dimension Six Operators I: Formalism and lambda Dependence*, *JHEP* **10** (2013) 087, [1308.2627].

[63] E. E. Jenkins, A. V. Manohar and M. Trott, *Renormalization Group Evolution of the Standard Model Dimension Six Operators II: Yukawa Dependence*, *JHEP* **01** (2014) 035, [1310.4838].

[64] R. Alonso, E. E. Jenkins, A. V. Manohar and M. Trott, *Renormalization Group Evolution of the Standard Model Dimension Six Operators III: Gauge Coupling Dependence and Phenomenology*, *JHEP* **04** (2014) 159, [1312.2014].

[65] S. Weinberg, *Baryon and Lepton Nonconserving Processes*, *Phys. Rev. Lett.* **43** (1979) 1566–1570.

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

[67] C. Arzt, M. B. Einhorn and J. Wudka, *Patterns of deviation from the standard model*, *Nucl. Phys.* **B433** (1995) 41–66, [hep-ph/9405214].

[68] D. Barducci et al., *Interpreting top-quark LHC measurements in the standard-model effective field theory*, 1802.07237.

[69] ATLAS collaboration, T. A. collaboration, *Search for the Associated Production of a Higgs Boson and a Top Quark Pair in Multilepton Final States with the ATLAS Detector*, ATLAS-CONF-2016-058.

[70] ATLAS collaboration, T. A. collaboration, *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, ATLAS-CONF-2016-080.

[71] ATLAS collaboration, T. A. collaboration, *Measurement of fiducial, differential and production cross sections in the $H \to \gamma\gamma$ decay channel with 13.3 fb^{-1} of 13 TeV proton-proton collision data with the ATLAS detector*, ATLAS-CONF-2016-067.

[72] ATLAS collaboration, T. A. collaboration, *Combination of the searches for Higgs boson production in association with top quarks in the $\gamma\gamma$, multilepton, and $b\bar{b}$ decay channels at $\sqrt{s}$=13 TeV with the ATLAS Detector*, ATLAS-CONF-2016-068.

[73] CMS collaboration, C. Collaboration, *Search for associated production of Higgs bosons and top quarks in multilepton final states at $\sqrt{s} = 13$ TeV*, CMS-PAS-HIG-16-022.

[74] CMS collaboration, C. Collaboration, *Search for $t\bar{t}H$ production in the $H \to b\bar{b}$ decay channel with 2016 pp collision data at $\sqrt{s} = 13$ TeV*, CMS-PAS-HIG-16-038.

[75] CMS collaboration, C. Collaboration, *Search for Higgs boson production in association with top quarks in multilepton final states at $\sqrt{s} = 13$ TeV*, CMS-PAS-HIG-17-004.

[76] D. Yu. Bardin, M. Grunewald and G. Passarino, *Precision calculation project report*, hep-ph/9902452.

[77] M. Gonzalez-Alonso, A. Greljo, G. Isidori and D. Marzocca, *Pseudo-observables in higgs decays*, *Eur. Phys J. C128* **75** (2015) , [1412.6038].

[78] A. Greljo, G. Isidori, J. M. Lindert and D. Marzocca, *Pseudo-observables in electroweak higgs production*, *Eur. Phys J. Cno.3, 158* **76** (2016) , [1512.06135].

[79] D. de Florian, *et al*, *[LHC Higgs Cross Section Working Group], "Handbook of LHC Higgs Cross Sections 4: Deciphering the Nature of the Higgs Sector,"* , [1610.07922].

[80] J. M. Butterworth, A. R. Davison, M. Rubin and G. P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, *Phys. Rev. Lett.* **100** (2008) 242001, [0802.2470].

[81] J. Thaler and K. Van Tilburg, *Identifying Boosted Objects with N-subjettiness*, *JHEP* **03** (2011) 015, [1011.2268].

[82] D. E. Kaplan, K. Rehermann, M. D. Schwartz and B. Tweedie, *Top Tagging: A Method for Identifying Boosted Hadronically Decaying Top Quarks*, *Phys. Rev. Lett.* **101** (2008) 142001, [0806.0848].

[83] T. Plehn, G. P. Salam and M. Spannowsky, *Fat jets for a light higgs*, *Phys. Rev Lett.* **104** (2010) , [0910.5472].

[84] T. Plehn, M. Spannowsky, M. Takeuchi and D. Zerwas, *Stop Reconstruction with Tagged Tops*, *JHEP* **10** (2010) 078, [1006.2833].

[85] M. R. Buckley, T. Plehn, T. Schell and M. Takeuchi, *Buckets of higgs and tops*, *JHEP* **1402** (2014) 130, [1310.6034].

[86] N. Moretti, P. Petrov, S. Pozzorini and M. Spannowsky, *Measuring the signal strength in $t\bar{t}h$ with $h?b\bar{b}$*, *Phys. Rev D.* **93** (2016) , [1510.08468].

[87] C. Degrande, J. M. Gerard, C. Grojean, F. Maltoni and G. Servant, *Probing top-higgs non-standard interactions at the lhc*, *JHEP* **1207** *(2012) 036 Erratum: [JHEP* **1303** *(2013) 032]* (2013) , [1205.1065].

[88] F. Maltoni, E. Vryonidou and C. Zhang, *Higgs production in association with a top-antitop pair in the standard model effective field theory at nlo in qcd*, *JHEP* **1610** (2016) 123, [1607.05330].

[89] C. Zhang and S. Willenbrock, *Effective-field-theory approach to top-quark production and decay*, *Phys. Rev* **83** (2011) , [1008.3869].

[90] P. L. Cho and E. H. Simmons, *Searching for g3 in $t\bar{t}$ production*, *Phys. Rev* **51** (1995) , [9408206].

[91] F. Krauss, S. Kuttimalai and T. Plehn, *Lhc multijet events as a probe for anomalous dimension-six gluon interactions*, *Phys. Rev D no.3, 035024* **95** (2017) , [1611.00767].

[92] R. Contino, A. Falkowski, F. Goertz, C. Grojean and F. Riva, *On the validity of the effective field theory approach to sm precision tests*, *JHEP* **1607** (2016) 144, [1604.06444].

[93] C. Zhang, *Constraining qqtt operators from four-top production: a case for enhanced EFT sensitivity*, *Chin. Phys.* **C42** (2018) 023104, [1708.05928].

[94] K. Rehermann and B. Tweedie, *Efficient Identification of Boosted Semileptonic Top Quarks at the LHC*, *JHEP* **03** (2011) 059, [1007.2221].

[95] ATLAS collaboration, T. A. collaboration, *Search for Higgs boson production via weak boson fusion and decaying to $b\bar{b}$ in association with a high-energy photon in the ATLAS detector*, ATLAS-CONF-2016-063.

[96] ATLAS collaboration, G. Aad et al., *Constraints on non-Standard Model Higgs boson interactions in an effective Lagrangian using differential cross sections measured in the $H \to \gamma\gamma$ decay channel at $\sqrt{s} = 8\,TeV$ with the ATLAS detector*, *Phys. Lett.* **B753** (2016) 69–85, [1508.02507].

[97] ATLAS collaboration, G. Aad et al., *Measurements of the Total and Differential Higgs Boson Production Cross Sections Combining the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ Decay Channels at $\sqrt{s}$=8 TeV with the ATLAS Detector*, *Phys. Rev. Lett.* **115** (2015) 091801, [1504.05833].

[98] CMS collaboration, V. Khachatryan et al., *Measurement of differential and integrated fiducial cross sections for Higgs boson production in the four-lepton decay channel in pp collisions at $\sqrt{s} = 7$ and 8 TeV*, *JHEP* **04** (2016) 005, [1512.08377].

[99] CMS collaboration, V. Khachatryan et al., *Measurement of differential cross sections for Higgs boson production in the diphoton decay channel in pp collisions at $\sqrt{s} = 8$ TeV*, *Eur. Phys. J.* **C76** (2016) 13, [1508.07819].

[100] CMS collaboration, C. Collaboration, *Search for Higgs boson pair production in the* $b\bar{b}l\nu l\nu$ *final state at* $\sqrt{s} = 13$ TeV, CMS-PAS-HIG-16-024.

[101] ATLAS collaboration, M. Aaboud et al., *Search for pair production of Higgs bosons in the* $b\bar{b}b\bar{b}$ *final state using proton–proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, *Phys. Rev.* **D94** (2016) 052002, [1606.04782].

[102] CMS collaboration, V. Khachatryan et al., *Search for two Higgs bosons in final states containing two photons and two bottom quarks in proton-proton collisions at 8 TeV*, *Phys. Rev.* **D94** (2016) 052012, [1603.06896].

[103] ATLAS collaboration, T. A. collaboration, *Search for the Standard Model Higgs boson produced in association with a vector boson and decaying to a* $b\bar{b}$ *pair in pp collisions at 13 TeV using the ATLAS detector*, ATLAS-CONF-2016-091.

[104] CMS collaboration, C. Collaboration, *Search for H to bbar in association with a single top quark as a test of Higgs boson couplings at 13 TeV*, CMS-PAS-HIG-16-019.

[105] ATLAS collaboration, T. A. collaboration, *Measurements of top-quark pair differential cross-sections in the lepton+jets channel in pp collisions at* $\sqrt{s} = 13$ *TeV using the ATLAS detector*, ATLAS-CONF-2016-040.

[106] CMS collaboration, C. Collaboration, *Measurement of particle level differential ttbar cross sections in the dilepton channel at sqrt(s) = 13 TeV*, CMS-PAS-TOP-16-007.

[107] ATLAS collaboration, M. Aaboud et al., *Measurement of top quark pair differential cross-sections in the dilepton channel in pp collisions at* $\sqrt{s} = 7$ *and 8 TeV with ATLAS*, *Phys. Rev.* **D94** (2016) 092003, [1607.07281].

[108] CMS collaboration, C. Collaboration, *Measurement of double differential cross sections for top quark pair production in pp collisions at sqrt[s] = 8 TeV*, CMS-PAS-TOP-14-013.

[109] ATLAS collaboration, T. A. collaboration, *Measurement of the cross-section of the production of a W boson in association with a single top quark with ATLAS at* $\sqrt{s} = 13\,\text{TeV}$, ATLAS-CONF-2016-065.

[110] CMS collaboration, C. Collaboration, *Combination of cross-section measurements of associated production of a single top quark and a W boson at* $\sqrt{s} = 8$ TeV *with the ATLAS and CMS experiments*, CMS-PAS-TOP-15-019.

[111] T. A. collaboration, *Measurement of the* $t\bar{t}Z$ *and* $t\bar{t}W$ *production cross sections in multilepton final states using 3.2* fb^{-1} *of pp collisions at 13 TeV at the LHC*, ATLAS-CONF-2016-003.

[112] CMS collaboration, C. Collaboration, *Measurement of the top pair-production in association with a W or Z boson in pp collisions at 13 TeV*, CMS-PAS-TOP-16-017.

[113] B. Lillie, J. Shu and T. M. P. Tait, *Top Compositeness at the Tevatron and LHC*, *JHEP* **04** (2008) 087, [0712.3057].

[114] A. Pomarol and J. Serra, *Top Quark Compositeness: Feasibility and Implications*, *Phys. Rev.* **D78** (2008) 074026, [0806.3247].

[115] K. Kumar, T. M. P. Tait and R. Vega-Morales, *Manifestations of Top Compositeness at Colliders*, *JHEP* **05** (2009) 022, [0901.3808].

[116] G. Cacciapaglia, R. Chierici, A. Deandrea, L. Panizzi, S. Perries and S. Tosi, *Four tops on the real projective plane at LHC*, *JHEP* **10** (2011) 042, [1107.4616].

[117] M. Perelstein and A. Spray, *Four boosted tops from a Regge gluon*, *JHEP* **09** (2011) 008, [1106.2171].

[118] J. A. Aguilar-Saavedra and J. Santiago, *Four tops and the $t\bar{t}$ forward-backward asymmetry*, *Phys. Rev.* **D85** (2012) 034021, [1112.3778].

[119] L. Beck, F. Blekman, D. Dobur, B. Fuks, J. Keaveney and K. Mawatari, *Probing top-philic sgluons with LHC Run I data*, *Phys. Lett.* **B746** (2015) 48–52, [1501.07580].

[120] P. S. Bhupal Dev and A. Pilaftsis, *Maximally Symmetric Two Higgs Doublet Model with Natural Standard Model Alignment*, *JHEP* **12** (2014) 024, [1408.3405].

[121] G. Bevilacqua and M. Worek, *Constraining BSM Physics at the LHC: Four top final states with NLO accuracy in perturbative QCD*, *JHEP* **07** (2012) 111, [1206.3064].

[122] M. Gerbush, T. J. Khoo, D. J. Phalen, A. Pierce and D. Tucker-Smith, *Color-octet scalars at the CERN LHC*, *Phys. Rev.* **D77** (2008) 095003, [0710.3133].

[123] B. S. Acharya, P. Grajek, G. L. Kane, E. Kuflik, K. Suruliz and L.-T. Wang, *Identifying Multi-Top Events from Gluino Decay at the LHC*, 0901.3367.

[124] T. Gregoire, E. Katz and V. Sanz, *Four top quarks in extensions of the standard model*, *Phys. Rev.* **D85** (2012) 055024, [1101.1294].

[125] D. Liu and R. Mahbubani, *Probing top-antitop resonances with $t\bar{t}$ scattering at LHC14*, *JHEP* **04** (2016) 116, [1511.09452].

[126] S. Gori, I.-W. Kim, N. R. Shah and K. M. Zurek, *Closing the Wedge: Search Strategies for Extended Higgs Sectors with Heavy Flavor Final States*, *Phys. Rev.* **D93** (2016) 075038, [1602.02782].

[127] J. H. Kim, K. Kong, S. J. Lee and G. Mohlabeng, *Probing TeV scale Top-Philic Resonances with Boosted Top-Tagging at the High Luminosity LHC*, *Phys. Rev.* **D94** (2016) 035023, [1604.07421].

[128] CMS collaboration, V. Khachatryan et al., *Search for Standard Model Production of Four Top Quarks in the Lepton + Jets Channel in pp Collisions at* $\sqrt{s} = 8$ *TeV*, *JHEP* **11** (2014) 154, [1409.7339].

[129] ATLAS collaboration, G. Aad et al., *Search for production of vector-like quark pairs and of four top quarks in the lepton-plus-jets final state in pp collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector*, *JHEP* **08** (2015) 105, [1505.04306].

[130] ATLAS collaboration, T. A. collaboration, *Search for four-top-quark production in final states with one charged lepton and multiple jets using 3.2* fb^{-1} *of proton-proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector at the LHC*, ATLAS-CONF-2016-020.

[131] CMS collaboration, C. Collaboration, *Search for standard model production of four top quarks in proton-proton collisions at 13 TeV*, CMS-PAS-TOP-16-016.

[132] ATLAS collaboration, T. A. collaboration, *Search for new physics using events with b-jets and a pair of same charge leptons in 3.2* fb^{-1} *of pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, ATLAS-CONF-2016-032.

[133] ATLAS collaboration, G. Aad et al., *Search for supersymmetry at* $\sqrt{s} = 13$ *TeV in final states with jets and two same-sign leptons or three leptons with the ATLAS detector*, *Eur. Phys. J.* **C76** (2016) 259, [1602.09058].

[134] CMS collaboration, C. Collaboration, *Search for SUSY with multileptons in 13 TeV data*, CMS-PAS-SUS-16-022.

[135] J. Thaler and L.-T. Wang, *Strategies to Identify Boosted Tops*, *JHEP* **07** (2008) 092, [0806.0023].

[136] N. Zhou, D. Whiteson and T. M. P. Tait, *Limits on Four-Top Production from the ATLAS Same-sign Top-quark Search*, *Phys. Rev.* **D85** (2012) 091501, [1203.5862].

[137] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, *JHEP* **07** (2003) 001, [hep-ph/0206293].

[138] *Expected performance of the ATLAS b-tagging algorithms in Run-2*, Tech. Rep. ATL-PHYS-PUB-2015-022, CERN, Geneva, Jul, 2015.

[139] D0 collaboration, V. M. Abazov et al., *Search for Higgs boson production in trilepton and like-charge electron-muon final states with the D0 detector*, *Phys. Rev.* **D88** (2013) 052009, [1302.5723].

[140] CDF collaboration, T. A. Aaltonen et al., *Search for new physics in trilepton events and limits on the associated chargino-neutralino production at CDF*, *Phys. Rev.* **D90** (2014) 012011, [1309.7509].

[141] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J.* **C71** (2011) 1554, [1007.1727].

[142] G. L. Kane, E. Kuflik, R. Lu and L.-T. Wang, *Top Channel for Early SUSY Discovery at the LHC*, *Phys. Rev.* **D84** (2011) 095004, [1101.1963].

[143] U. Haisch and J. F. Kamenik, *Searching for new spin-0 resonances at LHCb*, *Phys. Rev.* **D93** (2016) 055047, [1601.05110].

[144] S. Fajfer, J. F. Kamenik, I. Nisandzic and J. Zupan, *Implications of Lepton Flavor Universality Violations in B Decays*, *Phys. Rev. Lett.* **109** (2012) 161801, [1206.1872].

[145] 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*, *Phys. Lett.* **B764** (2017) 126–134, [1609.07138].

[146] M. J. Dolan, F. Kahlhoefer, C. McCabe and K. Schmidt-Hoberg, *A taste of dark matter: Flavour constraints on pseudoscalar mediators*, *JHEP* **03** (2015) 171, [1412.5174].

[147] C. Arina et al., *A comprehensive approach to dark matter studies: exploration of simplified top-philic models*, *JHEP* **11** (2016) 111, [1605.09242].

[148] C. B. Jackson, G. Servant, G. Shaughnessy, T. M. P. Tait and M. Taoso, *Higgs in Space!*, *JCAP* **1004** (2010) 004, [0912.0004].

[149] C. B. Jackson, G. Servant, G. Shaughnessy, T. M. P. Tait and M. Taoso, *Gamma Rays from Top-Mediated Dark Matter Annihilations*, *JCAP* **1307** (2013) 006, [1303.4717].

[150] F. Goertz, J. F. Kamenik, A. Katz and M. Nardecchia, *Indirect Constraints on the Scalar Di-Photon Resonance at the LHC*, *JHEP* **05** (2016) 187, [1512.08500].

[151] J. F. Kamenik and C. Smith, *FCNC portals to the dark sector*, *JHEP* **03** (2012) 090, [1111.6402].

[152] D. Dicus, A. Stange and S. Willenbrock, *Higgs decay to top quarks at hadron colliders*, *Phys. Lett.* **B333** (1994) 126–131, [hep-ph/9404359].

[153] M. Carena and Z. Liu, *Challenges and opportunities for heavy scalar searches in the $t\bar{t}$ channel at the LHC*, *JHEP* **11** (2016) 159, [1608.07282].

[154] N. Craig, J. Hajer, Y.-Y. Li, T. Liu and H. Zhang, *Heavy Higgs bosons at low $\tan\beta$: from the LHC to 100 TeV*, *JHEP* **01** (2017) 018, [1605.08744].

[155] P. J. Fox, J. Liu, D. Tucker-Smith and N. Weiner, *An Effective Z'*, *Phys. Rev.* **D84** (2011) 115006, [1104.4127].

[156] J. F. Kamenik, Y. Soreq and J. Zupan, *Lepton flavor universality violation without new sources of quark flavor violation*, *Phys. Rev.* **D97** (2018) 035002, [1704.06005].

[157] P. J. Fox, I. Low and Y. Zhang, *Top-philic Z' forces at the LHC*, *JHEP* **03** (2018) 074, [1801.03505].

[158] V. Barger, P. Langacker, M. McCaskey, M. Ramsey-Musolf and G. Shaughnessy, *Complex Singlet Extension of the Standard Model*, *Phys. Rev.* **D79** (2009) 015018, [0811.0393].

[159] M. Gonderinger, H. Lim and M. J. Ramsey-Musolf, *Complex Scalar Singlet Dark Matter: Vacuum Stability and Phenomenology*, *Phys. Rev.* **D86** (2012) 043511, [1202.1316].

[160] S. Fajfer and J. F. Kamenik, *On the flavor structure of the littlest Higgs model*, *JHEP* **12** (2007) 074, [0710.4293].

[161] J. C. Pati and A. Salam, *Lepton Number as the Fourth Color*, *Phys. Rev.* **D10** (1974) 275–289.

[162] H. Georgi and S. L. Glashow, *Unity of All Elementary Particle Forces*, *Phys. Rev. Lett.* **32** (1974) 438–441.

[163] H. Fritzsch and P. Minkowski, *Unified Interactions of Leptons and Hadrons*, *Annals Phys.* **93** (1975) 193–266.

[164] Super-Kamiokande collaboration, H. Nishino et al., *Search for Nucleon Decay into Charged Anti-lepton plus Meson in Super-Kamiokande I and II*, *Phys. Rev.* **D85** (2012) 112001, [1203.4030].

[165] B. Capdevila, A. Crivellin, S. Descotes-Genon, J. Matias and J. Virto, *Patterns of New Physics in $b \to s\ell^+\ell^-$ transitions in the light of recent data*, *JHEP* **01** (2018) 093, [1704.05340].

