EPFL

Thèse n° 9176

Effective Field Theories for New and Old Physics

Présentée le 30 septembre 2022

Faculté des sciences de base
Laboratoire de physique théorique des particules
Programme doctoral en physique

pour l'obtention du grade de Docteur ès Sciences

par

Alfredo GLIOTI

Acceptée sur proposition du jury

Prof. R. Houdré, président du jury
Prof. R. Rattazzi, directeur de thèse
Prof. R. Barbieri, rapporteur
Prof. G. Isidori, rapporteur
Prof. J. M. Penedones, rapporteur

École
polytechnique
fédérale
de Lausanne

2022

Acknowledgements

The first person to thank is of course my advisor Riccardo Rattazzi. I feel extremely lucky to have worked with him during the five years of my PhD. I have learned a lot about many different aspects of physics from our discussions which often ended with more new questions than answers. From him, I have also learned not to just take results at face value, but always look for weak points in my arguments and not to take established results for granted. Even though we have not worked together as much as either of us would have liked, primarily because of all the unfortunate events that happened during these years, I still feel extremely happy to have had this opportunity and to have the chance of continuing on this path. I hope that our collaboration will continue.

Secondly, I would like to thank the other collaborators that I have met and worked with during my doctorate. Starting chronologically from Luca Vecchi, who helped me a lot with the difficult transition from university student to researcher during my first year. Then to Andrea Wulzer, who has been almost a second advisor given the amount of time we have worked together on our projects. Finally, to Angelo Esposito with which I started working only this last year, but even in this short time, I have learned many things from our discussions. I hope I will be able to keep working with them in the future.

A huge thanks also goes to all the other students I have met during these years: Aditya, Adrien, Andrea, Eren, Filippo, Gabriel, Gil, Kamran, Kin, and Lorenzo. Among these, however, a special thanks goes to Siyu, who became a truly special friend even outside of the work environment and is the only reason why I managed to survive the months of lockdown without going insane. I also have to thank my friends from Perugia, especially Alessandro, Roberto, Lorenzo, Bernardo, and the two Luca, for the many game nights both remotely and in person.

Infine, ma non per ordine d'importanza, un grande ringraziamento va ai miei genitori. Questo dottorato è stata la mia prima esperienza lontano da loro, ma nonostante la distanza sono stati sempre presenti e pronti ad aiutarmi nei momenti di difficoltà, sia nelle piccole cose come traslochi o semplici consigli, sia quando sono stato male e avevo bisogno di aiuto anche per le più piccole cose. So di non averli mai ringraziati abbastanza per

tutto quello che hanno fatto per me e per i sacrifici che hanno fatto per farmi raggiungere questo risultato. Nonostante io non ve lo dica mai e forse non lo faccia neanche capire con il mio comportamento, vi voglio bene. Grazie.

Lausanne, September 28, 2022 Alfredo Glioti

Foreword

The material of this thesis is schematically organized as follows:

- Chapter 1 starts with a basic review of the main properties of Composite Higgs models. This is then followed by an original study on the constraints coming from flavor physics on Partial Compositeness models for different symmetry assumptions.

  A. Glioti, R. Rattazzi, L. Ricci and L. Vecchi, *To appear.*

- Chapter 2 is based on [1]:

  S. Chen, A. Glioti, G. Panico and A. Wulzer, *Parametrized classifiers for optimal EFT sensitivity*, *JHEP* **05** (2021) 247 [2007.10356].

- Chapter 3 is based on [2]:

  S. Chen, A. Glioti, R. Rattazzi, L. Ricci and A. Wulzer, *Learning from radiation at a very high energy lepton collider*, *JHEP* **05** (2022) 180 [2202.10509].

- Chapter 4 contains both a review and original studies on the $X(3872)$ resonance through Effective field Theory and standard Quantum Mechanics techniques.

  A. Esposito, A. Glioti, A.D. Polosa, R. Rattazzi and M. Tarquini, *To appear.*

- Chapter 5 is based on [3]:

  A. Glioti, R. Rattazzi and L. Vecchi, *Electroweak Baryogenesis above the Electroweak Scale*, *JHEP* **04** (2019) 027 [1811.11740].

Abstract

Effective Field Theories have changed our understanding of Quantum Field Theories. This thesis shows several applications of this powerful tool in the context of the Standard Model and for searches of New Physics.
The thesis starts with a review of the Standard Model and its open questions and is followed by an updated and systematic study of models of flavor in the context of Partial Compositeness in Composite Higgs theories. Following that, the question on how to measure the Wilson coefficients of the Standard Model effective operators at present and future experiments is addressed: first by using modern Machine Learning techniques by studying angular distributions for diboson production, followed then by a study on ElectroWeak radiation at a future Muon Collider and how to use it to better probe the new physics parameter space.
The fourth chapter deals instead with applying Non-Relativistic Effective Theories to the study of exotic mesons in the Standard Model. The two competing interpretations, a molecule formed of two mesons or a compact tetraquark state, and their consequences are studied. In particular this study is done on the X(3872) exotic charmonium and the consequences of the two accidental tunings of this system are discussed.
The last chapter addresses the problem of baryogenesis from the ElectroWeak phase transition. A new scalar sector is introduced that decouples the physics responsible for the generation of the baryon asymmetry from the weak scale. This helps solving the main problems that ElectroWeak baryogenesis models face, namely the large modifications to the Higgs physics and the need of large CP violating new effects.

Keywords: Effective Field Theory, Beyond the Standard Model, Composite Higgs, Flavor Physics, Machine Learning, Muon Collider, Tetraquark, ElectroWeak Baryogenesis.

Riassunto

Le Teorie di Campo Efficaci hanno cambiato il nostro modo di comprendere le Teorie di Campo Quantistiche. Questa tesi mostra varie applicazioni di questo potente mezzo sia nel contesto del Modello Standard che per le ricerche di Nuova Fisica.
La tesi inizia con una review del Modello Standard e dei suoi problemi ancora aperti ed è poi seguito da uno studio aggiornato e sistematico dei modelli di flavor nelle teorie di Higgs composto con Partial Compositeness. Di seguito, viene trattata la questione di come misurare, in esperimenti attuali e futuri, i coefficienti di Wilson degli operatori efficaci del Modello Standard: prima usando metodi basati su Machine Learning per studiare le distribuzioni angolari nella produzione di dibosoni, poi studiando la radiazione Elettrodebole ad un possible futuro acceleratore di muoni e come utilizzarla per esplorare efficacemente lo spazio dei parametri della nuova fisica.
Il quarto capitolo invece riguarda l'applicazione di Teorie Efficaci Non-Relativistiche allo studio di mesoni esotici nel Modello Standard. Vengono studiate le due possibili interpretazioni, una molecola formata da due mesoni o uno stato compatto di quattro quark, e le loro conseguenze. In particolare questo studio viene svolto sul charmonium esotico X(3872) e vengono discusse le conseguenze dei due tuning accidentali di questo sistema.
L'ultimo capitolo affronta il problema della bariogenesi dalla transizione di fase Elettrodebole. Viene introdotto un nuovo settore scalare che disaccoppia la fisica responsabile della generazione dell'asimmetria barionica dalla scala Elettrodebole. Questo aiuta a risolvere i problemi principali che i modelli di bariogenesi Elettrodebole devono affrontare, vale a dire le grandi modifiche alla fisica dell'Higgs e la necessità di nuovi effetti che violano significativamente CP.

Parole chiave: Teorie di Campo Efficaci, Oltre il Modello Standard, Higgs Composto, Fisica del Flavor, Machine Learning, Muon Collider, Tetraquark, Bariogenesi Elettrodebole.

Contents

Introduction

This thesis is a collection of works that show different applications of Effective Field Theories to study physics Beyond the Standard Model ("New Physics"), but also less clear aspects of the Standard Model itself ("Old Physics"). The notion of Effective Field Theory dates back to the 70s with works by Wilson and Weinberg that lead to the modern understanding of Quantum Field Theory. Effective Theories formalize ideas on which physics is built upon since the early times, that when describing phenomena at a given length scale one does not need to care about the details about what happens at much shorter scales. This is the reason why it is possible to study classical mechanics without knowing about the structure and dynamics of molecules and atoms, why it is possible to do chemistry without knowing the details of the atomic nucleus, why it was possible to study particle physics before the Standard Model without knowing the details of the strong and weak interactions, but especially why it is possible to study the Standard Model today without knowing what is hidden at distances much shorter than what we can probe at LHC.

As a simple example that captures the main properties of effective theories we can consider the multipole expansion in classical electrodynamics. One can imagine for instance some object of finite size a with an arbitrarily complicated geometry and charge distribution. If we knew exactly these two properties we could compute the electrostatic potential generated by that object everywhere in space. However, assume that we are only interested at measurements at distances $R \gg a$. In this case it is clear that all the intricate details of the object in question get blurred out. We can indeed expand the electrostatic potential in multipoles or, equivalently, in powers of a/R. At large distances the whole system can be accurately described by only a few parameters, for example the monopole moment. When R gets closer and closer to a we need to introduce more and more moments until we hit the point $R \sim a$ where this description breaks down and we need to introduce the full charge distribution in the theory. This example shows a few important properties of effective theories:

- *Universality*: the long distance expansion in terms of multipoles has the same form regardless of the microscopic details of the charge distribution. Furthermore, these details are encoded in some effective parameters, the multipole moments, of which

only a finite number are relevant at a given distance R.

- *Reductionism*: when probing shorter and shorter distances, more and more parameters are needed until a new more fundamental description appears.

- *Accidental Symmetries*: at distances where the monopole dominates, the system acquires an approximate $SO(3)$ rotational invariance, regardless of the symmetries of the underlying distribution. When also the dipole moment becomes relevant, the symmetry is broken to a $SO(2)$ rotation along the dipole axis. Finally, the symmetry completely disappears when the next moments become important.

Effective Field Theories are a powerful tool both in the case where the fundamental theory is known, but studied in a limit where only a few degrees of freedom are relevant, and where the fundamental theory is not known but we would like to measure its long distance effects. In this thesis we will see examples of both these applications. For the first case we will study Effective Theories for QCD hadrons in a non-relativistic limit and the dynamics of Matsubara zero modes in Quantum Field Theories at finite temperature. For the second case instead, we will study the Standard Model as an Effective Theory and parameterize the effects of New Physics in terms of higher-dimensional operators involving Standard Model fields.

In order to properly introduce the concepts that will be studied later, the rest of this introduction is going to briefly review the main properties of the Standard Model and its problems.

The Standard Model Lagrangian

The Standard Model is the minimal model capable of explaining the vast majority of collider data collected so far. With the exception of a few anomalies, perhaps expected given the huge amount of data gathered so far, and some open problems that we will address later, it is surprising the accuracy to which some of its predictions have been verified since its first formulation roughly fifty years ago. The most extreme example is the prediction of the anomalous magnetic moment of the electron, that agrees with the measured value up to 9 digits.

The principles of Special Relativity and Quantum Mechanics, on which Quantum Field Theories and in particular the Standard Model are based on, strongly constrain the possible dynamics that are allowed, leading to a theory that consists of local quantized fields described by a Hermitean and Lorentz invariant Lagrangian. A consequence of this construction is that, once the particle content has been specified, i.e. the Lorentz and gauge quantum numbers of the elementary states, the Lagrangian is simply built as the most general Hermitean combination of gauge and Lorentz invariant local operators.

The further requests of "renormalizability" limits the number of allowed operators to only those with energy dimension less or equal to 4. In the Effective Field Theory language, this request emerges naturally: assuming that there is a more fundamental theory with some additional states of mass greater than some scale Λ, at energies $E < \Lambda$ the true theory can be matched to an effective theory expanded in inverse powers of Λ

$$\mathcal{L}_{\text{eff}} = \mathcal{L}^{d\leq 4} + \frac{1}{\Lambda}\mathcal{L}^{d=5} + \frac{1}{\Lambda^2}\mathcal{L}^{d=6} + \dots \,. \tag{1}$$

The $d \leq 4$ part of the Lagrangian is what is usually referred to as the Standard Model Lagrangian, while the higher dimensional terms parametrize correction to the Standard Model due to the heavier physics that we have integrated out. As long as Λ is much larger than the energies that we can probe experimentally, the first term of the equation dominates.

The first ingredient to build the Standard Model is its gauge group

$$SU(3)_c \times SU(2)_L \times U(1)_Y \tag{2}$$

describing respectively the strong and electroweak interactions. The theory is then built by specifying the field content. All the fermionic degrees of freedom come in three generations, and are divided into quarks or leptons based on their color quantum numbers. There exist three type of quarks based on their electroweak and Lorentz quantum numbers: the left-handed doublets and two right-handed singlets

$$q_L^i = \begin{pmatrix} u_L^i \\ d_L^i \end{pmatrix}, \quad u_R^i\,, \quad d_R^i\,. \tag{3}$$

Similarly the leptons also come with a left-handed doublet and a right-handed singlet

$$L_L^i = \begin{pmatrix} \nu_L^i \\ e_L^i \end{pmatrix}, \quad e_R^i\,. \tag{4}$$

The model is then completed by adding, as dictated by gauge invariance, one vector field for each generator of the gauge group

$$G_\mu^a\,, \quad W_\mu^i\,, \quad B_\mu \tag{5}$$

and a finally the Higgs doublet, a scalar field necessary to allow mass terms for some gauge bosons and fermions

$$H = \begin{pmatrix} \phi^+ \\ \frac{h+i\phi_Z}{\sqrt{2}} \end{pmatrix}. \tag{6}$$

The Lorentz and gauge quantum numbers of these fields are summarized in Table 1.

Field	Lorentz	$SU(3)_c$	$SU(2)_L$	$U(1)_Y$
q_L	$(1/2,0)$	**3**	**2**	$1/6$
u_R	$(0,1/2)$	**3**	**1**	$2/3$
d_R	$(0,1/2)$	**3**	**1**	$-1/3$
L_L	$(1/2,0)$	**1**	**2**	$-1/2$
e_R	$(0,1/2)$	**1**	**1**	-1
G	$(1/2,1/2)$	**8**	**1**	0
W	$(1/2,1/2)$	**1**	**3**	0
B	$(1/2,1/2)$	**1**	**1**	0
H	$(0,0)$	**1**	**2**	$1/2$

Table 1: Lorentz and Gauge quantum numbers of the Standard Model fields.

The Standard Model Lagrangian then reads

$$\begin{aligned}\mathcal{L}_{\text{SM}} =& -\frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a - \frac{1}{4}W^a_{\mu\nu}W^{\mu\nu}_a - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \\ &+ i\bar{q}^i_L \not{D} q^i_L + i\bar{u}^i_R \not{D} u^i_R + i\bar{d}^i_R \not{D} d^i_R + i\bar{L}^i_L \not{D} L^i_L + i\bar{e}^i_R \not{D} e^i_R \\ &+ (D_\mu H)^\dagger D^\mu H + \mu_H^2 H^\dagger H - \lambda (H^\dagger H)^2 \\ &- \left(Y^u_{ij} \bar{q}^i_L \widetilde{H} u^j_R + Y^d_{ij} \bar{q}^i_L H d^j_R + Y^e_{ij} \bar{L}^i_L \widetilde{H} e^j_R + h.c. \right) ,\end{aligned} \tag{7}$$

where we the field strength tensors read generically in terms of a gauge field A^a_μ, coupling constant g_A and structure constant f^{abc} as

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - i g_A f^{abc} A^b_\mu A^c_\nu \,, \tag{8}$$

the covariant derivatives are defined as

$$D_\mu = \partial_\mu - i g_s \lambda^a G^a_\mu - i g T^a W^a_\mu - i g' Y B_\mu \,, \tag{9}$$

and finally the conjugate Higgs field is

$$\widetilde{H} = i\sigma_2 H^* \,. \tag{10}$$

The Lagrangian written in this way is however not parametrized in terms of fields that correspond directly with the physical propagating degrees of freedom. For example it does not contain any explicit mass terms for either gauge bosons or fermions, while the mass term for the Higgs has the wrong sign. A more convenient parametrization, that loses however the explicit $SU(2)_L$ invariance, but approximates the SM spectrum, is obtained after expanding around the proper vacuum. The Higgs field obtains a vacuum expectation value

$$\langle H \rangle = \frac{v}{\sqrt{2}} \approx 246\,\text{GeV}, \qquad v = \frac{\mu_H}{\sqrt{\lambda}} \tag{11}$$

that breaks the electroweak gauge group down to the electromagnetic one $SU(2)_L \times$

$U(1)_Y \to U(1)_{\rm EM}$. The Higgs doublet can then be reparametrized as

$$H = e^{i\sigma^a \frac{\pi^a(x)}{v}} \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} + \frac{h(x)}{\sqrt{2}} \end{pmatrix} \tag{12}$$

where the fields factorized in front can be eliminated through a gauge transformation in favor of the longitudinal degrees of freedom for the W and Z bosons. The kinetic term of the Higgs gives rise to a mass term for these two vector bosons, defined as the linear combinations

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp i W_\mu^2)\,, \quad Z_\mu = \cos\theta_w W_\mu^3 - \sin\theta_w B_\mu\,, \tag{13}$$

where we defined the Weinberg angle

$$\tan\theta_w = \frac{g'}{g} \tag{14}$$

The W and Z boson thus acquire a mass of

$$m_W = \frac{1}{2} g\, v \approx 80\,\text{GeV}\,, \quad m_Z = \frac{1}{2}\sqrt{g^2 + g'^2}\, v \approx 91\,\text{GeV} \tag{15}$$

while the photon field

$$A_\mu = \sin\theta_w W_\mu^3 + \cos\theta_w B_\mu\,, \tag{16}$$

remains massless and thus couples to the fermions in a gauge invariant way with a coupling of

$$e = g \sin\theta_w \tag{17}$$

The physical Higgs boson also acquires a mass term

$$m_h = \sqrt{2\lambda} v \approx 125\,\text{GeV}\,. \tag{18}$$

Even though the $SU(2)_L$ gauge symmetry is no longer explicit, its signature can still be found in the structure of the interaction involving the Higgs and the W and Z boson. The Lagrangian of the Higgs sector can indeed be written in the compact form

$$\mathcal{L}_{\rm H} = \frac{1}{2}(\partial_\mu h)^2 - \frac{1}{2} m_h^2 h^2 - \lambda v h^2 - \frac{1}{4}\lambda h^4 + \left[m_W^2 W^{\mu+} W_\mu^- + \frac{1}{2} m_Z^2 Z^\mu Z_\mu \right] \left(1 + \frac{h}{v}\right)^2\,. \tag{19}$$

This structure implies strong correlations between the couplings of the Higgs with itself and with the gauge bosons. Violations in the relationships among these couplings are possible signatures of new physics in the electroweak sector.

Finally, this mechanism also leads to mass terms for the fermions of the form

$$\mathcal{L}_{\rm mass} = \frac{v}{\sqrt{2}} \left(Y_{ij}^u \bar{u}_L^i u_R^j + Y_{ij}^d d_L^i d_R^j + Y_{ij}^e \bar{e}_L^i e_R^j + h.c. \right) \tag{20}$$

Similarly to before, note that in the Standard Model the interaction between the Higgs and the fermions has exactly the same structure but with $v \leftrightarrow h$. This gives another set of relations that could be possibly violated by new physics.

The Yukawa matrices Y can be diagonalized through bi-unitary transformations. In the mass-diagonal basis, each fermion f gets a mass proportional to the Higgs expectation value v and the eigenvalues y_f of Y as in

$$m_f = \frac{v}{\sqrt{2}} y_f \,. \tag{21}$$

For the lepton sector this diagonalization leads to three massive leptons

$$m_e \approx 0.51\,\text{MeV}\,, \quad m_\mu \approx 106\,\text{MeV}\,, \quad m_\tau \approx 1.78\,\text{GeV}\,, \tag{22}$$

and three massless neutrinos. The quark sector instead can only be diagonalized up to a 3×3 unitary matrix that is usually redefined away to give a diagonal mass matrix, but a non diagonal coupling with the W boson. The quark masses are given by

$$\begin{aligned} m_d(\mu_{\overline{MS}} = 2\,\text{GeV}) &\approx 5\,\text{MeV}\,, \quad m_u(\mu_{\overline{MS}} = 2\,\text{GeV}) \approx 2\,\text{MeV}\,, \\ m_s(\mu_{\overline{MS}} = 2\,\text{GeV}) &\approx 90\,\text{MeV}\,, \quad m_c(\mu_{\overline{MS}} = m_c) \approx 1.3\,\text{GeV}\,, \\ m_b(\mu_{\overline{MS}} = m_b) &\approx 4.2\,\text{GeV}\,, \quad m_t(\text{on-shell}) \approx 173\,\text{GeV}\,, \end{aligned} \tag{23}$$

where in parenthesis we denote the scale at which the renormalized masses are defined [1]. The remaining unitary matrix, named the Cabibbo–Kobayashi–Maskawa (CKM) matrix, gives rise to flavor-changing transitions in the quark sector through the emission of a W boson. This matrix can be parametrized in terms of three real numbers and a complex phase. A convenient, although approximate, parametrization of this matrix that takes into account the strong hierarchy among its entries is the Wolfenstein parametrization

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(\rho + i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{24}$$

where the measured values are

$$\lambda \approx 0.23\,, \quad A \approx 0.8\,, \quad \rho \approx 0.14\,, \quad \eta \approx 0.35. \tag{25}$$

It is important to notice that this matrix is the only source of flavor violation in the quark sector.

[1] The exact definition of these masses is complicated since, except for the top, all the quarks only appear in strongly coupled bound states. We will not discuss their definitions in detail.

Global and approximate symmetries

Considering only the $d \leq 4$ piece of the theory gives rise to several accidental symmetries, that is symmetries that do not come from any principle but happen because the only operators that violate them appear at $d > 4$. In this section we will review the most important of these symmetries, may they be exact or only approximate, and consider their consequences. Since higher dimension operators can violate them, they provide targets to look for new physics.

The Standard Model Lagrangian has an exact $U(1)^4$ symmetry related to the conservation of the baryon number and the lepton family number. Under this symmetry all quark fields rotate with the same phase while the three leptons rotate all independently each with a different phase. This symmetry explicitly forbids transitions such as proton decay $p \to e^+\pi^0$ or the neutrinoless muon decays $\mu^+ \to e^+\gamma$ or $\mu^+ \to e^+e^+e^-$, none of which have been observed so far. This symmetry is however known to be broken by the presence of neutrino masses. The evidence of a non-zero neutrino masses indeed comes from the observation of neutrino flavor oscillations. Neutrino masses introduce a mixing matrix similar to the quark one, the Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix, that forces the mass eigenstates to be a linear combination of the flavor states.

The total lepton number symmetry in the Standard Model is accidental. It is indeed broken by a dimension-5 operator (actually the only possible operator of this dimensionality)

$$\mathcal{L}_{\text{dim}-5} = \frac{c_{ij}}{\Lambda}(\bar{L}_L^i \widetilde{H})(\widetilde{H}\widetilde{L}_L^j) + h.c.\,, \tag{26}$$

that gives the left-handed neutrinos a Majorana mass term after the Higgs takes an expectation value. The current value of the neutrino masses $\lesssim 0.1$ eV points to a scale that generates this operator of roughly $\Lambda \sim v^2/m_\nu \approx 10^{14}\,\text{GeV}$.

The baryon number symmetry is also accidental, but this time it is broken at dimension-6. For example by the operator

$$\frac{c}{\Lambda^2}\epsilon_{ijk}\left(\bar{q}_L^{i,\alpha}\varepsilon_{\alpha\beta}(q_L^c)^{j,\beta}\right)\left(\bar{q}_L^{k,\gamma}\varepsilon_{\gamma\delta}(L_L^c)^{\delta}\right)\,, \tag{27}$$

that breaks the combination $B + L$. This operator could induce the decay of the proton. Given the current bound on the proton lifetime one can estimate $\Lambda \gtrsim 10^{15}\,\text{GeV}$. The lepton and baryon number symmetries are also broken by quantum anomalies within the Standard Model. However, baryon and number violating transitions due to this anomaly are extremely suppressed. Such transitions become active at temperatures higher than m_W, as we will see in chapter 5.

Several symmetries also arise in the limit in which the Yukawa couplings go to zero. The first one is an approximate $SO(3)$ symmetry of the Higgs sector, broken by the hypercharge gauging, dubbed custodial symmetry. The Higgs potential can indeed be

rewritten in terms of a 2×2 matrix

$$\mathcal{H} = \frac{1}{\sqrt{2}} \left[\widetilde{H} \; H \right] \tag{28}$$

as

$$\mathcal{L}_{\text{Higgs}} = \text{Tr}[(D_\mu \mathcal{H})^\dagger D^\mu \mathcal{H}] - \mu^2 \text{Tr}[\mathcal{H}^\dagger \mathcal{H}] + \lambda (\text{Tr}[\mathcal{H}^\dagger \mathcal{H}])^2 \tag{29}$$

that is invariant under a $SU(2)_L \times SU(2)_R$ transformation

$$\mathcal{H} \to L \mathcal{H} R^\dagger \, . \tag{30}$$

The left group coincides with the gauge group, while the symmetry under the right transformation is accidental. The hypercharge group $U(1)_Y$ gauges a subgroup of $SU(2)_R$ breaking it explicitly. In the limit $g' \to 0$ this symmetry is however exact and forces the mass of the W and Z bosons, that transform as a triplet under this symmetry, to be equal. Restoring g' this leads to the famous relation

$$\rho \equiv \frac{m_W^2}{m_Z^2 \cos\theta_w} = 1 \, . \tag{31}$$

Since, the Yukawa couplings break this symmetry, this relation is corrected at loop level. In particular the largest contribution comes from a top loop that gives a $\sim 1\%$ correction

$$\rho = 1 + \frac{3 y_t^2}{4\pi^2} \, . \tag{32}$$

This relation is measured very precisely and can be easily spoiled already at tree level by new physics and it is thus another important possible probe of new dynamics. An example of an operator that would lead to a correction of the ρ parameter is the dimension six operator

$$\frac{c_T}{\Lambda^2} (H^\dagger \overleftrightarrow{D}^\mu H)(H^\dagger \overleftrightarrow{D}_\mu H) \, . \tag{33}$$

When the Yukawa couplings are turned off, the Standard Model also gains an $U(3)^5$ symmetry that prevents any mixing between different families and chiralities of the fermions. When the Yukawas are turned back on, this group is broken to the $U(1)^4$ mentioned above. The only flavor-breaking parameters are those of the CKM matrix and the complex phase in this matrix is also the only source of CP violation in the model. To be more precise, the Standard Model Lagrangian contains also the topological term

$$\mathcal{L}_{\text{topo}} = -\frac{\theta_s}{16\pi^2} G^a_{\mu\nu} \widetilde{G}^{\mu\nu}_a \, , \tag{34}$$

that would lead to CP violating effects in QCD. The θ parameter is however measured to be compatible with zero $\theta \lesssim 10^{-10}$ from the upper bounds on the neutron Electric Dipole Moment (EDM). The reason for why this parameter is so small is still not currently known, but several explanations have been proposed such as the axion or the Nelson-Barr

mechanism. Both solutions require some new dynamics outside the Standard Model.

The fact that the only source of CP violation in the Standard Model is the CKM matrix has an important consequence that leads to a large suppression of most CP violating observables. The only physical phase of the matrix must indeed involve the product of four components to form the Jarlskog invariant

$$\mathrm{Im}[V_{ij}V_{kl}V_{il}^*V_{kj}^*] = J\sum_{m,n}\epsilon_{ikm}\epsilon_{jln}\,. \tag{35}$$

This quantity is the only imaginary part that is invariant under field redefinitions. Numerically it turns out to be quite small

$$J \approx A^2\lambda^6\eta \approx 3\times 10^{-5} \tag{36}$$

and furthermore it involves four insertions of the CKM matrix spanning through all three quark families. This has as a consequence that CP violating observables in the Standard Model usually arise at loop level, making the effect of CP violation extremely small.