[166] W. Altmannshofer, P. Stangl and D. M. Straub, *Interpreting Hints for Lepton Flavor Universality Violation*, *Phys. Rev.* **D96** (2017) 055008, [1704.05435].

[167] L.-S. Geng, B. Grinstein, S. Jäger, J. Martin Camalich, X.-L. Ren and R.-X. Shi, *Towards the discovery of new physics with lepton-universality ratios of $b \to s\ell\ell$ decays*, *Phys. Rev.* **D96** (2017) 093006, [1704.05446].

[168] M. Ciuchini, A. M. Coutinho, M. Fedele, E. Franco, A. Paul, L. Silvestrini et al., *On Flavourful Easter eggs for New Physics hunger and Lepton Flavour Universality violation*, *Eur. Phys. J.* **C77** (2017) 688, [1704.05447].

[169] G. D'Amico, M. Nardecchia, P. Panci, F. Sannino, A. Strumia, R. Torre et al., *Flavour anomalies after the R_{K^*} measurement*, *JHEP* **09** (2017) 010, [1704.05438].

[170] G. Hiller and I. Nisandzic, *R_K and R_{K^*} beyond the standard model*, *Phys. Rev.* **D96** (2017) 035003, [1704.05444].

[171] M. Bordone, G. Isidori and A. Pattori, *On the Standard Model predictions for R_K and R_{K^*}*, *Eur. Phys. J.* **C76** (2016) 440, [1605.07633].

[172] A. Celis, J. Fuentes-Martin, A. Vicente and J. Virto, *Gauge-invariant implications of the LHCb measurements on lepton-flavor nonuniversality*, *Phys. Rev.* **D96** (2017) 035026, [1704.05672].

[173] M. Bauer and M. Neubert, *Minimal Leptoquark Explanation for the $R_{D^{(*)}}$, R_K , and $(g-2)_g$ Anomalies*, *Phys. Rev. Lett.* **116** (2016) 141802, [1511.01900].

[174] T. Blake, G. Lanfranchi and D. M. Straub, *Rare B Decays as Tests of the Standard Model*, *Prog. Part. Nucl. Phys.* **92** (2017) 50–91, [1606.00916].

[175] D. Becirevic, O. Sumensari and R. Zukanovich Funchal, *Lepton flavor violation in exclusive $b \to s$ decays*, *Eur. Phys. J.* **C76** (2016) 134, [1602.00881].

[176] LHCb collaboration, R. Aaij et al., *Measurement of the $B_s^0 \to \mu^+\mu^-$ branching fraction and effective lifetime and search for $B^0 \to \mu^+\mu^-$ decays*, *Phys. Rev. Lett.* **118** (2017) 191801, [1703.05747].

[177] C. Bobeth, M. Gorbahn, T. Hermann, M. Misiak, E. Stamou and M. Steinhauser, *$B_{s,d} \to l^+l^-$ in the Standard Model with Reduced Theoretical Uncertainty*, *Phys. Rev. Lett.* **112** (2014) 101801, [1311.0903].

[178] F. Feruglio, P. Paradisi and A. Pattori, *On the Importance of Electroweak Corrections for B Anomalies*, *JHEP* **09** (2017) 061, [1705.00929].

[179] J. Aebischer, A. Crivellin, M. Fael and C. Greub, *Matching of gauge invariant dimension-six operators for $b \to s$ and $b \to c$ transitions*, *JHEP* **05** (2016) 037, [1512.02830].

[180] ALEPH, DELPHI, L3, OPAL, SLD, LEP Electroweak Working Group, SLD Electroweak Group, SLD Heavy Flavour Group collaboration, S. Schael et al., *Precision electroweak measurements on the Z resonance*, *Phys. Rept.* **427** (2006) 257–454, [hep-ex/0509008].

[181] W. Altmannshofer, C. Niehoff, P. Stangl and D. M. Straub, *Status of the $B \to K^*\mu^+\mu^-$ anomaly after Moriond 2017*, *Eur. Phys. J.* **C77** (2017) 377, [1703.09189].

[182] A. J. Buras, J. Girrbach-Noe, C. Niehoff and D. M. Straub, *$B \to K^{(*)}\nu\overline{\nu}$ decays in the Standard Model and beyond*, *JHEP* **02** (2015) 184, [1409.4557].

[183] F. del Aguila, J. de Blas and M. Perez-Victoria, *Electroweak Limits on General New Vector Bosons*, *JHEP* **09** (2010) 033, [1005.3998].

[184] R. Alonso, B. Grinstein and J. Martin Camalich, *Lepton universality violation and lepton flavor conservation in B-meson decays*, *JHEP* **10** (2015) 184, [1505.05164].

[185] C. Bobeth and A. J. Buras, *Leptoquarks meet ε'/ε and rare Kaon processes*, *JHEP* **02** (2018) 101, [1712.01295].

[186] CMS collaboration, A. M. Sirunyan et al., *Search for standard model production of four top quarks with same-sign and multilepton final states in proton–proton collisions at* $\sqrt{s} = 13$ *TeV*, *Eur. Phys. J.* **C78** (2018) 140, [1710.10614].

[187] E. Alvarez, D. A. Faroughy, J. F. Kamenik, R. Morales and A. Szynkman, *Four Tops for LHC*, *Nucl. Phys.* **B915** (2017) 19–43, [1611.05032].

[188] W. Altmannshofer, S. Gori, M. Pospelov and I. Yavin, *Neutrino Trident Production: A Powerful Probe of New Physics with Neutrino Beams*, *Phys. Rev. Lett.* **113** (2014) 091801, [1406.2332].

[189] CCFR collaboration, S. R. Mishra et al., *Neutrino tridents and W Z interference*, *Phys. Rev. Lett.* **66** (1991) 3117–3120.

[190] T. Mandal, S. Mitra and S. Seth, *Pair Production of Scalar Leptoquarks at the LHC to NLO Parton Shower Accuracy*, *Phys. Rev.* **D93** (2016) 035018, [1506.07369].

[191] J. Blumlein, E. Boos and A. Kryukov, *Leptoquark pair production in hadronic interactions*, *Z. Phys.* **C76** (1997) 137–153, [hep-ph/9610408].

[192] A. Belyaev, C. Leroy, R. Mehdiyev and A. Pukhov, *Leptoquark single and pair production at LHC with CalcHEP/CompHEP in the complete model*, *JHEP* **09** (2005) 005, [hep-ph/0502067].

[193] CMS collaboration, C. Collaboration, *Constraints on models of scalar and vector leptoquarks decaying to a quark and a neutrino at* $\sqrt{s} = 13$ TeV, CMS-PAS-SUS-18-001.

[194] B. Henning, X. Lu and H. Murayama, *How to use the Standard Model effective field theory*, *JHEP* **01** (2016) 023, [1412.1837].

[195] S. Fajfer, N. Kosnik and L. Vale Silva, *Footprints of leptoquarks: from* $R_{K^{(*)}}$ *to* $K \to \pi\nu\bar{\nu}$, *Eur. Phys. J.* **C78** (2018) 275, [1802.00786].

[196] A. Greljo and D. Marzocca, *High-*p_T *dilepton tails and flavor physics*, *Eur. Phys. J.* **C77** (2017) 548, [1704.09015].

[197] H. Georgi and Y. Nakai, *Diphoton resonance from a new strong force*, *Phys. Rev.* **D94** (2016) 075005, [1606.05865].

[198] B. Diaz, M. Schmaltz and Y.-M. Zhong, *The leptoquark Hunter's guide: Pair production*, *JHEP* **10** (2017) 097, [1706.05033].

[199] L. Di Luzio, A. Greljo and M. Nardecchia, *Gauge leptoquark as the origin of B-physics anomalies*, *Phys. Rev.* **D96** (2017) 115011, [1708.08450].

[200] L. Di Luzio, J. Fuentes-Martin, A. Greljo, M. Nardecchia and S. Renner, *Maximal Flavour Violation: a Cabibbo mechanism for leptoquarks*, *JHEP* **11** (2018) 081, [1808.00942].

[201] W. Buchmuller, R. Ruckl and D. Wyler, *Leptoquarks in Lepton - Quark Collisions*, *Phys. Lett.* **B191** (1987) 442–448.

[202] L. Di Luzio, M. Kirk and A. Lenz, *Updated B_s-mixing constraints on new physics models for $b \to s\ell^+\ell^-$ anomalies*, *Phys. Rev.* **D97** (2018) 095035, [1712.06572].

[203] M. Bordone, C. Cornella, J. Fuentes-Martin and G. Isidori, *A three-site gauge model for flavor hierarchies and flavor anomalies*, *Phys. Lett.* **B779** (2018) 317–323, [1712.01368].

[204] CMS collaboration, C. Collaboration, *Search for standard model production of four top quarks in final states with same-sign and multiple leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV*, CMS-PAS-TOP-18-003.

[205] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, *Phys. Lett.* **B641** (2006) 57–61, [hep-ph/0512210].

[206] D. Curtin, J. Galloway and J. G. Wacker, *Measuring the $t\bar{t}h$ coupling from same-sign dilepton $+2b$ measurements*, *Phys. Rev.* **D88** (2013) 093006, [1306.5695].

[207] CMS collaboration, A. M. Sirunyan et al., *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*, *JHEP* **07** (2017) 121, [1703.03995].

[208] CMS collaboration, A. M. Sirunyan et al., *Search for third-generation scalar leptoquarks decaying to a top quark and a τ lepton at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J.* **C78** (2018) 707, [1803.02864].

[209] CMS collaboration, C. Collaboration, *Search for pair production of second generation leptoquarks at sqrt(s)=13 TeV*, CMS-PAS-EXO-17-003.

[210] J. E. Camargo-Molina, A. Celis and D. A. Faroughy, *Anomalies in Bottom from new physics in Top*, *Phys. Lett.* **B784** (2018) 284–293, [1805.04917].

[211] CMS collaboration, C. Collaboration, *Search for leptoquarks coupling to third generation quarks*, CMS-PAS-B2G-16-027.

[212] I. Dorsner, S. Fajfer and A. Greljo, *Cornering Scalar Leptoquarks at LHC*, *JHEP* **10** (2014) 154, [1406.4831].

[213] ATLAS collaboration, M. Aaboud et al., *Search for additional heavy neutral Higgs and gauge bosons in the ditau final state produced in 36 fb^1 of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **01** (2018) 055, [1709.07242].

[214] ATLAS collaboration, M. Aaboud et al., *Search for new phenomena in different-flavour high-mass dilepton final states in pp collisions at $\sqrt{s} = 13$ Tev with the ATLAS detector*, *Eur. Phys. J.* **C76** (2016) 541, [1607.08079].

Appendix A

Four-top Backgrounds

Category	Backgrounds	FS	σ [fb]	decay mode	$\sigma \times BR$ [fb]	comments
$t\bar{t}W$	$t\bar{t}\ W^{\pm}$	5	350.4	$W_{\ell^{\pm}}\ W_{\ell^{\pm}}\ W_{\rm had}$	16.84	
	$t\bar{t}\ W^{\pm}\ j$	5	167.8	$W_{\ell^{\pm}}\ W_{\ell^{\pm}}\ W_{\rm had}$	8.06	MLM
	$t\bar{t}\ W^{\pm}\ jj$	5	96.8	$W_{\ell^{\pm}}\ W_{\ell^{\pm}}\ W_{\rm had}$	4.65	MLM
	$t\bar{t}\ W^{\pm}\ jj$	5		$W_{\ell^{\pm}}\ W_{\ell^{\pm}}\ W_{\ell^{\mp}}$	1.58	MLM, lost ℓ
	$t\bar{t}\ W^{\pm}\ bjj$	5	2.3	$W_{\ell^{\pm}}\ W_{\ell^{\pm}}\ W_{\rm had}$	0.11	
	$t\bar{t}\ W^{\pm}\ b\bar{b}\ jj$	4	2.1	$W_{\ell^{\pm}}\ W_{\ell^{\pm}}\ W_{\rm had}$	0.10	
$t\bar{t}Z$	$t\bar{t}\ Z$	5	583.3	$W_{\ell^{\pm}}\ W_{\rm had}\ Z_{\ell}$	22.33	lost ℓ
	$t\bar{t}\ Z\ j$	5	404.7	$W_{\ell^{\pm}}\ W_{\rm had}\ Z_{\ell}$	15.50	MLM, lost ℓ
	$t\bar{t}\ Z\ jj$	5	194.9	$W_{\ell^{\pm}}\ W_{\rm had}\ Z_{\ell}$	7.46	MLM, lost ℓ
	$t\bar{t}\ Z\ jj$	5		$W_{\ell^{\pm}}\ W_{\ell^{\pm}}\ Z_{\ell}$	3.18	MLM, lost ℓ
$t\bar{t}h$	$t\bar{t}\ h$	4	397.6	$W_{\ell^{\pm}}\ W_{\rm had}\ W_{\ell^{\pm}}\ W_{\rm had}$	4.70	$h \to WW^*$
	$t\bar{t}\ h$	4		$W_{\ell^{\pm}}\ W_{\rm had}\ Z_{\ell}\ Z_{\rm had}$	0.37	$h \to ZZ^*$
	$t\bar{t}\ h$	5	401.3	$W_{\ell^{\pm}}\ W_{\rm had}\ \tau_{\ell^{\pm}}\ \tau_{\rm had}$	2.18	$h \to \tau^+\tau^-$
Others	$tZ\ bjj$	5	176.7	$W_{\ell^{\pm}}\ Z_{\ell}$	4.52	lost ℓ
	$t\bar{t}\ W^+W^-$	4	8.0	$W_{\ell^{\pm}}\ W_{\rm had}\ W_{\ell^{\pm}}\ W_{\rm had}$	0.57	
	$t\bar{t}\ W^+W^-$	4		$W_{\ell^{\pm}}\ W_{\rm had}\ W_{\ell^+}\ W_{\ell^-}$	0.39	lost ℓ
	$W^{\pm}W^{\pm}\ b\bar{b}jj$	4	1.25	$W_{\ell^{\pm}}\ W_{\ell^{\pm}}$	1.94	
	$ZZ\ b\bar{b}j$	4	30.2	$Z_{\ell}\ Z_{\ell}$	0.31	lost ℓ
Signal	$t\bar{t}t\bar{t}$	4	**9.2**	$W_{\ell^{\pm}}\ W_{\ell^{\pm}}\ W_{\rm had}\ W_{\rm had}$	**0.66**	

Table A.1: Irreducible backgrounds for the SS dilepton search. In the comment column, MLM indicates that the jet matching was performed. *lost ℓ implies that for this background to produce a SS dilepton one or more of the leptons in a multi-lepton final state is lost either by not satisfying isolation requirements or down the beam pipe. In the last row we have included for comparison the SM four-top signal in the SS dilepton decay mode.*

Category	Backgrounds	FS	σ [pb]	decay mode	$\sigma \times BR \times \epsilon$ [fb]
Fake	$t\bar{t}\,j$	5	301.6	$W_{\ell^\pm}W_{\rm had}$	11.43
	$t\bar{t}\,jj$	5	124.9	$W_{\ell^\pm}W_{\rm had}$	4.74
	$t\bar{t}\,bjj$	5	5.3	$W_{\ell^\pm}W_{\rm had}$	0.20
	$t\bar{t}\,b\bar{b}\,jj$	4	3.0	$W_{\ell^\pm}W_{\rm had}$	0.11
	$t\bar{t}\,b\bar{b}\,3j$	4	2.3	$W_{\ell^\pm}W_{\rm had}$	0.09
Q-flip	$t\bar{t}\,jj$	5	124.9	$W_{\ell^\pm}W_{\ell^\mp}$	8.03
	$t\bar{t}\,bjj$	5	5.3	$W_{\ell^\pm}W_{\ell^\mp}$	0.34
	$t\bar{t}\,b\bar{b}\,jj$	4	3.0	$W_{\ell^\pm}W_{\ell^\mp}$	0.19
	$t\bar{t}\,b\bar{b}\,3j$	4	2.3	$W_{\ell^\pm}W_{\ell^\mp}$	0.15
	$Z\,b\bar{b}\,2j$	4	26.3	Z_ℓ	2.66

Table A.2: Reducible backgrounds for the SS dilepton search. Here $\epsilon = 10^{-4}$ for the Fake category and $\epsilon = 10^{-3}$ for Q-flip category.

A.1 Breakdown of SS dilepton backgrounds

In this appendix, we describe the backgrounds used for the analysis of sec. 5.3.1. The backgrounds that mimic the four-top SS dilepton signature fall into two types: (i) *Irreducible backgrounds* coming from rare SM processes that contain one real SS dilepton pair in the final state, (ii) *Reducible backgrounds* of instrumental origin where a "fake" SS dilepton pair is produced in the detector.

All background simulations were performed at LO with `MadGraph` for samples with less than two extra partons in the final state or `AlpGen` for samples with more than two extra partons. We chose the NN23LO set for the parton distribution function (PDF) and `FastJet` [37, 205] for clustering jets using the anti-k_T algorithm with $R = 0.4$. Background samples with up to 2 extra partons at the matrix element (ME) level were merged and matched to the parton shower using the MLM matching scheme. When available [35], the LO cross-sections were rescaled with a K-factor to include NLO corrections.

Irreducible backgrounds The non-negligible background processes used in the SS dilepton search strategy are divided into background categories as shown in the two first columns of Table A.1. There, we considered as an irreducible background any process producing final states at the ME level: one SS dilepton and a minimum of 4 hard quarks of which at least 2 must be b-quarks. In the two following columns we give the production cross-section calculated either in the 4-flavor or 5-flavor schemes (FS)[1], subject to the partonic cuts $p_T^j > 25$ GeV and $\Delta R_{jj} > 0.2$ for jets of any flavor. All irreducible background processes eventually decay into an intermediate boson

[1]The definition of the jet j in the second column depends on the FS: j arises from the partons g, u, d, s, c in the 4FS and from g, u, d, s, c, b in the 5FS.

state $V = W, Z$ (either on or off the mass shell), denoted by V_ℓ or $V_{\rm had}$ depending on whether these decay leptonically or hadronically into final states fermions. The decay modes into the final state fermions are shown in the fifth column, while on the sixth column we give a rough estimate of the total cross-section of the background process based on the corresponding branching ratios BR. In the last column, we give additional comments. We have checked that other backgrounds, such as tVV and VVV, etc ($V = W, Z, h$) are negligible when demanding $N_b \geq 3$ and did not include them in the classification.

Reducible backgrounds Due to a large contribution from fake leptons and charged-flipped lepton pairs, any four-top analysis at the LHC in the multi-lepton channel must include a correct estimation of these background processes. In this work, we estimate the fake lepton and charge-flip (Q-flip) backgrounds by simulating all potential SM sources giving rise to one fake SS dilepton at the detector level via jet-to-lepton faking $j \to \ell^\pm$ or a Q-flip for electron/positrons only $e^\pm \to e^\mp$. These misidentifications are mainly produced through the following mechanisms:

- Fake leptons: Can originate from incorrectly tagging as a lepton a charged meson related to a final state parton j. Another main source included in this category are non-prompt (real) leptons originating from the leptonic decay of a heavy meson inside a final state jet j. In this way, an initial $\ell^\pm j$ pair can be detected as a SS dilepton via $j \to \ell^\pm$ misidentification.

- Q-flip: Can occur when a real OS dilepton pair coming from some underlying process is mistaken in the detector as a SS dilepton by misidentifying the charge of one of the leptons. This effect is negligible for muons given that their charge is measured in both the inner silicon tracker and at the outer muon detector layer. Physically, the dominant contribution comes from trident conversion inside the detector: an incoming electron/positron suffers bremsstrahlung $e^\pm \to \gamma e^\pm \to e^\mp e^\pm e^\pm$ inside the silicon tracker. Since the probability of Q-flip is proportional to the flight length inside the medium, the central region of the tracker (e.g. $|\eta| < 2$) produce less $Q-$flipped electrons than regions near the end-caps.

The dominant sources generating both reducible backgrounds are $t\bar{t}$ + jets followed by $W^\pm$ + jets and Z + jets. For simplicity, it is enough to assume that the probabilities $\epsilon_{\rm fake}$ and $\epsilon_{\rm Qflip}$ for these misidentifications are flat in p_T and η within the kinematic regions of our search strategy. We also assume that the fake lepton or Q-flipped electron (positron) inherit exactly the same kinematic properties as the source jet or positron (electron). Under these simplifications, we just need to weight each of the simulated source samples with the corresponding misidentification probabilities, here fixed at the benchmark values $\epsilon_{\rm fake} = 10^{-4}$ and $\epsilon_{\rm Qflip} = 10^{-3}$ (see Appendix A.3). The complete list of all dominant background sources generating the SS dilepton reducible backgrounds are given in Table A.2.

A.2 Breakdown of trilepton backgrounds

We now turn to the dominant backgrounds for the trilepton channel used in Sec. 5.3.2. The classification is very similar to the SS dilepton and is given in Table A.3 for the

irreducible backgrounds and Table A.4 for the reducible backgrounds. All simulations were performed using the same tools and settings as described in Appendix A.1.

A.3 Fake lepton and Q-flip mis-identification probabilities

The usual way to estimate fake lepton and Q-flip backgrounds is by directly extracting the probability of object mis-identification as a function of p_T and η from control data samples in signal regions relevant to the proposed search. Electron Q-flip rates are estimated using a likelihood fit to $Z/\gamma^* \to ee$ data samples [6], while lepton fake rates are extracted using the matrix method applied to semi-leptonic $t\bar{t}$ data samples [133]. The extracted values are typically in the range $\epsilon_{\text{fake}} \sim \mathcal{O}(10^{-4})$ and $\epsilon_{\text{Qflip}} \sim \mathcal{O}(10^{-3})$ or less. For this reason many phenomenological analysis adopt as benchmark values:

$$\epsilon_{\text{fake}} = 10^{-4}\,, \qquad \epsilon_{\text{Qflip}} = 10^{-3}\,. \tag{A.1}$$

In order to better assess the sensitivity of our four-top search, it is important to know if this benchmark choice is either too conservative or too optimistic. The data-driven techniques used by the experimental collaborations are however difficult to implement without access to the 13 TeV control data sets. Nonetheless, we can

Category	Backgrounds	FS	σ [fb]	decay mode	$\sigma \times BR$ [fb]	comments
$t\bar{t}W$	$t\bar{t}\,W^{\pm}\,jj$	5	96.8	$W_{\ell^\pm}\,W_{\ell^\pm}\,W_{\ell^\pm}$	1.58	MLM
	$t\bar{t}\,W^{\pm}\,bjj$	5	2.3	$W_{\ell^\pm}\,W_{\ell^\pm}\,W_{\ell^\pm}$	0.04	
	$t\bar{t}\,W^{\pm}\,b\bar{b}\,jj$	4	2.1	$W_{\ell^\pm}\,W_{\ell^\pm}\,W_{\ell^\pm}$	0.03	
$t\bar{t}Z$	$t\bar{t}\,Z$	5	583.3	$W_{\ell^\pm}\,W_{\text{had}}\,Z_{\ell}$	22.33	
	$t\bar{t}\,Z\,j$	5	404.7	$W_{\ell^\pm}\,W_{\text{had}}\,Z_{\ell}$	15.50	MLM
	$t\bar{t}\,Z\,jj$	5	194.9	$W_{\ell^\pm}\,W_{\text{had}}\,Z_{\ell}$	7.46	MLM
	$t\bar{t}\,Z\,jj$	5		$W_{\ell^\pm}\,W_{\ell^\pm}\,Z_{\ell}$	3.18	MLM, lost ℓ
$t\bar{t}h$	$t\bar{t}\,h$	4	397.6	$W_{\ell^\pm}\,W_{\ell^\pm}\,W_{\ell^\mp}\,W_{\text{had}}$	1.60	$h \to WW^*$
	$t\bar{t}\,h$	4		$W_{\ell^\pm}\,W_{\ell^\mp}\,Z_{\ell}\,Z_{\text{had}}$	0.06	$h \to ZZ^*$
	$t\bar{t}\,h$	5	401.3	$W_{\ell^\pm}\,W_{\ell^\mp}\,\tau_{\ell^\pm}\,\tau_{\text{had}}$	0.74	$h \to \tau^+\tau^-$
Others	$t\,Z\,bjj$	5	176.7	$W_{\ell^\pm}\,Z_{\ell}$	4.52	
	$W^{\pm}Z\,b\bar{b}\,jj$	4	70.3	$W_{\ell^\pm}\,Z_{\ell}$	1.80	
	$t\bar{t}\,W^+W^-$	4	8.0	$W_{\ell^\pm}\,W_{\ell^\pm}\,W_{\ell^\mp}\,W_{\text{had}}$	0.39	
	$ZZ\,b\bar{b}j$	4	30.2	$Z_{\ell}\,Z_{\ell}$	0.31	
Signal	$t\bar{t}t\bar{t}$	4	**9.2**	$W_{\ell^\pm}\,W_{\ell^\pm}\,W_{\ell^\mp}\,W_{\text{had}}$	**0.45**	

Table A.3: Irreducible backgrounds for the trilepton search. In the comment column, MLM indicates that the jet matching was performed. *lost ℓ implies that for this background to produce a trilepton one lepton from a four-lepton final state is lost either by not satisfying isolation requirements or down the beam pipe. In the last row we have included for comparison the SM four-top signal in the trilepton decay mode.*

Category	Backgrounds	FS	σ [pb]	decay mode	$\sigma \times BR \times \epsilon$ [fb]
Fake	$t\bar{t}\,jj$	5	124.9	$W_{\ell^\pm}W_{\ell^\mp}$	0.80
	$t\bar{t}\,bjj$	5	5.3	$W_{\ell^\pm}W_{\ell^\mp}$	0.03
	$t\bar{t}\,b\bar{b}\,jj$	4	3.0	$W_{\ell^\pm}W_{\ell^\mp}$	0.02
	$t\bar{t}\,b\bar{b}\,3j$	4	2.3	$W_{\ell^\pm}W_{\ell^\mp}$	0.01
	$Z\,b\bar{b}\,2j$	4	26.3	Z_ℓ	0.27

Table A.4: Reducible backgrounds for the trilepton search. Here $\epsilon = 10^{-4}$ for the Fake category.

still estimate $\epsilon_{\text{fake,Qflip}}$ by combining MC simulations and existing LHC searches as proposed in Ref. [206]. For this we rely on the SUSY search [6] by ATLAS at 13 TeV and 13.2 fb^{-1}, where SS dileptons and trileptons were used as final states in different signal regions similar to the one in our multi-lepton search (for example **SR3b**). There, fake lepton and Q-flip backgrounds were estimated using data-driven techniques. We make use of their results (Tables 4 and 5 in Ref. [6]) to match MC samples by fitting the parameters $\epsilon_{\text{fake,Qflip}}$ that scale the total number of events in each signal region. For the fit we used samples of $t\bar{t}$ + jets, $W^\pm$ + jets and Z + jets followed by a $j \to \ell^\pm$ or $e^\pm \to e^\mp$ mis-identification. The best fit values we obtain are:

$$\epsilon_{\text{fake}} = 7.2 \times 10^{-5}\,, \qquad \epsilon_{\text{Qflip}} = 2.2 \times 10^{-4}\,. \tag{A.2}$$

Post-fit results of our fake lepton and Q-flip background simulations for this search, given by the colored bars in figure A.1, agree very well within error bars with the ATLAS data-driven estimations in each signal region (black dots). Our fit suggests that the benchmark probabilities in Eq. (A.1) are a conservative choice for our four-top search and that the number reducible backgrounds events presented in Table 5.1 and Table. 5.2 between parenthesis are over-estimated. For this reason we present the results in Sec. 5.3.3 based on the estimation in Eq. (A.2) instead. In principle, a better fit to the ATLAS results could be achieved by including p_T and η dependence for each probability and if more fitting parameters are included [206], but this should not alter considerably the values in Eq. (A.2).

A.4 Non-prompt lepton cuts

Here we give a few comments on the cut in Eq. (5.2) applied to leptons passing the mini-isolation requirement. Most analyses by ATLAS and CMS veto reconstructed leptons too close to jets (typically with a distance $\Delta R_{\ell j} < 0.4$ from any jet) in order to reduce (isolated) non-prompt lepton backgrounds originating from heavy meson decays inside jets. A slightly looser cut has been adopted by CMS in Ref. [1], based on rejecting isolated leptons satisfying:

$$p_{T_\ell} < \alpha\; p_{T_{\text{jet}}} \tag{A.3}$$

with $\alpha = 0.7$ for electrons and $\alpha = 0.85$ for muons. With these values, CMS claims a high background rejection while keeping more than 50% of the signal inside the

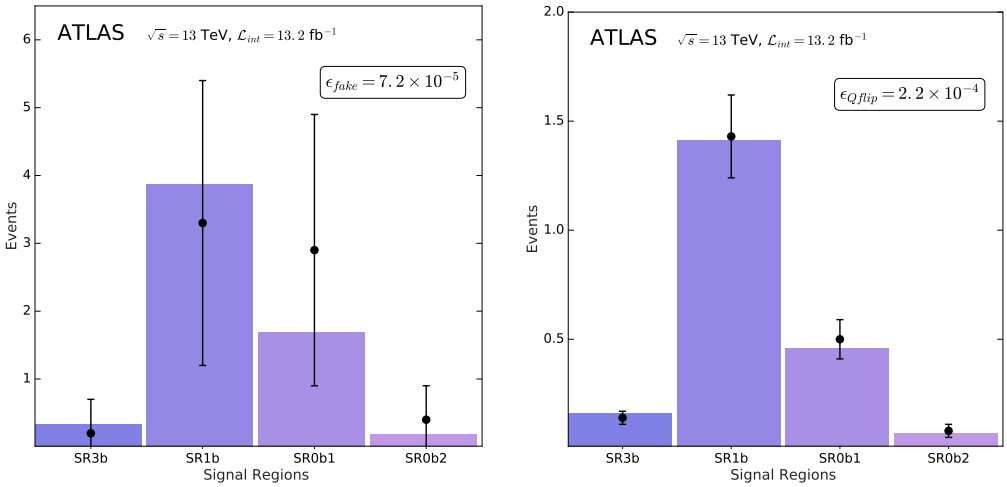


Figure A.1: Post-fit validation plots of our fake lepton and Q-flip background simulations (colored bars) for the best fit values given in Eq. (A.2) compared to ATLAS data-driven estimations (black dots with error bars). The definitions of the signal regions are given in Ref. [6].

h!]

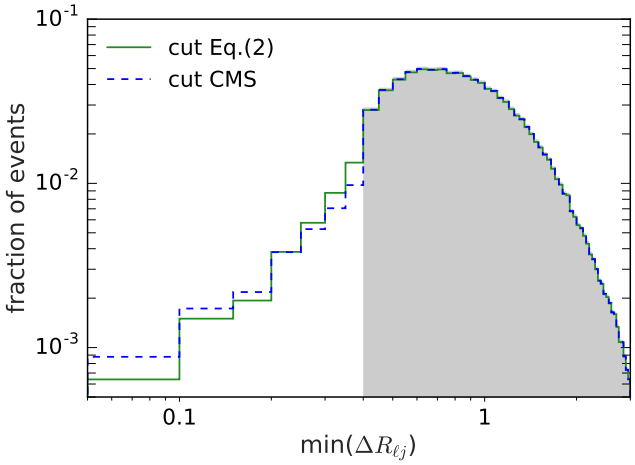


Figure A.2: Comparison between different non-prompt lepton cuts for the four-top signal: $\Delta R_{\ell j} < 0.4$ veto (shaded gray), the cut adopted by CMS in Ref. [1] (dashed blue) and the cut in Eq. (5.2) adopted in our analysis (green).

$\Delta R_{\ell j} < 0.4$ region. The non-prompt lepton cut in Eq. (5.2) used in our search has an additional dependence on the distance of the lepton from the jet, making the cut tighter (looser) at closer (larger) distances than the one in Eq. (A.3). In figure A.2 we show the fractions of four-top events after implementing the non-prompt lepton cuts: $\Delta R_{\ell j} < 0.4$ veto (shaded gray), the cut adopted by CMS in Ref. [1] (dashed blue) and the cut in Eq. (5.2) adopted in our analysis (green).

Appendix B

Collider limits for Leptoquarks

B.1 Direct limits on pair-produced LQs

An efficient way to set limits on LQs is to directly search for them at hadron colliders. At the LHC one of the most significant example of such a processes is the pair production $gg\,(q\bar{q}) \to \mathrm{LQ}^\dagger\mathrm{LQ}$, shown in Fig. B.1 (a). In both ATLAS and CMS the searches for this process in different decay channels into second and/or third generation quarks and leptons, $\mathrm{LQ}^\dagger\mathrm{LQ} \to q\bar{q}\ell\bar{\ell}$, $q\bar{q}\nu\bar{\nu}$, have been made. The results of these searches lead to model-independent bounds on both the mass and branching fractions of the LQ.

In Table B.1 we list the most recent lower limits on the masses of second/third generation scalar and vector LQs relevant to this work, for benchmark branching ratios set to $\beta = 1\,(0.5)$. These limits assume the following: (i) pair production is dominated by QCD interactions, and (ii) for vector LQs (V^μ) the LQ-gluon interaction term, $\mathcal{L} \supset -\kappa g_s V^\mu G_{\mu\nu} V^\nu$, is taken with $\kappa = 1$. The first assumption is in general true for LQ-fermion couplings of order ~ 1 or smaller [212]. In this regime, contributions to $q\bar{q} \to \mathrm{LQ}^\dagger\mathrm{LQ}$ with a t-channel lepton (where the amplitude is proportional to the squared LQ-fermion coupling) are subleading compared to QCD induced production. The assumption on the value of κ, instead, depends on the UV origin of the vector LQ [191]. If V^μ is a fundamental gauge boson of a new non-abelian gauge group then the gauge symmetry completely fixes the choice $\kappa = 1$. The possibility of having $|\kappa| < 1$ may arise in a UV theory where the vector LQ is

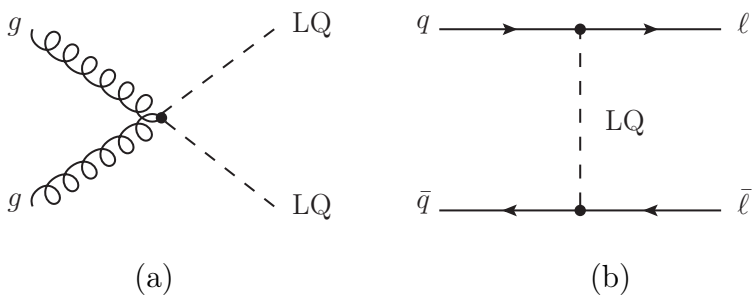


Figure B.1: (a) Representative Feynman diagram for LQ pair production via QCD interactions. (b) Feynman diagram for LQ t-channel exchange in $pp \to \ell\bar{\ell}$ production at the LHC. The dashed propagator represents either a scalar or vector LQ state.

Decays	LQs	Scalar LQ limits	Vector LQ limits	$\mathcal{L}_{\text{int}}$ / Ref.
$jj\,\tau\bar{\tau}$	S_1, R_2, S_3, U_1, U_3	–	–	–
$b\bar{b}\,\tau\bar{\tau}$	R_2, S_3, U_1, U_3	850 (550) GeV	1550 (1290) GeV	12.9 fb^{-1} [207]
$t\bar{t}\,\tau\bar{\tau}$	S_1, R_2, S_3, U_3	900 (560) GeV	1440 (1220) GeV	35.9 fb^{-1} [208]
$jj\,\mu\bar{\mu}$	S_1, R_2, S_3, U_1, U_3	1530 (1275) GeV	2110 (1860) GeV	35.9 fb^{-1} [209]
$b\bar{b}\,\mu\bar{\mu}$	R_2, U_1, U_3	1400 (1160) GeV	1900 (1700) GeV	36.1 fb^{-1} [198]
$t\bar{t}\,\mu\bar{\mu}$	S_1, R_2, S_3, U_3	1420 (950) GeV	1780 (1560) GeV	36.1 fb^{-1} [210, 211]
$jj\,\nu\bar{\nu}$	R_2, S_3, U_1, U_3	980 (640) GeV	1790 (1500) GeV	35.9 fb^{-1} [193]
$b\bar{b}\,\nu\bar{\nu}$	S_1, R_2, S_3, U_3	1100 (800) GeV	1810 (1540) GeV	35.9 fb^{-1} [193]
$t\bar{t}\,\nu\bar{\nu}$	R_2, S_3, U_1, U_3	1020 (820) GeV	1780 (1530) GeV	35.9 fb^{-1} [193]

Table B.1: Summary of the current limits from LQ pair production searches at the LHC. In the first column we give the searched final states and in the second column the LQs for which this search is relevant. In the next two columns, we present the current limits on the mass for scalar and vector LQs, respectively, for $\beta = 1$ ($\beta = 0.5$). In the last column, we display the value of the LHC luminosity for each search along with the experimental references. Note that j denotes any jet originating from a charm or a strange quark. Entries marked with $(-)$ indicate that no recast or search in this channel has been performed up to this date.

a composite particle, therefore giving rise to LHC limits weaker than for the gauge vector LQ presented in Table. B.1.

B.2 Limits from high-pT dilepton tails

As shown in Refs. [145, 196], a contribution arising from the t-channel exchange of LQs to $pp \to \ell\bar{\ell}$ ($\ell = \mu, \tau$) can be directly probed in the high-p_T tails of Drell-Yan processes at the LHC. In particular, larger values of Yukawa couplings, that are often needed to accommodate the B-anomalies, could modify the tail of the differential cross section of $pp \to \ell\bar{\ell}$. In the following we use LHC data from pp-collisions at 13 TeV to set limits on each LQ model. For this we have recast two recent searches by ATLAS at 36.1 fb^{-1} for a Z' decaying to $\mu\bar{\mu}$ [3] and $\tau\bar{\tau}$ [213], respectively.

For the di-tau analysis, we focused on the fully inclusive channel with hadronic taus (τ_{had}) in the final state, given that these perform considerably better at high p_T than the leptonic tau decay channels. For each search we counted the number of observed and background events above different threshold values of the invariant mass distributions $m_{\mu\mu}$ for the di-muon search and the total transverse mass distribution m_{tot} for the di-tau search (see ref. [213] for the definition of m_{tot}). An upper bound at 95% C.L. on the number of allowed signal events above each mass threshold was extracted for each search by minimizing the Log-Likelihood ratio with nuisance parameters for the background uncertainties, as described in [141]. Besides the current luminosity limits, we also estimated projected limits at a higher LHC luminosity of $\mathcal{L}_{\text{int}} = 300$ fb^{-1} by scaling the data and background events with the luminosity ratio and the background uncertainties with $\sqrt{\mathcal{L}_{\text{int}}}$ assuming that the data in the distribution tails are statistically dominated. This assumption holds

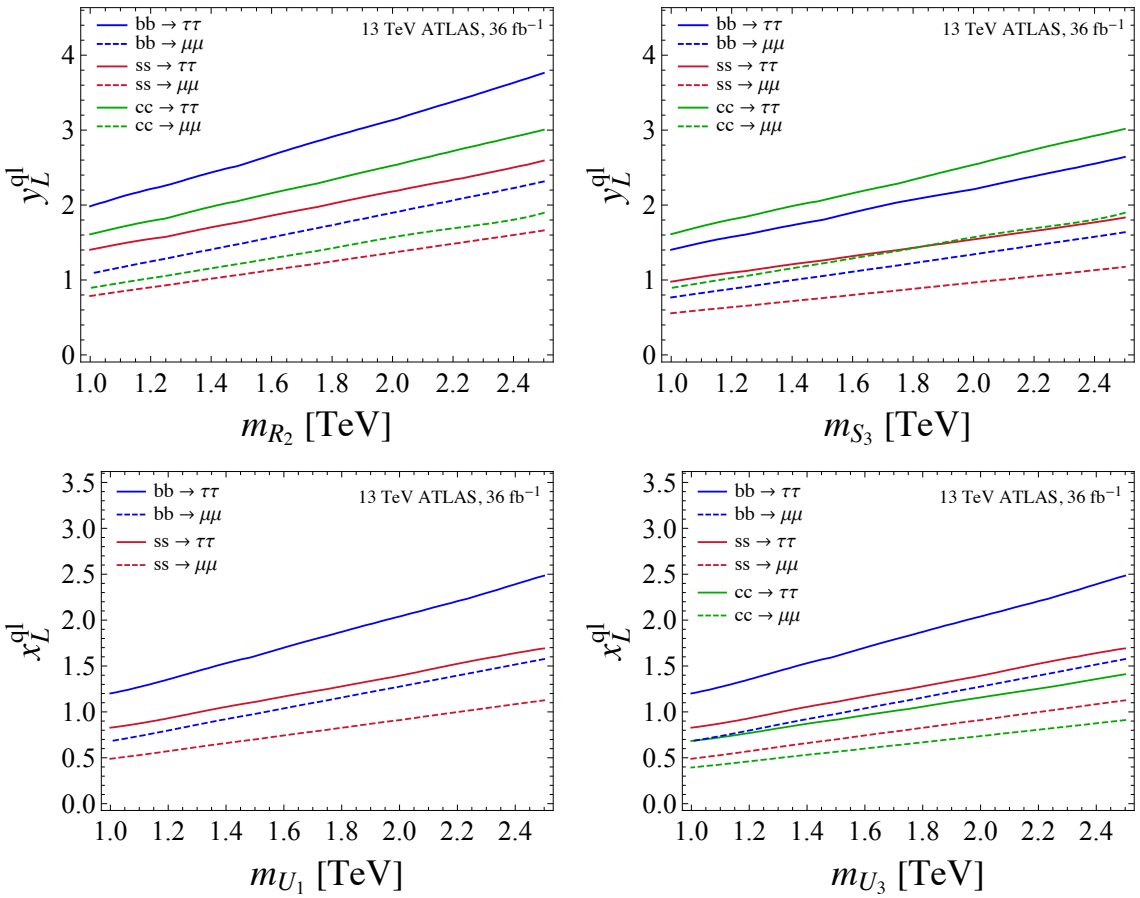


Figure B.2: The top panel (lower panel) shows current limits in the coupling vs mass plane for several scalar LQ (vector LQ) models from LHC searches in $pp \to \ell\bar{\ell}$ high-p_T tails at 13 TeV with $36\,\text{fb}^{-1}$ of data. The solid and dashed lines represent limits from di-tau and di-muon searches, respectively, for different initial quarks while turning one scalar (vector) LQ coupling y_L^{ql} (x_L^{ql}) at a time.

well for the leading backgrounds such as SM Drell-Yan production or fake τ_{had} from QCD jet mistagging since these are estimated using experimental data from control regions that improve with more statistics. Additionally, systematics in the tails of the di-muon and di-tau searches are well under control and only dominate the lower bins where the search is insensitive to the massive LQs.