CP violation in the Standard Model has been measured in rare processes involving hadrons. One important class of precision observables are the Flavor Changing Neutral Currents (FCNC). In the Standard Model those are forbidden at tree-level, since both the Z and γ couplings are flavor diagonal, and only appear at loop level. CP violation can be measured in such observables, for example in the mixing between neutral mesons and their decays. These effects are measured with great precision and are another great probe of flavor- and CP-violating new physics, as we will see in chapter 1. New physics can indeed contribute to these effects through dimension six operators such as the four fermion ones

$$\frac{c^{ij}}{\Lambda^2}(\bar{f}^i\gamma_\mu f^j)(\bar{f}^i\gamma^\mu f^j)\,, \tag{37}$$

with f^i some Standard Model quark. From the $K-\bar{K}$ mixing for example, assuming $c\sim 1$ one finds a bound on Λ of at least 10^5-10^6 TeV. A weaker coupling and some symmetry assumptions on the dynamics that generates this operator can however lead to a weaker bound and so a scale for new physics that is accessible to current or near future experiments.

What is missing?

There are still several open questions regarding the Standard Model, ranging from practical difficulties in extracting predictions, to observational facts that the model does not capture. In this section we will schematically review the relevant ones, some of which will serve as motivation for the works presented in this thesis.

Starting from the Standard Model itself, even though we can write the theory in just a few lines, it does not mean that we know how to extract all the possible predictions in a accurate and systematic way. For example, due to its strongly coupled nature at low energies, we do not have a first principles systematic way of computing the QCD hadron spectrum or the parton distribution functions for the proton. Another example of this is the existence of exotic QCD resonances, whose nature has not yet been understood. One of such resonances in the charmonia sector, the famous $X(3872)$, will be studied in chapter 4 through non-relativistic effective theories.

Even in the perturbative regime there is no systematic way of computing observables up to arbitrarily high precision. For example, even for the ElectroWeak theory the problem of IR effects at high energy and their resummation has not been completely solved and might become an issue for precision physics programs at future colliders. This will be addressed in chapter 3.

The ones just mentioned are essentially technical and computational problems within the Standard Model. There are however questions of a more fundamental nature that suggest that we have to modify the theory and go beyond. We can broadly divide these problems in two classes: the first one contains those for which we have observational evidences that cannot be explained within the model itself, while the second one contains the problems that appear at the theory level mainly because of difficulties in imagining how the the model could emerge from a more fundamental theory.

Several of the problems of the first class are related to astrophysical and cosmological observations. There is a large number of experimental evidences for the existence of dark matter, from the galaxy rotational curves to the CMB spectrum, but the Standard Model does not seem to contain any particle candidate that could explain it. Similarly there is the problem of baryogenesis: after inflation the number of baryons and anti-baryons in the universe should have been equal and, considering only Standard Model processes alone, this equality would have remained constant as the universe cooled down. It is however clear that this is not the case, since our universe contains only baryons and not anti-baryons. There must then exist new physics capable of generating this asymmetry during the cosmological evolution of the universe. Baryogenesis and dark matter will be studied in chapter 5.

Also in this class we can put the already mentioned question of addressing the neutrino masses. Finally there is of course the question of gravity. The Standard Model does not include gravitational interactions. Extending the model with a massless spin-2 field, the graviton, as to reproduce General Relativity, leads to an effective theory that breaks down at energies of the order of the Planck scale $\sim 10^{19}\,\text{GeV}$. So far, the only promising attempt at a quantum theory of gravity seems to be String Theory.

The second class of problems contains puzzles about why the parameters of the Standard

Model have the value that they have. One example already mentioned is the puzzle of the θ_s QCD angle for which, since CP is violated in the electroweak sector, there is no fundamental reason to expect it to be zero. Similarly, there are two more parameters in the Standard Model Lagrangian whose value is smaller than expected: the vacuum energy and the Higgs mass. These two are the only dimensionful parameters of the theory and, since there is no additional symmetry that appears when either parameter is zero, we would expect that any energy scale of the theory would directly contribute to them.

Let us take the Higgs mass as a more concrete example. We have seen that the Standard Model is an effective theory valid up to some scale Λ. The fact that we have not seen any violations of the fragile structure of the Standard Model would make us believe that Λ is very high, for example $\Lambda \sim 10^{14}\,\text{GeV}$ as suggested by the neutrino masses. However, if we believe that the Standard Model emerges from some new physics at that scale, building a theory in which the Higgs mass is calculable and 12 order of magnitudes smaller than Λ is extremely hard. Indeed, as with any quantum mechanical observables, the Higgs mass receives contributions from states of all possible energies. Without knowing the full fundamental theory, we can only estimate part of the contribution that comes from $E \lesssim \Lambda$, while the contributions coming from $E \gtrsim \Lambda$ is unknown. Explicitly we have

$$m_H^2 \sim \frac{3y_t^2}{4\pi^2}\Lambda^2 + \left[\delta m_H^2\right]_{E\gtrsim\Lambda} , \tag{38}$$

where the first term was estimated as the Standard Model contribution, coming from a single top quark loop. To reproduce the observed Higgs mass, the two terms should cancel out with a precision of $\sim (100\,\text{GeV}/\Lambda)^2$. The question on how to generate a scale separation between the weak scale and Λ is known as the Naturalness or Hierarchy problem.

A way out is to imagine that there is some TeV scale dynamics that intervenes and screens the Higgs mass from high energy quantum corrections. In chapter 1 we will review the solution offered by Composite Higgs models, in which the Higgs is seen as a composite particle that arises from a new QCD-like dynamics that confines at a few TeV. An alternative solution to the Higgs mass problem is offered by Supersymmetry. In these kinds of models a new space-time symmetry that links bosons and fermions is introduced. The fermionic partner of the Higgs then gains a chiral symmetry in the $m_H^2 \to 0$ limit. This implies that corrections to the Higgs mass must be proportional to m_H^2 itself, making it possible to have the Higgs naturally light.

Models of new physics should also try to explain the generation of the fermion mass hierarchies and the CKM mixing angles. As we will see in chapter 1 this is not a trivial task. At least for what regards Composite Higgs models, it seems that there is a tradeoff between dynamically explaining the origin of these hierarchies and making sure not to generate large contributions the processes that are extremely suppressed in the Standard Model, such as the EDMs and FCNCs.

To conclude this discussion, it is clear that the Standard Model itself cannot be complete and it must be extended by adding new dynamics above the weak scale. If the new physics scale is higher (but hopefully not too high) than what we can directly probe with the current or future collider experiments, this new dynamics will manifest itself as modifications of the Standard Model interactions through some higher dimensional effective operators as in equation 1. Since these operators can give effects that grow with the energy of the process, they could be measured as deviation in the high energy event distributions. Chapters 2 and 3 will talk about how to measure and in some case amplify such deviations to better probe new physics.

It must however be stressed that a description of new physics just based on the Effective Field Theory point of view is incomplete and it should always be corroborated by an explicit BSM model. It is indeed only with a model based on concrete dynamical hypotheses and symmetries that we can properly estimate the scale of new physics, find correlations between different operators and understand which operators are relevant for a given observable and which are not.

1 Flavor in Composite Higgs models

In this chapter we will study the properties of models of Higgs Compositeness (see [4] for a complete review), in particular regarding their flavor structure in the quark sector. Composite Higgs models are interesting since they offer a possible solution to the Higgs mass naturalness problem. In these models the Higgs is not an elementary particle, but a composite state emerging from some new dynamics that confines at a scale m_* not too far from the TeV scale. This scale is generated by the running of the couplings in the new dynamics and can thus be naturally smaller than other scales in the theory. This is similar to the Standard Model, where QCD confines at a scale $\Lambda_{\text{QCD}} \sim 1\,\text{GeV}$ quite far from the ElectroWeak scale. Let's see how this works in a little more detail. Suppose that this new dynamics, that we will call the *composite sector* or *strong sector* interchangeably, is created at a scale $\Lambda_{\text{UV}} \gg$ TeV. In order not to introduce another Naturalness problem, we assume that at this scale the theory is close to a fixed point and does not contain any strongly relevant deformations. This fact makes it possible to generate a large separation between the UV scale and the IR scale where the theory confines. Indeed, taking a QCD-like dynamics as a concrete example, the RG running gives

$$m_* \approx \Lambda_{\text{UV}}\, e^{-\frac{2\pi}{\beta_0}\frac{1}{\alpha_s(\Lambda_{\text{UV}})}}\,, \tag{1.1}$$

where $\beta_0 > 0$ is a constant that depends on the number of degrees of freedom of the theory, while $\alpha_s(\Lambda_{\text{UV}})$ is the coupling of this sector at the UV scale. The smallness of $\alpha_s(\Lambda_{\text{UV}})$ gets amplified by the exponential creating a large separation of scales.

At the scale m_*, the strong sector confines and creates some hadron-like composite states among which the lightest ones will form the Standard Model Higgs doublet. To separate these new states, that have not been seen experimentally, from the Higgs, we assume that the Higgs is an approximate Goldstone Boson of a global symmetry that is spontaneously broken by the confinement dynamics. This is similar to what happens in QCD where the pions and kaons are lighter than the rest of the hadrons because they are approximate Goldstone Bosons of the spontaneously broken chiral simmetry. The symmetry that

protects the Higgs mass is then explicitly broken by the gauge and fermion interactions allowing the generation of a potential for the Higgs at loop level. It should then be clear that this mechanism can solve the Higgs naturalness problem only as long as the scale m_* is not too far from the ElectroWeak scale.

It is important to note, however, that this QCD analogy is not completely appropriate for general Compostite Higgs models. Indeed, no explicit realization of such models are known in four dimensions. The only known examples of Composite Higgs theories are holographic models in five dimensions or models based on dimensional deconstruction.

In this chapter we will explore this idea in slightly more detail, first by reviewing the main assumptions, ideas and properties of these models, and then by analyzing in details the consequences in flavor physics, especially in the quark sector. The topics in the review section are mainly based on the two reviews [4, 5], while the analysis on flavor is original to this thesis work.

1.1 Minimal Model

The minimal Composite Higgs model [6] contain two weakly interacting sectors, that we will refer to as the *composite sector* and the *elementary sector*. Both sectors, as explained before, are generated at some high UV scale $\Lambda_{\rm UV} \gg$ TeV by some non-specified dynamics.

The composite sector is characterized by an exact global symmetry group $\mathcal{G}$. Around the confinement scale m_*, this group is spontaneously broken to a subgroup $\mathcal{H}$. These two groups are assumed to be such that the Higgs doublet can be a Goldstone boson of such a breaking, that is explicitly $H \in \mathcal{G}/\mathcal{H}$.

The elementary sector instead contains all the Standard Model particles, both fermions and gauge bosons, except for the Higgs doublet. As we will see better later, these elementary fermions do not exactly coincide with the Standard Model ones, but they mix in the IR with some heavy composite resonances. Since we know that the Higgs carries electroweak quantum numbers, the elementary electroweak gauge fields W and B must gauge a $SU(2)_L \times U(1)_Y$ subgroup of $\mathcal{H}$.

These two sectors interact via gauge interactions that couple the W and B to the conserved currents of the composite sector under the Standard Model gauge group. Furthermore we will assume some mixing interaction between the elementary fermions and the composite sector in order to reproduce the Standard Model Yukawa interactions once the heavy strong sector resonances are integrated out. The specifics of this interaction will be thoroughly explored later and in this section we will focus on the Higgs and electroweak sector of the theory.

As we said, the composite sector is characterized by a spontaneously broken global symmetry $\mathcal{G} \to \mathcal{H} \supset \mathcal{G}_{EW}$. The so-called minimal model assumes $\mathcal{G} = SO(5) \times U(1)_X$ and $\mathcal{H} = SO(4) \times U(1)_X$. The Abelian $U(1)_X$ is needed in order to correctly reproduce the hypercharges of the Standard Model fermions. The breaking $\mathcal{G} \to \mathcal{H}$ leads to four real Goldstone bosons that match the component of the Standard Model Higgs doublet. $SO(4) \times U(1)_X$ also contains the Standard Model ElectroWeak group: $SO(4)$ has indeed the same algebra as $SU(2)_L \times SU(2)_R$ and the ElectroWeak group can be identified as the product of $SU(2)_L$ and the $U(1)_Y$ generated by the diagonal generator T^3_R of $SU(2)_R$ and the abelian $U(1)_X$ as $Y = T^3_R + X$. The choice of $SO(4)$ also guarantees the presence of the custodial symmetry and thus avoids large corrections to the ρ parameter.

A complete construction of the Effective Lagrangian for the Standard Model degrees of freedom below the scale m_* that implements the non-linear realization of the symmetry group $\mathcal{G}$ would require the CCWZ construction [7, 8]. We will not go into details on this topic, but only consider a simple Sigma-Model. The Lagrangian for the minimal model can be built starting from a field Φ that transforms as a **5** of $SO(5)$ described by the Lagrangian [6, 9]

$$\mathcal{L} = \frac{1}{2}(D_\mu \Phi)^2 - \frac{g_*^2}{8}(|\Phi|^2 - f^2)^2 \,. \tag{1.2}$$

The covariant derivative is defined as

$$D_\mu \Phi = (\partial_\mu - igW^\alpha_\mu T^\alpha_L - ig; B_\mu T^3_R)\Phi \,, \tag{1.3}$$

that explicitly breaks the $SO(5)$ global symmetry to the Standard Model ElectroWeak one due to the gauging. The field Φ is then parametrized in terms of the non-linearly transforming Goldstone fields Π^i

$$\Phi(x) = \exp\left[i\frac{\sqrt{2}}{f}\Pi_i \hat{T}^i\right] F \,, \qquad F = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ f \end{pmatrix} \tag{1.4}$$

where F is the vacuum state, annihilated by all the $SO(4)$ generators and f is the expectation value of Φ that breaks $SO(5)$ to $SO(4)$. The generators $\hat{T}^i$ are the four "broken" generators, that is the generators of $SO(5)$ that are non zero on the vacuum state. The vev f is related to the typical resonance mass m_* and the coupling g_* by

$$m_* = g_* f \,, \tag{1.5}$$

as one could see for example by considering the radial fluctuations of Φ. As already mentioned, the fermion sector of the theory will be analyzed later. However, it is clear that the Standard Model fermions do not correspond to representations of $SO(4)$, so also their interactions explicitly break the Goldstone symmetry. Integrating out the Standard

Model fermions and gauge bosons at one loop, then creates a non-trivial potential for the Goldstones that triggers ElectroWeak Symmetry Breaking. The Lagrangian (1.2) can then be rewritten in terms of the Higgs doublet

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} \Pi^2 + i\Pi^2 \\ \Pi^4 - i\Pi^3 \end{pmatrix} , \tag{1.6}$$

that, after ElectroWeak Symmetry Breaking and in the unitary gauge, becomes

$$H = \begin{pmatrix} 0 \\ \frac{V+h}{\sqrt{2}} \end{pmatrix} , \tag{1.7}$$

where V is the vacuum expectation value of the Higgs doublet. In terms of h and the ElectroWeak bosons, the Lagrangian of the model (1.2) reads

$$\begin{aligned} \mathcal{L} =& \frac{1}{2}(\partial_\mu h)^2 + \frac{g^2}{8} f^2 \sin^2 \frac{V+h}{f} \left(g^2 (W^1)^2 + g^2 (W^2)^2 + (gW^3 - g'B)^2 \right) \\ =& \frac{1}{2}(\partial_\mu h)^2 + \frac{g^2}{4} f^2 \sin^2 \frac{V+h}{f} \left(|W|^2 + \frac{1}{2\cos\theta_w} Z^2 \right) . \end{aligned} \tag{1.8}$$

Notice that the form of the term in parenthesis in the first line is forced by custodial symmetry, in particular in the limit $g' \to 0$. In this model the mass of the W and Z vector bosons read

$$m_W = \cos\theta_w m_Z = \frac{1}{2} g f \sin\frac{V}{f} \equiv \frac{1}{2} g v , \tag{1.9}$$

where $v \approx 246\,\text{GeV}$ is the Standard Model ElectroWeak breaking scale. In the large f limit, the Lagrangian (1.8) can be expanded around $h = 0$

$$\mathcal{L} \supset \frac{g^2 v^2}{4} \left(|W|^2 + \frac{1}{2\cos\theta_w} Z^2 \right) \left[2\sqrt{1-\xi}\frac{h}{v} + (1-2\xi)\frac{h^2}{v^2} - \frac{4}{3}\xi\sqrt{1-\xi}\frac{h^3}{v^3} + \dots \right] . \tag{1.10}$$

The parameter $\xi = v^2/f^2$ parametrizes the separation of scale between the $\mathcal{G} \to \mathcal{H}$ breaking and the ElectroWeak symmetry breaking. In particular as $\xi \to 0$, the Higgs becomes effectively elementary and the Lagrangian coincides with the Standard Model.

The size of the parameter ξ is determined by the explicit form of the Higgs potential. This potential can only be computed in explicit models, such as holographic or multi-site models. It can however be estimated thanks to the symmetries of the model and the power counting rules that will be introduced in the next section. In general the potential has the form

$$V(h) \approx -\alpha f^2 \sin^2 \frac{V+h}{f} + \beta f^2 \sin^4 \frac{V+h}{f} , \tag{1.11}$$

where the parameter α receives contribution from both the gauge sector and the top-quark, while the parameter β receives contribution from the top-quark only. In terms of

these two parameters one finds that

$$\xi = \alpha/2\beta \tag{1.12}$$
$$m_H^2 = 8\xi(1-\xi)\beta\,. \tag{1.13}$$

Separating the v and f scale requires that $\beta \gg \alpha$, however in all known cases the two parameters appear either at the same order with respect to the composite theory parameters or with $\beta \ll \alpha$. The situation $\xi \ll 1$ thus usually requires some tuning between the two α and β terms. The amount of tuning clearly increases the further m_* is from the ElectroWeak scale. This roughly scales as

$$\frac{3y_t}{4\pi^2}\left(\frac{m_*}{m_H}\right)^2 \sim \left(\frac{m_*}{500\,\text{GeV}}\right)^2\,. \tag{1.14}$$

A natural Composite Higgs model would require $m_* \sim 500\,\text{GeV}$, but as we will see later, this value has been excluded by direct searches and Electroweak precision tests. Current flavor-universal bounds require $m_* \gtrsim 3\,\text{TeV}$, while, when considering also flavor observables, this bound can be pushed to even higher values. How to keep the minimum allowed m_* as low as possible will be the topic discussed in the rest of the chapter.

1.1.1 Adding Fermions

Introducing the elementary Standard Model fermions into Composite Higgs models is not a trivial task. Several things have to be accomplished in order to build a realistic model. First of all the interaction between the elementary sector and the composite strong dynamics must be able to generate the strongly hierarchical structure of the Yukawa couplings. Ideally this would also come with a dynamical explanation on how such a hierarchy is generated. However at the same time one must be careful not to generate large contributions to the many processes that are suppressed by the delicate flavor structure of the Standard Model. We will address the flavor problem in detail in section 1.2.

The standard way to couple the elementary and strong sector in Composite Higgs theories is through the mechanism of Partial Compositeness [10, 6]. According to this paradigm, the elementary fermions are coupled through linear mixings to some composite operators

$$\mathcal{L}_{\text{mix}} \sim \lambda_q^i \bar{q}_L^i O_q^i + \lambda_u^i \bar{O}_u^i u_R^i + \lambda_d^i \bar{O}_u^i d_R^i + h.c.\,, \tag{1.15}$$

where the ratios $\epsilon \equiv \lambda/g_*$ determine the amount of compositeness of each SM fermions. Indeed, the physical states will be a mixture of the elementary and composite states with

a mixing angle given by

$$\sin\theta^f = \frac{\lambda^f}{\sqrt{g_*^2 + (\lambda^f)^2}} = \frac{\epsilon^f}{\sqrt{1 + (\epsilon^f)^2}}\,. \tag{1.16}$$

The maximal value of the mixings λ is g_*, which corresponds to the limit where the elementary field is effectively composite. At energies smaller than the compositeness scale m_*, the strong dynamics can be integrated out and the SM interactions are recovered. In particular the Yukawa couplings are expected to be roughly

$$Y^u \sim \frac{\lambda_q^i \lambda_u^j}{g_*}\,, \qquad\qquad Y^d \sim \frac{\lambda_q^i \lambda_d^j}{g_*}\,, \tag{1.17}$$

as we will see better later.

The structure of partial compositeness allows for a dynamical generation of the flavor hierarchies by the RG running of the theory from the UV scale $\Lambda_{\rm UV} \gg$ TeV where the composite and elementary sectors are first generated to the IR scale m_*. Qualitatively, the energy scaling of λ is dictated by the anomalous dimension of the composite operators $\gamma = \dim[O] - 5/2$, where $\dim[O]$ is the conformal dimension of the operator O. For $\gamma > 0$ the running has the form

$$\lambda(m_*) \sim \lambda(\Lambda_{\rm UV}) \left(\frac{m_*}{\Lambda_{\rm UV}}\right)^\gamma\,, \tag{1.18}$$

leading to a small mixing with the composite sector that could reproduce the Yukawas of the lighter families, while a large mixing can be generated for $\gamma \lesssim 0$ that could reproduce the top Yukawa.

1.1.2 Power-counting

An alternative to eq. (1.8) would have been to work with an the explicitly $SU(2) \times U(1)$ invariant effective Lagrangian written in terms of the doublet H. In this case, expanding for large f, corresponds to a derivative expansion in term of higher dimensional operators

$$\mathcal{L} = D_\mu H^\dagger D^\mu H - \frac{2}{3f^2}|H|^2 D_\mu H^\dagger D^\mu H + \frac{1}{6f^2}\partial_\mu (H^\dagger H)\partial^\mu (H^\dagger H) + \dots\,. \tag{1.19}$$

This parametrization makes it explicit that the Standard Model is obtained for $f \to \infty$. This will be the parametrization that we will use in the following, since it corresponds to treating the Standard Model as an effective theory in which we have integrated out states with a mass of the order of m_*.

Since we do not possess an explicit calculable model for Composite Higgs, the only way we can extract some prediction is through power-counting rules based on some assumptions on the composite dynamics. Such rules allow us to estimate the size of the coefficients of

the higher-dimensional operators of the effective theory at the scale m_*. The simplest example of power-counting that can be considered in Composite Higgs models is the SILH power-counting [11]. The strong dynamics of heavy composite resonances is assumed to contain only one mass scale m_* and one coupling g_*. This scenario is an extension of the Naive Dimensional Analysis that is known to work quite well for QCD [12].

The origin of the two parameters m_* and g_* depend of course on the underlying model. For example, in holographic models the role of m_* is played by the mass of the Kaluza-Klein modes $m_* \sim 1/R$, with R the compactification radius, while the role of g_* is played by $g_* = 4\pi/\sqrt{N}$ where the "number of colors" in the 4d theory is written in terms of the 5d gauge coupling as $1/N \equiv g_5^2/(16\pi^2 R)$.

The form of the Lagrangian is then fixed, up to dimensionless numbers, by dimensional analysis as

$$\mathcal{L}_{\rm EFT} = \frac{m_*^4}{g_*^2}\hat{\mathcal{L}}\left[\frac{\partial}{m_*}, \frac{g_* H}{m_*}, \frac{g_*\sigma}{m_*}, \frac{g_*\Psi}{m_*^{3/2}}, \frac{gA}{m_*}, \frac{\lambda\psi}{m_*^{3/2}}\right] . \tag{1.20}$$

In this formula $\hat{\mathcal{L}}$ is a generic dimensionless polynomial of the fields in the parenthesis and it can be expanded to generate the Effective Lagrangian up to whatever order is desired. The dimensionless-coefficients of this expansion are in general expected to be $O(1)$.

This formula deserves an explanation. The Higgs doublet H, scalar resonances σ and fermionic resonances Ψ enter in the same way since they come from the same strong dynamics. The factors of g_* and m_* in the expansion are dictated by dimensional analysis [1]. For the elementary vector fields, collectively called A in the formula, the coupling accompanying them is given by their gauge coupling g, since these fields have always to appear in the form

$$\mathcal{L}_{\rm int} \supset g A_\mu J^\mu ,$$

as dictated by gauge invariance. Similarly, for the elementary fermions ψ the relevant coupling that enters is the mixing λ, that is the coupling that connects them with the strong sector.

While equation (1.20) is true in general, it can be supplied with further rules based on the assumptions on the UV dynamics. For example, in the minimal model operators that violate custodial symmetry, such as

$$O_T = \frac{c_T}{2f^2}(H^\dagger \overleftrightarrow{D}_\mu H)^2 , \tag{1.21}$$

are only generated by loops of elementary states. When this happens, the coefficient is multiplied by a factor of $g^2/16\pi^2$, where g is the relevant coupling needed to generate

[1] Even though g_* and the other couplings appear dimensionless in natural units, they actually carry a dimension in terms of powers of $\hbar$. Explicitly $[g] = [\hbar]^{-1/2}$. Since the action has units of $\hbar$, this forces the way in which couplings must appear in the effective Lagrangian.

that operator at loop level.

1.2 Flavor in Composite Higgs Models

1.2.1 Partial Compositeness

The most basic models that implement Partial Compositeness [10] are usually characterized by three sets of mixings

$$\mathcal{L}_{\rm mix} = \lambda_q^{ia}\bar{q}_L^i O_q^a + \lambda_u^{aj}\bar{O}_u^a u_R^j + \lambda_d^{aj}\bar{O}_d^a d_R^j + h.c.\,, \tag{1.22}$$

where we are assuming one set of partners (O_u, O_d, O_q) for each SM fermions. Notice that in some realization of partial compositeness a second mixing term with the left-handed doublet is introduced. This will be necessary for some of the scenarios we will consider in the following, but in the general case we study here it is unnecessary. We assume that the operators O come from the strongly coupled composite dynamics that we take to be characterized by one typical mass scale m_* and one typical coupling g_*.

We now consider the so-called anarchic partial compositeness scenario (see [4] for a review) where no flavor symmetries are assumed on the strong sector, with the only requirement for eq. (1.22) to reproduce the SM Yukawas at low energy. The mixing matrices λ can written as

$$\lambda_q^{ia} = C_q^{ia}\lambda_q^a\,, \quad \lambda_u^{ai} = \lambda_u^a C_u^{ai}\,, \quad \lambda_d^{ai} = \lambda_d^a C_d^{ai} \qquad \text{(no sum)}\,, \tag{1.23}$$

with all the C^{ij} a $O(1)$ anarchic complex matrices and λ^j hierarchical vectors $\lambda^1 \ll \lambda^2 \ll \lambda^3$.[2] Such a structure could be for example naturally generated [6] by the RG running from the scale at which the composite dynamic is created Λ_{UV} down to the confinement scale m_*, thanks to different scaling dimensions of the operators O^a.

In this scenario, at the scale m_* the operators O^a of a given mixing do not have any distinguishable quantum numbers and can be rotated in flavor space together with rotations in the elementary sector in order to rewrite the mixing matrices λ in a more convenient form

$$\mathcal{L}_{\rm mix} = \lambda_q^i\bar{q}_L^i O_q^i + \lambda_u^i\bar{O}_u^i u_R^i + \lambda_d^i\bar{O}_d^i d_R^i + h.c.\,. \tag{1.24}$$

At energies smaller than m_*, integrating out the composite resonances gives rise to the

[2] Notice that some realization of anarchic partial compositeness require two mixings for the SM doublets [13]. However we remark that this is not needed to reproduce the SM and furthermore it requires specific assumption on the UV model.

Standard Model Yukawa couplings as

$$Y_u^{ij} = \frac{\lambda_q^i \lambda_u^j}{g_*} c_{ij}, \quad Y_d^{ij} = \frac{\lambda_q^i \lambda_d^j}{g_*} c'_{ij}\,, \tag{1.25}$$

where c and c' are some $\mathcal{O}(1)$ complex anarchic matrices.