For our simulations we created the Universal FeynRules UFO files using `FeynRules` [33] for each LQ mediator (S_1, S_3, R_2, U_1 and U_3, defined in Ref. [7]) coupling exclusively to second and third generation of quarks and leptons. For R_2 and U_1 we have only considered, for simplicity, non-zero LH coupling matrices and omitted RH couplings (see Ref. [7] for the corresponding Lagrangians of each leptoquark). After exporting the UFOs to `MadGraph5` [35] we generated for each LQ mediator a statistically significant set of t-channel Drell-Yan event samples $q\bar{q} \to \mu\bar{\mu}$, $\tau\bar{\tau}$ for initial sea quarks, $q \in \{s, c, b\}$, and for vector (scalar) LQs at different masses in the 1-6 TeV (0.6-3 TeV) range. Each sample was subsequently showered and hadronized using `Pythia8` [36]. Final state hadronic taus and isolated muons where reconstructed and smeared using `Delphes3` [38] with the parameters set according to each experimental scenario. In order to illustrate the current reach of the LHC for each LQ

state we show in Fig. B.2 results from the $pp \to \mu\bar{\mu}$ and $pp \to \tau\bar{\tau}$ searches by ATLAS for scalar (vector) LQs in the y_L^{ql} (x_L^{ql}) coupling versus mass plane. [1] Each line corresponds to the 95% upper limit for the process $q\bar{q} \to \ell\bar{\ell}$ turning on each flavor coupling one at a time, with $q \in \{s, c, b\}$ and $\ell \in \{\mu, \tau\}$. Similar bounds of the same order can be extracted for the coupling products $y_L^{ql} y_L^{q'l}$ and $x_L^{ql} x_L^{q'l}$ with $q \neq q'$ from the quark flavor violating process $qq' \to \ell\bar{\ell}$.

Besides producing deviations in the di-lepton tails, LQ mediators that couple simultaneously to differently charged leptons may also produce measurable effects in LFV observables at the LHC. In particular, searching for the process $pp \to \ell\ell'$ with a LQ exchanged in the t-channel may provide an additional handle for setting constraints on the flavor structure of these LQ models. Existing searches for a massive Z' with LFV couplings have been presented by the LHC collaborations in the $Z' \to e\mu$, $e\tau$, $\mu\tau$ channels. In order to determine the sensitivity of the LHC to the t-channel process $pp \to \mu\tau$ we recast the LFV Z' search by ATLAS [214] at $3.2\,\text{fb}^{-1}$. We find, however, that the bounds on the LQs extracted from the high-p_T $m_{\mu\tau}$ tails are always weaker than the combined bounds from the flavor conserving di-muon and di-tau tails described above.

[1] We did not present plots for S_1 since these bounds are identical to the bounds for $S_3^{1/3}$.

Appendix C

Fermion Mixing

In the U(1)$'$ and SU(4)$_R$ UV models, the vector-like fermions mix with the SM fermions. We now describe this mixing in general terms without assuming any flavor structure in the UV theory. To keep things shorter we focus on the quark sector, denoting the light SM quarks as q^i with generation index $i = 1, 2, 3$ and the heavy vector-like quark partner as $\mathcal{Q}$. Right after the onset of spontaneous symmetry breaking, the mass sector in the interaction basis takes the form

$$-\mathcal{L}_{\text{mass}} \;=\; \left(\overline{q'_L}\,,\,\overline{\mathcal{Q}'_L}\right)\begin{pmatrix} m_q & 0 \\ \mu^\dagger & m_{\mathcal{Q}} \end{pmatrix}\begin{pmatrix} q'_R \\ \mathcal{Q}'_R \end{pmatrix} \;+\; \text{h.c.}\,, \tag{C.1}$$

where we have defined two 3×3 and one 1×3 mass matrices

$$m_q^{ij} \;\equiv\; \frac{v}{\sqrt{2}}\,y_q^{ij}\,, \tag{C.2}$$

$$\mu^i \;\equiv\; \frac{w}{\sqrt{2}}\,\lambda^i_{\mathcal{Q}}\,, \tag{C.3}$$

respectively, with w given by the VEV of some, here unspecified, scalar field and $m_{\mathcal{Q}}$ is the Dirac mass of the vector-like quark assumed to satisfy $m_{\mathcal{Q}} \gg v$. The "primed" fields in (C.1) denote that these are in the interaction eigenbasis, while "unprimed" fields denote fields in the mass eigenbasis.

C.1 Block-parametrization of 4×4 mixing matrices

We now rotate the fermions into the mass eigenbasis, given by the physical quark fields q^i, and the physical vector-like fermion $\mathcal{Q}$. For this we diagonalize the 4×4 mass matrix in (C.1) with a bi-unitary transformation defined by two 4×4 matrices $\mathcal{V}_{L,R}$:

$$\begin{pmatrix} q'_1 \\ q'_2 \\ q'_3 \\ \mathcal{Q}' \end{pmatrix}_{L,R} \;=\; \mathcal{V}^\dagger_{L,R}\begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ \mathcal{Q} \end{pmatrix}_{L,R} \tag{C.4}$$

We decompose each 4×4 unitary matrix $\mathcal{V}_{L,R}$ into four block matrices: 3×3 blocks $U_{L,R}$, 3×1 vectors $\Theta_{L,R}$ and $\Xi_{L,R}$, and two real numbers $a_{L,R}$ such that

$$\mathcal{V}_{L,R} \;\equiv\; \left(\begin{array}{c|c} U_{L,R} & \Theta_{L,R} \\ \hline \Xi_{L,R}^{\dagger} & a_{L,R} \end{array}\right) \tag{C.5}$$

From the unitary constraints of the matrices $\mathcal{V}_{L,R}$ we can easily solve for the two lower blocks in (C.5):

$$\Xi_X = \frac{U_X\,\Theta_X}{\sqrt{1-{\Theta_X}^{\dagger}\Theta_X}}\,, \qquad a_X = \sqrt{1-{\Theta_X}^{\dagger}\Theta_X}\,. \tag{C.6}$$

where X is a chirality index $X=\{L,R\}$. Without losing any generality we can now express the full unitary matrices as a product of a matrix $\mathcal{U}_X$ and a block diagonal matrix $\mathcal{D}_X$:

$$\mathcal{V}_X \;=\; \mathcal{U}_X\cdot\mathcal{D}_X \tag{C.7}$$

where each are "almost" unitary matrices, i.e. $\mathcal{U}_X^{\dagger}\mathcal{U}_X$ and $\mathcal{D}_X^{\dagger}\mathcal{D}_X$ are a diagonal perturbation of the unit matrix of order $\mathcal{O}(\Theta_X{\Theta_X}^{\dagger})$. Explicitly, these are given by

$$\mathcal{U}_X \;=\; \left(\begin{array}{c|c} \mathbb{1}_{3\times 3} & \frac{1}{\cos\theta_X}\,\Theta_X \\ \hline -\frac{1}{\cos\theta_X}\,{\Theta_X}^{\dagger} & 1 \end{array}\right) \tag{C.8}$$

$$\mathcal{D}_X \;=\; \left(\begin{array}{c|c} U_X & \mathbb{0} \\ \hline \mathbb{0} & \cos\theta_X \end{array}\right) \tag{C.9}$$

where we have defined one angle θ_X such that

$$\sin^2(\theta_X) \;\equiv\; {\Theta_X}^{\dagger}\Theta_X\,. \tag{C.10}$$

Notice that the block decomposition of the mixing matrices $\mathcal{V}_X$ described above for the moment depends exclusively on the light–heavy mixing parameters Θ_X and the light–light mixing matrices U_X. Of course, these mixings are not completely independent since they satisfy one additional unitary condition:

$${U_X}^{\dagger}U_X \;=\; \mathbb{1} \;-\; \Theta_X{\Theta_X}^{\dagger}\,. \tag{C.11}$$

At this stage it is convenient to decompose the light–light mixings U_X, which are non-unitary 3×3 matrices, into its *polar representation*:

$$U_X \;=\; R_X\cdot V_X^q \tag{C.12}$$

where V_X^q are unitary matrices $(V_X^q)^{\dagger}V_X^q = V_X^q(V_X^q)^{\dagger} = \mathbb{1}$ and R_X is a semi-positive Hermitian *polar* matrix $R_X = R_X^{\dagger}$. This last matrix quantifies the departure of the

mixing matrix U_X from being a unitary matrix. The polar decomposition (C.12) is unique and corresponds to the matrix version of the polar form of a complex number $z = r\exp(i\phi)$. Replacing the decomposition (C.12) into the unitary condition (C.11) uniquely determines the polar matrix as a function of the light–heavy mixing parameters to be

$$R_X = \sqrt{\mathbb{1} \; - \; \Theta_X {\Theta_X}^\dagger} \, . \tag{C.13}$$

This expression involves the square-root of a 3×3 matrix and is far from being a desirable expression. In fact, square-roots of matrices are in general difficult to compute in closed form for dimension 3 or larger. Fortunately, the mixing parameters satisfy $|\Theta_X^i \Theta_X^j| < 1$. This implies that the Taylor series for the square root function of the matrix $\mathbb{1} - \Theta_X {\Theta_X}^\dagger$ converges towards R_X. Hence, for practical purposes, we can use the well-known MacLauren series to express the square-root matrix

$$R_X \;=\; \mathbb{1} \;+\; \sum_{n=1}^{\infty} (-1)^n \left(\Theta_X {\Theta_X}^\dagger\right)^n \prod_{k=1}^{n} \left[\frac{1/2 - (k-1)}{k}\right] . \tag{C.14}$$

This expression can be substantially simplified if one notices that

$$\left(\Theta_X {\Theta_X}^\dagger\right)^n \;=\; \left(\sin^2\theta_X\right)^{n-1} \, \Theta_X {\Theta_X}^\dagger \tag{C.15}$$

holds true. Plugging this into (C.14) and performing some algebra provides the following solution for the square-root $\sqrt{\mathbb{1} \; - \; \Theta_X {\Theta_X}^\dagger}$:

$$R_X \;=\; \mathbb{1} \;-\; \frac{\Theta_X {\Theta_X}^\dagger}{1 + \cos\theta_X} \, , \tag{C.16}$$

with its inverse given by

$$(R_X)^{-1} \;=\; \mathbb{1} \;+\; \frac{\Theta_X {\Theta_X}^\dagger}{\cos\theta_X \, (1 + \cos\theta_X)} \, . \tag{C.17}$$

Here it is important two stress two things: (i) this result is exact and (ii) this results is unique. Square-roots of matrices are in general not unique (similar to real numbers which are usually ambiguous up to a sign) for an arbitrary matrix. Nevertheless, matrices that are positive semi-definite, i.e. with non-negative eigenvalues, have a unique square-root. In our case the matrix $\mathbb{1} - \Theta_X {\Theta_X}^\dagger$ is indeed positive semi-definite since the three real eigenvalues are given by $\{1, 1, \cos^2\theta_X\}$ for any Θ_X.

In summary, we have shown that each 4 × 4 mixing matrix can be parametrized in

a general way as:

$$\mathcal{V}_X = \left(\begin{array}{c|c} \mathbb{1} & \frac{\Theta_X}{\cos\theta_X} \\ \hline -\frac{\Theta_X{}^\dagger}{\cos\theta_X} & 1 \end{array}\right) \cdot \left(\begin{array}{c|c} \left[\mathbb{1} - \frac{\Theta_X\Theta_X{}^\dagger}{1+\cos\theta_X}\right] V_X^q & \mathbb{0} \\ \hline \mathbb{0} & \cos\theta_X \end{array}\right)$$

$$= \left(\begin{array}{c|c} \left[\mathbb{1}_{3\times3} - \frac{\Theta_X(\Theta_X)^\dagger}{1+\cos\theta_X}\right] V_X^q & \Theta_X \\ \hline -(\Theta_X)^\dagger V_X^q & \cos\theta_X \end{array}\right) \tag{C.18}$$

The 3×3 mixing matrices V_X^q, which are unitary, are independent of the 3×1 column matrix Θ_X. The six light–light mixing angles parametrizing $V_{L,R}$ and the six light–heavy mixing parameters in $\Theta_{L,R}$ are fundamental parameters of these models that have to be extracted directly from experimental measurements.

With our block parametrization, the fermion mass eigenstates can now be expressed on general grounds as the following mixtures of interaction eigenstates:

$$q_X = \left[\mathbb{1} - \frac{\Theta_X\Theta_X{}^\dagger}{1 + \cos\theta_X}\right] V_X^q \cdot q'_X + \Theta_X \mathcal{Q}'_X \tag{C.19}$$

$$\mathcal{Q}_X = \cos\theta_X \mathcal{Q}'_X + \frac{1}{\cos\theta_X}\left[\mathbb{1} - \frac{\Theta_X\Theta_X{}^\dagger}{1 + \cos\theta_X}\right] V_X^q \cdot q'_X \tag{C.20}$$

C.2 Left-Handed mixings are small

In the following we show that for the mass texture in (C.1), the light–heavy LH mixings Θ_L are always small when the physical mass of the vector-like quark is much larger than the electroweak scale. We also provide general formulas for the LH parameters as a function of the mass scales of the model and the RH mixing parameters.

When diagonalizing the mass matrix (C.1), one finds the following relation between the LH light–heavy mixing Θ_L and its RH counterpart Θ_R:

$$\frac{\Theta_L}{\cos\theta_L} = \left(\frac{m_q}{m_\mathcal{Q}}\right) \Delta\hat{m} \cdot \frac{\Theta_R}{\cos\theta_R}\,. \tag{C.21}$$

where $m_q \equiv m_{q_3}$ is the physical mass of the heaviest quark in q^i (for up-type quarks this corresponds to the top-quark mass, for down-type quarks to the bottom-quark mass), $m_\mathcal{Q}$ the bare vector-like quark mass and $\Delta\hat{m}$ is the dimensionless diagonal mass matrix accounting for the SM quark mass hierarchy with respect to m_q, defined by

$$\Delta\hat{m} \equiv \text{diag}\left(\frac{m_{q_1}}{m_q}, \frac{m_{q_2}}{m_q}, 1\right) \tag{C.22}$$

where m_{q_1} and m_{q_2} are the masses of the first and second generation quarks, respectively. The bare mass $m_\mathcal{Q}$ can be written in terms of the physical vector-like fermion mass $M_\mathcal{Q}$ as

$$m_{\mathcal{Q}} \;=\; M_{\mathcal{Q}}\,\frac{\cos\theta_R}{\cos\theta_L} \tag{C.23}$$

From the diagonalization of the mass matrix we can solve $\cos\theta_L$ and Θ_L as a function of the RH mixing parameters and the physical masses of the model. A straightforward computation yields the following expressions:

$$\cos\theta_L \;=\; \left[1+\left(\frac{m_q}{M_{\mathcal{Q}}\cos\theta_R}\right)^2 \Theta_R{}^\dagger \Delta\hat{m}^2 \Theta_R\right]^{-1/2}, \tag{C.24}$$

$$\Theta_L \;=\; \left(\frac{m_q}{M_{\mathcal{Q}}\cos\theta_R}\right)\frac{\Delta\hat{m}\cdot\Theta_R}{\sqrt{1+\left(\frac{m_q}{M_{\mathcal{Q}}\cos\theta_R}\right)^2 \Theta_R{}^\dagger \Delta\hat{m}^2 \Theta_R}}\,. \tag{C.25}$$

This last (exact) relations can be used to express the the 4×4 LH matrix $\mathcal{U}_L$ in (C.8) exclusively as a function of the RH mixings:

$$\mathcal{U}_L \;=\; \left(\begin{array}{c|c} \mathbb{1} & \frac{m_q}{M_{\mathcal{Q}}\cos\theta_R}\,\Delta\hat{m}\cdot\Theta_R \\ \hline -\frac{m_q}{M_{\mathcal{Q}}\cos\theta_R}\,\Theta_R{}^\dagger\cdot\Delta\hat{m} & 1 \end{array}\right) \tag{C.26}$$

and the block-diagonal matrix as

$$\mathcal{D}_L \;=\; \left(\begin{array}{c|c} \left[\mathbb{1}-\frac{\left(\frac{m_q}{m_{\mathcal{Q}}}\right)^2(\Delta\hat{m}\,\Theta_R\Theta_R{}^\dagger\Delta\hat{m})}{1+\sqrt{1-\left(\frac{m_q}{m_{\mathcal{Q}}}\right)^2\Theta_R^\dagger\Delta\hat{m}^2\Theta_R}}\right] V_L^q & \mathbb{0} \\ \hline \mathbb{0} & \left[1+\left(\frac{m_q}{M_{\mathcal{Q}}\cos\theta_R}\right)^2 \Theta_R{}^\dagger \Delta\hat{m}^2 \Theta_R\right]^{-1/2} \end{array}\right) \tag{C.27}$$

The bare mass appearing in the upper left block can also be expressed in terms of the physical mass and RH mixing parameters via

$$m_{\mathcal{Q}} \;=\; M_{\mathcal{Q}}\cos\theta_R\sqrt{1+\left(\frac{m_q}{M_{\mathcal{Q}}\cos\theta_R}\right)^2 \Theta_R{}^\dagger \Delta\hat{m}^2 \Theta_R}\,. \tag{C.28}$$

Notice that the main expressions (C.26), (C.27) and (C.28) determine the full $\mathcal{V}_L$ mixing matrix as a function of Θ_R, the physical quark masses m_{q^i}, the physical vector-like quark partner mass $M_{\mathcal{Q}}$ and the unitary matrix V_L^q. These formulas are exact and do not rely on an approximation.

The expression we found for $\mathcal{V}_L$ suggests that the three LH light–heavy mixing parameters Θ_L^i are suppressed in the decoupling limit $M_{\mathcal{Q}} \gg m_q$ for generic values of the RH mixings. This is quite clear from expression (C.25). Indeed, we have checked numerically that Θ_L^i have small values if the mass of the bare mass term $M_{\mathcal{Q}}$ of the vector-like quark is above the $\mathcal{O}(1)$ TeV scale. For masses $m_{\mathcal{Q}} > 1$ TeV the largest mixing angle Θ_L^3 is of order $\mathcal{O}(10^{-1})$ while the other two, Θ_L^1 and Θ_L^2,

are completely negligible because of the small up and charm masses in the up-quark sector.

We now take the limit of large vector-like fermion mass $M_{\mathcal{Q}} \gg m_q$ which we assume to be around a few TeV. The exact expressions given above can be expanded up to second order in $m_q/M_{\mathcal{Q}} \cos\theta_R$. For instance:

$$\cos\theta_L \;=\; 1 \;-\; \frac{1}{2}\left(\frac{m_q}{M_{\mathcal{Q}}\cos\theta_R}\right)^2 \Theta_R{}^\dagger \Delta\hat{m}^2 \Theta_R \;+\; \mathcal{O}\left(\frac{m_q}{M_{\mathcal{Q}}\cos\theta_R}\right)^3 \tag{C.29}$$

$$\Theta_L \;=\; \frac{m_q}{M_{\mathcal{Q}}\cos\theta_R}\Delta\hat{m}\cdot\Theta_R \;+\; \mathcal{O}\left(\frac{m_q}{M_{\mathcal{Q}}\cos\theta_R}\right)^3 \tag{C.30}$$

$$R_L \;=\; \mathbb{1} \;-\; \frac{1}{2}\left(\frac{m_q}{M_{\mathcal{Q}}\cos\theta_R}\right)^2 \Delta\hat{m}\,\Theta_R\Theta_R{}^\dagger\Delta\hat{m} \;+\; \mathcal{O}\left(\frac{m_q}{M_{\mathcal{Q}}\cos\theta_R}\right)^3 \tag{C.31}$$

The LH mixing matrix finally reduces in the $M_{\mathcal{Q}} \gg m_q$ limit to the more tractable expression:

$$\mathcal{V}_L \;=\; \mathcal{V}_L^{(0)} \;+\; \left(\frac{m_q}{M_{\mathcal{Q}}\cos\theta_R}\right)\mathcal{V}_L^{(1)} \;-\; \frac{1}{2}\left(\frac{m_q}{M_{\mathcal{Q}}\cos\theta_R}\right)^2 \mathcal{V}_L^{(2)} \;+\; \mathcal{O}\left(\frac{m_q}{M_{\mathcal{Q}}\cos\theta_R}\right)^3 \tag{C.32}$$

where the first term $\mathcal{V}_L^{(0)}$ is a unitary matrix while $\mathcal{V}_L^{(1,2)}$ are not. These are given by the expressions:

$$\mathcal{V}_L^{(1)} \;=\; \left(\begin{array}{c|c} V_L^q & \mathbb{0} \\ \hline \mathbb{0} & 1 \end{array}\right) \tag{C.33}$$

$$\mathcal{V}_L^{(1)} \;=\; \left(\begin{array}{c|c} \mathbb{0} & \Delta\hat{m}\,\Theta_R \\ \hline -(\Theta_R)^\dagger\Delta\hat{m}\,V_L^q & 0 \end{array}\right) \tag{C.34}$$

$$\mathcal{V}_L^{(2)} \;=\; \left(\begin{array}{c|c} \Delta\hat{m}\,\Theta_R(\Theta_R)^\dagger\Delta\hat{m}\,V_L^q & \mathbb{0} \\ \hline \mathbb{0} & (\Theta_R)^\dagger\Delta\hat{m}^2\Theta_R \end{array}\right) \tag{C.35}$$

It can be checked that this expansion for $\mathcal{V}_L$ is in fact unitary up to order $\mathcal{O}(m_q/M_{\mathcal{Q}}\cos\theta_R)^2$. With this approximation the LH physical quarks can now be expressed as a mixture of the weak eigenstate quark fields

$$q_L \;\simeq\; \left[\mathbb{1} - \frac{1}{2}\left(\frac{m_q}{M_{\mathcal{Q}}}\right)^2 \frac{\Delta\hat{m}\,\Theta_R\Theta_R{}^\dagger\Delta\hat{m}}{\cos^2\theta_R}\right] V_L^q\cdot q_L' \;+\; \left(\frac{m_q}{M_{\mathcal{Q}}}\right)\frac{\Delta\hat{m}\cdot\Theta_R}{\cos\theta_R}\,\mathcal{Q}_L' \tag{C.36}$$

$$\mathcal{Q}_L \;\simeq\; \left[1 - \frac{1}{2}\left(\frac{m_q}{M_{\mathcal{Q}}}\right)^2 \frac{\Theta_R{}^\dagger\Delta\hat{m}^2\Theta_R}{\cos^2\theta_R}\right]\mathcal{Q}_L' \;-\; \left(\frac{m_q}{M_{\mathcal{Q}}}\right)\frac{\Theta_R{}^\dagger\cdot\Delta\hat{m}}{\cos\theta_R}\,V_L^q\cdot q_L' \tag{C.37}$$

As can be seen because of the small LH light-heavy mixing, the SM LH up-quarks are mostly LH light quarks while the exotic fermion is mostly comprised of LH heavy vector fermion. On the other, the RH fields in mix arbitrarily in this model.

Appendix D

Electroweak interactions in $\mathrm{U}(1)'$

The vector-like fermions in the top-philic $\mathrm{U}(1)'$ modify the SM electro-weak interactions giving rise to non-standard interactions for the W, Z and Higgs bosons with light quarks. This will produce deviations from unitarity in the CKM matrix and give rise to FCNC in the Z and Higgs couplings. We now write down explicitly these new interactions and give general formulas.

D.1 CKM non-unitarity

We first focus on charged currents, namely the coupling of the $W^{\pm}$ boson to LH fermions. The Lagrangian describing these interactions can be generically written in the eigenmass basis as

$$\mathcal{L}_W \;=\; \frac{g}{\sqrt{2}}\, W_\mu^{\pm} \left(\begin{array}{cc} \overline{u_L} & \overline{\mathcal{T}_L} \end{array} \right) \gamma^\mu \cdot V_L \cdot d_L \;+\; \text{h.c.} \tag{D.1}$$

where V_L is a 4×3 mixing matrix given by

$$V_L \;=\; \left(\begin{array}{c} V_{\mathrm{CKM}} \\ \\ V_L^{(d\mathcal{T})} \end{array} \right) \;=\; \mathcal{V}_L \left(\begin{array}{c} \mathbb{1} \\ \mathbb{0} \end{array} \right) (V_L^d)^\dagger \,. \tag{D.2}$$

One finds after a direct computation that

$$V_{\mathrm{CKM}} \;=\; \left[\mathbb{1} - \frac{\Theta_L {\Theta_L}^\dagger}{1+\cos\theta_L} \right] V_L^u (V_L^d)^\dagger \tag{D.3}$$

$$V_L^{(d\mathcal{T})} \;=\; -\frac{{\Theta_L}^\dagger}{\cos\theta_L} \left[\mathbb{1} - \frac{\Theta_L {\Theta_L}^\dagger}{1+\cos\theta_L} \right] V_L^u (V_L^d)^\dagger \tag{D.4}$$

Here the 3×3 matrix V_{CKM} is analogous to the SM CKM matrix connecting up and down quarks while $V_L^{(d\mathcal{T})}$ is a new 1×3 matrix governing the interaction between down, strange and bottom quarks with the new vector-like quark $\mathcal{T}$ via W boson exchange. We can rewrite these expression as a function of the RH mixing parameters and the physical masses as:

$$V_{\rm CKM} = \left[\mathbb{1} - \left(1 - \frac{1}{\sqrt{1+(\frac{m_t}{M_{\mathcal{T}}})^2 \frac{\Theta_R{}^\dagger \Delta\hat{m}^2 \Theta_R}{\cos\theta_R}}}\right) \frac{\Delta\hat{m}\,\Theta_R \Theta_R{}^\dagger \Delta\hat{m}}{\Theta_R{}^\dagger \Delta\hat{m}^2 \Theta_R}\right] V_L^u (V_L^d)^\dagger \quad \text{(D.5)}$$

$$V_L^{(d\mathcal{T})} = -\frac{m_t}{M_{\mathcal{T}}} \frac{\Theta_R{}^\dagger}{\cos\theta_R} V_{\rm CKM} \quad \text{(D.6)}$$

It is convenient to expand the resulting CKM matrix in the $M_{\mathcal{T}} \gg m_t$ limit up to order $\mathcal{O}(m_t/M_{\mathcal{T}})^2$. We find the following expression

$$V_{\rm CKM} \simeq (V_L^u)^\dagger V_L^d - \frac{1}{2}\left(\frac{m_t}{M_{\mathcal{T}} \cos\theta_R}\right)^2 \Delta\hat{m}\,\Theta_R \Theta_R{}^\dagger \Delta\hat{m} \cdot V_L^u (V_L^d)^\dagger \quad \text{(D.7)}$$

The first term in (D.7) is a product of two unitary matrices hence it is unitary, while the second term in (D.7) spoils unitarity of the CKM matrix at order $\mathcal{O}(m_t/M_{\mathcal{T}})^2$. In the decoupling limit taking $M_{\mathcal{T}} \to \infty$ and $\Theta_R \to \mathbb{0}$ we, in fact, recover the SM definition of the CKM matrix $V_{\rm CKM} = (V_L^u)^\dagger V_L^d$, i.e. a measure of the misalignment between the up and down quark Yukawa sectors. The spoiling of unitarity is less severe for the first two generations given that these are protected by the large hierarchy of masses in $\Delta\hat{m}^2$. Again, the largest deviations of CKM unitarity in our U(1)′ model are expected in the third row and column for the top-quark couplings to $W^\pm$.

D.2 Z boson couplings

We now turn to the neutral currents coupling to the Z boson. Since the vector-like quark $\mathcal{T}$ has the same quantum numbers as the RH up-quarks the Z boson will only have modified couplings to the LH fermion fields, parametrized by the 4×4 matrix X_L. In the physical eigenbasis the Lagrangian reads

$$\mathcal{L}_Z = \frac{g}{2c_W} Z_\mu \left[\frac{1}{2}\begin{pmatrix} \overline{u_L} & \overline{\mathcal{T}_L} \end{pmatrix} \gamma^\mu X_L \begin{pmatrix} u_L \\ \mathcal{T}_L \end{pmatrix} - s_W^2 Q_{em} \begin{pmatrix} \overline{u_X} & \overline{\mathcal{T}_X} \end{pmatrix} \gamma^\mu \begin{pmatrix} u_X \\ \mathcal{T}_X \end{pmatrix}\right]. \quad \text{(D.8)}$$

where Q_{em} is the electromagnetic charge and θ_W is the Weinberg angle and the Hermitian mixing matrix X_L is given by

$$X_L = \left(\begin{array}{c|c} X_L^{(u)} & X_L^{(u\mathcal{T})} \\ \hline X_L^{(u\mathcal{T})\dagger} & X_L^{(\mathcal{T})} \end{array}\right) = \mathcal{V}_L \begin{pmatrix} \mathbb{1} & \mathbb{0} \\ \mathbb{0} & 0 \end{pmatrix} \mathcal{V}_L^\dagger, \quad \text{(D.9)}$$

and each block matrix has the following form:

$$X_L^{(u)} = \mathbb{1} - \Theta_L \Theta_L{}^\dagger \tag{D.10}$$

$$X_L^{(u\mathcal{T})} = -(\cos\theta_L)\,\Theta_L\,, \tag{D.11}$$

$$X_L^{(\mathcal{T})} = \sin^2\theta_L\,. \tag{D.12}$$

Solving these blocks in the RH mixings and masses yields the LH couplings of the Z boson:

$$X_L^{(u)} = \mathbb{1} - \left(\frac{m_t}{M_{\mathcal{T}}\cos\theta_R}\right)^2 \left[\frac{\Delta\hat{m}\,\Theta_R\Theta_R{}^\dagger\Delta\hat{m}}{1+(\frac{m_t}{M_{\mathcal{T}}\cos\theta_R})^2\,\Theta_R{}^\dagger\Delta\hat{m}^2\Theta_R}\right] \tag{D.13}$$

$$X_L^{(u\mathcal{T})} = -\left(\frac{m_t}{M_{\mathcal{T}}\cos\theta_R}\right) \left[\frac{\Delta\hat{m}\,\Theta_R}{1+(\frac{m_t}{M_{\mathcal{T}}\cos\theta_R})^2\,\Theta_R{}^\dagger\Delta\hat{m}^2\Theta_R}\right] \tag{D.14}$$

$$X_L^{(\mathcal{T})} = \left(\frac{m_t}{M_{\mathcal{T}}\cos\theta_R}\right)^2 \left[\frac{\Theta_R{}^\dagger\Delta\hat{m}^2\Theta_R}{1+(\frac{m_t}{M_{\mathcal{T}}\cos\theta_R})^2\,\Theta_R{}^\dagger\Delta\hat{m}^2\Theta_R}\right]. \tag{D.15}$$

The expression for the Z boson coupling in (D.13) gives rise to very dangerous tree-level FCNC in the Z coupling to up quark currents. These are however suppressed to all quarks by the large cutoff, and even more suppressed for the first two generations because of the small up and charm quark masses.

Now that we have written the general expressions for all the gauge boson couplings, we can see how these are related between themselves. One can derive the following useful *sum rules* for the U(1)′ model relating the three massive gauge bosons interacting with LH light–light charged and neutral currents:

$$X_L^{(u)} = V_{\rm CKM}V_{\rm CKM}{}^\dagger = \mathbb{1} - K_L^{(u)} \tag{D.16}$$

One nice way to interpret these relations is the following: *the deviations from flavor universality of the Z boson interactions with LH up-quark currents match the expected departures from CKM unitarity in charged currents and determine exactly the Z′ couplings to LH up-quarks FCNC, all proportional to* $(m_u^i/M_{\mathcal{T}})^2$. Of course this statement holds precisely at tree level.

D.3 The Higgs interactions

We now give expressions for the couplings of the SM Higgs h. As we show below, the U(1)′ model will modify the SM interactions of the Higgs boson with the upper quarks giving rise to FCNC at tree level. Mixing also introduces new Yukawa interactions between the Higgs and the vector-like fermion.

The relevant Lagrangian containing the Higgs interactions in the mass eigenbasis is given by

$$-\mathcal{L}_h \;=\; h\,(\overline{u_L}\,,\overline{\mathcal{T}_L})\cdot Y\cdot\begin{pmatrix} u_R \\ \mathcal{T}_R\end{pmatrix} \;+\; h\,\overline{d_L}\,\frac{m_d}{v}\,d_R \;+\; \text{h.c.} \tag{D.17}$$

where h is the Higgs boson, m_d is the 3×3 down quark mass matrix and v is the electroweak breaking VEV $v\simeq 246$ GeV. The matrix Y is a 4×4 Yukawa matrix defined here to be the expressions:

$$Y \;=\; \left(\begin{array}{c|c} Y^{\,(u)} & Y^{\,(u\mathcal{T})} \\ \hline Y^{\,(\mathcal{T}u)^\dagger} & Y^{\,(\mathcal{T})}\end{array}\right) \;=\; \frac{1}{v}\,\mathcal{V}_L\begin{pmatrix} m_u & \mathbb{0} \\ \mathbb{0} & 0\end{pmatrix}\mathcal{V}_R^\dagger\,, \tag{D.18}$$

where m_u is the up-quark 3×3 mass matrix (C.2), $Y^{\,(u)}$ a 3×3 block describing the coupling of the Higgs to light up-quarks, $Y^{\,(u\mathcal{T})}$ and $Y^{\,(\mathcal{T}u)^\dagger}$ are 3×1 blocks for the mixed couplings of the Higgs to light–heavy bilinears, and $Y^{\,(\mathcal{T})}$ the 1×1 coupling of the Higgs to the vector-like quark bilinear. We find that these matrices are given by

$$Y^{\,(u)} \;=\; \frac{m_t}{v}\,\left(\mathbb{1}\;-\Theta_L{\Theta_L}^\dagger\right)\,\Delta\hat{m} \tag{D.19}$$

$$Y^{\,(u\mathcal{T})} \;=\; \left(\frac{1}{\cos\theta_R}\right)\,Y^{\,(u)}\Delta\hat{m}\,\Theta_R \tag{D.20}$$

$$Y^{\,(\mathcal{T}u)} \;=\; \left(\frac{m_t}{M_\mathcal{T}\cos\theta_R}\right)\,\Delta\hat{m}\,Y^{\,(u)}\,\Theta_R \tag{D.21}$$

$$Y^{\,(\mathcal{T})} \;=\; \left(\frac{m_t}{M_\mathcal{T}\cos^2\theta_R}\right)\,{\Theta_R}^\dagger\,\Delta\hat{m}\,Y^{\,(u)}\,\Theta_R \tag{D.22}$$

Expressing the Yukawa couplings of light states as a function of RH mixing parameters gives the expected result

$$Y^{\,(u)} \;=\; \frac{\Delta m_u}{v} \;-\; \frac{m_t}{v}\left(\frac{m_t}{M_\mathcal{T}\cos\theta_R}\right)^2\left[\frac{\Delta\hat{m}\,\Theta_R{\Theta_R}^\dagger\Delta\hat{m}}{1+(\frac{m_t}{M_\mathcal{T}\cos\theta_R})^2\,{\Theta_R}^\dagger\Delta\hat{m}^2\Theta_R}\right]\,\Delta\hat{m} \tag{D.23}$$

where we remind that $\Delta m_u\equiv\text{diag}(m_u,m_c,m_t)$. The first term of this expression coincides with the prediction of the SM, i.e. that the Higgs couples in a flavor diagonal fashion with couplings proportional to the mass of the fermion. The second term in (D.23), on the other hand, violates explicitly this crucial SM prediction, introducing couplings of the Higgs to FCNC in the up-quark sector. Nevertheless, such FCNC are heavily suppressed by the mass ratio $(m_t/M_\mathcal{T})^2$. The mass hierarchy of up-quarks predicts that the largest flavor changing effects will be between the charm and the top-quark that can be probed in anomalous decays of the top quark of the form $t\to hc$. Flavor diagonal deviations of the top-quark Yukawa are also expected to be potentially the largest.

Above we showed by the exact relation (D.16) that the deviations from flavor universality in Z boson interactions and the violation of CKM unitarity can be

parametrized by the Z' couplings to FCNC for LH up-quarks. For the Higgs in the U(1)$'$ we obtain the following useful sum-rule

$$Y^{(u)} \;=\; X_L^{(u)}\,\frac{\Delta m_u}{v} \;=\; \left[\mathbb{1} \;-\; K_L^{(u)}\right]\,\frac{\Delta m_u}{v} \tag{D.24}$$

Here we find again that the deviations from the SM behavior in Higgs couplings are equivalent to those in the Z boson couplings (or equivalently to the violation of CKM unitarity) and are determined by the Z' couplings to LH up-quarks.

Appendix E

Features of the SU(4) models

E.1 Gauge boson spectrum

The mass eigenstates of the gauge bosons can be directly extracted from the kinetic terms of the condensed scalars $\langle\Omega_{1,3}\rangle$. For the leptoquark U_μ, the coloron g' and the SM gluons g one obtains

$$U_\mu^{1,2,3} = \frac{1}{\sqrt{2}}(H_\mu^{9,11,13} - iH_\mu^{10,12,14}) \tag{E.1}$$

$$g_\mu'^a = \frac{1}{\sqrt{g_4^2 + g_3^4}}\,(g_4\,H_\mu^a - g_3\,G_\mu'^a) \tag{E.2}$$

$$g_\mu^a = \frac{1}{\sqrt{g_4^2 + g_3^4}}\,(g_3\,H_\mu^a + g_4\,G_\mu'^a). \tag{E.3}$$

with mass spectrum given by

$$m_U^2 = \frac{1}{4}g_4^2(v_1^2 + v_3^2) \tag{E.4}$$

$$m_{g'}^2 = \frac{1}{2}(g_3^2 + g_4^2)v_3^2 \tag{E.5}$$

$$m_g^2 \;= 0\,. \tag{E.6}$$

In general, the spectrum of the colorless gauge bosons Z' and B_μ depends on the still unspecified Y' charges $\omega_{1,3}$ of the scalars. The squared mass matrix M_0^2 for these neutral vector bosons is found to be

$$M_0^2 = \begin{pmatrix} \frac{1}{8}g_4^2\,(3v_1^2 + v_3^2) & \sqrt{\frac{3}{8}}g_1g_4(\omega_1v_1^2 - \omega_3v_3^2) \\ \sqrt{\frac{3}{8}}g_1g_4(\omega_1v_1^2 - \omega_3v_3^2) & g_1^2(\omega_1^2v_1^2 + 3\omega_3^2v_3^2) \end{pmatrix}. \tag{E.7}$$

We now demand that the SM hypercharge Y remains unbroken after the first stage symmetry breaking. This gives rise to a massless boson B_μ and implies that the determinant of the mass matrix M_0 vanishes. We find that this condition is fulfilled only if the primed hypercharges for the Ω_3 and Ω_1 scalars satisfy the relation

$$\omega_3 \;=\; -\frac{\omega_1}{3}. \tag{E.8}$$

After imposing this condition one finds the mass eigenstates for the two neutral gauge bosons to be

$$Z'_\mu = \frac{1}{\sqrt{24\,\omega^2\,g_1^2+g_4^2}}\left(2\sqrt{6}\,g_1\,\omega\,B'_\mu - g_4 H^{15}_\mu\right) \tag{E.9}$$

$$B_\mu = \frac{1}{\sqrt{24\,\omega^2\,g_1^2+g_4^2}}\left(g_4 B'_\mu + 2\sqrt{6}\,g_1\,\omega\,H^{15}_\mu\right) \tag{E.10}$$

with masses given by

$$m^2_{Z'} = \frac{1}{8}\,(24\,\omega^2\,g_1^2+g_4^2)\,(3\,v_1^2+v_3^2) \tag{E.11}$$

$$m^2_B = 0\,. \tag{E.12}$$

All these relations coincide with those derived in Ref. [200] for the particular case $\omega = 1/6$.

E.2 The scalar sector

We now comment on the scalar sector of this model which has a rich structure. One can decompose the Ω_3 and Ω_1 scalars into the $\mathcal{G}_{SM}$ subgroup:

$$\Omega_3 \sim (\mathbf{8},\mathbf{1},0)\,\oplus\,(\mathbf{3},\mathbf{1},4\,\omega)\,\oplus\,(\mathbf{1},\mathbf{1},0) \tag{E.13}$$

$$\Omega_1 \sim (\bar{\mathbf{3}},\mathbf{1},-4\,\omega)\,\oplus\,(\mathbf{1},\mathbf{1},0)\,. \tag{E.14}$$

The imaginary part of the octet in (E.13) corresponds to the would-be Nambu-Goldstone bosons that are eaten away by the longitudinal components of the colorons g'. The real part of this octet gives rise to a propagating degree of freedom, a massive scalar color octet S_O. Next, there are two complex color triplets $S_T^{3,1} \sim (\mathbf{3},\mathbf{1},4\,\omega)$, one in Ω_3 and one in Ω_1. One particular linear combination of S_T^3 and S_T^1 will be absorbed by the vector leptoquark U_R and provide it with mass, while the orthogonal combination will be a physical degree of freedom. Finally what remains are two complex scalar singlets $S^{3,1} \sim (\mathbf{1},\mathbf{1},0)$ each embedded in $\Omega_{3,1}$. One linear combination is absorbed by the Z' gauge boson while the other three are propagating degrees of freedom.