The requirement of reproducing the SM masses and the CKM matrix, fixes the hierarchy among the different λ. Namely we have

$$\frac{m_u}{m_c} \sim \frac{\lambda_q^1}{\lambda_q^2}\frac{\lambda_u^1}{\lambda_u^2}\,, \quad \frac{m_c}{m_t} \sim \frac{\lambda_q^2}{\lambda_q^3}\frac{\lambda_u^2}{\lambda_u^3}\,, \quad \frac{m_d}{m_s} \sim \frac{\lambda_q^1}{\lambda_d^2}\frac{\lambda_d^1}{\lambda_d^2}\,, \quad \frac{m_d}{m_b} \sim \frac{\lambda_q^2}{\lambda_d^3}\frac{\lambda_d^2}{\lambda_d^3}\,, \tag{1.26}$$

$$V_{\rm CKM}^{ij} \sim V_{\rm CKM}^{ji} \sim \frac{\lambda_q^i}{\lambda_q^j} \qquad \text{for } i < j\,, \tag{1.27}$$

where the $\pm$ sign denotes an arbitrary linear combination of the two terms that could also have opposite sign. The conditions in eq. (1.25) and eq. (1.26) fix the parameters of the model up to one undetermined quantity, that in these models is conventionally taken to be

$$x_t = \lambda_q^3/\lambda_u^3\,, \qquad \text{with } \; y_t/g_* \lesssim x_t \lesssim g_*/y_t\,. \tag{1.28}$$

The total lack of a structure in flavor space leads to large tree-level contributions to many of the SM flavor observables that are accidentally suppressed in the SM, especially CP violating ones. For example, the composite dynamics generates an electric dipole moment (EDM) for the neutron. To be compatible with the strong experimental constraint [14], the scale m_* must be really large. If the dipole is generated at tree-level in the strong dynamics, we can estimate this effect as

$$d_n \sim \frac{m_{u/d}}{m_*^2} \leq 1.8 \times 10^{-26}\ \text{cm} \implies m_* \gtrsim 45 - 75\ \text{TeV}\,. \tag{1.29}$$

In models where the dipoles are only generated at loop level, the bound on m_* would instead be suppressed by a power of $g_*/4\pi$. Assuming a similar anarchic structure also for the lepton sector, gives an even stronger constraint on m_* coming from the much more precisely measured EDM of the electron. In that case the bound becomes

$$d_e \leq 1.1 \times 10^{-29}\ \text{cm} \implies m_* \gtrsim 1300\ \text{TeV}\,. \tag{1.30}$$

Again, the bound on the mass gets reduced by a factor $g_*/4\pi$ when the operator is induced at 1-loop level.

Large tree-level contributions also appear to many other flavor observables. For example, from the $K - \bar{K}$ mixing we find $m_* \gtrsim 25\,\text{TeV}$, while from $b \to s\gamma$ we have $m_* \gtrsim 29\,\text{TeV}$.

Finally it is important to note that in models in which we do not make any assumptions on the composite sector, contributions to the neutron EDM might also come from flavor-

universal corrections to the Standard Model dynamics. For example, if the composite sector violates CP also in the limit where all the mixing are turned off, we generate the so-called Weinberg operator

$$\mathcal{L} \supset c(m_*)\frac{g_s^3(m_*)}{g_*^2 m_*^2}\frac{1}{3!}f^{abc}G_\mu^{a\nu}G_\nu^{b\rho}\widetilde{G}_\rho^{c\mu}\,, \tag{1.31}$$

that contributes to the neutron EDM [15]. Indeed, we can estimate from NDA up to a dimensionless order one number

$$\frac{d_n}{e} \approx c(1\,\text{GeV})\frac{g_s^3(m_*)}{g_*^2 m_*^2}\frac{\Lambda_{\text{QCD}}}{4\pi}\,, \tag{1.32}$$

where $c(1\,\text{GeV})$ is the value of the Wilson Coefficient $c(m_*)$ at the $\Lambda_{\text{QCD}} = 1\,\text{GeV}$ scale.[3] Such a running can be computed by first running from the scale m_* to 1 TeV and then down to ~ 1 GeV using the results of [16]

$$c(1\,\text{GeV}) \approx 0.3\,c(1\,\text{TeV}) \approx 0.3\,c(m_*)\left(\frac{\alpha_s(m_*)}{\alpha_s(1\,\text{TeV})}\right)^{15/14}. \tag{1.33}$$

Imposing the current experimental constraint, this translates to a bound on the composite parameters of

$$m_* g_* \gtrsim 110\,\text{TeV}\,. \tag{1.34}$$

Even though this bound is weaker than the one from the quark dipole moments for large values of g_*, it is completely independent on the flavor structure of the model and only comes from the assumption of CP violation in the composite sector.

1.2.2 Symmetric strong sector

In the previous section we have seen that the hypothesis of anarchic partial compositeness is incompatible with a small ($\lesssim$ few TeV) value for m_* as would be required to address the hierarchy problem. The main issues are related to CP odd observables and flavor transitions, especially the ones in the down sector, that are very well experimentally constrained. A possible way out is to invoke some symmetries in the strong dynamics. However, the price to pay if we follow this path is that the flavor hierarchy cannot be explained via RG evolution as it is done in anarchic models. In this study we adopt this paradigm and we show how, by motivated and structured assumptions on the strong sector flavor symmetry group, we can maintain $m_* \lesssim 10\,\text{TeV}$.

The first hypothesis is directly suggested by eq. (1.34): a TeV scale dynamics which generate the CP odd Weinberg operator $GG\widetilde{G}$ unavoidably produces a large neutron

[3] Arguments based on QCD sum-rules [15] estimate a smaller matching coefficient in eq. (1.32). In particular in [15] it was found $d_n/e = c(1\,\text{GeV})\frac{g_s^3(m_*)}{g_*^2 m_*^2}$ ((10 − 30)MeV).

dipole moment. To avoid this, we then have to assume that the composite dynamics respects CP and the only source of its breaking comes from the various mixings of eq. (1.24), that is of course needed to reproduce the SM. This is however not enough to avoid the large CP violating effects mentioned before. Indeed, the $\mathcal{O}(1)$ phases coming from the mixings λ, in general, imply again the strong bounds of eqs. (1.29) and (1.30) and from the other flavor violating observables. Some flavor symmetry is necessary.

A very efficient solution to this problem is given by Minimal Flavor Violation (MFV) [17]. The idea is to assume that the only sources of violation of the flavor symmetries of the SM quarks (or in general fermions) $SU(3)_q \times SU(3)_u \times SU(3)_d$ are given by the SM Yukawas, transforming as spurions as

$$\begin{array}{cccc} & SU(3)_q & SU(3)_u & SU(3)_d \\ Y^u \sim & 3 & \bar{3} & 1 \\ Y^d \sim & 3 & 1 & \bar{3} \end{array} \tag{1.35}$$

and that any new flavor violating contribution is proportional to them. Notice that this hypothesis, together with a CP invariant strong dynamics, not only protect dipoles but also screens (at least partially) flavor transitions, as we will see in detail later on.

In partial compositeness this idea is realized imposing flavor symmetries on the strong sector [18, 19]. Starting from the double mixing Lagrangian

$$\mathcal{L}_{\rm mix} = \lambda_{q_u}^{ia}\bar{q}_L^i O_{q_u}^a + \lambda_{q_d}^{ia}\bar{q}_L^i O_{q_d}^a + \lambda_u^{aj}\bar{O}_u^a u_R^j + \lambda_d^{aj}\bar{O}_d^a d_R^j + h.c.\,, \tag{1.36}$$

in the absence of the mixings λ, the elementary are symmetric under the flavor group $\mathcal{F}_e = U(3)_q \times U(3)_u \times U(3)_d$. Similarly we assume a flavor group $\mathcal{F}_c$ under which the O operators transform. To reproduce the SM Yukawas

$$Y_u^{ij} \simeq (\lambda_{q_u}\lambda_u)^{ij}\,, \qquad\qquad Y_d^{ij} \simeq (\lambda_{q_d}\lambda_d)^{ij}\,, \tag{1.37}$$

it is clear that we need O_{q_u} and O_u (O_{q_d} and O_d) to transform under the same group. At most $\mathcal{F}_c$ can be $U(3)_U \times U(3)_D$, with an obvious notation. The MFV paradigm requires the λs of eq. (1.36) to be spurions of the breaking of $\mathcal{F} \equiv \mathcal{F}_e \times \mathcal{F}_c$ to $U(1)_{\mathcal{B}}$ according to the same pattern of the SM Yukawas. This last point, together with eq. (1.37) selects

only few options, summarized by the following breaking scheme

$$\mathcal{F}_e \times (\mathcal{F}_c \equiv U(3)_U \times U(3)_D) \xrightarrow{\lambda_u \propto \lambda_d \propto \mathbb{1}} U(3)_q \times U(3)_{U+u} \times U(3)_{D+d} \xrightarrow[\lambda_{q_d} \propto Y_d]{\lambda_{q_u} \propto Y_u} U(1)_{\mathcal{B}}\,, \tag{1.38}$$

$$\mathcal{F}_e \times (\mathcal{F}_c \equiv U(3)_U \times U(3)_D) \xrightarrow{\lambda_{q_u} \propto \lambda_{q_d} \propto \mathbb{1}} U(3)_{U+D+q} \times U(3)_u \times U(3)_d \xrightarrow[\lambda_d \propto Y_d]{\lambda_u \propto Y_u} U(1)_{\mathcal{B}}\,, \tag{1.39}$$

$$\mathcal{F}_e \times (\mathcal{F}_c \equiv U(3)_{U+D}) \xrightarrow{\lambda_{q_u} \propto \lambda_{q_d} \propto \mathbb{1}} U(3)_{U+D+q} \times U(3)_u \times U(3)_d \xrightarrow[\lambda_d \propto Y_d]{\lambda_u \propto Y_u} U(1)_{\mathcal{B}}\,, \tag{1.40}$$

$$\mathcal{F}_e \times (\mathcal{F}_c \equiv U(3)_U \times U(3)_D) \xrightarrow{\lambda_{q_u} \propto \lambda_d \propto \mathbb{1}} U(3)_{U+q} \times U(3)_u \times U(3)_{D+d} \xrightarrow[\lambda_d \propto Y_d]{\lambda_{q_u} \propto Y_u} U(1)_{\mathcal{B}}\,, \tag{1.41}$$

$$\mathcal{F}_e \times (\mathcal{F}_c \equiv U(3)_U \times U(3)_D) \xrightarrow{\lambda_u \propto \lambda_{q_d} \propto \mathbb{1}} U(3)_{U+u} \times U(3)_{D+q} \times U(3)_d \xrightarrow[\lambda_{q_d} \propto Y_d]{\lambda_u \propto Y_u} U(1)_{\mathcal{B}}\,, \tag{1.42}$$

that we will now review in turns.

Right-handed compositeness

The first pattern, in eq. (1.38), assumes both right handed quarks mixings to be completely aligned with the composite fermions. Consequently the left handed mixings need to be aligned with the SM Yukawas. This scenario, that we denote $U(3)^2_{\rm RC}$, has been widely explored in the literature [13, 18]. Here we will briefly review it and present updated bounds.

The model is characterized by the following form of the mixing parameters

$$U(3)^2_{\rm RC}: \quad \lambda_u^{ia} \equiv \varepsilon_u g_* \delta^{ia}\,, \qquad \lambda_d^{ia} \equiv \varepsilon_d g_* \delta^{ai}\,, \qquad \lambda_{q_u}^{ai} \simeq \frac{Y_u^{ai}}{\varepsilon_u}\,, \qquad \lambda_{q_d}^{ai} \simeq \frac{Y_d^{ai}}{\varepsilon_d}\,. \tag{1.43}$$

We will assume that $\lambda/g_* < 1$ such that a perturbative expansion of the Wilson coefficients can be obtained by counting the insertions of the mixing the parameters. From this assumption follows that $y_t/g_* \lesssim \varepsilon_u < 1$ and $y_b/g_* \lesssim \varepsilon_d < 1$.

As we have already anticipated, this scenario trivially protects the neutron EDM. The reason is that the flavor quantum numbers of the dipole interactions are the same as the Yukawa couplings. Therefore, at leading order in the insertion of the λ's the coefficient of the dipole is necessarily aligned with the SM fermion mass matrix and so, under our assumption of CP-invariance up to λ, in the mass basis it is diagonal and real. Higher order insertions of λ will induce a misaligned correction to the dipole. However, this effect is parametrically suppressed by factors of order $Y^2/16\pi^2$.

In this model the strongest constraints come from the interplay of the measurements of the different four-fermion operators. Particularly important are the bounds from the compositeness of the light families and the flavor transition involving down quarks, as we will see in a moment. The $U(3)$ symmetries of this scenario forces all up (down) right-handed quarks to be equally composite, with a degree of compositeness measured by the parameter ε_u (ε_d). The compositeness of the light families can be tested at the (HL-)LHC by studying the dijet angular distributions [20]. In particular we focus on the following two operators

$$\mathcal{O}^{(1)}_{uu} = (\bar{u}_R\gamma^\mu u_R)(\bar{u}_R\gamma_\mu u_R)\,, \qquad \mathcal{O}^{(1)}_{dd} = (\bar{d}_R\gamma^\mu d_R)(\bar{d}_R\gamma_\mu d_R)\,, \tag{1.44}$$

whose coefficients are expected to be of order $g_*^2\varepsilon_u^4/m_*^2$ and $g_*^2\varepsilon_d^4/m_*^2$ respectively. The projected 95%-bound for the 14 TeV LHC at 300 fb^{-1} integrated luminosity is [21][4]

$$m_* \gtrsim 12.7\, g_*\varepsilon_u^2\ \text{TeV}\,, \qquad m_* \gtrsim 5.8\, g_*\varepsilon_d^2\ \text{TeV}\,. \tag{1.45}$$

Regarding the flavor transitions instead, the most important bounds come from transition involving the B quark. We firstly consider the rare leptonic decays of B_s mesons into muons. In our scenarios, if we assume the leptons have a negligible mixing with the composite sector, the dominant transitions come from modifications of the Z couplings:

$$[\mathcal{O}_Z]^{ij} = \bar{f}^i\gamma^\mu f^j J^{(Z)}_\mu\,, \tag{1.46}$$

where $J^{(Z)}_\mu$ is the Standard Model current that couples to the Z. In our framework these are induced by

$$\left[\mathcal{O}^{(1)}_{Hq}\right]^{ij} = \left(H^\dagger i\overleftrightarrow{D}_\mu H\right)\bar{q}^i_L\gamma^\mu q^j_L\,, \qquad \left[\mathcal{O}^{(3)}_{Hq}\right]^{ij} = \left(H^\dagger i\overleftrightarrow{D}_\mu \tau^a H\right)\bar{q}^i_L\gamma^\mu \tau^a q^j_L\,, \tag{1.47}$$

where $\tau^a \equiv \sigma^a/2$ with σ^a the three Pauli matrices and the quarks have free flavor indices. Integrating out the Higgs and the electroweak bosons we get the effective operators of the form of eq. (1.46).

The B_s decay effective Hamiltonian is conventionally parameterized as

$$\mathcal{H}_{\text{eff}} = -\frac{4G_F}{\sqrt{2}}V^{tb}_{\text{CKM}}(V^{ts}_{\text{CKM}})^*\frac{e^2}{16\pi^2}\left[C_{10}(\bar{s}_L\gamma^\mu b_L)(\bar{\ell}\gamma_\mu\gamma^5\ell) + C'_{10}(\bar{s}_R\gamma^\mu b_R)(\bar{\ell}\gamma_\mu\gamma^5\ell)\right]\,. \tag{1.48}$$

Notice that eq. (1.46) also contains the operators C_9 and C'_9, of the same form of the previous ones but coupled with the leptonic current without the γ^5. The latter are suppressed by a factor $(1-4\sin^2\theta_w)\sim 0.08$ compared to the ones in (1.48).

[4] CMS [22] and ATLAS [23] only released 13 TeV results at 37 fb^{-1} at the time of writing. Moreover, both the analyses consider only operators universal on the fermion family. We thus decided to use the projections for the end of Run 3 from [21] for our bounds.

The current neutral anomalies in the rare B decays point towards non-zero values for the $C^{(\prime)}_{9/10}$ Wilson coefficients involving muons. Moreover the present fit prefers new physics aligned to the direction $C_9 = -C_{10}$ (see [24]), even if some improvement with the data can be obtained including the effect of only one operator among C_9 and C_{10}. We report the results from [24], for the best fit in some new physics direction

$$C_9 = -0.73 \pm 0.15\,, \qquad C_{10} = 0.54 \pm 0.12\,, \qquad C_9 = -C_{10} = +0.39 \pm 0.07\,, \tag{1.49}$$

where for each case all the other operators are set to zero.

The $U(3)^2_{\rm RC}$ model we are considering here generates C_{10} with a coefficient that can be estimated as

$$C_{10} = \frac{\sqrt{2}}{G_F}\frac{4\pi^2}{e^2}\,\frac{y_t^2}{m_*^2\varepsilon_u^2} \sim 36.6\frac{1}{m_*^2\varepsilon_u^2}\,. \tag{1.50}$$

The right-handed sector is instead diagonal and thus C'_{10} is not generated at leading order. Requiring C_{10} not to exceed twice the 1σ error on the single operator best fit in eq. (1.49) we get the following bound on m_*

$$m_* \gtrsim \frac{12.3}{\varepsilon_u}\,\text{TeV}\,. \tag{1.51}$$

With an appropriate choice of the composite group and the representation of the composite operators coupled to the elementary fermions, this bound could become weaker. Indeed, a P_{LR} symmetry [25] can be invoked to suppress the deviations of the Z couplings to the left-handed components of the down-type quarks arising from (1.104). In such a case the leading contribution would come from operators of the type

$$\bar{q}^i_L\gamma^\mu q^j_L\;\partial^\nu B_{\nu\mu}\,, \qquad\qquad \bar{q}^i_L\tau^a\gamma^\mu q^j_L\;(D^\nu W_{\nu\mu})^a\,, \tag{1.52}$$

and the resulting bound would be parametrically suppressed by a power of $g^{(Z)}_{SM}/g_* = g/(c_W g_*)$

$$m_* \gtrsim \frac{9.2}{g_*\varepsilon_u}\,\text{TeV}\,. \tag{1.53}$$

The last important bounds come from the $\Delta F = 2$ transitions in the B meson system. The right-composite MFV model only generates the $\mathcal{O}_1$ operator of eq. (1.124) at tree-level. Indeed in this model, the spurions accompanying fermion bilinears involving two quarks of different chiralities (LR) or two right-handed quarks (RR) are clearly real and diagonal in flavor since they are proportional to a Yukawa matrix or its square respectively. This is not true however for bilinears involving only left-handed quarks (LL), since they are proportional to a linear combination of $Y_uY_u^\dagger$ and $Y_dY_d^\dagger$ that is in general not diagonal.

The strongest bound comes from the B_d for which we estimate

$$\mathrm{Abs}\, C_1^{bd} \sim \frac{A^2 y_t^4 \lambda^6}{m_*^2 g_*^2 \varepsilon_u^4}\,, \tag{1.54}$$

whose corresponding bound on m_* is

$$m_* \gtrsim \frac{6.5}{g_* \varepsilon_u^2}\,. \tag{1.55}$$

A complete analysis of all the bounds of this model, including also other observables, is done in sec. 1.4.

Putting together the bounds just discussed, we find that this model requires quite a large value for m_*. If we do not assume the P_{LR} symmetry the strongest bounds are in eq. (1.45) from quark compositeness, which tends to require small ε_u, and (1.51) from $\Delta F = 1$ transitions, which require sizable ε_u. Marginalizing over ε_u we see that the lower bound on the new physics mass scale reads $m_* \gtrsim 12.4 g_*^{1/3} \gtrsim 12.4$ TeV. If one instead assumes P_{LR}, the bound from the B_s decay becomes subdominant. The minimum value for m_* can then be found by combining eq. (1.45) with eq. (1.55) to give

$$m_* \gtrsim 9\,\mathrm{TeV}\,. \tag{1.56}$$

This bound is obtained for $\varepsilon_u = 0.85/\sqrt{g_*}$, which for not so large g_* is compatible with the hypothesis of composite u-type quarks.

It is clear that the origin of this problem is that the mixing in the up-sector ε_u is the same for both the light families (that control the quark compositeness bound) and the top-quark (which gives the dominant contribution to flavor-violating processes). In sec. 1.3 we will quantify the improvement obtain when separating the mixing parameters of the first two families from that of the top quark.

Left-handed compositeness

The second and third patterns, in eqs. (1.39) and (1.40), feature SM doublets completely aligned with the composite ones. From the flavor point of view, these two scenarios are almost equivalent as long as the left handed mixings are comparable $\lambda_{q_u} \sim \lambda_{q_d}$. The main qualitative difference is that imposing the larger symmetry $\mathcal{F}_c = U(3)_U \times U(3)_D$ suppresses tree-level transitions between right-handed up and down quarks. In other words, the model (1.40) allows the flavor-violating structure $(Y_u^\dagger Y_d)_{ij} \bar{u}_i \gamma^\mu d_j$, whereas these can only arise combined with a loop suppression $g_*^2/16\pi^2$ in (1.39). These transitions are, however, not significantly constrained experimentally, as we see in detail in sec. 1.4. For this reason we will therefore only consider the case of (1.40) that can be realized with a single mixing partner of the left-handed doublets. This scenario, that we denote

$U(3)_{\rm LC}$, has been previously explored in [13, 18].

This model thus contains the following mixing parameters

$$U(3)_{\rm LC}: \qquad \lambda_q^{ia} \equiv \varepsilon_q g_* \delta^{ia}\,, \qquad \lambda_u^{ai} \simeq \frac{Y_u^{ai}}{\varepsilon_q}\,, \qquad \lambda_d^{ai} \simeq \frac{Y_d^{ai}}{\varepsilon_q}\,, \tag{1.57}$$

where $y_t/g_* \lesssim \varepsilon_q < 1$.

For the same reasons as $U(3)^2_{\rm RC}$, there is no contribution to the neutron EDM. Yet, this implementation of MFV significantly suppresses flavor-changing transitions at tree-level. This is readily understood since the spurion that controls LL structures is the identity, while for the RL and RR structures we have respectively the Yukawa and its square, which in the mass basis are diagonal. The only exception is the weakly-constrained $\bar{u}_R - d_R$ structures mentioned above.

The most relevant constraints arise from flavor conserving processes. For example, consider the operator

$$\mathcal{O}_{qq}^{(1)} = (\bar{q}_L \gamma^\mu q_L)(\bar{q}_L \gamma_\mu q_L)\,, \tag{1.58}$$

for left-handed transitions in dijet searches. Using again the bounds from [21], we find

$$m_* \gtrsim 14.2\, g_* \varepsilon_q^2\, \text{TeV}\,, \tag{1.59}$$

that clearly favors a small value for ε_q.

The strongest bound for this model however comes from the electroweak sector. The operators in eq. (1.104) indeed generate a flavor universal correction to the W coupling of the left-handed quarks as

$$\frac{g}{\sqrt{2}}(1 + \delta g_W^L)\bar{u} V_{\rm CKM} \gamma^\mu P_L d W_\mu^+\,, \tag{1.60}$$

where we estimate

$$\delta g_W^L \sim v^2 \frac{\varepsilon_q^2 g_*^2}{m_*^2}\,. \tag{1.61}$$

This correction is constrained by the measured unitarity condition on $V_{\rm CKM}$. Indeed, the experimental condition [26] $(1+\delta g_W^L)^2 \sum_i |V_{\rm CKM}^{ui}|^2 - 1 = (1.5 \pm 0.7) \times 10^{-3}$ in our model (where $V_{\rm CKM}$ is in fact unitary) reads $(1+\delta g_W^L)^2 - 1 \approx 2\delta g_W^L = (1.5 \pm 0.7) \times 10^{-3}$. By requiring that this modification does not exceed twice the 1-σ error, we derive the bound

$$m_* \gtrsim 9.3\, g_* \varepsilon_q\, \text{TeV}\,. \tag{1.62}$$

Since to reproduce the top Yukawa ε_q cannot be smaller than y_t/g_*, this gives an absolute lower bound on m_*:

$$m_* \gtrsim 9.3\, y_t\, \text{TeV} \approx 8\, \text{TeV}\,. \tag{1.63}$$

Similarly to $U(3)^2_{\rm RC}$, also in this model the origin of the strong lower bound on the new physics scale is the universality of the ε_q parameter among the three families.

Mixed compositeness

Finally the two last scenarios of eqs. (1.41) and (1.42) are somehow hybrid between the pattern previously described [18] and they maintain the same weak spots of the latter. In particular the scenario described in eq. (1.42) has the same strong bounds of $U(3)^2_{\rm RC}$, from $b \to s$ transitions and compositeness of the light up families. The model of eq. (1.41), regardless of the flavor structure is excluded up to $m_* > 8$ TeV from the unitarity of the CKM. For these reasons we will not analyze them further.

1.3 Less symmetric strong sector

We now go beyond MFV violation and we show how a smaller strong sector flavor group $\mathcal{F}_c$ allows to keep m_* relatively low. The logic we follow is to start from a MFV set-up, identify its weak spots and then try to reduce the tension with the data by downgrading $\mathcal{F}_c$ to a smaller group. We stress that our goal is not to find the composite Higgs model less constrained by flavor measurements, in fact we can expect that ad-hoc constructions and specific assumptions for the spurions λ can have weaker bounds than the ones we will find in the rest of the section. Our procedure, on the contrary, is systematic and based on two well defined hypotheses: the composite sector respects some flavor symmetry $\mathcal{F}_c$ and some of the SM quarks belong to the strong sector. The amount of compositeness of the latter are measured by some parameters ε and we can check a posteriori if $\varepsilon \sim 1$ or if $\varepsilon \ll 1$, meaning that our hypothesis is inconsistent. Furthermore we take as starting point the Lagrangian of eq. (1.36), including one partner for each right-handed field and at most two partners for the SM doublets.

1.3.1 Right-handed compositeness

We first start from the scenario where the right-handed quarks are aligned with the composite sector. In sec. 1.2.2 we analyzed in detail how MFV can be realized in this hypothesis and why this model is only consistent with data for large values of m_*. In particular we saw that flavor transitions are enhanced by very composite right-handed top. At the same time composite top means composite up and charm which are very well tested and excluded by LHC measurements. The way out we want to explore is to separate the top from the two lighter families reducing the $U(3)_U \times U(3)_D$ flavor group to $U(2)_U \times U(1)_T \times U(3)_D$. In this section we show how this scenario, that we denote as $U(2) \times U(3)_{\rm RC}$, has enough freedom to have the right-handed top fully composite, maintaining at the same time mildly composite light quarks and not generating other dangerous (CP odd) effects as electric dipoles. A previous study of this scenario is

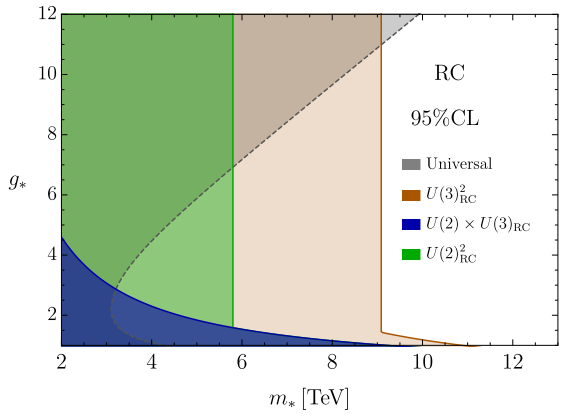


Figure 1.1: Summary plot for the three right-handed compositeness models under consideration.

presented in [27], that we update and extend in the following. Then later on we go beyond that and we try to separate also the heaviest family in the down sector. We will see that this scenario, where $\mathcal{F}_c = U(2)_U \times U(1)_T \times U(2)_D \times U(1)_B$, does not improve significantly over $U(2) \times U(3)_{\rm RC}$.