The vacuum structure in (6.57) can be found by minimizing the most general scalar potential consistent with gauge symmetry:

$$-V(H,\Omega) = V_H + V_{\Omega_1} + V_{\Omega_3} + V_{\text{portal}} + V_{\text{mix}} \tag{E.15}$$

where φ is the standard Higgs doublet and

$$V_\varphi = \mu^2(\varphi^\dagger\varphi) - \lambda(\varphi^\dagger\varphi)^2 \tag{E.16}$$

$$V_{\Omega_1} = \mu_1^2\,(\Omega_1^\dagger\Omega_1) - \lambda_1\,(\Omega_1^\dagger\Omega_1)^2 \tag{E.17}$$

$$V_{\Omega_3} = \mu_3^2\,\mathrm{Tr}(\Omega_3^\dagger\Omega_3) - \lambda_3\,\mathrm{Tr}(\Omega_3^\dagger\Omega_3)^2 \tag{E.18}$$

$$V_{\text{portal}} = \lambda_{\varphi 1}\,(\varphi^\dagger\varphi)(\Omega_1^\dagger\Omega_1) + \lambda_{\varphi 3}\,(\varphi^\dagger\varphi)\mathrm{Tr}(\Omega_3^\dagger\Omega_3) + \lambda_{13}\,(\Omega_1^\dagger\Omega_1)\mathrm{Tr}(\Omega_3^\dagger\Omega_3) \tag{E.19}$$

$$V_{\text{mix}} = \lambda_5\,\Omega_1^\dagger(\Omega_3\Omega_3^\dagger)\Omega_1 + \lambda_6\,\epsilon_{ABCD}\epsilon^{abc}\,(\Omega_1)^A(\Omega_3)^B_a(\Omega_3)^C_b(\Omega_3)^D_c \tag{E.20}$$

For a specific choice of the potential parameters one can find that the vacuum configuration given in (6.57) corresponds to a local minimum. For details see [200].

E.2.1 SU(4) generators

The 15 4×4 generators of the SU(4) algebra are $\Gamma^A \equiv T^A/2$ defined by:

$$T^1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad , \quad T^2 = \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{E.21}$$

$$T^3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad , \quad T^4 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{E.22}$$

$$T^5 = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad , \quad T^6 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{E.23}$$

$$T^7 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad , \quad T^8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{E.24}$$

$$T^9 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad , \quad T^{10} = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix} \tag{E.25}$$

$$T^{11} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad , \quad T^{12} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix} \tag{E.26}$$

$$T^{13} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad , \quad T^{14} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \end{pmatrix} \tag{E.27}$$

$$T^{15} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix} \tag{E.28}$$

These satisfy the normalization $\mathrm{Tr}(T^A T^B) = 2\delta^{AB}$.

Razširjeni povzetek v slovenskem jeziku

8.1 Uvod

Leta 2012 je bil odkrit zadnji in najbolj razvpit delec, ki ga napoveduje Standardni Model (SM). Odkritje je bilo potrjeno simultano s strani eksperimentov ATLAS in CMS, kjer oba potekata na največjem pospeševalniku na svetu, velikem hadronskem trkalniku LHC (ang. *Large Hadron Collider*) v CERN-u. Odkritje Higgsovega bozona je zaznamovalo konec 60-letnega iskanja tega delca, hkrati pa tudi začetek novega obdobja iskanja fizike izven SM, imenovane tudi fizika BSM (ang. *beyond Standard Model*). Poleg Higgsovega in umeritvenih bozonov je kvark t še edini preostali delec z maso na elektrošibki skali, ki lahko igra pomembno vlogo pri iskanju fizike BSM. Dosedanje fizikalne omejitve iz podatkov LHC Run-I, LHC Run-II in Tevatrona ter elektrošibke in nizko-energijske omejitve v fiziki okusa so še vedno najšibkeje določene za fermione tretje generacije. To omogoča, da se potencialni pojavi fizike BSM pri kvarkih t lahko skrivajo znotraj faznega prostora parametrov, ki so še vedno neraziskani na energijski skali, dostopni na LHC. V tej tezi raziskujemo fenomenologijo fizike BSM pri pospeševalnikih, ki je v glavnem povezana s kvarki t v sektorju kvarkov, kjer ni resonančnih učinkov na trkalniku LHC.

Eden od zanimivejših procesov, ki so ga pred kratkim odkrili pri eksperimentih ATLAS in CMS, je produkcija tth. Ta proces je bistvenega pomena za karakterizacijo lastnosti Higgsovega bozona in je tudi občutljiv na več efektivnih operaterjev fizike BSM. Razvili smo nov način proučevanja produkcije tth, oziroma katerega koli podobnega procesa z veliko mnogoterostjo težkih resonanc Standardnega Modela (t, h, $W^{\pm}$, Z) v končnem stanju. V ta namen uvajamo množico t.i. *idealiziranih sipalnih presekov*, ki temeljijo na zelo naravni razdelitvi faznega prostora v različne kinematične regije. Razdelitev je bila navdihnjena s strani podstrukturnih tehnik potisnjenih *debelih* curkov (ang. *fat jets*), ki jih je mogoče zlahka uporabiti v eksperimentalni postavitvi. Kot dokaz načela smo izračunali idealizirane sipalne preseke za produkcijo tth in prikazali, kako so le-ti zelo občutljivi na različne scenarije fizike BSM. To izrecno pokažemo za reprezentativno množico efektivnih operaterjev dimenzije 6, ki vstopajo v produkcijo tth.

Eden od ciljev te disertacije je pokazati, da je nova fizika (ang. *new physics* ali NP) lahko tudi relativno lahka in skrita v eksotičnih signalih na trkalniku LHC. Zlasti nova lahka stanja, ki so povezana izključno s kvarki t, ne bodo vplivala na tipične procese na nivoju drevesnih diagramov. Simultana produkcija štirih kvarkov t na lupini, $pp \to t\bar{t}t\bar{t}$, na hadronskem trkalniku, je eden od ultimativnih fizikalnih izzivov za LHC v naslednjih letih. V tem delu predlagamo zelo preprosto strategijo

iskanja SM procesa $pp \to t\bar{t}t\bar{t}$, ki temelji na najbolj občutljivih razpadnih kanalih, in sicer na enako nabitem dileptonskem in na trileptonskem razpadnem kanalu. Naše rezultate uporabljamo, da postavimo prve omejitve na ne-resonančne t-filične Z' in t-filične skalarne modele. Prav tako prikažemo, da bo LHC lahko v bližnji prihodnosti enostavno izmeril sipalni presek za SM produkcijo 4-t stanja s to večleptonsko strategijo.

V zadnjem delu teze preskočimo v področje fizike nizkih energij. Prvi znaki nove fizike v povezavi s kvarki t bi se lahko prvič pojavili kot majhna odstopanja pri natančnostih meritvah fizikalnih opazljivk v redkih procesih na skalah veliko nižjih od elektrošibke. Zanimivo je tudi dejstvo, da je bilo pri nizko-energijskih eksperimentih razpadov mezonov B v t.i. tovarnah mezonov B poročanih več indikacij kršitve leptonske univerzalnosti (ang. *Lepton Flavor Universality* ali LFU), ki presegajo vrednosti, pričakovane v SM. Predvsem nas zanimajo neskladja pri prehodih $b \to s$, ki spreminjajo okus, izmerjena s strani kolaboracije LHCb v CERN-u pri semi-leptonskih razpadih $B \to K\ell^+\ell^-$ in $B \to K^*\ell^+\ell^-$ [14, 15, 16, 17, 18]. V zadnjih letih so pri eksperimentu LHCb napravili vrsto natančnih meritev razmerij razpadov, ki kršijo LFU, kjer primerjajo muonske in elektronske razpade:

$$R_{K^{(*)}} \equiv \frac{\mathrm{Br}(B \to K^{(*)}\mu^+\mu^-)}{\mathrm{Br}(B \to K^{(*)}e^+e^-)}\,. \tag{8.29}$$

Za ti dve razmerji je bilo izmerjeno, da sta od napovedi SM oddaljeni za signifikanco približno 2.5σ. V našem delu raziskujemo možnosti pojasnjevanja odstopanj te vrste na ravni ene zanke skupaj z novo fiziko in pretežno v relaciji s kvarki t in mioni, vse skupaj znotraj opisa efektivne teorije in poenostavljenih dinamičnih modelov. Analiziramo tudi LHC fenomenologijo različnih t-filičnih rešitev anomalij $R_{K^{(*)}}$ in predlagamo nove iskalne kanale za LHC. Nazadnje, razlaga $R_{K^{(*)}}$ z masivnimi vektorskimi bozoni pomeni, da mora obstajati več prostostnih stopenj, zato da lahko imamo renormalizabilno teorijo, ki je konsistentna pri visokih energijah. V delu pokažemo, da je možno zgraditi razred ultra-vijoličnih (UV) zaključkov, ki temeljijo na umeritveni množici SU(4), tako da je katerikoli znan vektorski leptokvark v literaturi (ki je $\mathrm{SU}(2)_L$ singlet) predstavljen kot umeritveni bozon. Motivirani s strani anomalijam fizike B mezonov, konstruiramo t-filičen UV model, kjer umeritveni leptokvark in skalar povzročita kršitev LFU v prehodih $b \to s$ pri nizkih energijah.

8.2 Idealizirane opazljivke v produkciji tth

Naš prvotni cilj je karakterizacija nove fizike ki vstopa v generične $2 \to N$ delčne več-resonančne razpadne procese pri hadronskih trkalnikih

$$pp \to \; \mathcal{X}^1\mathcal{X}^2 ... \mathcal{X}^N \; \to \; x^1x^2 ... x^n\,, \tag{8.30}$$

kjer $\mathcal{X}^j$ predstavlja katerokoli od znanih težkih SM resonanc $\mathcal{X} = \{W^\pm, Z, t, h\}$ in x^j predstavljajo katerekoli od možnih brezmasnih končnih stanj $x = \{\gamma, \ell^\pm, \nu_\ell, u, d, s, c, b\}$. To nato uporabimo za analizo produkcije tth. Metoda, ki jo tukaj opisujemo, temelji na definiciji množice opazljivk o_i, ki jih imenujemo *idealizirane opazljivke*. Formalno gledano, so te opazljivke $\{o_1, o_2, ..., o_k\}$ sestavljene iz *bazne opazljivke* o in množice *kategorij* $\mathcal{K}$. Idealizirane opazljivke $o_i(\mathcal{A})$ so definirane kot omejitev opazljivke o na kategorijo $\kappa_i \in \mathcal{K}$, t.j. $o_i \equiv o|_{\kappa_i}$, tako da velja $o = \sum_{i=1}^k o_i$.

Kinematične kategorije za več-resonančne procese Sedaj definiramo množico kategorij $\mathcal{K}$ za več-resonančne procese SM, ki temeljijo na binarni klasifikaciji, $\kappa = \{B, R\}$, za vsako posamezno resonanco $\mathcal{X}^j$. Tukaj predpostavimo, da je $\mathcal{X}^j$ mogoče najti v enemu izmed *dveh* kinematičnih stanj: v t.i. potisnjenem ($\kappa = B$) ali pa v *mirnem* ($\kappa = R$) stanju. Vsako resonanco označimo s kategoričnimi spremenljivkami κ_j kot $\mathcal{X}^j_{\kappa_j}$ z $\kappa_i = \{B, R\}$ in vse kategorije v $\mathcal{K}$ z združitvijo $\mathcal{X}^1_{\kappa_1}\mathcal{X}^2_{\kappa_2}...\mathcal{X}^N_{\kappa_N}$. Pomniti velja, da če so vse resonance v $pp \to \mathcal{X}^1\mathcal{X}^2...\mathcal{X}^N$ različne (t.j. z različnimi kvantnimi števili in masami), bo v $\mathcal{K}$ možnih $k = 2^N$ kategorij. Množico kinematičnih kategorij nato zapišemo kot

$$\mathcal{K} \;=\; \left\{\mathcal{X}^1_R\mathcal{X}^2_R\mathcal{X}^3_R...\mathcal{X}^N_R\,,\;\mathcal{X}^1_B\mathcal{X}^2_R\mathcal{X}^3_R...\mathcal{X}^N_R\,,\;\mathcal{X}^1_R\mathcal{X}^2_B\mathcal{X}^3_R...\mathcal{X}^N_R\,,\;...\,,\;\mathcal{X}^1_B\mathcal{X}^2_B\mathcal{X}^3_B...\mathcal{X}^N_B\right\}\,. \tag{8.31}$$

$\mathcal{K}$ ima samo eno *mirno stanje*, $\mathcal{X}^1_R\mathcal{X}^2_R...\mathcal{X}^N_R$, kjer so vse nastale resonance v mirovanju. Pričakujemo, da bo večina dogodkov spadala v to kategorijo zaradi prevladujočega praga za produkcijo. Za preostalih 2^N-1 *potisnjenih kategorij* je pričakovana manjša zasedenost v SM in večja občutljivost na NP.

Na nivoju partonov lahko definiramo kategoriji R in B za vsako resonanco na podlagi tega, ali vsi brezmasni razpadni produkti partonske resonance padejo znotraj stožca s fiksnim radijem ΔR ali ne:

- *Klasifikator na osnovi stožca: Resonanca $\mathcal{X}$ je klasificirana kot potisnjena ($\mathcal{X}_B$), če je za vse pare hčerinskih delcev x^i in x^j v razpadu $\mathcal{X} \to x^1x^2...x^n$ izpolnjen pogoj $\max(R_{ij}) < \Delta R_{\mathcal{X}}$ drugače je resonanca v mirovanju ($\mathcal{X}_R$). Tukaj je $R_{ij} \equiv R(x^i, x^j) = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}$ ravninska razdalja med parom hčerinskih delcev in $\Delta R_{\mathcal{X}}$ je $\mathcal{O}(1)$ parameter, specifičen za resonanco.*

Parameter stožca $\Delta R_{\mathcal{X}}$ je prost parameter, ki ga lahko optimiziramo za vsako težko resonanco. Opaziti velja, da je klasifikator na osnovi stožca geometrijskega izvora. Popolnoma je neodvisen od narave ali mnogoterosti razpadajočih hčerinskih delcev uporabljenih v postopku. Lahko se ga izračuna za tako leptonske kot za hadronske $1 \to n$ razpade. Druga, morda preprostejša možnost, je razvrstitev resonanc na podlagi transverzalne gibalne količine, ki jo nosijo, ne da bi uporabili kinematiko razpada:

- *Klasifikator na podlagi transverzalne gibalne količine p_T: Resonanca $\mathcal{X}$ je klasificirana kot potisnjena ($\mathcal{X}_B$), če je izpolnjen pogoj $p_T^{\mathcal{X}} > q_T^{\mathcal{X}}$, drugače je resonanca v mirovanju ($\mathcal{X}_R$). $p_T^{\mathcal{X}}$ predstavlja partonsko transverzalno gibalno količino resonance in $q_T^{\mathcal{X}}$ prestavlja parameter mase reda velikosti $\mathcal{O}(10^2)$ GeV, specifičnega za resonanco.*

Sedaj se osredotočimo na produkcijo znotraj *tth* znotraj SM na trkalniku LHC in v efektivni teoriji polja znotraj SM (SMEFT). V tem več-resonančnem postopku ni možno razlikovati med kvarki t in $\bar{t}$, kar pomeni, da obstaja šest inekvivalentnih kinematičnih kategorij $t_Xt_Yh_Z$, kjer je $X, Y, Z \in \{B, R\}$: ena kategorija, ki ustreza stanju mirovanja $t_Rt_Rh_R$ in pet potisnjenih kategorij, ki razdeljujejo celoten fazni prostor. To so

$$\mathcal{K}_{t\bar{t}h} \;=\; \{t_Rt_Rh_R,\; t_Rt_Rh_B,\; t_Rt_Bh_R,\; t_Bt_Rh_B,\; t_Bt_Bh_R,\, t_Bt_Bh_B\}\,. \tag{8.32}$$

Za osnovno opazljivko o uporabimo totalni sipalni presek za produkcijo $o = \sigma_{\text{tot}}(pp \to t\bar{t}h)$. Na račun tega dobimo šest *idealiziranih sipalnih presekov* $o_i = \sigma(t_Xt_Yh_Z)$. Po

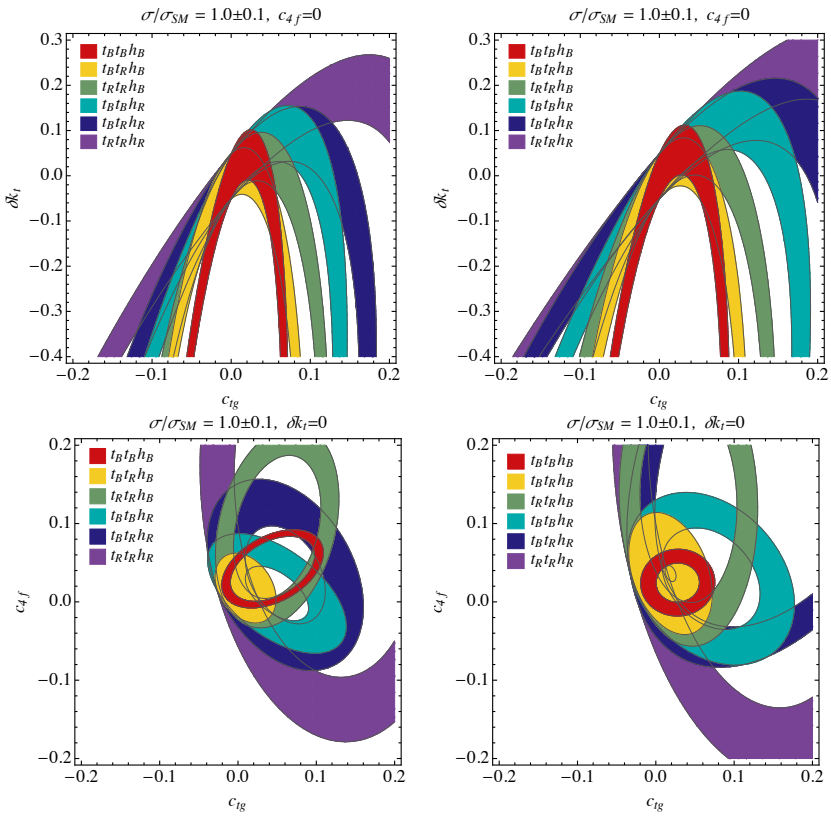


Figure 8.1: (Zgoraj) Občutljivost na operatorje EFT dimenzije 6 $(\delta\kappa_t, c_{tg})$ vsake kinematične kategorije $t_X t_Y h_Z$ za kategorizacijo na podlagi stožca (levo) in kategorizacija na podlagi p_T (desno). Nastavili smo $c_{4f} = 0$. Barvni pasovi prikazujejo $\mu^{\text{EFT}}_{XYZ} = 1 \pm 0.1$. (Spodaj) Občutljivost na operatorje EFT dimenzije 6 (c_{tG}, c_{4f}) vsake kinematične kategorije $t_X t_Y h_Z$ za kategorizacijo na podlagi stožca (levo) in kategorizacija na podlagi p_T (desno). Nastavili smo $\delta k_T = 0$.

definiciji zadostujejo

$$\sigma_{\text{tot}}(pp \to t\bar{t}h) \;=\; \sum_{X,Y,Z\in\{B,R\}} \sigma(t_X t_Y h_Z)\,. \tag{8.33}$$

Prav tako je priročno normalizirati idealizirane sipalne preseke glede na totalni sipalni presek za produkcijo $\mu_{XYZ} \equiv \sigma(t_X \bar{t}_Y h_Z)/\sigma_{\text{tot}}(pp \to t\bar{t}h)$.

Občutljivost $\sigma(t_X t_Y h_Z)$ na ne-resonančne procese NP Z namenom, da bi pokazali, kako je šest kategorij občutljivih na različne smeri v faznem prostoru parametrov procesov NP, izračunamo $\sigma(t_X t_Y h_Z)$ z naslednjo podmnožico SMEFT operaterjev dimenzije 6:

$$\begin{aligned}\mathcal{L}^{\text{EFT}} \;\supset\; & -\; \delta k_t \frac{y_t}{v^2}\bar{Q}t\tilde{\varphi}\left(\varphi^\dagger\varphi - \frac{v^2}{2}\right) \;+\; \text{h.c.} \\ & -\; c_{tg}\frac{g_s y_t}{4v^2}\bar{Q}\sigma^{\mu\nu}T^A t G^A_{\mu\nu}\tilde{\varphi} \;+\; \text{h.c.} \\ & +\; \frac{c_{4f}}{v^2}\sum_{i=1,2}\left[(\bar{Q}u^i_R)(\bar{u}^i_R Q) + (\bar{q}^i_L t)(\bar{t}q^i_L)\right],\end{aligned} \tag{8.34}$$

kjer je $Q = (t_L, b_R)^T$ tretja generacija levo-ročnega (LH) dubleta kvarkov, $q_L = (u_L^i, d_R^i)^T$ sta prvi in drugi generacijski LH dublet kvarkov, $t = t_R$ je desno-ročni (RH) kvark t, G je jakost polja QCD, y_t predstavlja sklopitev Yukawe za kvark t, g_s je močna sklopitev, $v = 256$ GeV predstavlja Higgsovo vakuumsko pričakovano vrednost in φ Higgsov dublet. Relevantni Wilsonovi koeficienti so δk_t, c_{tg} in c_{4f}. Slika 8.1 prkazuje območja $\mu_{XYZ}^{\text{EFT}} = 1 \pm 0.1$ za vsako kategorijo znotraj $(\delta k_t - c_{tg})$ in $(c_{4f} - c_{tg})$ projiciranih ravnin, za katere je bil tretji Wilsonov koeficient postavljen na vrednost 0. V vsaki sliki podamo rezultate za kategorizacijo na podlagi stožca (levo) in kategorizacijo na podlagi p_T (desno) idealiziranih opazljivk. Zanimivo je, da združena množica idealiziranih opazljivk odstrani vse ravne smeri. Vsi pasovi se sekajo v zaprtem območju v bližini napovedi SM (ničelni Wilsonovi koeficienti). Najbolj ortogonalne kategorije so kategorije mirovanja $t_R t_R h_R$ (vijolične regije) in popolnoma potisnjene kategorije $t_B t_B h_B$ (rdeče regije) z vsemi ostalimi kategorijami vmes. Iz slik je tudi razvidno, da idealizirani sipalni preseki, ki so bili definirani s kategorizacijo na podlagi stožca, v primerjavi z opazljivkami, definiranimi s kategorizacijo na podlagi p_T, nekoliko bolje določajo različne smeri v faznem prostoru parametrov EFT.