Our results are summarized in Fig. 1.1. We can see that MFV is excluded up $m_* \sim 9\,\text{TeV}$, independently of the other parameters of the model. This bound is obtained by combining the constraints from compositeness and $\Delta F = 2$ transitions in the B sector, while the tail at small values of g_* comes from the $\Delta F = 1$ leptonic B decays. The minimal m_* value is roughly a factor 3 larger than the minimum m_* allowed by universal constraints. The assumptions of $U(2)^2_{\rm RC}$ slightly improve the situation as shown by the green contour, that it is dominated by the $b \to s\gamma$ transition. Yet, among the models we studied in this section the mildest flavor-dependent constraints are found in the $U(2) \times U(3)_{\rm RC}$ scenario, where the most relevant bound comes from rare B decays. In some sense this model seems to have the ideal amount of symmetries necessary to avoid indirect searches; more or less symmetries will just increase the tension with the data.

$U(2) \times U(3)_{\mathbf{RC}}$

This model is characterized by the flavor group $\mathcal{F}_c = U(2)_U \times U(1)_T \times U(3)_D$ and is described by the following Lagrangian

$$\begin{aligned}\mathcal{L}_{\rm mix}^{U(3)\times U(2)_{RC}} =& \lambda_{q_u}^{iA}\bar{q}_L^i O_{q_u}^A + \lambda_{q_d}^{ia}\bar{q}_L^i O_{q_d}^a + \lambda_{q_t}^i \bar{q}_L^i O_{q_t} \\ &+ g_*\varepsilon_u \bar{O}_u^A u_R^i \delta_{Ai} + g_*\varepsilon_d \bar{O}_d^a d_R^i \delta_{ai} + g_*\varepsilon_t \bar{O}_t t_R + h.c.\,.\end{aligned} \tag{1.64}$$

where $A = 1, 2$ is the index of the $U(2)_U$ group, while $a = 1, 2, 3$ is the index of the $U(3)_D$ group. The mixing parameters λ and ε can be identified as spurions of the breaking of the flavor group $\mathcal{F}$, transforming as in the table in App. 1.A. While in the down sector

this scenario is completely analogous to $U(3)^2_{\rm RC}$, it differs in the up sector, having the top separated from the other two quarks forming a doublet of $U(2)_U$. A scheme of the symmetry breaking pattern for this and the other right-handed scenarios is reported in Fig. 1.4.

The Yukawa couplings for the up and down sector are written in terms of the mixings as

$$Y_u \sim (\lambda_{q_u}\varepsilon_u|\lambda_{q_t}\varepsilon_t)\,, \quad Y_d \sim \lambda_{q_d}\varepsilon_d \tag{1.65}$$

where the vertical bar means that the 3×2 matrix and the 3×1 vector should be stacked together to form a 3×3 matrix. We parametrize the mixings λ in the most general way that can reproduce the SM Yukawas. We only assume that $\lambda_{q_t} \gg \lambda_{q_u}$, which may be explained by RG effects similarly to anarchic Partial Compositeness. Under this hypothesis $\lambda_{q_t} \sim y_t$ whereas λ_{q_u} has the largest size of order y_c. In order to generate a Yukawa for the up-quark, we have to assume that one component is accidentally small and of the order of y_u. Using the flavor symmetries of the system, we can reduce them in the following way

$$\begin{aligned} U(2)\times U(3)_{\rm RC}: \qquad & \lambda_{q_u} = \frac{1}{\varepsilon_u}\begin{pmatrix} y_u & 0 \\ 0 & y_c \\ a\,y_c & b\,y_c \end{pmatrix}, \quad \lambda_{q_t} = \frac{1}{\varepsilon_t}\begin{pmatrix} 0 \\ 0 \\ y_t \end{pmatrix}, \\ & \lambda_{q_d} = \frac{1}{\varepsilon_d}\widetilde{U}_d \begin{pmatrix} y_d & 0 & 0 \\ 0 & y_s & 0 \\ 0 & 0 & y_b \end{pmatrix}. \end{aligned} \tag{1.66}$$

The coefficients a and b are assumed to be $\mathcal{O}(1)$ complex numbers with an arbitrary phase and the matrix $\widetilde{U}$ is a 3×3 unitary matrix. Notice that, by field redefinition we can remove all but one phase on $\widetilde{U}_d$ meaning that in this model there are in total three physical phases.

In this basis the Yukawa matrices are not diagonal and can be diagonalized with bi-unitary transformations to go to the mass basis. In particular the two matrices U_u and V_u that diagonalize Y_u can be computed explicitly in the limit $y_u \ll y_c \ll y_t$ as shown in app. 1.A. The matrix $\widetilde{U}_d$ is instead identified in terms of the CKM matrix as

$$V_{\rm CKM} = U_u^\dagger \widetilde{U}_d\,. \tag{1.67}$$

This model contains in total five additional real free parameters, namely ε_u, ε_d, ε_t, $|a| \sim 1$ and $|b| \sim 1$, and two additional CP violating phases, $\arg[a]$ and $\arg[b]$, compared to the SM. Therefore, even if we assume a CP even strong sector, there are two additional phases giving rise to new potentially measurable effects in electric dipole or CP odd flavor transition. Moreover, notice that, from the Lagrangian in eq. (1.64) it is possible to recover the $U(3)^2_{\rm RC}$ model simply by setting $a = b = 0$ and $\varepsilon_t = \varepsilon_u$. This point is also summarized in Fig. 1.4.

We now inspect the bounds on this scenario in details.

Electric dipole cancellation

As we mentioned before, the new phases in the mixing of eq. (1.64) can in general produce new observables effects. A remarkable property of all the models under exam in this section, is the absence of electric dipole moments at leading order in the spurion insertions. In the specific case of $U(2) \times U(3)_{\rm RC}$, moreover, non-zero electric dipoles are generate at 1-loop only for the up sector and they happen to be extremely suppressed, while contributions from the down quarks come at at least 2-loops.

Contributions to the electric and chromoelectric dipole moments of the neutron are induced by the imaginary part of the Wilson coefficients $C_{ff\gamma}$ and C_{ffg} of the following operators

$$\mathcal{O}_{ff\gamma} = e\bar{f}_L \sigma^{\mu\nu} \gamma^5 f_R F_{\mu\nu} \,, \qquad\qquad \mathcal{O}_{ffg} = g_s \bar{f}_L \sigma^{\mu\nu} \gamma^5 f_R G_{\mu\nu} \,. \tag{1.68}$$

In MFV, these are obviously aligned with the SM Yukawas at tree-level and so they are real and diagonal in the mass basis. In the case of $U(2) \times U(3)_{\rm RC}$ the same is clearly true for the down sector, which is completely equivalent to MFV, but the cancellation of dipoles in the up sector is less trivial and requires a specific discussion.

Let start from the spurions in eq. (1.64), the flavour structure of the Wilson coefficient of the operators in eq. (1.68) is

$$[C_{uu\gamma}]^{ij} \propto [C_{uug}]^{ij} \propto (\lambda_{q_u} \varepsilon_u | x \lambda_{q_t} \varepsilon_t)^{ij} \,, \tag{1.69}$$

where x is an $\mathcal{O}(1)$ real number with $x-1$ measuring the departure from the SM Yukawas, and i, j denote the SM flavor. To prove the absence of electric dipoles we must show that the coefficients in eq. (1.69), once rotated in the mass basis, have real diagonal entries $i = j$. A way to see it is to start from the explicit form for the spurions in eq. (1.66) and remove all the phases from λ_{q_u}, λ_{q_t}, ε_u and ε_t. This can be done, for instance, rotating the SM fields, in the Lagrangian of eq. (1.64), as

$$u_R^1 \to u_R^1 \, e^{-i\arg[a]} \,, \qquad q_L^1 \to q_L^1 \, e^{-i\arg[a]} \,, \quad u_R^2 \to u_R^2 \, e^{-i\arg[b]} \,, \qquad q_L^2 \to q_L^2 \, e^{-i\arg[b]} \,. \tag{1.70}$$

This field rotation removes the phases from the two parameters a and b and moves them to the down sector. In this way the new up Yukawa in eq. (1.65) is now free of complex entries and can be diagonalized by orthogonal matrices only. As a consequence the dipole Wilson coefficient in eq. (1.69) is build out just starting from real objects and it is clearly real itself. Moreover any additional rotation of the fields which keeps the Yukawas real and diagonal in the mass basis, cannot modify the diagonal entries of eq. (1.69) resulting

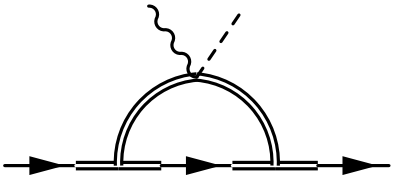


Figure 1.2: A 1-loop contribution to the neutron EDM in our theory. The double line represents states in the composite dynamics, while the single lines represent elementary fermions.

in no-tree level electric dipoles.

There is another way to prove the absence of electric dipoles at tree level. Starting again from eq. (1.69) we can see that, at tree level, $[C_{uu\gamma/g}]^{ij}$ is built out of $\lambda_{u,t}$ and $\epsilon_{u/t}$. Since electric dipoles are observables, and so basis independent, we can search for flavor invariants constructed as combinations of $\lambda_{u,t}$ and $\epsilon_{u/t}$. However, a basis for such invariants is given by

$$\lambda_{q_t}^\dagger \lambda_{q_t}\,, \quad \mathrm{Tr}[\lambda_{q_u}\lambda_{q_u}^\dagger]\,, \quad \mathrm{Tr}[(\lambda_{q_u}\lambda_{q_u}^\dagger)^2]\,, \quad \lambda_{q_t}^\dagger(\lambda_{q_u}\lambda_{q_u}^\dagger)\lambda_{q_t}\,, \quad \lambda_{q_t}^\dagger(\lambda_{q_u}\lambda_{q_u}^\dagger)^2\lambda_{q_t}\,, \quad |\varepsilon_t|^2\,, \quad |\varepsilon_u|^2\,, \tag{1.71}$$

and it is manifestly real. Therefore, the only way to have a complex invariant is to include also the λ_{q_d} spurion, which can only occur at next to leading order.

In the $U(2)\times U(3)_{\mathrm{RC}}$ model the first diagonal imaginary contribution for the up-sector requires at least one loop of the elementary fields, generated via a process similar to Fig. 1.2. The Wilson coefficient of eq. (1.68) is given by

$$[C_{uu\gamma}]^{ij} = \frac{1}{16\pi^2}\frac{v}{\sqrt{2}}\left[\lambda_{q_d}\lambda_{q_d}^\dagger(\lambda_{q_u}\varepsilon_u|x\lambda_{q_t}\varepsilon_t)\right]^{ij} \tag{1.72}$$

giving the following leading contribution to the electric dipole moment

$$d_n \sim e\frac{y_b^2}{16\pi^2}\frac{\Re(a)A\eta\lambda^3}{m_*^2\varepsilon_d^2}m_c \implies m_* > \frac{0.05}{\varepsilon_d}\ \mathrm{TeV}\,. \tag{1.73}$$

The latter is of the same order of the bound on Zb_Lb_L from LEP measurements, reported in Tab. 1.1 and it is clearly under control for $\varepsilon_d \gtrsim 0.05$. Notice that in the $a \to 0$ limit the previous contributions goes to zero.

We also remark that 1-loop contributions to down quarks electric dipole moments are zero, as can be seen by a straightforward calculation. Indeed the contribution reads

$$[C_{dd\gamma}]^{ij} = \frac{1}{16\pi^2}\frac{v}{\sqrt{2}}\left[(\lambda_{q_u}\lambda_{q_u}^\dagger + x_1\lambda_{q_t}\lambda_{q_t}^\dagger + x_2\lambda_{q_d}\lambda_{q_d}^\dagger)\lambda_{q_d}\right]^{ij}\,, \tag{1.74}$$

with x_1 and x_2 some $O(1)$ real numbers. It is clear that the term proportional to x_2 is real and diagonal in the mass basis. The first two terms in the mass basis instead read

$$(\lambda_{q_u}\lambda_{q_u}^\dagger + x_1\lambda_{q_t}\lambda_{q_t}^\dagger)\lambda_{q_d}^\dagger \to V_{\rm CKM}^\dagger \left[\frac{\bar{Y}_u^2}{\varepsilon_u^2} + \left(\frac{x_1}{\varepsilon_t^2} + \frac{1}{\varepsilon_u^2}\right) U_u^\dagger \widetilde{Y}_t^2 U_u\right] V_{\rm CKM} \frac{\bar{Y}_d}{\varepsilon_d} \tag{1.75}$$

with the notation of App. 1.A. Since this contribution is written as the product of a Hermitean matrix and a real and diagonal matrix, the diagonal entries are real as well, proving that also at 1-loop there is no contribution to dipoles.

Constraints

We just showed that the $U(2) \times U(3)_{\rm RC}$ scenario is safe from electric dipoles. Moreover, since the down sector is aligned with the Yukawas, this model is also protected from dipole operators mediating $b \to s\gamma$ transitions, which usually give rise to strong bounds. However, this model has a less trivial flavor structure than MFV and the less amount of symmetries in the strong sector open the way to new potentially measurable effects. All the constrains on the model are detailed in sec. 1.4, in the following we just discuss the most important signatures of the model.

The strongest constraints on $U(2) \times U(3)_{\rm RC}$ come from the same observables as $U(3)^2_{\rm RC}$. The bound from light fermion compositeness clearly coincides with eq. (1.45), i.e. we have

$$m_* \gtrsim 12.7\, g_* \varepsilon_u^2\, \text{TeV}\,, \qquad\qquad m_* \gtrsim 5.8\, g_* \varepsilon_d^2\, \text{TeV}\,. \tag{1.76}$$

The bounds from $\Delta F = 2$ and $\Delta F = 1$ transitions in the B sector of eqs. (1.55) and (1.51) are instead now enhanced by a small ε_t. They are respectively

$$m_* \gtrsim \frac{6.5}{g_* \varepsilon_t^2}\,, \tag{1.77}$$

and

$$m_* \gtrsim \frac{12.3}{\varepsilon_t}\,\text{TeV}\,,\ (\text{w/o } P_{LR}) \qquad\qquad m_* \gtrsim \frac{9.2}{g_* \varepsilon_t}\,\text{TeV}(\text{w/ } P_{LR})\,. \tag{1.78}$$

It is clear that in this model we can have a small value for ε_u and ε_d in order to minimize the bounds from compositeness, still keeping $\varepsilon_t \sim 1$ to minimize the flavor bounds. The ideal scenario has a maximal top compositeness $\varepsilon_t = 1$ and the P_{LR} protection, in which case we find $m_* \gtrsim 10.8\, \varepsilon_u$ TeV. An order one value $\varepsilon_u \sim 0.3$ is perfectly compatible with the assumption of composite up-type quarks and would imply the relatively weak bound $m_* \gtrsim 3$ TeV. Much smaller values of ε_u would render the compositeness assumption less plausible. In addition if $\varepsilon_u \lesssim 0.04\, \varepsilon_t$ stronger bounds than the ones considered here would arise from s-to-d $\Delta F = 2$ transitions (see Tab. 1.4).

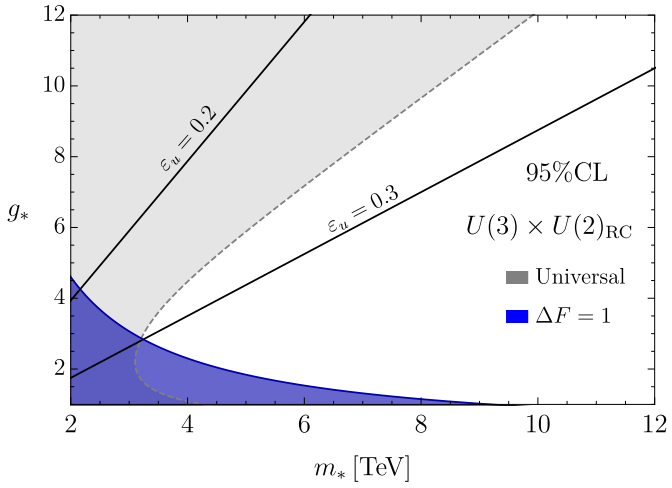


Figure 1.3: Current constraints on $U(2) \times U(3)_{\rm RC}$ model at 95% CL. The blue line is obtained setting $\varepsilon_t = 1$. The black lines are the bounds from up quark compositeness for two different values of ε_u.

Apart from what we just mentioned, no other large effects are generated in this model. The strongest bound comes from eq. (1.78) that, in presence of P_{LR} protection is relatively under control. Moreover, the latter, has to be compared with the universal signatures discussed in sec. 1.4.1, as shown by the summary plot in Fig. 1.3.

$U(2)^2_{\mathbf{RC}}$

We would like now to see if anything is gained by considering a smaller strong flavor group $\mathcal{F}_c$, studying the case where both the right-handed top and bottom possess only an abelian flavor symmetry. The step of separating the bottom is not motivated by the bounds previously discussed, which will not be relaxed. Considering this model is rather aimed to see how much freedom we can leave to the strong sector before it results in unwanted flavor signatures.

The scenario we consider now has so $\mathcal{F}_c = U(2)_U \times U(1)_T \times U(2)_D \times U(1)_B$ that is then broken by the right-handed mixings to $U(2)_{u+U} \times U(1)_{t+T} \times U(2)_{d+D} \times U(1)_{b+B}$. The mixing Lagrangian reads

$$\begin{aligned}\mathcal{L}_{\rm mix}^{U(2)^2_{RC}} =& \lambda_{q_u}^{iA} \bar{q}_L^i O_{q_u}^A + \lambda_{q_d}^{iB} \bar{q}_L^i O_{q_d}^B + \lambda_{q_t}^i \bar{q}_L^i O_{q_t} + \lambda_{q_b}^i \bar{q}_L^i O_{q_b} \\ &+ g_* \varepsilon_u \bar{O}_u^A u_R^i \delta_{Ai} + g_* \varepsilon_d \bar{O}_d^B d_R^i \delta_{Bi} + g_* \varepsilon_t t_R \bar{O}_t + g_* \varepsilon_b b_R \bar{O}_b + h.c.\,,\end{aligned} \tag{1.79}$$

where the indices $A, B = 1, 2$ transform respectively under $U(2)_U$ and $U(2)_D$. Analogously to the $U(2) \times U(3)_{\rm RC}$ scenario the Yukawas take the following form

$$Y_u \sim (\lambda_{q_u} \varepsilon_u | \lambda_{q_t} \varepsilon_t)\,, \quad Y_d \sim (\lambda_{q_d} \varepsilon_d | \lambda_{q_b} \varepsilon_b)\,, \tag{1.80}$$

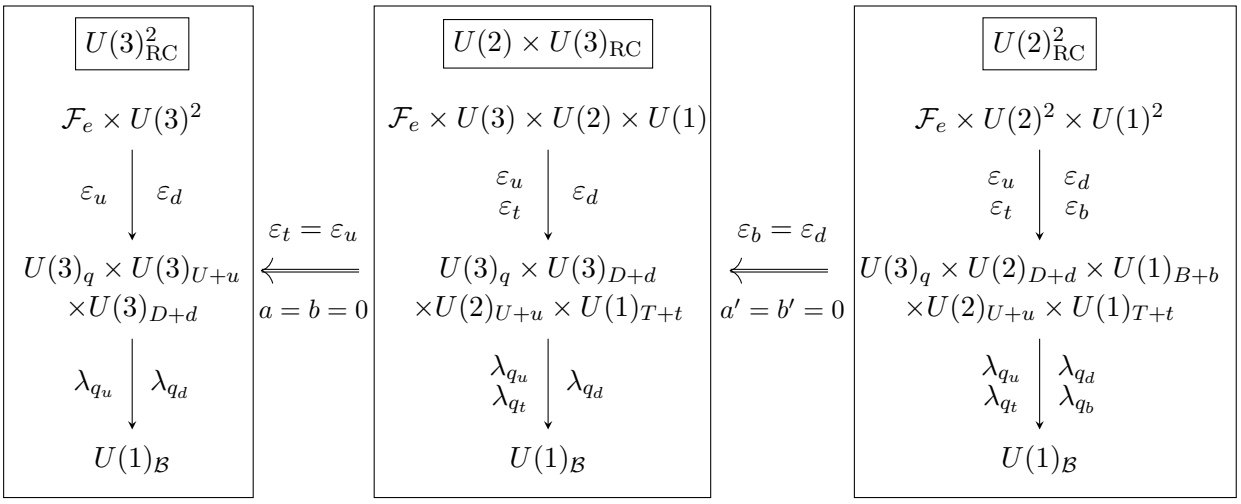


Figure 1.4: Scheme of the various right-handed compositeness models under consideration. Starting from the less symmetric $U(2)^2_{\rm RC}$ on the right, it is possible to recover $U(3)^2_{\rm RC}$ switching off some of the parameters.

and we can parameterize the mixings λ as

$$U(2)^2_{\rm RC}: \quad \begin{aligned} \lambda_{q_u} &= \frac{1}{\varepsilon_u}\begin{pmatrix} y_u & 0 \\ 0 & y_c \\ a\, y_c & b\, y_c \end{pmatrix}, & \lambda_{q_t} &= \frac{1}{\varepsilon_t}\begin{pmatrix} 0 \\ 0 \\ y_t \end{pmatrix}, \\ \lambda_{q_d} &= \frac{1}{\varepsilon_d}\widetilde{U}_d\begin{pmatrix} y_d & 0 \\ 0 & y_s \\ a'\, y_s & b'\, y_s \end{pmatrix}, & \lambda_{q_b} &= \widetilde{U}_d\frac{1}{\varepsilon_b}\begin{pmatrix} 0 \\ 0 \\ y_b \end{pmatrix}, \end{aligned} \tag{1.81}$$

with $\widetilde{U}_d$ a 3×3 unitary matrix.[5] Notice that, even in this model, we can remove all the phases but one from $\widetilde{U}_d$, with a total of 5 phases in eq. (1.81). The matrices that diagonalize the Yukawas are once again reported in app. 1.A and $\widetilde{U}_d$ is rewritten in terms of the CKM as

$$V_{\rm CKM} = U_u^\dagger \widetilde{U}_d U_d\,. \tag{1.82}$$

Similarly to before, for this model we assume the parameters a, b, a' and b' to be $O(1)$ with arbitrary phases. The model and its relation to the other Right-handed compositeness scenarios is summarazied in Fig. 1.4.

[5]We only need a single $\widetilde{U}_d$ to reach this form: we start from a 3×3 unitary matrix that aligns λ_b in the third direction and then we multiply it with a 2×2 block diagonal unitary matrix that leaves λ_b invariant but allow to diagonalize the top 2×2 block of λ_d together with a $U(2)_D$ matrix on the right.

Constraints

This model inherits many features of the $U(2) \times U(3)_{\rm RC}$ case just examined. The most important aspect is the absence of the electric dipoles at tree-level. The proof simply follows from the arguments presented for the $U(2) \times U(3)_{\rm RC}$. In that model we showed that one cannot build physical observables using only spurion in the up-sector, the same arguments trivially extends also to the $U(2)^2_{\rm RC}$ case for both the down and up sectors. Moreover, notice that, differently from $U(2) \times U(3)_{\rm RC}$ now we have one-loop dipoles for both up and down quarks. The latter are totally analogous to eq. (1.73) and are negligible.

The compositeness bound of eq. (1.76) and the flavor bound of eqs. (1.78) and (1.77) apply also here. In this model however we also generate additional flavor transitions. The first one is the $\Delta F = 2$ transition for right handed quarks. These bounds come from the operator $\widetilde{\mathcal{O}}_1$ of eq. (1.124). The complete list of bounds can be found in sec. 1.4, here we report only the strongest one, related to B_d system and we get

$$m_* \gtrsim 19.3\, g_* \varepsilon_b^2 \,\text{TeV}\,. \tag{1.83}$$

Additional effects are also present in $\Delta F = 1$ (see sec. 1.4.2 for details), in fact, contrary to $U(3)^2_{\rm RC}$ and $U(2) \times U(3)_{\rm RC}$ the model $U(2)^2_{\rm RC}$ has a less trivial structure in the down sector and generate a non-suppressed C'_{10} in eq. (1.48)

$$C'_{10}\big|_{U(2)^2_{\rm RC}} = \frac{\sqrt{2}}{G_F}\frac{4\pi^2}{e^2}\frac{1}{V^{tb}_{\rm CKM}(V^{ts}_{\rm CKM})^*}\frac{g_*^2 y_s(\varepsilon_b^2 - \varepsilon_d^2)}{m_*^2 y_b} \sim 23.9\frac{g_*^2(\varepsilon_b^2 - \varepsilon_d^2)}{m_*^2}\,. \tag{1.84}$$

Still from [24] we can read the bound on C'_{10}, which implies the following disequality on m_*

$$|C'_{10}| \lesssim 0.2 \implies m_* \gtrsim 10.9\, g_* \varepsilon_b \text{ TeV}\,. \tag{1.85}$$

The previous equation clearly tells that a small ε_b is preferred. Moreover there are two effects that are enhanced by a small ε_b. The first is the anomalous Zb_Lb_L coupling (see Tab. 1.1), the second is the 1-loop contribution to neutron dipole moment. The latter can be simply red from eq. (1.73) replacing $\varepsilon_d \to \varepsilon_b$. Combining eq. (1.85) with the bound from Zb_Lb_L anomalous coupling in Tab. 1.1 we find

$$m_* \gtrsim 0.9\,\sqrt{g_*}\,\text{TeV}\,, \tag{1.86}$$

obtained marginalizing on ε_b. This effect is less relevant than universal constraints on C_H, as described in sec. 1.103. The lowest value of m_* outlined in the previous formula is obtained for $\varepsilon_b = 0.086/\sqrt{g_*}$. However, values of ε_b much smaller than unity are incompatible with the hypothesis of b-quark compositeness. To defend this hypothesis

much larger values of ε_b are necessary, and as a result a stronger constraint on m_* follows.

The $U(2)^2_{\rm RC}$ model also generates another $\Delta F = 1$ effect, the $B \to X_s\gamma$ transitions, via the so-called EW dipoles. These interactions are described according to the following effective operators

$$\bar{f}^i_{L(R)} \sigma_{\mu\nu} f^j_{R(L)} F^{\mu\nu}_{SM} \,, \tag{1.87}$$

with $F^{\mu\nu}_{SM}$ the field strength of the SM gauge bosons and i, j the flavor indexes of the fermions.

The strongest constraints on these kind of interactions come from the radiative decay of B mesons. The effective Hamiltonian relevant for $b \to s\,\gamma$ transition is

$$\mathcal{H}_{\rm eff} = -\frac{4G_F}{\sqrt{2}} V^{tb}_{\rm CKM}(V^{ts}_{\rm CKM})^* \frac{m_b e}{16\pi^2} F^{\mu\nu} \left(C_7 \bar{s}_L \sigma_{\mu\nu} b_R + C_7' \bar{s}_R \sigma_{\mu\nu} b_L \right) \,, \tag{1.88}$$

and the new physics contribution to the previous Wilson coefficient is constrained to be [28] roughly

$$\mathrm{Re}\, C_7 \lesssim 0.03 \,, \qquad\qquad |C_7'| \lesssim 0.07 \tag{1.89}$$

at 95% CL. In our model we estimate

$$C_7'|_{U(2)^2_{\rm RC}} \sim \frac{m_s}{m_b} \frac{b' 4\sqrt{2}\pi^2}{G_F V^{tb}_{\rm CKM}(V^{ts}_{\rm CKM})^* m_*^2} \sim 2.4 \left(\frac{1\,\mathrm{TeV}}{m_*^2} \right)^2 \,, \tag{1.90}$$

while C_7 is suppressed by a further power or m_s/m_b. This translates into the much stronger bound

$$m_* \gtrsim 5.8\,\mathrm{TeV} \,. \tag{1.91}$$

Combining all these bounds the largest possible value for ε_b is roughly $\varepsilon_b \lesssim 0.5/g_*$. The constraints on ε_d are much milder and read $\varepsilon_d \lesssim 1/\sqrt{g_*}$. Overall, scenarios with composite d-type quarks seem to require $g_* \sim 1$.

To conclude, separating the third family of the down-sector does not seem to offer a significant improvement. Compared to the $U(2) \times U(3)_{\rm RC}$ model there are important new signatures. The $\Delta F = 2$ transition and the new correction to C_{10}' may be suppressed by taking ε_b small. However, the bound in eq. (1.91) gives a robust absolute lower bound on m_* independent on the other model parameters.

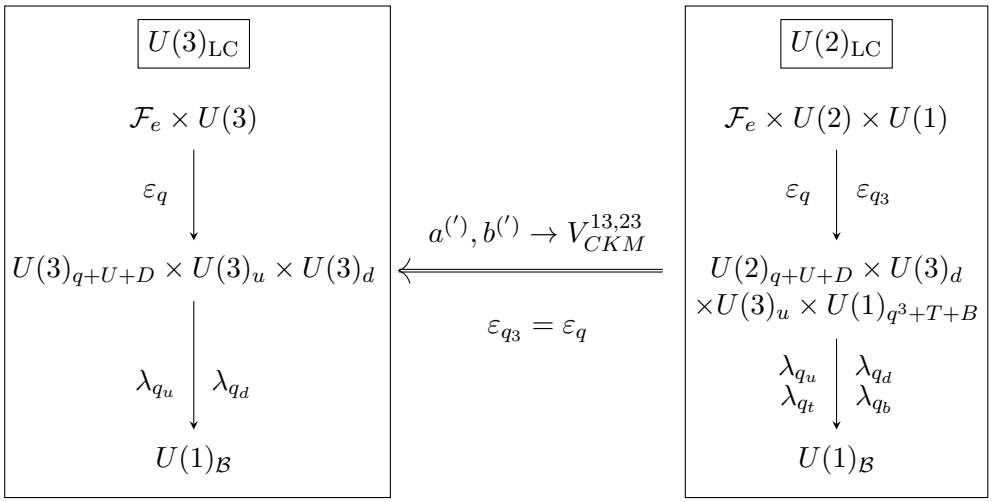


Figure 1.5: Same as Fig. 1.4 for the left-handed compositeness case. We are assuming only one mixing for the SM doublets. There is not an easy way to reproduce MFV by fixing the values of the four parameters $a^{(\prime)}$ $b^{(\prime)}$. In this limit they will be expressed as some functions of the CKM angles and phase.

1.3.2 Left-handed compositeness

We now move on and we discuss the left-compositeness models. Our starting point is the $U(3)_{\rm LC}$ model of sec. 1.2.2, where for concreteness we assume only one mixing for each SM doublet. We have seen that MFV model requires at least $m_* \gtrsim 8\,\text{TeV}$, due to the tension with the precise measurement on the CKM matrix unitarity, see eq. (1.60). We now try to relax this constrain assuming a smaller $\mathcal{F}_c$.

$U(2)_{\mathbf{LC}}$

Following the same logic we adopted so far, we can start from $U(3)_{\rm LC}$ and then separate the third family reducing the strong sector flavor symmetries. More explicitly, we take the left-handed mixing to break the composite group $\mathcal{F}_c = U(2)_{U+D} \times U(1)_{T+B}$ and the elementary flavor group to the diagonal $U(2)_{q+U+D} \times U(1)_{q^3+T+B}$, according to the following Lagrangian

$$\mathcal{L}_{\rm mix}^{U(2)_{LC}} = g_* \varepsilon_q \bar{q}_L^i O_q^A \delta_{iA} + g_* \varepsilon_{q_3} \bar{q}_L^3 O_{q_3} + \lambda_u^{Ai} \bar{O}_u^A u_R^i + \lambda_t^i \bar{O}_t u_R^i + \lambda_d^{Ai} \bar{O}_d^A d_R^i + \lambda_b^i \bar{O}_b d_R^i + h.c.\,, \tag{1.92}$$

where the index $A = 1, 2$ is the index of $U(2)_{U+D}$. The SM Yukawas are then obtained as

$$Y_u \sim \left(\frac{\lambda_u \varepsilon_q}{\lambda_t \varepsilon_{q_3}} \right), \quad Y_d \sim \left(\frac{\lambda_d \varepsilon_q}{\lambda_b \varepsilon_{q_3}} \right), \tag{1.93}$$

where this time the 2×3 and 1×3 blocks are stacked on top of each other to form the two 3×3 matrices. Similarly to before, we can use the flavor symmetries to reduce the λ mixings in the following form

$$U(2)_{\rm LC}: \quad \begin{aligned} \lambda_u &= \frac{1}{\varepsilon_q}\begin{pmatrix} y_u & 0 & a\,y_c \\ 0 & y_c & b\,y_c \end{pmatrix}, \quad \lambda_t = \frac{1}{\varepsilon_{q_3}}\begin{pmatrix} 0 & 0 & y_t \end{pmatrix}, \\ \lambda_d &= \frac{1}{\varepsilon_q}\widetilde{U}_d^{(2)}\begin{pmatrix} y_d & 0 & a'\,y_s \\ 0 & y_s & b'\,y_s \end{pmatrix}, \quad \lambda_b = \frac{1}{\varepsilon_{q_3}}\begin{pmatrix} 0 & 0 & y_b \end{pmatrix}, \end{aligned} \tag{1.94}$$

where a, b, a' and b' are complex numbers with arbitrary phases and $\widetilde{U}_d^{(2)}$ is a 2×2 orthogonal matrix, since it is always possible to remove phases from $\widetilde{U}_d^{(2)}$ by rephasing the elementary quarks.

The matrices U_u, V_u, U_d and V_d that diagonalize the two Yukawas are reported in app. 1.A and the CKM matrix is readily obtained as

$$V_{\rm CKM} = U_u^\dagger \widetilde{U}_d U_d\,, \qquad\qquad \widetilde{U}_d \equiv \begin{pmatrix} \widetilde{U}_d^{(2)} & 0 \\ 0 & 1 \end{pmatrix}. \tag{1.95}$$

An important difference with the previous model is that $\widetilde{U}_d$ is now determined in terms of a 2×2 orthogonal matrix and thus contains only a single real parameter. This means that three combinations of a, b, a' and b' are determined as function of two remaining CKM angles and the CKM phases Thus, among a, b, a' and b' there are only 3 phases and 2 real parameters. Given the form of U_u and U_d, we can see that for a generic 2×2 rotation matrix of angle θ, $V_{\rm CKM}$ is roughly

$$V_{\rm CKM} \sim \begin{pmatrix} \cos\theta & \sin\theta & a\frac{y_c}{y_t} + a'\frac{y_s}{y_b} \\ \sin\theta & \cos\theta & b\frac{y_c}{y_t} + b'\frac{y_s}{y_b} \\ a\frac{y_c}{y_t} + a'\frac{y_s}{y_b} & b\frac{y_c}{y_t} + b'\frac{y_s}{y_b} & 1 \end{pmatrix}. \tag{1.96}$$

For all $O(1)$ coefficients, the $(1,3)$ and $(3,1)$ components are one order of magnitude larger than the SM measured value because of the ratio $y_s/y_b \sim 0.02$. Thus, to reproduce the CKM mixing angles we assume $|a| \sim |b| \sim |b'| \sim 1$, $\theta \sim \lambda$ and $|a'| \sim 0.1$.

From the previous discussion it emerges that there is not an easy way to recover minimal flavor violation just setting to zero some of the parameters. In general it is necessary to set $\varepsilon_q = \varepsilon_{q_3}$ and to fix a, b, a' and b' as functions of the various CKM entries. A cartoon summarizing the model can be found in Fig. 1.5.

Constraints

The spurion structure of this model is analogous to the one of $U(2)^2_{\rm RC}$. Again this scenario is free from dipole moment at leading order in the mixings λ. The proof follows

from the same argument presented for $U(2) \times U(3)_{\rm RC}$. Interestingly this model also does not generate dipoles even at 1-loop. In fact, even at this order it is impossible to build a combination of spurions that contributes to the dipoles and involves both up- and down-type mixings. As we explained this is a necessary requirement to access one of the physical phases of out model.

A first constraint to this model comes from the compositeness bound of eq. (1.59) that gives

$$m_* \gtrsim 14.2\, g_* \varepsilon_q^2\, \text{TeV}\,, \tag{1.97}$$

which can however be satisfied by taking a small value for ε_q. The model however faces several constraints from flavor-violating processes, in particular a strong bound comes from the $b \to s\gamma$ transitions induced by operators as in eq. (1.87). This effect, in analogy to $U(2)^2_{\rm RC}$, gives

$$C_7|_{U(2)^2_{\rm LC}} \sim \frac{m_s}{m_b} \frac{4\sqrt{2}\pi^2 b'}{G_F V^{tb}_{\rm CKM} (V^{ts}_{\rm CKM})^* m_*^2} \sim 2.4 \left(\frac{1\,\text{TeV}}{m_*^2}\right)^2\,, \tag{1.98}$$

while C_7' is suppressed by a further power or m_s/m_b. This translates into an absolute lower bound on m_* given by

$$m_* \gtrsim 8.8\,\text{TeV}\,. \tag{1.99}$$

Moreover, the (semi-)leptonic B_s decays (see sec. 1.4.2) are also enhanced in this model, giving

$$C_{10}|_{U(2)_{\rm LC}} = \frac{\sqrt{2}}{G_F} \frac{4\pi^2}{e^2} \frac{g_*^2(\varepsilon_{q_3}^2 - \varepsilon_q^2)}{m_*^2} \sim 48.7\, \frac{g_*^2 \varepsilon_{q_3}^2}{m_*^2}\,, \tag{1.100}$$

while the operator C_{10}' is generated with a negligible coefficient. Again the C_{10} coefficient can be suppressed in presence of P_{LR} protections. Imposing the experimental bound we find

$$m_* \gtrsim 14.2\, g_* \varepsilon_{q_3}\,\text{TeV}\,,\ (\text{w/o } P_{LR}) \qquad m_* \gtrsim 10.6\, \varepsilon_{q_3}\,\text{TeV}(\text{w/ } P_{LR})\,, \tag{1.101}$$

pushing toward the small ε_{q_3} direction.

From the previous discussion and from the various bounds reported in sec. 1.4 we find that the strongest constraint is given by eq. (1.99), assuming P_{LR} protection. In fact for each value of m_* above 8.8 TeV we can always find values of ε_{q_3} compatible. Concretely, the strongest constraint on ε_{q_3} comes from the right-handed top coupling in table 1.1.

To conclude, in this section we have found that both $U(3)_{\rm LC}$ and $U(2)_{\rm LC}$ not allowed to have m_* in the few TeV range. While MFV is severely constrained by the CKM unitarity, the less symmetric $U(2)_{\rm LC}$ model suffers from large corrections to flavor-violating dipole operators.

1.3.3 Other possibilities

In the previous sections we just considered a minimalist set of scenarios of partial compositieness and we need to comment on few more possibilities.

First of all, we remark that our logic was to start from the MFV hypothesis and then lower down the symmetries of the strong sector. We have seen in sec. 1.2.2 that 5 realization of MFV, (summarized in eqs. (1.38, 1.39, 1.40, 1.41, 1.42)), are possible and we considered explicitely how to improve the ones of eq. (1.38) and eq. (1.40), respectively in sec. 1.3.1 and sec. 1.3.2. Nevertheless also the other scenarios might be improved by the right choice of $\mathcal{F}_c$ and they deserve further investigations. For instance, given that the strongest bounds are related to the up sector, we expect that the model in eq. (1.42) can be improved following the same pattern we specified for eq. (1.38) resulting in similar constraints.

A more careful discussion it is needed for the models in eqs. (1.41,1.39). In both of them the strongest constraints are related to the unitarity of the CKM put under stress by a large and equal compositeness of all the light families $\varepsilon_{q_u} \gtrsim y_t/g_*$. This can be relaxed downgrading $U(3)_U \to U(2)_U \times U(1)_T$ and considering a smaller ε_{q_u} to avoid large LEP constrains. At the same time both of these scenarios, differently with $U(2)_{\rm LC}$ discussed before, are expected not to produce dangerous effects in the down sector, being completely aligned with the SM Yukawas. Yet, we still have to quantify the values of the different ε that makes this model compatibles with the various bound.
Moreover, we have to remark that in partial compositeness there can be different realization from the ones we have studied only relying in a small set of hypotheses. For instance one can start from the single mixing Lagrangian in eq. (1.22) and assume an $U(3)$ symmetry in the strong sector and $\lambda_q \propto \lambda_d$ [29].
Another possibility has been considered in [30, 13] and relies on an approximated $U(2)^2$ symmetry. We still need a detailed analysis of this scenarios and a quantitative comparison with the models we presented. We just mention that the minimal model considered in [30] has a structure very similar to our $U(2)^2_{\rm RC}$ model but on the contrary needs stronger assumptions, namely the absence of some spurions that might cause additional effects. Moreover the $U(2)^3$ flavor symmetry cannot be realized according to our set of hypothesis where the maximal amount of symmetry of the strong sector is $U(3)^2$ and, up to the mixings, the SM quarks always respect $U(3)^3$.

1.4 Experimental constraints

In this section we summarize and extend the analysis performed so far on the various bounds for the different models under consideration. We remark that this section partially overlap with the previous discussion, moreover it aims to collect all the relevant constraints.

In order to make the discussion more systematic we define the tree-level flavor structures $[S]^{ij}$ as the coefficient that involves the smallest number of mixing λ and ε for a given flavor-dependent operators. In the rest of the section we often refer to these tree-level flavor structures, whose explicit expression can be found in App. 1.A.

1.4.1 Universal constraints

The first class of effects we consider are the universal ones [31], where all the new interactions among the SM fields can be described through higher dimensional operators involving only SM bosons.

In composite Higgs models we can parameterize the leading effects according to the following Lagrangian [11]

$$\begin{aligned}\mathcal{L}_{SILH} = &\frac{c_H}{2}\left(\partial_\mu (H^\dagger H)\right)^2 + \frac{c_T}{2}\left(H^\dagger \overleftrightarrow{D}_\mu H\right)\left(H^\dagger \overleftrightarrow{D}^\mu H\right) + ig\frac{c_W}{2}\left(H^\dagger \overleftrightarrow{D}^a_\mu H\right) D_\nu W^{a\,\mu\nu} \\ &+ig\frac{c_B}{2}\left(H^\dagger \overleftrightarrow{D}_\mu H\right)\partial_\nu B^{\mu\nu} - g^2\frac{c_{2W}}{2}(D^\mu W^a_{\mu\nu})(D_\rho W^{a\rho\nu}) - g'^2\frac{c_{2B}}{2}(\partial^\mu B_{\mu\nu})(\partial_\rho B^{\rho\nu})\,,\end{aligned} \tag{1.102}$$

where the size of the Wilson coefficient, as function on the new physics parameters, can estimated according to the so-called "SILH" power-counting. In particular, apart from c_T, we expect the other Wilson coefficients in eq. (1.102) to be general for composite Higgs models, i.e. to not depend on the detailed flavor structure, and to scale as

$$c_H \sim \frac{g_*^2}{m_*^2}\,, \qquad c_W \sim c_B \sim \frac{1}{m_*^2}\,, \qquad c_{2W} \sim c_{2B} \sim \frac{1}{m_*^2 g_*^2}\,, \tag{1.103}$$

as functions of g_* and m_*.

The operator c_W and c_B can be related to the S parameter[32] which has been constrained by EWPT (see for instance [33]). The operator c_{2W} and c_{2B} are very well measured at (HL)-LHC. W is expected to be stronger than Y and we can already read the bound on W from recent run-2 results [34]. We extract the value for c_H from [35].We report the results for the current constraints in the left panel of Fig. 1.6. In the right panel we report the projections for HL-LHC. The bounds are obtained from [36, 37] and we updated the W constraint according to the projections of [38].[6]

The size of c_T, on the contrary, is less generic and strongly depends on the specific assumptions on the strong dynamics. Moreover the c_T coefficient, usually related to the $\hat{T}$ oblique parameter, is protected by the custodial symmetry that we assume in all our discussion. The main sources of custodial breaking, which generate a non-zero $\hat{T}$, are the

[6] Notice that the estimate in [38] assumes conservative hypothesis for the systematic uncertainties, given the recent measurements at 100 fb^{-1} in [34].

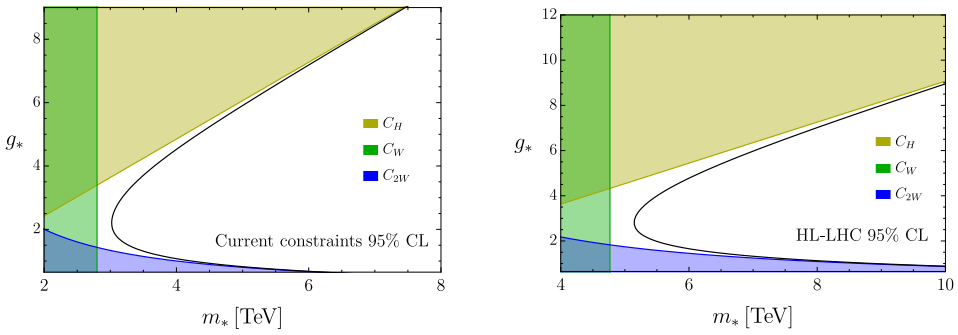


Figure 1.6: Current and future universal constraints on CH-models at 95% CL. See the main text for details.

exchanges of composite fermions, in the case the latter are not singlet under $SU(2)_R$.[7] These contributions have been computed in [13] for a defined fermion content. In the various scenarios we consider we don't find them competitive with the other constraints.

1.4.2 Anomalous couplings to SM gauge bosons

The first class of non universal effects we consider are the corrections to SM vector bosons couplings. We focus on the following set of operators

$$\begin{aligned} \left[\mathcal{O}_{Hq}^{(1)}\right]^{ij} &= \left(H^\dagger i\overleftrightarrow{D}_\mu H\right)\bar{q}_L^i\gamma^\mu q_L^j\,, \quad \left[\mathcal{O}_{Hq}^{(3)}\right]^{ij} = \left(H^\dagger i\overleftrightarrow{D}_\mu \tau^a H\right)\bar{q}_L^i\gamma^\mu\tau^a q_L^j\,, \\ \left[\mathcal{O}_{Hu}\right]^{ij} &= \left(H^\dagger i\overleftrightarrow{D}_\mu H\right)\bar{u}_R^i\gamma^\mu u_R^j\,, \quad \left[\mathcal{O}_{Hd}\right]^{ij} = \left(H^\dagger i\overleftrightarrow{D}_\mu H\right)\bar{d}_R^i\gamma^\mu d_R^j\,, \end{aligned} \tag{1.104}$$

where $\tau^a \equiv \sigma^a/2$ with σ^a the three Pauli matrices and the quarks have free flavor indices. The previous operators can be further classified in two categories: the flavor diagonal corrections to Z and W bosons couplings and the $\Delta F = 1$ flavor transitions, that we discuss respectively in sec. 1.4.2 and sec. 1.4.2.

$\Delta F = 0$

We already discussed in eqs. (1.60) and below the corrections to the W couplings, particularly important in the case of left-handed compositeness. Other important bounds come from bottom and top couplings to the Z. The former is very well constrained by LEP measurements, resulting in the following 95% bounds [39]

$$|[C_{Hq^3}^{(1)} + C_{Hq^3}^{(3)}]^{33}| \lesssim 5.5^{-2}\,\text{TeV}^{-2}\,, \qquad |[C_{Hd}]^{33}| \lesssim 1.1^{-2}\,\text{TeV}^{-2}\,. \tag{1.105}$$

[7] This contribution usually exceed the so-called IR effects [9].

	$[C^{(1)}_{Hq}+C^{(3)}_{Hq}]^{33}$		$[C_{Hd}]^{33}$		$[C_{Hu}]^{33}$	
$\mathcal{F}_c$	P.C.	Bound	P.C.	Bound	P.C.	Bound
$U(3)^2_{RC}$	$\frac{y_t^2}{m_*^2\varepsilon_u^2}\left(+\frac{y_b^2}{m_*^2\varepsilon_d^2}\right)$	$\frac{4.8}{\varepsilon_u}(\frac{0.08}{\varepsilon_d})$	$\frac{\varepsilon_d^2 g_*^2}{m_*^2}$	$1.1 g_*\varepsilon_d$	$\frac{\varepsilon_u^2 g_*^2}{m_*^2}$	$0.6\, g_*\varepsilon_u$
$U(3)\times U(2)_{RC}$	$\frac{y_t^2}{m_*^2\varepsilon_t^2}\left(+\frac{y_b^2}{m_*^2\varepsilon_d^2}\right)$	$\frac{4.8}{\varepsilon_t}(\frac{0.08}{\varepsilon_d})$	$\frac{\varepsilon_d^2 g_*^2}{m_*^2}$	$1.1 g_*\varepsilon_d$	$\frac{\varepsilon_t^2 g_*^2}{m_*^2}$	$0.6\, g_*\varepsilon_t$
$U(2)^2_{RC}$	$\frac{y_t^2}{m_*^2\varepsilon_t^2}\left(+\frac{y_b^2}{m_*^2\varepsilon_b^2}\right)$	$\frac{4.8}{\varepsilon_t}(\frac{0.08}{\varepsilon_b})$	$\frac{\varepsilon_b^2 g_*^2}{m_*^2}$	$1.1 g_*\varepsilon_b$	$\frac{\varepsilon_t^2 g_*^2}{m_*^2}$	$0.6\, g_*\varepsilon_t$
$U(3)_{LC}$	$\frac{\varepsilon_q^2 g_*^2}{m_*^2}$	$5.5\, g_*\varepsilon_q$	$\frac{y_b^2}{\varepsilon_q^2 m_*^2}$	—	$\frac{y_t^2}{\varepsilon_q^2 m_*^2}$	$\frac{0.5}{\varepsilon_q}$
$U(2)_{LC}$	$\frac{\varepsilon_{q_3}^2 g_*^2}{m_*^2}$	$5.5\, g_*\varepsilon_{q_3}$	$\frac{y_b^2}{\varepsilon_{q_3}^2 m_*^2}$	—	$\frac{y_t^2}{\varepsilon_{q_3}^2 m_*^2}$	$\frac{0.5}{\varepsilon_{q_3}}$

Table 1.1: Constraints from anomalous Z/W coupling for the different scenarios under consideration. For each operator and scenario we report the Naive Power Counting (PC), assuming the operator is produced at TL, and the lower bound on m_* expressed in TeV.

The top coupling instead is less precisely constrained by LHC measurements and we estimate it from the individual operators reach on $O^{(3)}_{Hq^3}$, $O^{(1)}_{Hq^3}$ and O_{Ht}

$$|[C^{(3)}_{Hq}]^{33}| \lesssim 0.8^{-2}\,\text{TeV}^{-2}\,, \quad |[C^{(1)}_{Hq}]^{33}| \lesssim 1.1^{-2}\,\text{TeV}^{-2}\,, \quad |[C_{Hu}]^{33}| \lesssim 0.6^{-2}\,\text{TeV}^{-2}\,, \tag{1.106}$$

extracted at 95% CL from [40]. In the previous equation while we are taking the single operator reach for $[O_{Hu}]^{33}$, we are considering the marginalized ones for $O^{(3)}_{Hq^3}$, $O^{(3)}_{Hq^3}$ since in the two dimensional plane of their Wilson coefficients the direction corresponding to the b couplings is the only very well tested (eq. (1.105)).

We report in table 1.1 the resulting power-counting and bounds on the Composite Higgs parameters space for the various scenarios under consideration. We can notice that the strongest bounds, for all the models, are given by the coupling the left bottom to the Z. These effects are partially screened in models featuring P_{LR} protection.[8] As an example if in our model the up doublets partners transform in the $(\mathbf{2},\mathbf{2})_{2/3}$ of $SU(2)_L\times SU(2)_R\times U(1)_X$ (see [6] for details), then the sum $C^{(3)}_{Hq^3}+C^{(1)}_{Hq^3}$ is not generated at leading order from the exchanges of O_{q_u}. Notice that, in this case, additional contributions come from the exchanges of the down doublets partners O_{q_d}. These effects are proportional to y_b^2 and are typically small, they are reported in parenthesis in table 1.1. The leading contribution to Zb_Lb_L coupling is then generated through operators of the form

$$[\mathcal{O}^{(1)}_{qD}]^{ij} = \bar{q}^i_L\gamma^\nu q^j_L D^\nu B_{\mu\nu}\,, \qquad [\mathcal{O}^{(3)}_{qD}]^{ij} = \bar{q}^i_L\tau^a\gamma^\nu q^j_L D^\nu W^a_{\mu\nu}\,, \tag{1.107}$$

giving rise to an anomalous Z coupling suppressed of g^2_{SM}/g^2_* with respect to the $\mathcal{O}^{(3/1)}_{qH}$

[8] Notice that in some minimal models P_{LR} arises as an accidental symmetry in the lowest order derivative expansion [41].

$\mathcal{F}_c$	$[C_{Hq}^{(3)}]^{33}$	$[C_{qD}^{(1)} + C_{qD}^{(3)}]^{33}$	$[C_{Hd}]^{33}$	$[C_{Hu}]^{33}$
$U(3)^2_{RC}$	$\frac{0.7}{\varepsilon_u}$	$\frac{3.5}{g_*\varepsilon_u}$	$1.1 g_*\varepsilon_d$	$0.6\, g_*\varepsilon_u$
$U(3)\times U(2)_{RC}$	$\frac{0.7}{\varepsilon_t}$	$\frac{3.5}{g_*\varepsilon_t}$	$1.1 g_*\varepsilon_d$	$0.6\, g_*\varepsilon_t$
$U(2)^2_{RC}$	$\frac{0.7}{\varepsilon_t}$	$\frac{3.5}{g_*\varepsilon_t}$	$1.1 g_*\varepsilon_d$	$0.6\, g_*\varepsilon_b$
$U(3)_{LC}$	$0.8\, g_*\varepsilon_q$	$4.1\, \varepsilon_q$	—	$\frac{0.5}{\varepsilon_q}$
$U(2)_{LC}$	$0.8\, g_*\varepsilon_{q_3}$	$4.1\, \varepsilon_{q_3}$	—	$\frac{0.5}{\varepsilon_{q_3}}$

Table 1.2: Lower bound on m_* in TeV from anomalous Z/W coupling for the different scenarios under consideration. We assume custodial protection for the Zb_Lb_L coupling.

contributions. We stress that in most of the models P_{LR} protection is mandatory to avoid large effects from $\Delta F = 1$ b-to-s transitions. Moreover in presence of these P_{LR} protection the bounds get weaker, we report them in Tab. 1.2.

We also need to remark that in this section we are considering anomalous coupling only to the third family even if Z coupling to light families have comparable experimental bounds [39]. However we find that in LC models the strongest constraints arise from $V_{\rm CKM}$ unitarity and are slightly stronger than anomalous Z coupling effects. In RC models, moreover, the compositeness of the light families is better constrained by compositeness test at the LHC [18].

Finally, in $U(3)_{\rm LC}$ and $U(2)_{\rm LC}$ with only one partner for each SM doublet anomalous right-handed $DF = 1$ effects are mediated by the operator

$$[O_{Hq}]^{ij} = \left(H_c^\dagger i D_\mu H\right) \bar{u_R}^i \gamma^\mu d_R^j \,. \tag{1.108}$$

Furthermore, the only relevant effect is expected from the third family, generating a Wilson Coefficient scaling as

$$[C_{Hq}]^{33} \sim \frac{y_b y_t}{\varepsilon_q^2} \frac{1}{m_*^2} \,, \tag{1.109}$$

in case of MFV or $U(2)_{\rm LC}$ through the replacement $\varepsilon_q \to \varepsilon_{q_3}$. Given the present experimental constraints [40], $|[C_{Hq}]^{33}| \lesssim 3.6$ TeV$^-2$ at 95% CL, this effect is cleary negligible.