8.3 Produkcija 4-t stanja v SM in izven

Sedaj predlagamo strategijo iskanja procesa $pp \to t\bar{t}t\bar{t}$ pri LHC energiji 13 TeV v več-leptonskih razpadnih kanalih. V SM je sipalni presek $t\bar{t}t\bar{t}$ pretežno dominiran na račun interakcij QCD in je reda velikosti $\mathcal{O}(\alpha_S^4)$. Poleg tega je minimalna zahtevana partonska energija pri trčenju v težiščnem koordinatnem sistemu enaka $4m_t \sim 692$ GeV, kar povzroči zelo majhen sipalni presek na trkalniku LHC.

Strategija iskanja Eden od glavnih znakov dogodka 4-t je veliko število curkov kvarkov b, ki prihajajo iz šibkega razpada vsakega kvarka t. Pričakuje se, da ima prevladujoče $t\bar{t}$ ozadje manj curkov kvarka b na dogodek, zaradi česar mnogoterost curkov kvarkov b, N_b, najpomembnejši signal za diskriminacijo ozadja. Da lahko to dejstvo izkoristimo, bomo v analizi uporabili visoko efektivno delovno točko za algoritem za označevanje curkov kvarkov b in tako uporabili rez na tem številu. To bi moralo zadostovati, da se razmerje med signalom in ozadjem znatno izboljša. Naše iskanje temelji na več-leptonskih razpadnih kanalih z dvema ali več leptonsko-razpadajočima kvarkoma t. Ti kanali imajo manjše razvejitveno razmerje kot hadronski, vendar je prisotnost ozadja QCD v tem primeru precej manjša. Da bi povečali občutljivost signala, se osredotočimo na dve različni signalni območji: eno za enako nabit (ang. *same sign* ali SS) dileptonski kanal $\ell^\pm\ell^\pm$, drugo pa za trileptonski kanal $\ell^\pm\ell^\pm\ell^\mp$.

Optimalno selekcijo za iskanje dogodkov enako-nabitih dileptonov sestavljajo pogoji: (i) en $\ell^\pm\ell^\pm$ par, ki zadošča pogojem mini-izolacije, (ii) točno 7 rekonstruiranih curkov ($N_j = 7$), (iii) or katerih morajo vsaj trije curki biti curki kvarka b ($N_b \geq 3$). Optimalno selekcijo za iskanje dogodkov trileptonov sestavljajo pogoji: (i) trije leptoni z naboji $\ell^\pm\ell^\pm\ell^\mp$, ki zadoščajo pogojem mini-izolacije [94], (ii) točno 5 rekonstruiranih curkov ($N_j = 5$), (iii) od katerih morajo vsaj trije curki biti curki kvarka b ($N_b \geq 3$), in (iv) veto na oknu mase bozona Z 70 GeV$< m_\ell <$ 105 GeV, za redukcijo ozadja z leptonskim razpadom bozona Z.

Ozadje za obe signalni območji je bilo klasificirano v večji meri natančnosti. Prevladujoča in ireducibilna ozadja so $t\bar{t}Z$, $t\bar{t}W^{\pm}$ in $t\bar{t}h$ z dodatnimi curki. Število dogodkov za te procese je neposredno določeno preko Monte Carlo (MC) simulacij. Prevladujoča reducibilna ozadja so več-leptonski dogodki, ki prihajajo iz $t\bar{t}$ procesov z dodatnimi curki in z napačno prepoznanimi leptoni iz napačno rekonstruiranih curkov ($j \to \ell^{\pm}$) ali pa, v primeru SS dileptonske kategorije, z elektroni, ki so spremenili predznak naboja ($e^{\pm} \to e^{\mp}$). Verjetnost za napačno prepoznan lepton ali elektron s spremenjenim predznakom naboja je zelo težko oceniti na podlagi prvotnih načel. Le-ti morajo biti izluščeni na podlagi kombinacij MC simulacij in tehnik na podlagi podatkov (ki jih zagotavljajo eksperimentalne kolaboracije).

Izluščena zgornja meja za jakost signala, ki smo jo določili pri iskanju 4-t procesa znotraj SM, je

$$\mu_{t\bar{t}t\bar{t}}^{\mathrm{SM}} = \frac{\sigma_{t\bar{t}t\bar{t}}}{\sigma_{t\bar{t}t\bar{t}}^{\mathrm{SM}}} \;\leq\; 1.87 \quad \mathrm{at95\%CL}, \tag{8.35}$$

pri količini podatkov, ki ustreza integrirani luminoznosti $\mathcal{L}_{\mathrm{int}} = 300\,\mathrm{fb}^{-1}$. Ta rezultat nakazuje, da je trkalnik LHC občutljiv na produkcijo procesa 4-t znotraj SM.

Aplikacija na iskanje nove fizike Zgoraj opisano metodo za iskanje znotraj SM lahko uporabimo, da dobimo limite za ne-resonančne t-filične modele v NP. Osredotočimo se na: (1) nevtralen t-filičen vektorski bozon Z' in (2) nevtralen skalar ϕ z Yukawa sklopitvijo na kvark t. V prvem modelu zaradi enostavnosti predpostavimo, da se Z', z maso $m_{Z'}$, signifikantno sklaplja samo z RH kvarki t. Relevanten interakcijski Lagrangian se potem zapiše kot

$$\mathcal{L}_{Z'} \;=\; -\,g_{tZ'}\,\bar{t}_R \not{Z}' t_R\,. \tag{8.36}$$

V drugem modelu so relevantne $\phi - t$ interakcije opisane z

$$\mathcal{L}_{\phi} \;=\; -\,y_{t\phi}\,\bar{t}_L \phi t_R + \mathrm{h.c.}\,. \tag{8.37}$$

Za oba modela velja, da je vpliv ostalih interakcij na produkcijo 4-t procesov v večji meri neodvisen, saj nove prostostne stopnje nikoli niso ustvarjene na lupini. Pridobljena območja izključenosti na stopnji zaupanja 95% v faznem prostoru parametrov NP (masa proti sklopitvi) so prikazana na Sliki. 8.2, za oba poenostavljena modela.

8.4 t-filične gonilne sile: od razpadov mezonov B do trkalnikov

V primeru, da so deviacije od SM okrite pri nizkih energijah v tovarnah mezonov B ali pri eksperimentu LHCb, bi bilo smiselno iskati potencialne efekte NP v procesih s kvarki t pri visokih vrednostih p_T na trkalniku LHC. Tukaj prikažemo z enim dotičnim primerom, kako se NP pri kvarkih t lahko izrazi s kršitvijo LFU v redkih razpadih mezonov B kot prav tako preko deviacij v redkih procesih, ki vključujejo kvarke t, pri hadronskih trkalnikih. Še posebej to povezujemo s trenutnimi namigi kršitve LFU pri nedavnih eksperimentalnih meritvah procesov $b \to s\ell^+\ell^-$ s strani kolaboracije LHCb. Znanstveniki pri LHCb so izmerili LFU razmerje

$$R_{K^{(*)}} = \frac{\mathcal{B}(B \to K^{(*)}\mu^+\mu^-)}{\mathcal{B}(B \to K^{(*)}e^+e^-)}\,, \tag{8.38}$$

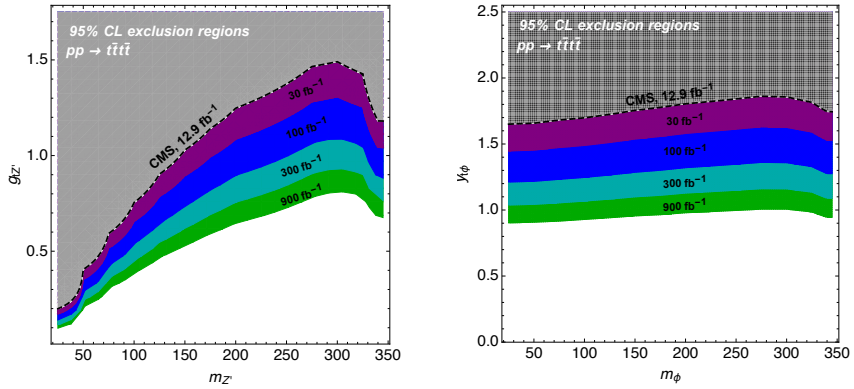


Figure 8.2: Izključitvena območja NP na trkalniku LHC za luminoznosti 30fb^{-1} (Vijolična), 100 fb^{-1} (Modra), 300 fb^{-1} (Temna Cian) in 900 fb^{-1} (Zelena), za model Z' (Levo) in za skalarni model ϕ (Desno). Že obstoječa omejitev, izluščena iz nedavne raziskave na eksperimentu CMS [1] je prikazana v sivem odtenku s črtkanimi konturami.

kjer $\mathcal{B}$ predstavlja delno razvejitveno razmerje. Za to opazljivko poročajo vrednosti, ki so $\approx 2.5\,\sigma$ *manjše* od tistih, ki jih napoveduje SM. Zanimiva razlaga NP, ki je predstavljena v tej tezi, sestoji iz novega težkega posrednika, ki se sklaplja pretežno s RH kvarki t in mioni. Zahtevani prispevki za razlago anomalij pri procesih $b \to s\ell^+\ell^-$ se nahajajo na nivoju diagramov z eno zanko.

t-filija To je najbolj minimalistična nastavitev, kjer, pri mejni vrednosti skale, dobimo samo podmnožico t-filičnih operatorjev z neničelnimi Wilsonovimi koeficienti. To so:

$$\mathcal{O}_{lt}^{ij} = (\bar{l}_i\gamma^\mu l_j)(\bar{t}\gamma_\mu t)\,, \tag{8.39}$$
$$\mathcal{O}_{et}^{ij} = (\bar{e}_i\gamma^\mu e_j)(\bar{t}\gamma_\mu t)\,, \tag{8.40}$$

kjer so $l_i = (\nu_L^i, \ell_L^i)^T$ LH leptonski dubleti in $e_i = \ell_R^i$ RH leptonski singleti. Pri strukturi okusa leptonskih tokov v (8.39)–(8.40) predpostavimo prevlado druge generacije, zato da dobimo sklopitve z mioni. Po zlomu elektrošibke simetrije, ti operatorji spremenijo sklopitve bozonov Z z mioni na kvantnem nivoju, preko diagrama, prikazanega na Sliki 8.3 (levo). To se zgodi na račun operatorjev $\mathcal{O}_{lt}$ in $\mathcal{O}_{et}$, ki se mešajo pod evolucijo renormalizacijskih grup z $(\varphi^\dagger i\overleftrightarrow{D}_\mu\varphi)(\bar{l}\gamma^\mu l)$ in $(\varphi^\dagger i\overleftrightarrow{D}_\mu\varphi)(\bar{e}\gamma^\mu e)$. Modifikacije hitrosti razpada procesa $Z \to \mu^+\mu^-$ so omejene s strani testov LFU v razpadih bozonov Z , ki so bili opravljeni na eksperimentu LEP-I, in parametra leptonske asimetrije, določenega s strani eksperimenta Aleph. Na podoben način podajo t-filični operatorji $\mathcal{O}_{lt}$ in $\mathcal{O}_{et}$ skladne prispevke na nivoju ene zanke za prehode $b \to s\ell^+\ell^-$ preko diagrama, prikazanega na sliki 8.3 (desno), ki so odgovorni za anomalije pri mezonih B. Globalna funkcija χ^2 je zgrajena iz vseh spremenljivk polov bozona Z, R_{K^*} in prav tako kotnih spremenljivk iz procesov $b \to s\mu^+\mu^-$. Rezultati prilagajanja krivulje ustreznim podatkom so povzeti na sliki 6.4 za dve merilni mejni vrednosti $\Lambda = 1$ TeV in $\Lambda = 1.8$ TeV. Globalni fit daje preferenco (negativni) vektorski strukturi za mionske tokove: $\mathcal{C}_{lt} \sim \mathcal{C}_{et} < 0$.

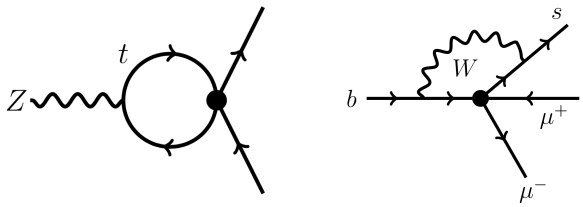


Figure 8.3: Eno-zančni prispevki v unitarni umeritvi za procese $Z \to \mu^+\mu^-$ (levo) in $b \to s\mu^+\mu^-$ (desno) iz 4-fermionskega operatoja, ki vključuje kvarke t kot prav tako mione.

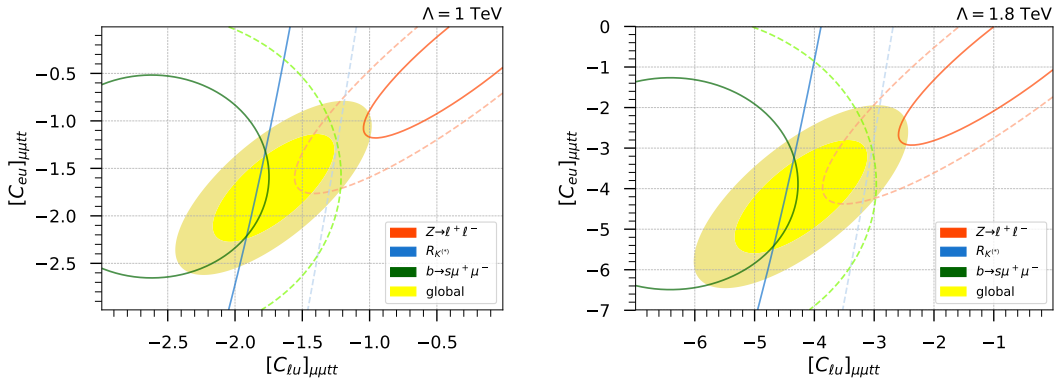


Figure 8.4: Prednostna območja pri stopnji zaupanja 68% in 95% v ravnini $(\mathcal{C}_{lt}, \mathcal{C}_{et})_{\mu\mu tt}$ iz globalnega χ^2 (rumena zapolnjena), iz $b \to s\mu^+\mu^-$ opazljivk (zeleno), $R_{K^{(*)}}$ (modro) in iz meritev LEP-I (rdeče). Izbrani sta bili dve merilni mejni vrednosti na višjem delu skale $\Lambda = 1, 1.8$ TeV.

mediator	$Z' \sim (\mathbf{1}, \mathbf{1}, 0)$	$S_1 \sim (\bar{\mathbf{3}}, \mathbf{1}, \frac{1}{3})$	$R_2 \sim (\mathbf{3}, \mathbf{2}, \frac{7}{6})$	$\widetilde{U}_1 \sim (\mathbf{3}, \mathbf{1}, \frac{5}{3})$	$\widetilde{V}_2 \sim (\bar{\mathbf{3}}, \mathbf{2}, -\frac{1}{6})$
$\mathcal{O}_{lt}$	✓	✗	✓	✗	✓
$\mathcal{O}_{et}$	✓	✓	✗	✓	✗
$\mathcal{C}_{lt} < 0$	✓	✗	✓	✓	✗
$\mathcal{C}_{et} < 0$	✓	✗	✓	✓	✗

Table 8.1: Tabela prikazuje možne mediatorje, ki na drevesnem nivoju generirajo dva relevantna operatorja. Bozon Z' predstavlja vektorski bozon v singletni reprezentaciji umeritvene grupe SM, medtem ko nomenklatura za leptokvarke ustreza tisti iz [7]. Zadnja vrstica prikazuje tiste, za katere so Wilsonovi koeficienti negativni, kot je zahtevano s strani nizko-energijskega fita.

Poenostavljeni t-filični modeli Sedaj bomo analizirali možne mediatorje NP, ki lahko ustvarijo t-filične operatorje $\mathcal{O}_{lt,et}$ v režimu UV. Te lahko klasificiramo v mediatorje barvnega singleta (Z') in v barvne mediatorje (leptokvarki). Vse možnosti so prikazane v tabeli 8.1. Edini mediator, ki lahko ustvari nizko-energijske strukture, ki so potrebne za anomalije mezonov B, je t-filičen Z', ki se sklaplja z mioni kot

prikazuje poenostavljen dinamični Lagrangian

$$\mathcal{L}_{Z'} = -\frac{1}{4} Z'_{\mu\nu} Z'^{\mu\nu} + \frac{1}{2} M^2_{Z'} Z'_\mu Z'^\mu + Z'_\mu \left(\mathcal{J}^\mu_t + \mathcal{J}^\mu_\ell \right) \tag{8.41}$$

s fermionskimi tokovi $\mathcal{J}^\alpha_t = \epsilon^{tt}_R (\bar{t}\gamma^\alpha t)$ in $\mathcal{J}^\alpha_\mu = \epsilon^{\mu\mu}_L (\bar{l}_2 \gamma^\alpha l_2) + \epsilon^{\mu\mu}_R (\bar{\mu}_R \gamma^\alpha \mu_R)$. Tri sklopitve ϵ^{tt}_R in $\epsilon^{\mu\mu}_{L,R}$ ter masa $M_{Z'}$ so prosti parametri tega poenostavljenega modela. Samo kombinacija dveh t-filičnih leptokvarkov, $\widetilde{U}_1 \sim \left(\mathbf{3}, \mathbf{1}, \frac{5}{3}\right)$ (vektorski leptokvark) in $R_2 \sim \left(\mathbf{3}, \mathbf{2}, \frac{7}{6}\right)$ (skalarni leptokvark), lahko ustvarita potrebno obnašanje pri nizkih energijah. Poenostavljen Lagrangian za ta $\widetilde{U}_1 - R_2$ model je

$$\mathcal{L}_{\text{LQ}} \;=\; \mathcal{L}_{\text{kin}} - \frac{1}{2} M^2_{\widetilde{U}} \widetilde{U}^{\mu\,\dagger} \widetilde{U}_\mu - \frac{1}{2} M^2_R R^\dagger_2 R_2 + \kappa_S \, \bar{t} \, R^T_2 (i\tau_2) l_2 + \kappa_V (\bar{t}_R \gamma^\alpha \mu_R) \widetilde{U}_{1\alpha} + \text{h.c.} \tag{8.42}$$

z dvema sklopitvama, κ_S in κ_V, in dvema masama, M_R in $M_{\widetilde{U}}$, ki predstavljajo proste parametre tega poenostavljenega modela.

Fenomenologija trkalnikov t-filični Z' in $\widetilde{U}_1 - R_2$ modeli imajo relativno majhen odtis v nizko-energijskih opazljivkah. Po drugi strani lahko pri visokih energijah uporabimo več procesov za omejitev obeh modelov na trkalniku LHC.

Z' povzroči deviacije pri dveh procesih, in sicer pri $pp \to t\bar{t}Z' \to t\bar{t}\mu^+\mu^-$ ter $pp \to t\bar{t}Z' \to t\bar{t}t\bar{t}$ pri trkalniku LHC, in pri nevtrinski *trident* produkciji $\nu_\mu \gamma^* \to \nu_\mu \mu^+ \mu^-$ zaznanih pri eksperimentih z nevtrinskimi žarki (Z' se v zlomljeni fazi sklaplja tudi z nevtrinskimi tokovi). Za produkcijo $t\bar{t}\mu^+\mu^-$ nastavimo limite tako, da ponovimo postopek kolaboracije ATLAS [3] za iskanje inkluzivne di-muonske resonance $pp \to Z' \to \mu^+\mu^- + X$, medtem ko za produkcijo $t\bar{t}t\bar{t}$ uporabimo limite, ki so izpeljane v (8.35) in nedavne CMS omejitve iz Ref. [186]. Limita tega modela za Z' z maso 1 TeV je podana na sliki 8.5 (levo), kjer smo nastavili muonsko sklopitev tako, da je v celoti vektorska preko $\epsilon^{\mu\mu}_V \equiv \epsilon^{\mu\mu}_L = \epsilon^{\mu\mu}_R$ (v skladu s preferiranim območjem nizko-energijskega fita). Črn črtkan pas ustreza najbolšemu nizko-energijskemu fitu s signifikanco 1σ. Omeniti velja, da kombinacija iskanja procesov na trkalniku LHC skupaj z iskanjem nevtrinskih trident procesov poda komplementarne rezultate. Podamo tudi projekcije na višje vrednosti integrirane luminoznosti 300 fb^{-1} pri LHC (barvne pikčaste konture). Ti rezultati nakazujejo, da bomo s tretjim zajemanjem podatkov na trkalniku LHC sposobni preveriti fazni prostor parametrov, ki je pomemben anomalije mezonov B pri tem poenostavljen modelu.

t-filičen model leptokvarkov $\widetilde{U}_1 - R_2$ ima manjši vpliv pri višjih vrednostih p_T. Edini relevanten proces, ki lahko testira ta model, je produkcija parov $pp \to \widetilde{U}^\dagger_1 \widetilde{U}_1,\ R^\dagger_2 R_2 \to t\mu^+ \bar{t}\mu^-$. Glede na to, da ni bilo opravljene še nobene meritve iskanja razpadov leptokvarkov v tem dotičnem razpadnem kanalu s strani eksperimentalnih kolaboracij, smo ponovili postopek za iskanje inkluzivnih di-muonov s strani kolaboracij ATLAS [3]. Rezultati s trenutno vrednostjo luminoznosti (modro območje) in njihova projekcija na 300 fb^{-1} (modra črtkana kontura) so podani na sliki 8.5 (desno). Dovoljeno območje s strani te raziskave nakazuje na visoke vrednosti mase leptokvatkov. Po drugi strani pa to implicira na močno sklopitev, kar pomeni, da se raziskava približuje nepertubativni limiti, zato da lahko dobimo anomalijo pri mezonih B.

Naproti kompletnim UV modelom Sedaj opišemo razred UV modelov, s katerim lahko zapišemo kateregakoli od singletnih SU(2)$_L$ vektorskih leptokvarkov kot

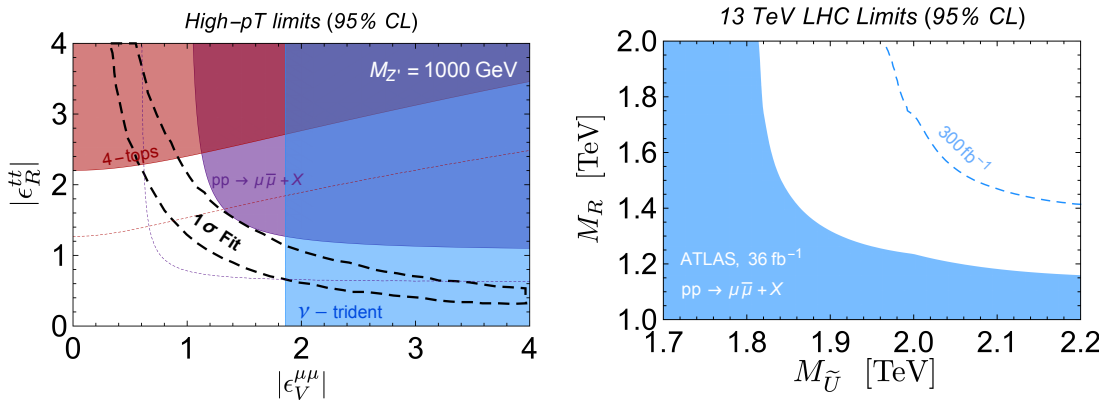


Figure 8.5: (Levo) Povzetek območij pri visokih vrednosti p_T za model Z'. Rdeče, vijolično in modro izključitveno območje pri 95% stopnji zaupanja ustrezajo iskanju 4-t pri LHC, iskanju di-muonskega repa in pa nevtrinskemu trident eksperimentu CCFR, v enakem vrstnem redu. Pikčaste barvne konture predstavljajo omejitve pri LHC projekciji na $300\,\mathrm{fb}^{-1}$. Črno črtkano območje ustreza signifikanci $1\,\sigma$ globalnega fita $R_{K^{(*)}}$ na podatke LEP-I in opazljivke procesa $b \to s\mu\mu$. (Desno) Izključitveno območje za maso pri modelu $\tilde{U}_1 - R_2$ iz raziskave di-muonskega repa pri LHC pri luminoznosti $36.1\,\mathrm{fb}^{-1}$. Črtkana kontura prikazuje projekcijo rezultata na $300\,\mathrm{fb}^{-1}$ LHC podatkov.

umeritveni bozon. Poleg tega se ti umeritveni bozoni še sklapljajo z ne-univerzalnimi interakcijami s polji iz SM in lahko imajo mase reda velikosti TeV. Nato to izrecno pokažemo za vektorski leptokvark $\widetilde{U}_\mu \sim (\mathbf{3}, \mathbf{1}, 5/3)$.

Začetna točka je naslednja ne-abelova razširitev umeritvene grupe SM:

$$\mathcal{G}(4) \;=\; \mathrm{SU}(4) \otimes \mathrm{SU}(3)_{c'} \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1)_{Y'}\,. \tag{8.43}$$

Grupi s črtico ($'$) $\mathrm{SU}(3)_{c'}$ in $\mathrm{U}(1)_{Y'}$ ne ustrezata običajnim grupam SM za reprezentacijo barve in hipernaboja. Slednje so v resnici vgrajene v $\mathcal{G}(4)$ in se pojavijo šele po spontani zlomitvi simetrije (ang. *spontaneous symmetry breaking* ali SSB) nad elektrošibko skalo. Običjani kvarki, leptoni[1] in Higgsovo poljeφ so vzeti kot singleti pod grupo SU(4) ampak so nabiti s konvencionalnimi reprezentacijami naboja pod grupo $\mathrm{SU}(3)_{c'} \otimes \mathrm{U}(1)_{Y'}$. Posledično se noben od umeritvenih bozono, ki so povezani s SU(4), ne sklaplja neposredno s polji SM. Te sklopitve se pojavijo šele v fazi zlomljene simetrije kot posledica fermionskega mešanja. Da se to lahko zgodi, uvedemo še en dodaten, težek, vektorju podoben fermion $\Psi \sim (\mathbf{4}, \mathbf{1}, \mathbf{1}, \psi)$ v reprezentaciji grupe (8.43). Hipernaboj s črtico ψ je parameter, ki je pomemben del strukture modela, in ga bomo kasneje postavili na fiksno vrednost. Fermionsko mešanje bo veliko med fermionom Ψ in med RH poljem SM, Kar privede do dejstva, da se bo vektorski leptokvark pretežno sklapljal s polji SM po spontani zlomitvi simetrije. Zaradi tega se nanašamo na ta model kot $\mathrm{SU}(4)_R$, zato da poudarimo desnoročno kiralnost prevladujoče sklopitve. Po SSB $\mathcal{G}(4) \to \mathcal{G}_{\mathrm{SM}}$ se pojavi bogat spekter novih težkih umeritvenih bozonov [197]. Natančneje so to vektorski leptokvarko bozon, ki ga označimo kot $U_R \sim (\mathbf{3}, \mathbf{1}, p)$ s hipernabojem p, koloron $g' \sim (\mathbf{8}, \mathbf{1}, 0)$ in vektorski singler $Z' \sim (\mathbf{1}, \mathbf{1}, 0)$. Brezmasni bozoni iz SM se tudi pojavijo na tej SSB,

[1] In our setup we also introduce RH neutrinos ν_R^i responsible for neutrino masses.

natančneje gluoni in nevtralen hipernabojni bozon B. Da dosežemo to spontano zlomitev simetrije, uvedemo dve novi skalarni polji:

$$\Omega_3 \sim (\overline{\mathbf{4}}, \mathbf{3}, \mathbf{1}, \omega) \qquad \text{in} \qquad \Omega_1 \sim (\overline{\mathbf{4}}, \mathbf{1}, \mathbf{1}, -3\omega)\,, \tag{8.44}$$

kjer je hipernaboj s črtico ω še en parameter strukture modela. ω in ψ sta eventuelno oba nastavljena na različni vrednosti, kar vodi do različnih modelov vektorskih leptokvarkov v IR. Vzorec SSB je $\mathrm{SU}(4)\otimes\mathrm{SU}(3)_{c'}\otimes\mathrm{U}(1)_{Y'} \to \mathrm{SU}(3)_c\otimes\mathrm{U}(1)_Y$. Zlomitev grupe SM $\mathcal{G}_{\mathrm{SM}} \to \mathrm{SU}(3)_c \otimes \mathrm{U}(1)_{\mathrm{em}}$ je tako dosežena s singletom Higgsovega polja φ pod grupo SU(4) in nabita pod $\mathcal{G}_{\mathrm{SM}'}$ na običajen način $\varphi \sim (\mathbf{1}, \mathbf{1}, \mathbf{2}, 1/2)$ z elektrošibkim zlomom vakuumske pričakovane vrednosti pri $\langle\varphi\rangle = v \approx 256$ GeV. Fermionski Lagrangian je podan s tremi členi $\mathcal{L}_{\mathrm{light}} + \mathcal{L}_{\mathrm{heavy}} + \mathcal{L}_{\mathrm{mix}}$, kjer sta prva dva Lagrangiana v interakcijski bazi definirana kot

$$\begin{aligned}
\mathcal{L}_{\mathrm{light}} &= \overline{Q^i} i \not{D} Q^i + \overline{u^i} i \not{D} u^i + \overline{d^i} i \not{D} d^i - (y^d_{ij} \overline{Q^i} \varphi d^j + y^u_{ij} \overline{Q^i} \tilde{\varphi} u^j + \text{h.c.}) \\
&\quad \overline{L^i} i \not{D} L^i + \overline{e^i} i \not{D} e^i + \overline{\nu^i_R} i \not{D} \nu^i_R - (y^e_{ij} \overline{L^i} \varphi e^j + y^\nu_{ij} \overline{L^i} \tilde{\varphi} \nu^j_R + \text{h.c.}) \qquad (8.45) \\
\mathcal{L}_{\mathrm{heavy}} &= \overline{\Psi} \not{D} \Psi \; - \; m_\Psi (\overline{\Psi}_L \Psi_R + \overline{\Psi}_R \Psi_L)\,. \qquad (8.46)
\end{aligned}$$

Tukaj sta Q^i in L^i LH kvarkovsko in leptonsko polje, u^i, d^i RH kvarka tipa u in d in e^i ter ν^i_R RH polji leptona in nevtrina z družinskim indeksom $i = 1, 2, 3$, H pa je skalarni dublet SM. Mešani lahek-težak Lagrangian Yukawe v isti bazi lahko v splošnem zapišemo kot

$$\mathcal{L}_{\mathrm{mix}} = -\, \lambda^i_{\mathsf{q}} \, (\overline{\Psi}_L \, \Omega_3^\dagger \, \mathsf{q}^i_R) \; - \; \lambda^i_{\mathsf{e}} \, (\overline{\Psi}_L \, \Omega_1^\dagger \, \mathsf{e}^i_R) \; + \; \text{h.c.} \tag{8.47}$$

kjer so $\lambda^i_{\mathsf{q},\mathsf{e}}$ Yukawa sklopitve RH kvarkovskega polja $\mathsf{q}_R = \{u, d\}$ in RH leptonskega polja $\mathsf{e}_R = \{e, \nu_R\}$. Vsi možni Yukawa členi v $\mathcal{L}_{\mathrm{mix}}$ niso dovoljeni zaradi umeritvene invariance $\mathrm{SU}(3)_{c'} \otimes \mathrm{U}(1)_{Y'}$. Možnosti so odvisne od izbire črtkanih hipernabojev ω in ψ skalarnega in vektorju podobnega fermiona. Ločimo le med naslednjimi štirimi možnimi načini sestave modela:

$$\omega = \frac{5}{12}\,,\ \psi = \frac{1}{4} \quad \Longrightarrow \quad \mathcal{L}_{\mathrm{mix}} \supset -\lambda^i_u(\overline{\Psi}_L\Omega_3^\dagger u^i) - \lambda^i_e(\overline{\Psi}_L\Omega_1^\dagger e^i) \tag{8.48}$$

$$\omega = \frac{1}{6}\,,\ \psi = \frac{1}{2} \quad \Longrightarrow \quad \mathcal{L}_{\mathrm{mix}} \supset -\lambda^i_u(\overline{\Psi}_L\Omega_3^\dagger u^i) - \lambda^i_\nu(\overline{\Psi}_L\Omega_1^\dagger \nu^i_R) \tag{8.49}$$

$$\omega = \frac{1}{6}\,,\ \psi = -\frac{1}{2} \quad \Longrightarrow \quad \mathcal{L}_{\mathrm{mix}} \supset -\lambda^i_d(\overline{\Psi}_L\Omega_3^\dagger d^i) - \lambda^i_e(\overline{\Psi}_L\Omega_1^\dagger e^i) \tag{8.50}$$

$$\omega = -\frac{1}{12}\,,\ \psi = -\frac{1}{4} \quad \Longrightarrow \quad \mathcal{L}_{\mathrm{mix}} \supset -\lambda^i_d(\overline{\Psi}_L\Omega_3^\dagger d^i) - \lambda^i_\nu(\overline{\Psi}_L\Omega_1^\dagger \nu^i_R) \tag{8.51}$$

Pod umeritveno grupo SM $\mathcal{G}_{\mathrm{SM}}$ lahko vektorju podoben fermion Ψ razbijemo na komponento barvnega tripleta Ψ^{q} in brezbarvno komponento Ψ^{e}, tako da velja $\Psi_{L,R} = (\Psi^{\mathsf{q}}, \Psi^{\mathsf{e}})_{L,R}$. Hipernaboj umeritvenega leptokvarka je povezan s parametrom ω. Reprezentacijo leptokvarka pod grupo SM zapišemo kot

$$U^\mu_R \;\sim\; (\mathbf{3}, \mathbf{1}, 4\,\omega), \tag{8.52}$$

To implicira, da za tri možne izbire vrednosti ω v enačbah (8.48)–(8.51) dobimo tri dobro znane modele vektorskih leptokvark singletov:

- $U_R = \widetilde{U}_1 \sim (\mathbf{3}, \mathbf{1}, 5/3)$, če je $\omega = 5/12$ z leptokvarkovskimi interakcijami (8.48).

- $U_R = U_1 \sim (\mathbf{3}, \mathbf{1}, 2/3)$, če je $\omega = 1/6$ z leptokvarkovskimi interakcijami(8.49) ali (8.50).
- $U_R = \overline{U}_1 \sim (\mathbf{3}, \mathbf{1}, -1/3)$, če je $\omega = -1/12$ z leptokvarkovskimi interakcijami (8.51).

Ti trije bozoni $\widetilde{U}_1^\mu$, U_1^μ in $\overline{U}_1^\mu$ (informacije o notaciji so pod referenco [7]) so edini trije možni vektorski leptokvarki, ki so singleti pod grupo $\mathrm{SU}(2)_L$. Vektorski leptokvark se sklaplja neposredno s tokovi $\bar{\Psi}^{\mathsf{q}}\gamma^\mu\Psi^{\mathsf{e}}$. Po spontanemu zlomu simetrije se težki fermioni mešajo s polji SM, kar privede do efektivne sklopitve leptokvarka s poljem SM oblike

$$\mathcal{L}_{\mathrm{LQ}} \;=\; \frac{g_4}{\sqrt{2}}\, U_R^\mu\; \bar{\mathsf{q}}_R\, \gamma_\mu \cdot W_R^{(\mathsf{qe})} \cdot \mathsf{e}_R + \;\mathrm{h.c.} \tag{8.53}$$

kjer je $W_R^{(\mathsf{qe})}$ matrika mešanja velikosti 3×3, ki meri neujemanje med kvarki in leptoni znotraj istega multipleta. Ta matrika je eksplicitno podana preko $W^{(\mathsf{qe})} = \Theta_R^{\mathsf{q}}\Theta_R^{\mathsf{e}\,\dagger}$,kjer so $\Theta_R^{\mathsf{q,e}}$ 3×1 matrike, ki nadzorujejo mešanje med vektojem podobnimi fermioni in lahkimi stanji SM. Ko določamo naboje $\{\psi, \omega\}$ v vsakem scenariju gradnje modela v (8.48)–(8.51), se izraz (8.53) zreducira v enega of naslednjih modelov za leptokvarke iz literature:

$$\mathcal{L}_{\widetilde{U}} \;=\; \frac{g_4}{\sqrt{2}}\,(\Theta_R^u\Theta_R^{e\,\dagger})_{ij}\,\widetilde{U}_1^\mu\,(\overline{u}_R^i\gamma_\mu e_R^j) \;+\; \mathrm{h.c.} \tag{8.54}$$

$$\mathcal{L}_{U} \;=\; \frac{g_4}{\sqrt{2}}\,(\Theta_R^u\Theta_R^{\nu\,\dagger})_{ij}\,U_1^\mu\,(\overline{u}_R^i\gamma_\mu \nu_R^j) \;+\; \mathrm{h.c.} \tag{8.55}$$

$$\mathcal{L}_{U} \;=\; \frac{g_4}{\sqrt{2}}\,(\Theta_R^d\Theta_R^{e\,\dagger})_{ij}\,U_1^\mu\,(\overline{d}_R^i\gamma_\mu e_R^j) \;+\; \mathrm{h.c.} \tag{8.56}$$

$$\mathcal{L}_{\overline{U}} \;=\; \frac{g_4}{\sqrt{2}}\,(\Theta_R^d\Theta_R^{\nu\,\dagger})_{ij}\,\overline{U}_1^\mu\,(\overline{d}_R^i\gamma_\mu \nu_R^j) \;+\; \mathrm{h.c.} \tag{8.57}$$

Do sedaj še nismo diskutirali okusne strukture modela. Do te točke so bile Yukawa sklopitve iz (8.47) povesem generične, kar pomeni, da so mešalni vzorci še vedno nedoločeni. Od štirih različnih modelov $\mathrm{SU}(4)_R$, ki smo jih opisali zgoraj, smo v glavnem zainteresirani v tistega, ki ustvari umeritveni leptokvark $\widetilde{U}_1$ s hipernabojem $p = 5/3$, ki se sklaplja pretežno z RH kvarki tipa u in nabitimi leptoni. V tem primeru se težki fermion v grupni reprezentaciji SM razstavi kot $\Psi_{L,R} = (\mathcal{T}, \mathcal{E})^T$, kjer je $\mathcal{T} \sim (\mathbf{3}, \mathbf{1}, 2/3)$ RH partner kvarka t in $\mathcal{E} = (\mathbf{1}, \mathbf{1}, -1)$ je RH partner leptona. Če v (8.47) predpostavimo prevlado tretje generacije kvarkov ($\lambda_u^3 \gg \lambda_u^{1,2}$) in druge generacije leptonov ($\lambda_e^2 \gg \lambda_e^{1,3}$), to privede do interakcij RH vektorskih leptokvarkov, prikazanih v (8.42), ki so potrebne za animalije mezonov B. Poleg tega, če v ta opis uvedemo še dodatno skalarno stanje $\Omega_2 \sim (\mathbf{4}, \mathbf{1}, \mathbf{2}, -3/4)$, ki se sklaplja pretežno z mionskimi partnerji, po SSB dobimo skalarni leptokvark R_2, ki je potreben za generacijo $\mathcal{O}_{lt}$.

List of publications

1. **Confronting lepton flavor universality violation in B decays with high-p_T tau lepton searches at LHC**
   Darius A. Faroughy, A. Greljo, J. F. Kamenik
   Phys. Lett. **B** 764, 126 (2017)
   arxiv:1609.07138

2. **Four tops for LHC**
   E. Alvarez, Darius A. Faroughy, J. F. Kamenik, R. Morales, A. Szynkman
   Nucl. Phys. **B** 915 (2017) 19 – 43
   arXiv:1611.05032

3. **The role of S_3 GUT leptoquark in flavor universality and collider searches**
   I. Doršner, S. Fajfer, Darius A. Faroughy, N. Košnik
   JHEP 1710 (2017) **188**
   arXiv:1706.07779

4. **Anomalies in Bottom from New Physics in Top**
   J. E. Camargo-Molina, A. Celis, Darius A. Faroughy
   Phys. Lett. **B** 784 (2018) 284-293
   arXiv:1805.04917

5. **Scalar leptoquarks from GUT to accommodate the B-physics anomalies**
   D. Bečirević, I. Doršner, S. Fajfer, Darius A. Faroughy, N. Košnik, O. Sumensari
   Phys. Rev. D98 (2018) no.5, 055003
   arXiv:1806.05689

6. **Closing the window on single leptoquark solutions to the B-physics anomalies**
   A. Angelescu, D. Bečirević, Darius A. Faroughy, O. Sumensari
   JHEP 1810 (2018) 183
   arXiv:1808.08179

7. **Opportunities in Flavour Physics at the HL-LHC and HE-LHC**
   WG4: Opportunities in Flavour Physics
   Section 8 The top quark and flavor physics,
   Subsection 8.1 Determinations of V_{tx}
   CERN-LPCC-2018-06
   arXiv:1812.07638

8. **Direct searches motivated by recent B-anomalies**
   Darius A. Faroughy
   SciPost Phys.Proc. 1 (2019) 021
   arXiv:1811.07582

9. **Uncovering latent jet substructure**
   Barry M. Dillon, Darius A. Faroughy, Jernej F. Kamenik
   Submitted to PRL
   arXiv:1904.04200