$\Delta F = 1$

Rare B-decays

We turn to rare (semi-)leptonic B decays. As already expained in the main text, we focus on the following effective interaction Hamiltonian

$$\mathcal{H}_{\text{eff}} = -\frac{4G_F}{\sqrt{2}} V^{tb}_{\text{CKM}}(V^{ts}_{\text{CKM}})^* \frac{e^2}{16\pi^2}\Big[C_{10}(\bar{s}_L\gamma^\mu b_L)(\bar{\ell}\gamma_\mu\gamma^5\ell) + C'_{10}(\bar{s}_R\gamma^\mu b_R)(\bar{\ell}\gamma_\mu\gamma^5\ell) + C_9(\bar{s}_L\gamma^\mu b_L)(\bar{\ell}\gamma_\mu\ell) + C'_9(\bar{s}_R\gamma^\mu b_R)(\bar{\ell}\gamma_\mu\ell)\Big]. \tag{1.110}$$

In all our models $C^{(\prime)}_{9/10}$ are generated through the anomalous Z coupling operator of eq. (1.46) and $C^{(\prime)}_9$ is suppressed with respect to $C^{(\prime)}_{10}$.

We can express C_{10} and C'_{10} through the tree-level flavor structure reported in App. 1.A

$$C_{10} \simeq \frac{\sqrt{2}}{G_F}\frac{1}{V^{tb}_{\text{CKM}}(V^{ts}_{\text{CKM}})^*}\frac{4\pi^2}{e^2}\left[\frac{g_*}{m_*^2}S^{d_L d_L}\right]^{sb}, \quad C'_{10} \simeq \frac{\sqrt{2}}{G_F}\frac{1}{V^{tb}_{\text{CKM}}(V^{ts}_{\text{CKM}})^*}\frac{4\pi^2}{e^2}\left[\frac{g_*}{m_*^2}S^{d_R d_R}\right]^{sb}, \tag{1.111}$$

where we remind that the letter on the side of the brackets denote the quark type in the mass basis. Again, in the presence of P_{LR} protections, the C_{10} coefficient in eq. (1.111) is suppressed by $(g^{(Z)}_{SM}/g_*)^2$. Notice that in App. 1.A we report the factors S only for $U(3)^2_{\text{RC}}$, $U(2)\times U(3)_{\text{RC}}$ and $U(2)_{\text{LC}}$. The case of $U(3)^2_{\text{RC}}$ can be obtained simply by the replacements of figs. 1.4 and $U(3)_{\text{LC}}$ do not generate neither $C^{(\prime)}_{10}$ nor $C^{(\prime)}_9$.

The scaling of the $C^{(\prime)}_{10}$ Wilson coefficients are summarized in the following equations

$$\begin{aligned}
&U(3)^2_{\text{RC}}: && C_{10} = \tfrac{\sqrt{2}}{G_F}\tfrac{4\pi^2}{e^2}\tfrac{y_t^2}{m_*^2\varepsilon_u^2}, && C'_{10} = \times,\\
&U(2)\times U(3)_{\text{RC}}: && C_{10} = \tfrac{\sqrt{2}}{G_F}\tfrac{4\pi^2}{e^2}\tfrac{y_t^2}{m_*^2\varepsilon_t^2}, && C'_{10} = \times,\\
&U(2)^2_{\text{RC}}: && C_{10} = \tfrac{\sqrt{2}}{G_F}\tfrac{4\pi^2}{e^2}\tfrac{y_t^2}{m_*^2\varepsilon_t^2}, && C'_{10} = \tfrac{\sqrt{2}}{G_F}\tfrac{4\pi^2}{e^2}\tfrac{1}{V^{tb}_{\text{CKM}}(V^{ts}_{\text{CKM}})^*}\tfrac{y_s}{y_b}\tfrac{g_*^2(\varepsilon_b^2-\varepsilon_d^2)}{m_*^2}\\
&U(3)_{\text{LC}}: && C_{10} = \times, && C'_{10} = \times,\\
&U(2)_{\text{LC}}: && C_{10} = \tfrac{\sqrt{2}}{G_F}\tfrac{4\pi^2}{e^2}\tfrac{g_*^2(\varepsilon_{q3}^2-\varepsilon_q^2)}{m_*^2}, && C'_{10} = \tfrac{\sqrt{2}}{G_F}\tfrac{4\pi^2}{e^2}\tfrac{1}{V^{tb}_{\text{CKM}}(V^{ts}_{\text{CKM}})^*}\tfrac{y_s^2}{m_*^2\varepsilon_{q3}^2},
\end{aligned} \tag{1.112}$$

where the $\times$ means that no tree-level effects are present. Imposing the following constraints [24] on the Wilson Coefficient

$$|C_{10}| \lesssim 0.24, \qquad |C'_{10}| \lesssim 0.2, \tag{1.113}$$

we find the bounds on m_* in table 1.3.

$\mathcal{F}_c$	C_{10}	C_{10} (w/ P_{LR})	C'_{10}
$U(3)^2_{RC}$	$\frac{12.3}{\varepsilon_u}$	$\frac{9.2}{g_*\varepsilon_u}$	$\times$
$U(3)\times U(2)_{RC}$	$\frac{12.3}{\varepsilon_t}$	$\frac{9.2}{g_*\varepsilon_t}$	$\times$
$U(2)^2_{RC}$	$\frac{12.3}{\varepsilon_t}$	$\frac{9.2}{g_*\varepsilon_t}$	$10.9\, g_*\varepsilon_b$
$U(3)_{LC}$	$\times$	$\times$	$\times$
$U(2)_{LC}$	$14.2\, g_*\varepsilon_{q_3}$	$10.6\varepsilon_{q_3}$	—

Table 1.3: Constraints from the leptonic and semi-leptonic B decays. For each operator and scenario we report the lower bound on m_* expressed in TeV. With — and $\times$ we indicate respectively that the bound is negligible or the operator is not generated at leading order.

1.4.3 Dipole operators

All the model considered in this study generate negligible EDMs, as discussed in the main text. Here we summarize the effects related to the so-called EW dipoles, i.e. interactions described according to the effective operators of the form

$$\bar{f}^i_{L(R)}\sigma_{\mu\nu}f^j_{R(L)}F^{\mu\nu}_{SM}\,, \tag{1.114}$$

with $F^{\mu\nu}_{SM}$ the field strength of the SM gauge bosons and i, j the flavor indexes of the fermions.

The B system

The effective Hamiltonian relevant for $b\to s\,\gamma$ transition is

$$\mathcal{H}_{\rm eff} = -\frac{4G_F}{\sqrt{2}}V^{tb}_{\rm CKM}(V^{ts}_{\rm CKM})^*\frac{m_b e}{16\pi^2}F^{\mu\nu}\left(C_7\bar{s}_L\sigma_{\mu\nu}b_R + C'_7\bar{s}_R\sigma_{\mu\nu}b_L\right)\,. \tag{1.115}$$

and the Wilson coefficient can be estimated as

$$C_7 = \frac{4\sqrt{2}\pi^2}{G_F V^{tb}_{\rm CKM}(V^{ts}_{\rm CKM})^*}\frac{v}{\sqrt{2}m_b}\left[\frac{S^{u_L d_R}}{m_*^2}\right]^{sb}, C'_7 = \frac{4\sqrt{2}\pi^2}{G_F V^{tb}_{\rm CKM}(V^{ts}_{\rm CKM})^*}\frac{v}{\sqrt{2}m_b}\left[\frac{S^{u_R d_L}}{m_*^2}\right]^{sb}. \tag{1.116}$$

From the explicit expression of App. 1.A we find

$$\begin{aligned}
&U(3)^2_{\rm RC}: && C_7 = \times\,, && C_7' = \times\,,\\
&U(2)\times U(3)_{\rm RC}: && C_7 = \times\,, && C_7' = \times\,,\\
&U(2)^2_{\rm RC}: && C_7 = \left(\frac{m_s}{m_b}\right)^2 \frac{4\sqrt{2}\pi^2}{G_F V^{tb}_{\rm CKM}(V^{ts}_{\rm CKM})^* m_*^2}\,, && C_7' = \frac{m_s}{m_b}\frac{4\sqrt{2}\pi^2}{G_F V^{tb}_{\rm CKM}(V^{ts}_{\rm CKM})^* m_*^2}\,,\\
&U(3)_{\rm LC}: && C_7 = \times\,, && C_7' = \times\,,\\
&U(2)_{\rm LC}: && C_7 = \left(\frac{m_s}{m_b}\right) \frac{4\sqrt{2}\pi^2}{G_F V^{tb}_{\rm CKM}(V^{ts}_{\rm CKM})^* m_*^2}\,, && C_7' = \left(\frac{m_s}{m_b}\right)^2 \frac{4\sqrt{2}\pi^2}{G_F V^{tb}_{\rm CKM}(V^{ts}_{\rm CKM})^* m_*^2}\,.
\end{aligned} \tag{1.117}$$

Imposing the constraints from [28]

$$\mathrm{Re}C_7 < 0.03\,, \qquad\qquad |C_7'| < 0.07 \tag{1.118}$$

we get the following bounds

$$U(2)^2_{\rm RC}: \quad m_* > 5.8\,\mathrm{TeV} \qquad\qquad U(2)_{\rm LC}: \quad m_* > 8.8\,\mathrm{TeV}\,. \tag{1.119}$$

The D system

Regarding the D mesons, the main $\Delta F = 1$ constraints come from the direct CP violation in the hadronic decays, usually encapsulated in the observable

$$\Delta a_{\rm CP} \equiv a_{K^+K^-} - a_{\pi^+\pi^-}\,. \tag{1.120}$$

The latest measured value is [42]

$$\Delta a_{\rm CP} = (-15.4 \pm 2.9)\times 10^{-4}\,, \tag{1.121}$$

roughly compatible with the expected SM prediction. In our models, such transitions could arise at tree-level from flavor-changing dipole operators

$$\mathcal{H}_{\rm eff} = \frac{G_F}{\sqrt{2}}\frac{m_c}{4\pi^2}\left(C_8 \bar{u}_L \sigma^{\mu\nu} g_s G_{\mu\nu} c_R + C_8' \bar{c}_L \sigma^{\mu\nu} g_s G_{\mu\nu} u_R\right)\,. \tag{1.122}$$

These operators are generated at tree-level by the three models in sec. 1.3, while for MFV models they are zero. We estimate for all three models

$$\begin{aligned}
&U(2)\times U(3)_{\rm RC}\\
&U(2)^2_{\rm RC} \qquad\qquad : \qquad\qquad \mathrm{Im}\, C_8 \sim \frac{4\pi^2}{m_*^2}\frac{2}{G_F}\frac{y_c^2}{y_t^2} \sim 1.5\times 10^{-4}\left(\frac{1\,\mathrm{TeV}}{m_*}\right)^2\,,\\
&U(2)_{\rm LC}
\end{aligned} \tag{1.123}$$

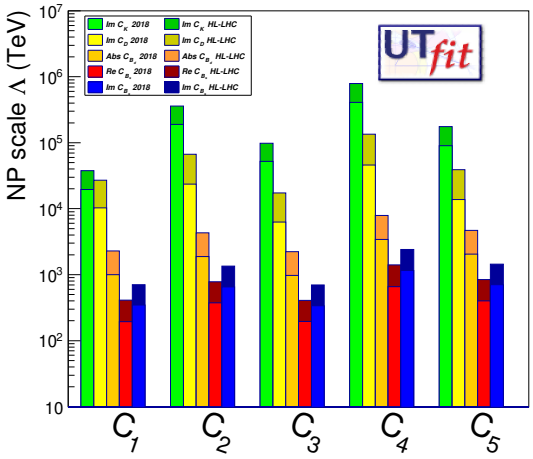


Figure 1.7: 95% constraints on $\Delta F = 2$ operators, from [44]

while C_8' is further suppressed by y_u/y_c. The imaginary part of C_8 enters at $O(1)$ in $\Delta a_{\rm CP}$ as shown in [43]. However we see that for $m_* \gtrsim$ TeV this contribution is well below the experimental accuracy, leading to no relevant bounds on the models.

1.4.4 Four-fermions operators

An important set of constraints comes from the $\Delta F = 2$ transitions. Adopting the convention of [4], the short-distance contact operators mediating the $\Delta F = 2$ flavor transitions can be reduced to

$$\begin{aligned}
\mathcal{O}_1^{ij} &= (\bar{f}_{iL}^{\alpha}\gamma^{\mu} f_{jL}^{\alpha})(\bar{f}_{iL}^{\beta}\gamma_{\mu} f_{jL}^{\beta})\,, & \widetilde{\mathcal{O}}_1^{ij} &= (\bar{f}_{iR}^{\alpha}\gamma^{\mu} f_{jR}^{\alpha})(\bar{f}_{iR}^{\beta}\gamma_{\mu} f_{jR}^{\beta})\,,\\
\mathcal{O}_2^{ij} &= (\bar{f}_{iR}^{\alpha} f_{jL}^{\alpha})(\bar{f}_{iR}^{\beta} f_{jL}^{\beta})\,, & \widetilde{\mathcal{O}}_2^{ij} &= (\bar{f}_{iL}^{\alpha} f_{jR}^{\alpha})(\bar{f}_{iL}^{\beta} f_{jR}^{\beta})\,,\\
\mathcal{O}_3^{ij} &= (\bar{f}_{iR}^{\alpha} f_{jL}^{\beta})(\bar{f}_{iR}^{\beta} f_{jL}^{\alpha})\,, & \widetilde{\mathcal{O}}_3^{ij} &= (\bar{f}_{iL}^{\alpha} f_{jR}^{\beta})(\bar{f}_{iL}^{\beta} f_{jR}^{\alpha})\,,\\
\mathcal{O}_4^{ij} &= (\bar{f}_{iR}^{\alpha} f_{jL}^{\alpha})(\bar{f}_{iL}^{\beta} f_{jR}^{\beta})\,, & \mathcal{O}_5^{ij} &= (\bar{f}_{iR}^{\alpha} f_{jL}^{\beta})(\bar{f}_{iL}^{\beta} f_{jR}^{\alpha})\,,
\end{aligned} \tag{1.124}$$

where i and j denote the flavor indices while α and β denote the color indices. In the following we will neglect the Q_3, $\widetilde{Q}_3$ and Q_5 operators, since they have the same power counting in terms of the new physics parameters, but a weaker experimental bound. The bounds we use can be seen in Fig. 1.7 and are taken from [44].

The Kaon system

The strongest bounds for the Kaon system come from the $K^0 - \bar{K}^0$ mixings and in particular from the observables Δm_K and ϵ_K. In terms of the operators defined in (1.124),

$\mathcal{F}_c$	C_1^{sd}	$\widetilde{C}_1^{sd}$	C_2^{sd}	$\widetilde{C}_2^{sd}$	C_4^{sd}
$U(3)^2_{RC}$	$\frac{4.3(8.6)}{g_*\varepsilon_u^2}$	$\times$	$\times$	$\times$	$\times$
$U(3)\times U(2)_{RC}$	$\frac{4.3(8.6)}{g_*\varepsilon_t^2}\left[\frac{0.4(0.8)}{g_*\varepsilon_t\varepsilon_u}\right]$	$\times$	$\times$	$\times$	$\times$
$U(2)^2_{RC}$	$\frac{4.3(8.6)}{g_*\varepsilon_t^2}\left[\frac{0.4(0.8)}{g_*\varepsilon_t\varepsilon_u}\right]$	$7.5(15)g_*\varepsilon_b^2$	—	—	—
$U(3)_{LC}$	$\times$	$\times$	$\times$	$\times$	$\times$
$U(2)_{LC}$	$5.7(11.4)g_*\varepsilon_{q_3}^2$	—	—	—	—

Table 1.4: Constraints from the Kaon system $\Delta F = 2$ transitions. For each operator and scenario we report the lower bound on m_* expressed in TeV. In parenthesis the projected bounds for HL-LHC. In square brackets an additional contribution that become large for small values of ε_u. With — and $\times$ we indicate respectively that the bound is negligible or the operator is not generated at leading order.

the current and projected for HL-LHC bounds (in parenthesis) are

$$\begin{aligned} |\mathrm{Im}\, C_1^{sd}| \sim |\mathrm{Im}\, \widetilde{C}_1^{sd}| &\lesssim 2.5\times 10^{-9}\,(6.3\times 10^{-10})\,\mathrm{TeV}^{-2}\,,\\ |\mathrm{Im}\, C_2^{sd}| \sim |\mathrm{Im}\, \widetilde{C}_2^{sd}| &\lesssim 2.5\times 10^{-11}\,(6.3\times 10^{-12})\,\mathrm{TeV}^{-2}\,,\\ |\mathrm{Im}\, \widetilde{C}_4^{sd}| &\lesssim 6.3\times 10^{-12}\,(1.6\times 10^{-12})\,\mathrm{TeV}^{-2}\,. \end{aligned} \tag{1.125}$$

Using the flavor structures of App. 1.A we find for the various models

$$\begin{aligned} U(3)^2_{\rm RC}: &\quad \mathrm{Im}\, C_1^{sd} = 2\tfrac{A^4 y_t^4 \eta\lambda^{10}}{m_*^2 g_*^2 \varepsilon_u^4}\,, &&\quad \mathrm{Im}\, \widetilde{C}_1^{sd} = \times\,,\\ U(2)\times U(3)_{\rm RC}: &\quad \mathrm{Im}\, C_1^{sd} = 2\tfrac{A^4 y_t^4 \eta\lambda^{10}}{m_*^2 g_*^2 \varepsilon_t^4}\,, &&\quad \mathrm{Im}\, \widetilde{C}_1^{sd} = \times\,,\\ U(2)^2_{\rm RC}: &\quad \mathrm{Im}\, C_1^{sd} = 2\tfrac{A^4 y_t^4 \eta\lambda^{10}}{m_*^2 g_*^2 \varepsilon_t^4}\,, &&\quad \mathrm{Im}\, \widetilde{C}_1^{sd} = \tfrac{y_s^4}{y_b^4}\tfrac{g_*^2\varepsilon_b^4}{m_*^2}\,,\\ U(3)_{\rm LC}: &\quad \mathrm{Im}\, C_1^{sd} = \times\,, &&\quad \mathrm{Im}\, \widetilde{C}_1^{sd} = \times\,,\\ U(2)_{\rm LC}: &\quad \mathrm{Im}\, C_1^{sd} = 2A^4\eta\lambda^{10}\tfrac{g_*^2\varepsilon_{q3}^4}{m_*^2}\,, &&\quad \mathrm{Im}\, \widetilde{C}_1^{sd} = (a')^2\,\tfrac{y_d^2 y_s^6}{y_b^4}\tfrac{1}{g_*^2\varepsilon_q^4}\,. \end{aligned} \tag{1.126}$$

The $U(2)\times U(3)_{\rm RC}$ and $U(2)^2_{\rm RC}$ models also generate a contribution to C_1 that is enhanced for small values of ε_u. We estimate this contribution to be

$$\begin{aligned} &U(2)\times U(3)_{\rm RC}:\\ &U(2)^2_{\rm RC}: \end{aligned} \qquad \mathrm{Im}\, C_1^{sd} = 2\frac{A^4 y_c^2 y_t^2 \eta\lambda^6}{m_*^2 g_*^2 \varepsilon_u^2 \varepsilon_t^2} \tag{1.127}$$

The $U(2)^2_{\rm RC}$ and $U(2)_{\rm LC}$ models also generate the three mixed chirality operators $\mathcal{O}_2$, $\widetilde{\mathcal{O}}_2$ and $\mathcal{O}_4$, but with negligible coefficients compared to the experimental bound. The bound on m_* coming from all these operators is reported in Tab. 1.4.

The B_d and B_s system

Another important set of constraints from $\Delta F = 2$ transitions, come from the flavor violating processes involving the neutral B_d and B_s mesons. For these operators the current experimental bounds [44] are

$$\begin{aligned} |\mathrm{Abs}\, Q_1^{bd}| \sim |\mathrm{Abs}\, \widetilde{Q}_1^{bd}| &\lesssim 1 \times 10^{-6}\, (2.5 \times 10^{-7})\, \mathrm{TeV}^{-2}\,, \\ |\mathrm{Abs}\, Q_2^{bd}| \sim |\mathrm{Abs}\, \widetilde{Q}_2^{bd}| &\lesssim 2.5 \times 10^{-7}\, (6.3 \times 10^{-8})\, \mathrm{TeV}^{-2}\,, \\ |\mathrm{Abs}\, \widetilde{Q}_4^{bd}| &\lesssim 1.1 \times 10^{-7}\, (2.8 \times 10^{-8})\, \mathrm{TeV}^{-2}\,, \end{aligned} \tag{1.128}$$

and

$$\begin{aligned} |\mathrm{Re}\, Q_1^{bs}| \sim |\mathrm{Re}\, \widetilde{Q}_1^{bs}| &\lesssim 2.5 \times 10^{-5}\, (6.3 \times 10^{-6})\, \mathrm{TeV}^{-2}\,, \\ |\mathrm{Im}\, Q_1^{bs}| \sim |\mathrm{Im}\, \widetilde{Q}_1^{bs}| &\lesssim 8 \times 10^{-6}\, (2 \times 10^{-6})\, \mathrm{TeV}^{-2}\,, \\ |\mathrm{Re}\, Q_2^{bs}| \sim |\mathrm{Re}\, \widetilde{Q}_2^{bs}| &\lesssim 8.1 \times 10^{-6}\, (1.6 \times 10^{-8})\, \mathrm{TeV}^{-2}\,, \\ |\mathrm{Im}\, Q_2^{bs}| \sim |\mathrm{Im}\, \widetilde{Q}_2^{bs}| &\lesssim 2.8 \times 10^{-6}\, (6.9 \times 10^{-7})\, \mathrm{TeV}^{-2}\,, \\ |\mathrm{Re}\, \widetilde{Q}_4^{bs}| &\lesssim 4 \times 10^{-6}\, (1 \times 10^{-6})\, \mathrm{TeV}^{-2}\,, \\ |\mathrm{Im}\, \widetilde{Q}_4^{bs}| &\lesssim 1 \times 10^{-6}\, (2.5 \times 10^{-7})\, \mathrm{TeV}^{-2}\,. \end{aligned} \tag{1.129}$$

For these observables we estimate

$$\begin{aligned} U(3)^2_{\mathrm{RC}}: &\quad \mathrm{Abs}\, C_1^{bd} = \tfrac{A^2 y_t^4 \lambda^6}{m_*^2 g_*^2 \varepsilon_u^4}\,, && \mathrm{Abs}\, \widetilde{C}_1^{bd} = \times\,, \\ U(2)\times U(3)_{\mathrm{RC}}: &\quad \mathrm{Abs}\, C_1^{bd} = \tfrac{A^2 y_t^4 \lambda^6}{m_*^2 g_*^2 \varepsilon_t^4}\,, && \mathrm{Abs}\, \widetilde{C}_1^{bd} = \times\,, \\ U(2)^2_{\mathrm{RC}}: &\quad \mathrm{Abs}\, C_1^{bd} = \tfrac{A^2 y_t^4 \lambda^6}{m_*^2 g_*^2 \varepsilon_t^4}\,, && \mathrm{Abs}\, \widetilde{C}_1^{bd} = \tfrac{y_s^2}{y_b^2}\tfrac{g_*^2 \varepsilon_b^4}{m_*^2}\,, \\ U(3)_{\mathrm{LC}}: &\quad \mathrm{Abs}\, C_1^{bd} = \times\,, && \mathrm{Abs}\, \widetilde{C}_1^{bd} = \times\,, \\ U(2)_{\mathrm{LC}}: &\quad \mathrm{Abs}\, C_1^{bd} = A^2\lambda^6 \tfrac{g_*^2 \varepsilon_{q_3}^4}{m_*^2}\,, && \mathrm{Abs}\, \widetilde{C}_1^{bd} = (a')^2\, \tfrac{y_d^2 y_s^2}{g_*^2 \varepsilon_q^4}\,, \end{aligned} \tag{1.130}$$

and

$$\begin{aligned} U(3)^2_{\mathrm{RC}}: &\quad \mathrm{Re/Im}\, C_1^{bs} = \tfrac{A^2 y_t^4 \lambda^4}{m_*^2 g_*^2 \varepsilon_u^4}(1 + i\eta\lambda^2)\,, && \mathrm{Re/Im}\, \widetilde{C}_1^{bs} = \times\,, \\ U(2)\times U(3)_{\mathrm{RC}}: &\quad \mathrm{Re/Im}\, C_1^{bs} = \tfrac{A^2 y_t^4 \lambda^4}{m_*^2 g_*^2 \varepsilon_t^4}(1 + i\eta\lambda^2)\,, && \mathrm{Re/Im}\, \widetilde{C}_1^{bs} = \times\,, \\ U(2)^2_{\mathrm{RC}}: &\quad \mathrm{Re/Im}\, C_1^{bs} = \tfrac{A^2 y_t^4 \lambda^4}{m_*^2 g_*^2 \varepsilon_t^4}(1 + i\eta\lambda^2)\,, && \mathrm{Re/Im}\, \widetilde{C}_1^{bs} = \tfrac{y_s^2}{y_b^2}\tfrac{g_*^2 \varepsilon_b^4}{m_*^2}\,, \\ U(3)_{\mathrm{LC}}: &\quad \mathrm{Re/Im}\, C_1^{bs} = \times\,, && \mathrm{Re/Im}\, \widetilde{C}_1^{bs} = \times\,, \\ U(2)_{\mathrm{LC}}: &\quad \mathrm{Re}\, C_1^{bs} = A^2\lambda^4 \tfrac{g_*^2 \varepsilon_{q_3}^4}{m_*^2}\,, && \mathrm{Re/Im}\, \widetilde{C}_1^{bs} = (a')^2\, \tfrac{y_s^4}{g_*^2 \varepsilon_q^4}\,. \end{aligned} \tag{1.131}$$

Also in this case the mixed chirality operators are generated only for the $U(2)^2_{\mathrm{RC}}$ and $U(2)_{\mathrm{LC}}$ models but give negligible bounds. The bounds on m_* can be found in table 1.5.

$\mathcal{F}_c$	C_1^{bd}	$\widetilde{C}_1^{bd}$	C_2^{bd}	$\widetilde{C}_2^{bd}$	C_4^{bd}
$U(3)^2_{RC}$	$\frac{6.5(13)}{g_*\varepsilon_u^2}$	$\times$	$\times$	$\times$	$\times$
$U(3)\times U(2)_{RC}$	$\frac{6.5(13)}{g_*\varepsilon_t^2}$	$\times$	$\times$	$\times$	$\times$
$U(2)^2_{RC}$	$\frac{6.5(13)}{g_*\varepsilon_t^2}$	$19.3(38.7)g_*\varepsilon_b^2$	—	—	—
$U(3)_{LC}$	$\times$	$\times$	$\times$	$\times$	$\times$
$U(2)_{LC}$	$8.7(17.3)g_*\varepsilon_{q_3}^2$	—	—	—	—

$\mathcal{F}_c$	C_1^{bs}	$\widetilde{C}_1^{bs}$	C_2^{bs}	$\widetilde{C}_2^{bs}$	C_4^{bs}
$U(3)^2_{RC}$	$\frac{5.9(11.8)}{g_*\varepsilon_u^2}$	$\times$	$\times$	$\times$	$\times$
$U(3)\times U(2)_{RC}$	$\frac{5.9(11.8)}{g_*\varepsilon_t^2}$	$\times$	$\times$	$\times$	$\times$
$U(2)^2_{RC}$	$\frac{5.9(11.8)}{g_*\varepsilon_t^2}$	$6.8(13.6)g_*\varepsilon_b^2$	—	—	—
$U(3)_{LC}$	$\times$	$\times$	$\times$	$\times$	$\times$
$U(2)_{LC}$	$7.9(15.8)g_*\varepsilon_{q_3}^2$	—	—	—	—

Table 1.5: Constraints from the B_d and B_s systems $\Delta F = 2$ transitions. For each operator and scenario we report the lower bound on m_* expressed in TeV. In parenthesis the projected bounds for HL-LHC. With — and $\times$ we indicate respectively that the bound is negligible or the operator is not generated at leading order.

The D system

The final set of $\Delta F = 2$ constraints we study come from the D system. These constraints are complementary to the previous ones since they involve up-type quarks. The current experimental bounds are [44]

$$\begin{aligned} |\mathrm{Im}\, Q_1^{cu}| \sim |\mathrm{Im}\, \widetilde{Q}_1^{cu}| &\lesssim 1\times 10^{-8}\,(1.6\times 10^{-9})\,\mathrm{TeV}^{-2}\,, \\ |\mathrm{Im}\, Q_2^{cu}| \sim |\mathrm{Im}\, \widetilde{Q}_2^{cu}| &\lesssim 2.5\times 10^{-9}\,(2\times 10^{-10})\,\mathrm{TeV}^{-2}\,, \\ |\mathrm{Im}\, \widetilde{Q}_4^{cu}| &\lesssim 5\times 10^{-10}\,(1\times 10^{-10})\,\mathrm{TeV}^{-2}\,. \end{aligned} \tag{1.132}$$

In this sector most contributions are negligible compared to the experimental bounds. We report here the only ones that could be potentially relevant

$$\begin{aligned} U(2)\times U(3)_{\mathrm{RC}}: &\quad \mathrm{Im}\, \widetilde{C}_1^{cu} = \frac{y_c^4}{y_t^4}\frac{g_*^2\varepsilon_t^4}{m_*^2}\,, \\ U(2)^2_{\mathrm{RC}}: &\quad \mathrm{Im}\, \widetilde{C}_1^{cu} = \frac{y_c^4}{y_t^4}\frac{g_*^2\varepsilon_t^4}{m_*^2}\,, \\ U(2)_{\mathrm{LC}}: &\quad \mathrm{Im}\, C_1^{cu} = \frac{y_c^4}{y_t^4} g_*^2\varepsilon_{q_3}^4\,. \end{aligned} \tag{1.133}$$

The bounds on m_* can be found in table 1.6.

$\mathcal{F}_c$	C_1^{cu}	$\widetilde{C}_1^{cu}$	C_2^{cu}	$\widetilde{C}_2^{cu}$	C_4^{cu}
$U(3)^2_{RC}$	—	×	×	×	×
$U(3) \times U(2)_{RC}$	—	$0.12(0.31) g_* \varepsilon_t^2$	—	—	—
$U(2)^2_{RC}$	—	$0.12(0.31) g_* \varepsilon_t^2$	—	—	—
$U(3)_{LC}$	×	×	×	×	×
$U(2)_{LC}$	$0.12(0.31) g_* \varepsilon_{q_3}^2$	—	—	—	—

Table 1.6: Constraints from the D system $\Delta F = 2$ transitions. For each operator and scenario we report the lower bound on m_* expressed in TeV. In parenthesis the projected bounds for HL-LHC. With — and × we indicate respectively that the bound is negligible or the operator is not generated at leading order.

1.5 Conclusions and outlook

In this work we have examined several possibilities for the flavor structure in Composite Higgs models that implement Partial Compositeness. The class of models that we studied have either the six right-handed SM quarks or the three left-handed ones strongly composite. The models are further classified according to the flavor group of the composite sector. The remaining mixings are instead taken to be generic, with the only request that the Standard Model can be reproduced. The models we considered are summarized in Fig. 1.4, for the right-handed compositeness models, and Fig. 1.5, for the left-handed compositeness ones. The characteristic that unites all these models is the exact absence of contributions to the neutron EDM at leading order in the number of spurion insertions, assuming no additional CP violating effects arise from the strong sector.

We started from models that realize Minimal Flavor Violation and we then tried to reduce the flavor symmetry group of the composite dynamics in order to relax the tension with the experimental bounds. Regarding the right-handed compositeness models we have seen that there is a clear advantage in going from $U(3)^2_{\rm RC}$ to $U(2) \times U(3)_{\rm RC}$ since this removes the strong constraints coming from the interplay of light quark compositeness bounds and $\Delta F = 2$ transitions in the B sector. Further reducing the symmmetry group to $U(2)^2_{\rm RC}$ is instead not so convenient, since this model introduces new flavor-violating processes in the down sector. Most of these can be taken under control for sufficiently small ϵ_b with the only exception being $b \to s\gamma$, which is independent on ϵ_b and gives an irreducible bound on m_* that was absent in $U(2) \times U(3)_{\rm RC}$.

We tried to apply the same logic to the left-handed compositeness models (with a single partner for the SM doublets), where the $U(3)_{\rm LC}$ MFV model faces strong bounds from combining light quark compositeness and the CKM unitarity constraints. This lead us to consider the $U(2)_{\rm LC}$ model that, however, introduces new flavor violating effects, absent in MFV. Even for this model, we found the strongest constraint to come from the $b \to s\gamma$ transitions. In Fig. 1.8 we report a summary of the bounds of all the models studied.

It is interesting to notice that from the flavor point of view, in the scenarios considered the

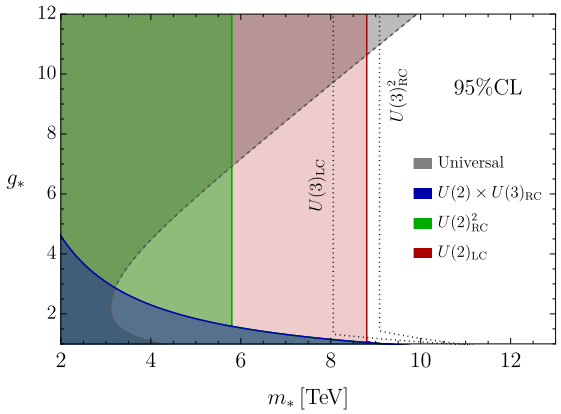


Figure 1.8: Summary of all the combined bounds for the various models examined in this work. The grey area is for flavor-universal bounds and are common to all models. The dotted lines are the two implementations of MFV for right- and left-handed compositeness models.

strongest bounds always come from the B mesons mixings and decays. This is different compared to the anarchic scenario where the strongest constraints are due to EDMs and to the K mesons system. We started to assess the projected sensitivity expected from the BELLE II [45] and HL-LHC [44] programs, but we still miss a complete study.[9]

In this direction we also plan to assess the impact of future collider on the different setup we considered. Preliminary projections for a universal and top-philic composite dynamics are already present in [2] but a systematic discussion on flavor is still missing.

Our study should also be extended to the lepton sector, already partially addressed in the literature [19, 47]. Moreover an unified discussion of the SM flavor is still missing and can be crucial in light of the foreseen experiments Mu3e [48] and Mu2e [49] that are expected to improve the current bounds by four orders of magnitude.

Our preliminary results show that flavor observables, especially in the B meson sector, play a remarkable role in the composite Higgs framework. Different flavor structures can indeed completely exclude a strong dynamics accounting for the SM masses or can give specific signatures of it. Research and progress in flavor physics is mandatory in order to exploit the full potential of the next experiments and this can play a major role in our understanding of fundamental physics in the next few years.

[9] See also [46].

1.A Additional details on the models

In this appendix we go into more details in the construction of the three models presented in section 1.3. In particular for each model we report the quantum number of the spurions, the matrices that diagonalize the Yukawas and the tree-level flavor structures used to built the Wilson coefficients in sec. 1.4. The diagonalizing matrices are computed in the $y_u \ll y_c \ll y_t$ and $y_d \ll y_s \ll y_b$ limit and for each entry of the matrix we only report the leading term of this expansion. The tree-level flavor structures are the coefficients in the mass basis of all the possible fermion bilinears that can appear in the dim-6 effective operators as fixed by spurion analysis. Explicitly if an operator contains the bilinear $\dots \bar{f}^i_\chi(\dots) g^j_\lambda \dots$ of quarks $f, g = u, d$ with chirality χ and λ and flavor i and j, by flavor symmetry it will must appear with a coefficient

$$[S^{f_\chi g_\lambda}]_{ij} \,.$$

By tree-level we mean that in building $[S^{f_\chi g_\lambda}]_{ij}$ we only have two spurion insertions with the composite index contracted. In our conventions we take this structures to be normalized in such a way that they have the units of measure of a single coupling.

1.A.1 $U(2) \times U(3)_{\mathbf{RC}}$

This model is characterized by three diagonal right-handed mixings ε_u, ε_d and ε_t, and three left-handed off-diagonal mixings λ_{q_u}, λ_{q_d} and λ_{q_t}. The mixing Lagrangian is reported in (1.64) These parameters break the flavor group that for this model is assumed to be

$$\mathcal{F} = U(3)_q \times U(3)_u \times U(3)_d \times U(2)_U \times U(3)_D \times U(1)_T \,.$$

The form that we assume for the λ mixing is reported in the main text in eq. (1.66), where we assume the numbers a and b to be $O(1)$ complex numbers with arbitrary phases. The spurion non-abelian quantum numbers for these mixing parameters, with respect to the full flavor group are

	$U(3)_q$	$U(3)_u$	$U(3)_d$	$U(2)_U$	$U(3)_D$	$U(1)_T$
λ_{q_u}	3	1	1	$\bar{2}$	1	0
λ_{q_d}	3	1	1	1	$\bar{3}$	0
λ_{q_t}	3	1	1	1	1	-1
$\varepsilon_u \delta_{Ai}$	1	$\bar{3}$	1	2	1	0
$\varepsilon_d \delta_{ai}$	1	1	$\bar{3}$	1	3	0
$\varepsilon_t \delta_{3i}$	1	$\bar{3}$	1	1	1	1

In this model, the Yukawa matrices are given by (1.65). The only non-trivial matrix that has to be diagonalized is Y_u. For Y_d the diagonalization matrix $\widetilde{U}_d$ can be rewritten in

terms of the CKM matrix as in eq.(1.67). We define the U_u and V_u matrices by

$$Y_u = U_u \bar{Y}_u V_u^\dagger \,, \tag{1.134}$$

where the bar indicates the diagonal and real matrix. The explicit form of U_u and V_u in the limit $y_u \ll y_c \ll y_t$ is

$$U_u = \begin{pmatrix} 1 & -a^* b \frac{y_u y_c}{y_t^2} & a^* \frac{y_u y_c}{y_t^2} \\ a b^* \frac{y_u y_c}{y_t^2} & 1 & b^* \frac{y_c^2}{y_t^2} \\ -a \frac{y_u y_c}{y_t^2} & -b \frac{y_c^2}{y_t^2} & 1 \end{pmatrix}, \quad V_u = \begin{pmatrix} 1 & -a^* b \frac{y_c^2}{y_t^2} & a^* \frac{y_c}{y_t} \\ a b^* \frac{y_u^2}{y_t^2} & 1 & b^* \frac{y_c}{y_t} \\ -a \frac{y_c}{y_t} & -b \frac{y_c}{y_t} & 1 \end{pmatrix}. \tag{1.135}$$

Finally we consider the tree-level flavor structures of this model in the mass basis.

$$S^{d_L d_R} = \bar{Y}_d \tag{1.136}$$
$$S^{u_L d_R} = V_{\rm CKM} S^{d_L d_R} \tag{1.137}$$
$$S^{d_R d_R} = g_* \varepsilon_d^2 \tag{1.138}$$
$$S^{u_L u_R} = \bar{Y}_u + (x-1) U_u^\dagger \widetilde{Y}_t V_u \tag{1.139}$$
$$S^{d_L u_R} = V_{\rm CKM}^\dagger S^{u_L u_R} \tag{1.140}$$
$$S^{u_R d_R} = 0 \tag{1.141}$$
$$S^{u_R u_R} = g_* V_u^\dagger {\rm diag}(\varepsilon_u^2, \varepsilon_u^2, x\varepsilon_t^2) V_u \tag{1.142}$$
$$S^{u_L u_L} = \frac{1}{g_* \varepsilon_u^2} \left[(\bar{Y}_u)^2 + \left(\frac{\varepsilon_u^2}{\varepsilon_t^2} x_1 - 1 \right) U_u^\dagger (\widetilde{Y}_t)^2 U_u \right] + \frac{x_2}{g_* \varepsilon_d^2} V_{\rm CKM} (\bar{Y}_d)^2 V_{\rm CKM}^\dagger \tag{1.143}$$
$$S^{u_L d_L} = S^{u_L u_L} V_{\rm CKM} \tag{1.144}$$
$$S^{d_L d_L} = V_{\rm CKM}^\dagger S^{u_L u_L} V_{\rm CKM} \tag{1.145}$$

where we have defined the $\widetilde{Y}_t$ matrix as

$$\widetilde{Y}_t = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & y_t \end{pmatrix}, \tag{1.146}$$

and the x parameters are $O(1)$ numbers that characterize the misalignment of these structure with respect to the Yukawas.

1.A.2 $U(2)^2_{\mathbf{RC}}$

This model is characterized by four diagonal right-handed mixings ε_u, ε_d, ε_t and ε_b, and four left-handed off-diagonal mixings λ_u, λ_{q_d}, λ_{q_t} and λ_{q_b}. The mixing Lagrangian is written in (1.79) These parameters break the flavor group that for this model is assumed to be

$$\mathcal{F} = U(3)_q \times U(3)_u \times U(3)_d \times U(2)_U \times U(2)_D \times U(1)_T \times U(1)_B \,.$$

The form that we assume for the λ mixing is reported in the main text in eq. (1.81), where we assume the numbers a, b, a' and b' to be $O(1)$ complex numbers with arbitrary phases. The spurion non-abelian quantum numbers for these mixing parameters, with respect to the full flavor group are

	$U(3)_q$	$U(3)_u$	$U(3)_d$	$U(2)_U$	$U(2)_D$	$U(1)_T$	$U(1)_B$
λ_{q_u}	3	1	1	$\bar{2}$	1	0	0
λ_{q_d}	3	1	1	1	$\bar{3}$	0	0
λ_{q_t}	3	1	1	1	1	-1	0
λ_{q_b}	3	1	1	1	1	0	-1
$\varepsilon_u \delta_{Ai}$	1	$\bar{3}$	1	2	1	0	0
$\varepsilon_d \delta_{Bi}$	1	1	$\bar{3}$	1	2	0	0
$\varepsilon_t \delta_{3i}$	1	$\bar{3}$	1	1	1	1	0
$\varepsilon_b \delta_{3i}$	1	1	$\bar{3}$	1	1	0	1

In this model, the Yukawa matrices are given by (1.80). We define the diagonalizing matrices by

$$Y_u = U_u \bar{Y}_u V_u^\dagger \,, \quad \widetilde{U}_d^\dagger Y_d = U_d \bar{Y}_d V_d^\dagger \,, \tag{1.147}$$

where the bar indicates the diagonal and real matrices. The explicit form these matrices in the limit $y_u \ll y_c \ll y_t$ and $y_d \ll y_s \ll y_b$ is

$$U_u = \begin{pmatrix} 1 & -a^* b \frac{y_u y_c}{y_t^2} & a^* \frac{y_u y_c}{y_t^2} \\ a b^* \frac{y_u y_c}{y_t^2} & 1 & b^* \frac{y_c^2}{y_t^2} \\ -a \frac{y_u y_c}{y_t^2} & -b \frac{y_c^2}{y_t^2} & 1 \end{pmatrix} , \quad V_u = \begin{pmatrix} 1 & -a^* b \frac{y_c^2}{y_t^2} & a^* \frac{y_c}{y_t} \\ a b^* \frac{y_u^2}{y_t^2} & 1 & b^* \frac{y_c}{y_t} \\ -a \frac{y_c}{y_t} & -b \frac{y_c}{y_t} & 1 \end{pmatrix} ,$$

$$U_d = \begin{pmatrix} 1 & -a'^* b' \frac{y_d y_s}{y_b^2} & a'^* \frac{y_d y_s}{y_b^2} \\ a' b'^* \frac{y_d y_s}{y_b^2} & 1 & b'^* \frac{y_s^2}{y_b^2} \\ -a' \frac{y_d y_s}{y_b^2} & -b' \frac{y_s^2}{y_b^2} & 1 \end{pmatrix} , \quad V_d = \begin{pmatrix} 1 & -a'^* b' \frac{y_s^2}{y_b^2} & a'^* \frac{y_s}{y_b} \\ a' b'^* \frac{y_d^2}{y_b^2} & 1 & b'^* \frac{y_s}{y_b} \\ -a' \frac{y_s}{y_b} & -b' \frac{y_s}{y_b} & 1 \end{pmatrix} . \tag{1.148}$$

Finally we consider the tree-level flavor structures of this model in the mass basis.

$$S^{d_L d_R} = \bar{Y}_d + (x-1) U_d^\dagger \widetilde{Y}_b V_d \tag{1.149}$$
$$S^{u_L d_R} = V_{\rm CKM} S^{d_L d_R} \tag{1.150}$$
$$S^{d_R d_R} = g_* V_d^\dagger \mathrm{diag}(\varepsilon_d^2, \varepsilon_d^2, x\varepsilon_b^2) V_d \tag{1.151}$$
$$S^{u_L u_R} = \bar{Y}_u + (x-1) U_u^\dagger \widetilde{Y}_t V_u \tag{1.152}$$
$$S^{d_L u_R} = V_{\rm CKM}^\dagger S^{u_L u_R} \tag{1.153}$$
$$S^{u_R d_R} = 0 \tag{1.154}$$
$$S^{u_R u_R} = g_* V_u^\dagger \mathrm{diag}(\varepsilon_u^2, \varepsilon_u^2, x\varepsilon_t^2) V_u \tag{1.155}$$
$$S^{u_L u_L} = \frac{1}{g_* \varepsilon_u^2} \left[(\bar{Y}_u)^2 + \left(\frac{\varepsilon_u^2}{\varepsilon_t^2} x_1 - 1 \right) U_u^\dagger (\widetilde{Y}_t)^2 U_u \right] \tag{1.156}$$
$$+ V_{\rm CKM} \frac{1}{g_* \varepsilon_d^2} \left[x_2 (\bar{Y}_d)^2 + \left(\frac{\varepsilon_d^2}{\varepsilon_b^2} x_3 - x_2 \right) U_d^\dagger (\widetilde{Y}_b)^2 U_d \right] V_{\rm CKM}^\dagger \tag{1.157}$$
$$S^{u_L d_L} = S^{u_L u_L} V_{\rm CKM} \tag{1.158}$$
$$S^{d_L d_L} = V_{\rm CKM}^\dagger S^{u_L u_L} V_{\rm CKM} \tag{1.159}$$

where we have defined the $\widetilde{Y}_t$ and $\widetilde{Y}_b$ matrices as

$$\widetilde{Y}_{t/b} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & y_{t/b} \end{pmatrix}, \tag{1.160}$$

and the x parameters are $O(1)$ numbers.

1.A.3 $U(2)_{\mathbf{LC}}$

This model is characterized by two diagonal left-handed mixings ε_q and ε_{q_3}, and four right-handed off-diagonal mixings λ_u, λ_d, λ_t and λ_b. The mixing Lagrangian is written in (1.92) These parameters break the flavor group that for this model is assumed to be

$$\mathcal{F} = U(3)_q \times U(3)_u \times U(3)_d \times U(2)_{U+D} \times U(1)_{T+B}$$

The form that we assume for the λ mixing is reported in the main text in eq. (1.94), where we assume the numbers a, b and b' to be $O(1)$ complex numbers with arbitrary phases, while a' is $O(0.1)$ with an arbitrary phase in order to reproduce the CKM. The spurion non-abelian quantum numbers for these mixing parameters, with respect to the full flavor group are

	$U(3)_q$	$U(3)_u$	$U(3)_d$	$U(2)_{U+D}$	$U(1)_{T+B}$
λ_u	1	$\bar{3}$	1	2	0
λ_d	1	1	$\bar{3}$	2	-1
λ_t	1	$\bar{3}$	1	1	-1
λ_b	1	1	$\bar{3}$	1	0
$\varepsilon_q \delta_{iA}$	3	1	1	$\bar{2}$	0
$\varepsilon_{q_3} \delta_{i3}$	3	1	1	1	1

In this model, the Yukawa matrices are given by (1.93). We define the diagonalizing matrices by

$$Y_u = U_u \bar{Y}_u V_u^\dagger \,, \quad \widetilde{U}_d^\dagger Y_d = U_d \bar{Y}_d V_d^\dagger \,, \tag{1.161}$$

where the bar indicates the diagonal and real matrices. The explicit form these matrices in the limit $y_u \ll y_c \ll y_t$ and $y_d \ll y_s \ll y_b$ is

$$\begin{aligned}
U_u &= \begin{pmatrix} 1 & -ab^* \frac{y_c^2}{y_t^2} & a\frac{y_c}{y_t} \\ a^* b \frac{y_u^2}{y_t^2} & 1 & b\frac{y_c}{y_t} \\ -a^* \frac{y_c}{y_t} & -b^* \frac{y_c}{y_t} & 1 \end{pmatrix}, \quad V_u = \begin{pmatrix} 1 & -ab^* \frac{y_u y_c}{y_t^2} & a\frac{y_u y_c}{y_t^2} \\ a^* b \frac{y_u y_c}{y_t^2} & 1 & b\frac{y_c^2}{y_t^2} \\ -a^* \frac{y_u y_c}{y_t^2} & -b^* \frac{y_c^2}{y_t^2} & 1 \end{pmatrix}, \\
U_d &= \begin{pmatrix} 1 & -a'b'^* \frac{y_s^2}{y_b^2} & a' \frac{y_s}{y_b} \\ a'^* b' \frac{y_d^2}{y_b^2} & 1 & b' \frac{y_s}{y_b} \\ -a'^* \frac{y_s}{y_b} & -b'^* \frac{y_s}{y_b} & 1 \end{pmatrix}, \quad V_d = \begin{pmatrix} 1 & -a'b'^* \frac{y_d y_s}{y_b^2} & a' \frac{y_d y_s}{y_b^2} \\ a'^* b' \frac{y_d y_s}{y_b^2} & 1 & b' \frac{y_s^2}{y_b^2} \\ -a'^* \frac{y_d y_s}{y_b^2} & -b'^* \frac{y_s^2}{y_b^2} & 1 \end{pmatrix}.
\end{aligned} \tag{1.162}$$

Finally we consider the tree-level flavor structures of this model in the mass basis.

$$S^{d_L d_R} = \bar{Y}_d + (x-1) U_d^\dagger \widetilde{Y}_b V_d \tag{1.163}$$

$$S^{u_L d_R} = V_{\rm CKM} S^{d_L d_R} \tag{1.164}$$

$$S^{d_R d_R} = \frac{1}{g_* \varepsilon_q^2} \left[(\bar{Y}_d)^2 + \left(\frac{\varepsilon_q^2}{\varepsilon_{q_3}^2} x - 1 \right) V_d^\dagger (\widetilde{Y}_b)^2 V_d \right] \tag{1.165}$$

$$S^{u_L u_R} = \bar{Y}_u + (x-1) U_u^\dagger \widetilde{Y}_t V_u \tag{1.166}$$

$$S^{d_L u_R} = V_{\rm CKM}^\dagger S^{u_L u_R} \tag{1.167}$$

$$S^{u_R d_R} = \frac{1}{g_* \varepsilon_q^2} \left[\bar{Y}_u V_{\rm CKM} \bar{Y}_d + \left(\frac{\varepsilon_q^2}{\varepsilon_{q_3}^2} x - 1 \right) V_u^\dagger \widetilde{Y}_t \widetilde{Y}_b V_d \right] \tag{1.168}$$

$$S^{u_R u_R} = \frac{1}{g_* \varepsilon_q^2} \left[(\bar{Y}_u)^2 + \left(\frac{\varepsilon_q^2}{\varepsilon_{q_3}^2} x - 1 \right) V_u^\dagger (\widetilde{Y}_t)^2 V_u \right] \tag{1.169}$$

$$S^{u_L u_L} = g_* U_u^\dagger {\rm diag}(\varepsilon_q^2, \varepsilon_q^2, x\varepsilon_{q_3}^2) U_u \tag{1.170}$$

$$S^{u_L d_L} = S^{u_L u_L} V_{\rm CKM} \tag{1.171}$$

$$S^{d_L d_L} = V_{\rm CKM}^\dagger S^{u_L u_L} V_{\rm CKM} \tag{1.172}$$

where we have defined the $\widetilde{Y}_t$ and $\widetilde{Y}_b$ matrices as

$$\widetilde{Y}_{t/b} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & y_{t/b} \end{pmatrix} , \tag{1.173}$$

and the x parameters are $O(1)$ numbers.

2 Quadratic Classifiers for EFT

The amazing richness of LHC data makes searching for new physics an extremely complex process. Three main steps can be identified, taking however into account that they are strongly interconnected and not necessarily sequential in time. First, we need a target new physics theory. In our case this is provided by the Standard Model (SM) itself, supplemented by operators of energy dimension $d > 4$ that encapsulate the indirect effects of heavy new particles and interactions. This setup is often dubbed the SM Effective Field Theory (EFT) in the context of high-energy physics (see e.g. Refs. [50, 11, 51]). However the EFT approach is extremely common and widely employed in many other domains, eminently in Flavor Physics. The methodologies discussed in this chapter could thus find applications also in other areas.

The second step is to turn the new physics theory into concrete predictions. These should be sufficiently accurate, since the EFT operator effects are often a small correction to the pure SM predictions. The predictions are provided by Monte Carlo generator codes, which produce event samples that are representative of the true particles momenta distributions. Accurate simulation of the detector response are further applied in order to obtain a representation of the distribution as a function of the variables that are actually observed in the experiment. It should be mentioned that this program occasionally fails. Namely it could be impossible for the Monte Carlo codes to provide a sufficiently accurate representation of all the components of the data distribution, for instance of reducible backgrounds from misidentification. In this case, the artificial Monte Carlo event sample should be supplemented with natural data collected in some control region, which model the missing component. We will not discuss this case explicitly, however it should be emphasized that our methodology would apply straightforwardly. Namely, the "Monte Carlo" samples we refer to in what follows might well not be the output of a Monte Carlo code, but rather have (partially) natural origin.

The final step, i.e. the actual comparison of the predictions with the data, is often further split in two, by identifying suitable high-level observables (e.g. cross sections in bins)

that are particularly sensitive to the EFT operators. These observables can be measured in an experimental analysis that does not target the EFT explicitly, and compared with the EFT predictions at a later stage. This is convenient from the experimental viewpoint because the results are model independent and thus potentially useful also to probe other new physics theories. If the measurements are performed at the truth (unfolded) level, this is also convenient for theorists because detector effects need not to be included in the predictions. However a strategy based on intermediate high-level observables is unavoidably suboptimal. It would approach optimality only if the fully differential distribution was measured for all the relevant variables, with sufficiently narrow binning. However there are often too many discriminating variables to measure their distribution fully differentially, and, even if this was feasible, one would not be able to predict accurately the cross section in too many bins. In this situation, the sensitivity to the presence (or absence) of the EFT operators could be strongly reduced and it could be impossible to disentangle the effect of different operators and resolve flat directions in the parameter space of the EFT Wilson coefficients. One should thus switch to the direct comparison of the EFT with the data, by employing more sophisticated unbinned multivariate data analysis techniques.

Several multivariate methods have been developed and applied to the EFT or to similar problems, including Optimal Observables [52, 53], the so-called "Method of Moments" [54–56] and similar approaches (see e.g. Ref. [57]) based on parametrizations of the scattering amplitude. The virtue of these techniques is that they are still based on high-level observables, making data/theory comparison simpler. The disadvantage is that they are intrinsically suboptimal and not systematically improvable towards optimality.

A potentially optimal approach, which is closely analog to the one based on Machine Learning we employ in this chapter, is the "Matrix Element Method" [58–62]. The main idea behind this construction is that optimal data analysis performances are unmistakably obtained by employing the likelihood $\mathcal{L}(c\,|\mathcal{D})$, i.e. the probability density of the observed dataset "$\mathcal{D}$" seen as a function of the free parameters "c" of the probability model. In our case, the free parameters coincide with the EFT Wilson coefficients.[1] The LHC data consist of independent repeated measurements of the variable "x" that describes the kinematical configuration of each observed event. Therefore the likelihood factorizes and evaluating it requires only the knowledge of the distribution of x. More precisely, since we are interested in $\mathcal{L}(c\,|\mathcal{D})$ only up to an overall c-independent normalization, it is sufficient to know the ratio $r(x, c)$ between the density as a function of c (and x) and the density at some fixed value $c = \overline{c}$. The SM point, $\overline{c} = 0$, can be conventionally chosen.

It should be emphasized that extracting $r(x, c)$ is a highly non-trivial task, as nicely

[1] The notion of "optimality" can be made fully rigorous and quantitative, both when the purpose of the analysis is to measure the free parameters of the EFT and when it is to test the EFT hypothesis ($c \neq 0$) against the SM ($c = 0$) one, and both from a Bayesian and from a frequentist viewpoint. The case of a frequentist hypothesis test is discussed in more details below.

explained in Ref. [63] in terms of latent variables. The Monte Carlo generator code does of course implement an analytic point-by-point representation of the density (and, in turn, of r), which is however expressed in terms of abstract variables "z" and not of the variables x that are actually observed. The analytic representation of r in the z variables can be used as a surrogate of $r(x, c)$ only if there is a faithful one-to-one correspondence between z and x. This is typically the case at leading order, if showering and detector effects are small, and if there are not undetected particles. However it is sufficient to have neutrinos in the final state, or to include Next to Leading Order (NLO) corrections to spoil the correspondence between z and x. Showering, hadronization and detector effects also wash out the correspondence. In the Matrix Element Method, $r(x, c)$ is obtained by a phenomenological parameterization of these effects in terms of transfer functions that translate the knowledge of the density at the "z" level into the one at the "x" level. The free parameters of the phenomenological modeling of the transfer functions are fitted to Monte Carlo samples.

The Matrix Element Method is potentially optimal and improvable towards optimality. However it is not "systematically" improvable, in the sense that a more accurate reconstruction of $r(x, c)$ requires a case-by-case optimization of the transfer function modeling. With the alternative employed in this chapter, based on the reconstruction of $r(x, c)$ using Machine Learning techniques rather than phenomenological modeling, systematic improvement is possible using bigger Neural Networks and larger training sets. Furthermore refining the reconstruction by including additional effects requires substantial effort and increases the computational complexity of the Matrix Element Method, while the complexity of the Machine Learning-based reconstruction is a priori independent of the degree of refinement of the simulations. Therefore it is important to investigate these novel techniques as an alternative and/or as a complement to the Matrix Element approach.

There is already a considerable literature on the reconstruction of $r(x, c)$ using Neural Networks [64–68, 63, 69] and several algorithms exist. Here we adopt the most basic strategy, mathematically founded on the standard Statistical Learning problem of classification (see Section 2.1.1 for a brief review), which we improve by introducing the notion of "Quadratic Classifier". The relation between our methodology and the existing literature, the possibility of integrating it in other algorithms and to apply it to different problems is discussed in details in Section 2.1.2 and in the Conclusions. However it is worth anticipating that, unlike simulator-assisted techniques [67], the Quadratic Classifier only exploits Monte Carlo data samples (in the extended sense outlined above) and no other information on the data generation process. It can thus be used as it is with any Monte Carlo generator.

Apart from describing the Quadratic Classifier, the main aim of this chapter is to investigate the potential impact of Machine Learning methods on LHC EFT searches, from two viewpoints.

The first question we address is if and to what extent statistically optimal sensitivity to the presence or absence of the EFT operators can be achieved. In order to answer, a rigorous quantitative notion of optimality is defined by exploiting the Neyman–Pearson lemma [70], namely the fact that the "best" (maximum power at fixed size, in the standard terminology of e.g. Ref. [71]) frequentist test between two simple hypotheses is the one that employs the likelihood ratio as test statistic. By regarding the EFT at each given value of the c Wilson coefficients as a simple hypothesis, to be compared with the SM $c = 0$ hypothesis, we would thus obtain the strongest expected 95% Confidence Level (CL) exclusion bounds on c (when the SM is true) if the true distribution ratio $r(x, c)$ was available and used for the test. This can be compared with the bound obtained by employing the approximate ratio $\hat{r}(x, c)$ reconstructed by the Neural Network, allowing us to quantify the approximation performances of the method in objective and useful terms.[2] Of course, the exact $r(x, c)$ is not available in a realistic EFT problem, therefore the comparison can only be performed on a toy problem. In order to make it as close as possible to reality, our "Toy" problem is defined in terms of an analytical approximation of the differential cross section of the process of interest (i.e. fully leptonic ZW, see below), implemented in a dedicated Monte Carlo generator.

The second aim of this chapter is to quantify the potential gain in sensitivity of Machine Learning techniques, compared with the basic approach based on differential cross section measurements in bins. The associated production of a Z and a W boson decaying to leptons, at high transverse momentum ($p_{T,\mathrm{Z}} > 300$ GeV) and with the total integrated luminosity of the High Luminosity LHC (HL-LHC), is considered for illustration. This final state has been selected to be relatively simple, but still described by a large enough number of variables to justify the usage of unbinned analysis techniques. Moreover it has been studied already quite extensively in the EFT literature (see e.g. Refs. [72–79]) and a number of variables have been identified, including those associated with the vector bosons decay products [80, 81, 76], with the potential of improving the sensitivity to the EFT operators.

The comparison with the binned analysis is performed on the Toy version of the problem mentioned above, on the exact tree-level (LO) modeling of the process and on NLO QCD plus parton showering Monte Carlo data. By progressively refining our modeling of the problem in these three stages, this comparison also illustrates the flexibility of the approach and the fact that increasingly sophisticated descriptions of the data are not harder for the machine to learn. This should be contrasted with the Matrix Element method, which would instead need to be substantially redesigned at each step. As an illustration, we will show that employing the analytical approximated distribution ratio, that was optimal on the Toy problem, leads to considerably worse performances than

[2] It should be emphasized that we adopt this specific notion of "optimality" only because the frequentist hypothesis test between two simple hypotheses is relatively easy to implement in a fully rigorous manner. The reconstructed likelihood ratio could be employed for any other purpose and/or relying on asymptotic approximations using standard statistical techniques.

the Neural Network already at LO. At NLO the performances further deteriorate and the reach is essentially identical to the one of the binned analysis.

The rest of the chapter is organized as follows. In Section 2.1 we introduce the Quadratic Classifier as a natural improvement of the standard Neural Network classifier for cases, like the one of the EFT, where the dependence of the distribution ratio on the "c" parameters is known. The fully leptonic ZW process, the EFT operators we aim at probing and the relevant kinematical variables, are discussed in Section 2.2. The Toy, the LO and the NLO Monte Carlo generators employed in the analysis are also described. The first set of results, aimed at assessing the optimality of the Quadratic Classifier on the Toy data, are reported in Section 2.3. The results obtained with the LO Monte Carlo are also discussed, showing the stability of the Neural Network performances as opposite to the degradation of the sensitivity observed with the Matrix Element and with the binned analysis methods. NLO results are shown in Section 2.4. We will see that the Quadratic Classifier methodology applies straightforwardly at NLO in spite of the fact that negative weights are present in the NLO Monte Carlo samples. The only complication associated with negative weights, which we discuss in Section 2.4.1, is not related with the reconstruction of the $\hat{r}(x, c)$ function by the Neural Network, but with the calculation of the distribution of the variable $\hat{r}(x, c)$ itself, which is needed for the hypothesis test. All the technical details on the Neural Network design and training are summarized in Section 2.5, and our conclusions are reported in Section 2.6. Appendices 2.A and 2.B contain the generalization of the Quadratic Classifier for an arbitrary number of Wilson coefficients and the proof of its asymptotic optimality.

2.1 Teaching new physics to a machine

Consider two hypotheses, H_0 and H_1, on the physical theory that describes the distribution of the variable x. In the concrete applications of the following sections, H_0 will be identified with the SM EFT and H_1 with the SM theory. The statistical variable $x \in \mathrm{X}$ describes the kinematical configuration in the search region of interest X. In the following, x will describe the momenta of 3 leptons and the missing transverse momentum, subject to selection cuts. Each of the two hypotheses (after choosing, if needed, their free parameters) characterizes the distribution of x completely. Namely they contain all the information that is needed to compute, in line of principle, the differential cross sections $d\sigma_0(x)$ and $d\sigma_1(x)$. The differential cross sections describe both the probability density function of x, which is obtained by normalization

$$\mathrm{pdf}(x|H_{0,1}) = \frac{1}{\sigma_{0,1}(\mathrm{X})} \frac{d\sigma_{0,1}}{dx}\,, \tag{2.1}$$

and the total number of instances of x (i.e. of events) that is expected to be found in the dataset, denoted as $\mathrm{N}(\mathrm{X}|H_{0,1})$. This is equal to the cross section integrated on X and multiplied by the luminosity of the experiment, namely $\mathrm{N}(\mathrm{X}|H_{0,1}) = \mathrm{L} \cdot \sigma_{0,1}(\mathrm{X})$.

The total number of observed events follows a Poisson distribution. Hence for a given observed dataset $\mathcal{D} = \{x_i\}$, with $\mathcal{N}$ observed events, the H_1/H_0 log likelihood ratio reads

$$\lambda(\mathcal{D}) \equiv \log \frac{\mathcal{L}(H_1|\mathcal{D})}{\mathcal{L}(H_0|\mathcal{D})} = \mathrm{N}(\mathrm{X}|H_0) - \mathrm{N}(\mathrm{X}|H_1) - \sum_{i=1}^{\mathcal{N}} \log \frac{d\sigma_0(x_i)}{d\sigma_1(x_i)} \,. \tag{2.2}$$

The statistic $\lambda(\mathcal{D})$ is known as the "extended" log likelihood ratio [82], and it is the central object for hypothesis testing (H_0 versus H_1) or for measurements (if H_0 contains free parameters), both from a Frequentist and from a Bayesian viewpoint. The "N" terms in eq. (2.2) can be computed as Monte Carlo integrals. What is missing in order to evaluate λ is thus the cross section ratio

$$r(x) \equiv \frac{d\sigma_0(x)}{d\sigma_1(x)} \,. \tag{2.3}$$

This should be known locally in the phase space as a function of x.

The physical knowledge of the H_0 and H_1 models gets translated into Monte Carlo generator codes, which allow us to estimate $\sigma_{0,1}(\mathrm{X})$ and to produce samples, $\mathrm{S}_{0,1}$, of artificial events following the $\mathrm{pdf}(x|H_{0,1})$ distributions. More precisely, the Monte Carlo generates weighted events $\mathrm{e} = (x_\mathrm{e}, w_\mathrm{e})$, with x_e one instance of x and w_e the associated weight. If the w_e's are not all equal, x_e does not follow the pdf of x and the expectation value of the observables $O(x)$ has to be computed as a weighted average. We choose the normalization of the weights such that they sum up to $\sigma_{0,1}(\mathrm{X})$ over the entire sample

$$\sum_{\mathrm{e}\in\mathrm{S}_{0,1}} w_\mathrm{e} = \sigma_{0,1}(\mathrm{X}) \,. \tag{2.4}$$

With this convention, the weighted sum of $O(x_\mathrm{e})$ approximates the integral of $O(x) \cdot d\sigma_{0,1}(x)$ on $x \in \mathrm{X}$. Namely

$$\sum_{\mathrm{e}\in\mathrm{S}_{0,1}} w_\mathrm{e} O(x_\mathrm{e}) \xrightarrow{\mathrm{LS}} \int_{x\in\mathrm{X}} d\sigma_{0,1}(x)\, O(x) = \sigma_{0,1}(\mathrm{X})\, \mathrm{E}\left[O|H_{0,1}\right] \,, \tag{2.5}$$

in the Large Sample (LS) limit where $\mathrm{S}_{0,1}$ are infinitely large. We will see below how to construct an estimator $\widehat{r}(x)$ for $r(x)$ (or, in short, to fit $r(x)$) using finite S_0 and S_1 samples.

For tree-level Monte Carlo generators the previous formulas could be made simpler by employing unweighted samples where all the weights are equal. However radiative corrections need to be included for sufficiently accurate predictions, at least up to NLO in the QCD loop expansion. NLO generators can only produce weighted events, and some of the events have a negative weight. Therefore the NLO Monte Carlo samples cannot be rigorously interpreted as a sampling of the underlying distribution. However provided they consistently obey the LS limiting condition in eq. (2.5), they are equivalent to

ordinary samples with positive weights for most applications, including the one described below.

2.1.1 The Standard Classifier

The estimator $\widehat{r}(x)$ can be obtained by solving the most basic Machine Learning problem, namely supervised classification with real-output Neural Networks (see Ref. [83] for an in-depth mathematical discussion). One considers a Neural Network acting on the kinematical variables and returning $f(x) \in (0,1)$. This is trained on the two Monte Carlo samples by minimizing the loss function

$$L[f(\cdot)] = \sum_{\mathrm{e}\in \mathrm{S}_0} w_{\mathrm{e}}[f(x_{\mathrm{e}})]^2 + \sum_{\mathrm{e}\in \mathrm{S}_1} w_{\mathrm{e}}[1-f(x_{\mathrm{e}})]^2\,, \tag{2.6}$$

with respect to the free parameters (called weights and biases) of the Neural Network. The trained Neural Network, $\widehat{f}(x)$, is in one-to-one correspondence with $\widehat{r}(x)$, namely

$$\widehat{f}(x) = \frac{1}{1+\widehat{r}(x)} \quad \Leftrightarrow \quad \widehat{r}(x) = \frac{1}{\widehat{f}(x)} - 1\,. \tag{2.7}$$

The reason why $\widehat{r}(x)$, as defined above, approximates $r(x)$ is easily understood as follows. If the Monte Carlo training data are sufficiently abundant, the loss function in eq. (2.6) approaches its Large Sample limit and becomes

$$L[f(\cdot)] \xrightarrow{\mathrm{LS}} \int_{x\in \mathrm{X}} d\sigma_0(x)[f(x)]^2 + \int_{x\in \mathrm{X}} d\sigma_1(x)[1-f(x)]^2\,. \tag{2.8}$$

Furthermore if the Neural Network is sufficiently complex (i.e. contains a large number of adjustable parameters) to be effectively equivalent to an arbitrary function of x, the minimum of the loss can be obtained by variational calculus. By setting to zero the functional derivative of L with respect to f one immediately finds

$$\widehat{f}(x) \simeq \frac{d\sigma_1(x)}{d\sigma_1(x)+d\sigma_0(x)} = \frac{1}{1+r(x)} \quad \Rightarrow \quad \widehat{r}(x) \simeq r(x)\,. \tag{2.9}$$

The same result holds for other loss functions such as the standard Cross-Entropy, which has been found in Ref. [66] to have better performances for EFT applications, or the more exotic "Maximum Likelihood" loss [84], which is conceptually appealing because of its connection with the Maximum Likelihood principle. We observed no strikingly different performances with the various options, but we did not investigate this point in full detail. In what follows we stick to the quadratic loss in eq. (2.6).

The simple argument above already illustrates the two main competing aspects that control the performances of the method and its ability to produce a satisfactory approximation of $r(x)$. One is that the Neural Network should be complex in order to attain

a configuration that is close enough to the (absolute) minimum, $f(x) = 1/(1 + r(x))$, of the loss functional in eq. (2.8). In ordinary fitting, this is nothing but the request that the fit function should contain enough adjustable parameters to model the target function accurately. On the other hand if the Network is too complex, it can develop sharp features, while we are entitled to take the Large Sample limit in eq. (2.8) only if f is a smooth enough function of x. Namely we need f to vary appreciably only in regions of the X space that contain enough Monte Carlo points. Otherwise the minimization of eq. (2.6) brings f to approach zero at the individual points that belong to S_0 sample, and to approach one at those of the S_1 sample. This phenomenon, called overfitting, makes that for a given finite size of the training sample, optimal performances are obtained by balancing the intrinsic approximation error of the Neural Network against the complexity penalty due to overfitting. A third aspect, which is extremely important but more difficult to control theoretically, is the concrete ability of the training algorithm to actually reach the global minimum of the loss function in finite time. This requires a judicious choice of the minimization algorithm and of the Neural Network activation functions.

The problem of fitting $r(x)$ is mathematically equivalent to a classification problem. A major practical difference however emerges when considering the level of accuracy that is required on $\widehat{r}(x)$ as an approximation of $r(x)$. Not much accuracy is needed for ordinary classification, because $\widehat{r}(x)$ (or, equivalently, $\widehat{f}(x)$) is used as a discriminant variable to distinguish instances of H_0 from instances of H_1 on an event-by-event basis. Namely, one does not employ $\widehat{r}(x)$ directly in the analysis of the data, but a thresholded version of $\widehat{r}(x)$ that isolates regions of the X space that are mostly H_0-like (r is large) or H_1-like (r is small). Some correlation between $\widehat{r}(x)$ and $r(x)$, such that $\widehat{r}(x)$ is large/small when $r(x)$ is large/small, is thus sufficient for good classification performances. Furthermore the region where $r(x) \simeq 1$ is irrelevant for classification.

The situation is radically different in our case because the EFT operators are small corrections to the SM. The regions where the EFT/SM distribution ratio is close to one cover most of the phase-space, but these regions can contribute significantly to the sensitivity if they are highly populated in the data sample. Mild departures of $r(x)$ from unity should thus be captured by $\widehat{r}(x)$, with good accuracy relative to the magnitude of these departures. Obviously the problem is increasingly severe when the free parameters of the EFT (i.e. the Wilson coefficients "c") approach the SM value $c = 0$ and $r(x)$ approaches one. On the other hand it is precisely when c is small, and the EFT is difficult to see, that a faithful reconstruction of $r(x)$ would be needed in order to improve the sensitivity of the analysis.

2.1.2 The Quadratic Classifier

Barring special circumstances, the EFT prediction for the differential cross section is a quadratic polynomial in the Wilson coefficients.[3] If a single operator is considered, so that a single free parameter c is present and the SM corresponds to the value $c = 0$, the EFT differential cross section reads

$$d\sigma_0(x;c) = d\sigma_1(x)\left\{[1 + c\,\alpha(x)]^2 + [c\,\beta(x)]^2\right\}\,, \tag{2.10}$$

where $\alpha(x)$ and $\beta(x)$ are real functions of x. An estimator $\widehat{r}(x,c)$ for the distribution ratio in the entire Wilson coefficients parameters space could thus be obtained as

$$\widehat{r}(x,c) = [1 + c\,\widehat{\alpha}(x)]^2 + [c\,\widehat{\beta}(x)]^2\,, \tag{2.11}$$

from estimators $\widehat{\alpha}(x)$ and $\widehat{\beta}(x)$ of the coefficient functions $\alpha(x)$ and $\beta(x)$. Notice that eq. (2.10) parametrizes, for generic $\alpha(x)$ and $\beta(x)$, the most general function of x and c which is quadratic in c, which is always positive (like a cross section must be) and which reduces to the SM cross section for $c = 0$. The equation admits a straightforward generalization for an arbitrary number of c parameters, which we work out in Appendix 2.A.

The estimators $\widehat{\alpha}(x)$ and $\widehat{\beta}(x)$ are obtained as follows. We first define a function $f(x;c) \in (0,1)$, in terms of two neural networks n_α and n_β with unbounded output (i.e. $\mathrm{n}_{\alpha,\beta} \in (-\infty,+\infty)$ up to weight-clipping regularization), with the following dependence on c:

$$f(x,c) \equiv \frac{1}{1 + [1 + c\,\mathrm{n}_\alpha(x)]^2 + [c\,\mathrm{n}_\beta(x)]^2}\,. \tag{2.12}$$

Next, we consider a set $\mathcal{C} = \{c_i\}$ of values of c and we generate the corresponding EFT Monte Carlo samples $\mathrm{S}_0(c_i)$. At least two distinct values of $c_i \neq 0$ need to be employed, however using more than two values is beneficial for the performances. Monte Carlo samples are also generated for the H_1 (i.e. $c = 0$) hypothesis, one for each of the $\mathrm{S}_0(c_i)$ samples. These are denoted as $\mathrm{S}_1(c_i)$ in spite of the fact that they are all generated according to the same $c = 0$ hypothesis. The samples are used to train the $\mathrm{n}_{\alpha,\beta}$ Networks, with the loss function

$$L[\mathrm{n}_\alpha(\cdot),\mathrm{n}_\beta(\cdot)] = \sum_{c_i\in\mathcal{C}}\left\{\sum_{\mathrm{e}\in\mathrm{S}_0(c_i)} w_\mathrm{e}[f(x_\mathrm{e},c_i)]^2 + \sum_{\mathrm{e}\in\mathrm{S}_1(c_i)} w_\mathrm{e}[1 - f(x_\mathrm{e},c_i)]^2\right\}. \tag{2.13}$$

We stress that in the second term in the curly brackets, the function f is evaluated on the same value of $c = c_i$ that is employed for the generation of the $\mathrm{S}_0(c_i)$ Monte Carlo

[3] The only exception is when the relevant EFT effects are modifications of the SM particles total decay widths. Also notice that the cross section is quadratic only at the leading order in the EFT perturbative expansion, which is however normally very well justified since the EFT effects are small. Higher orders could nevertheless be straightforwardly included as higher order polynomial terms.

sample which we sum over in the first term.

By taking the Large Sample limit for the loss function as in eq. (2.8), differentiating it with respect to n_α and n_β and using the quadratic condition (2.10), it is easy to show that the trained Networks $\widehat{\mathrm{n}}_\alpha$ and $\widehat{\mathrm{n}}_\beta$ approach α and β, respectively. Namely

$$\widehat{\alpha}(x) \equiv \widehat{\mathrm{n}}_\alpha(x) \simeq \alpha(x)\,, \qquad \widehat{\beta}(x) \equiv \widehat{\mathrm{n}}_\beta(x) \simeq \beta(x)\,. \tag{2.14}$$

More precisely, by taking the functional derivative one shows that the configuration $\mathrm{n}_\alpha = \alpha$ and $\mathrm{n}_\beta = \beta$ is a local minimum of the loss in the Large Sample limit. It is shown in Appendix 2.B that this is actually the unique global minimum of the loss.

It is simple to illustrate the potential advantages of the Quadratic Classifier, based on the analogy with the basic binned histogram approach to EFT searches. In that approach, the X space is divided in bins and the likelihood ratio is approximated as a product of Poisson distributions for the countings observed in each bin. Rather than $\widehat{r}(x, c)$, the theoretical input required to evaluate the likelihood are estimates $\widehat{\sigma}_0(\mathrm{b}; c)$ for the cross sections integrated in each bin "b". Employing the Standard Classifier approach to determine $\widehat{r}(x, c)$ would correspond in this analogy to compute $\widehat{\sigma}_0(\mathrm{b}; c)$ for each fixed value of c using a dedicated Monte Carlo simulation. By scanning over c on a grid, $\widehat{\sigma}_0(\mathrm{b}; c)$ would be obtained by interpolation. Every EFT practitioner knows that this is a is very demanding and often unfeasible way to proceed. Even leaving aside the computational burden associated with the scan over c, the problem is that the small values of c (say, $c = \bar{c}$) we are interested in probing typically predict cross sections that are very close to the SM value and it is precisely the small relative difference between the EFT and the SM predictions what drives the sensitivity. A very small Monte Carlo error, which in turn requires very accurate and demanding simulations, would be needed in order to be sensitive to these small effects. In the Standard Classifier method, the counterpart of this issue is the need of generating very large samples for training the Neural Network. Furthermore, this should be done with several values of c for the interpolation. This approach is computationally demanding even when a single Wilson coefficient is considered, and it becomes rapidly unfeasible if c is a higher-dimensional vector of Wilson coefficients to be scanned over.

The strategy that is normally adopted in standard binned analyses is closely analog to a Quadratic Classifier. One enforces the quadratic dependence of $\sigma_0(\mathrm{b}; c)$ on c as in eq. (2.10), and estimates the three polynomial coefficients (i.e. the SM cross section and the analog of α and β) in each bin by a χ^2 fit to $\widehat{\sigma}_0(\mathrm{b}; c)$, as estimated from the Monte Carlo simulations for several values of c. The values of c used for the fit are much larger that the reach of the experiment $c = \bar{c}$, so that their effects are not too small and can be captured by the Monte Carlo simulation. The Quadratic Classifier works in the exact same way. It can learn $\widehat{\alpha}(x)$ and $\widehat{\beta}(x)$ using training samples generated with large values of c, for which the difference between the $\mathrm{S}_0(c)$ and $\mathrm{S}_1(c)$ is sizable. The training

can thus recognize this difference, producing accurate estimates of $\widehat{\alpha}(x)$ and $\widehat{\beta}(x)$. This accurate knowledge results in an accurate estimate of $\widehat{r}(x, c)$ and of its departures from unity even at the small value $c = \bar{c}$, because our method exploits the exact quadratic relation in eq. (2.10).

It should be noted that the "Quadratic Classifier" introduced in eq. (2.12) is "Parametrized" in the sense that it encapsulates the dependence on the c parameters, but it is the exact opposite of the Parametrized Neural Network (or Parametrized Classifier) of Ref. [65]. In that case, the Wilson Coefficient c is given as an input to the Neural Network, which acts on an enlarged (x, c) features space. The purpose is to let the Neural Network learn also the dependence on c of the distribution ratio in cases where this is unknown. Here instead we want to enforce the quadratic dependence of the distribution ratio on c, in order to simplify the learning task.

An alternative strategy to exploit the analytic dependence on c is the one based on "morphing" [68]. Morphing consists in selecting one point in the parameter space for each of the coefficient functions that parametrize $d\sigma_0(x; c)$ as a function of c, and expressing $d\sigma_0(x; c)$ as a linear combination of the cross-sections computed at these points. For instance, a total of 3 "morphing basis points", $c_{1,2,3}$, are needed for a single Wilson coefficient and quadratic dependence, and $d\sigma_0(x; c)$ is expressed as a linear combination of $d\sigma_0(x; c_{1,2,3})$. This rewriting can be used to produce two distinct learning algorithms.

The first option is to learn the density ratios $d\sigma_0(x; c_{1,2,3})/d\sigma_1(x)$ individually (one-by-one or simultaneously), by using training data generated at the morphing basis points $c_{1,2,3}$, and to obtain $\widehat{r}(x, c)$ using the morphing formula. In the analogy with ordinary binning, this would correspond to extracting the dependence on c of the cross-sections by a quadratic interpolation of $\widehat{\sigma}_0(\mathrm{b}; c_{1,2,3})$ at the selected points. Of course it is possible to reconstruct the cross sections accurately also by using 3 very accurate simulations, rather than fitting less accurate simulations at several points. However a judicious choice of the values of $c_{1,2,3}$ is essential for a proper reconstruction of the quadratic and of the linear term of the polynomial. For the former, it is sufficient to take c very large, but for the latter a value of c should be selected that is neither too large, such that the quadratic term dominates by too many orders of magnitude, nor too small such that the constant SM term dominates. Notice that the optimal c depends on the analysis bin because the EFT effects relative to the SM (and the relative importance of the quadratic and linear terms) can be vastly different in different regions of the phase space. With "plain" morphing as described above, we are obliged to employ only few values of c, which might not be enough to cover the entire phase space accurately. With the Quadratic Classifier instead, all values of c that are useful to learn the distribution in some region of the phase space (see e.g. eq. (2.27)) can be included simultaneously in the training set.

Alternatively, one can use the morphing formula in place of eq. (2.10), producing a different parameterization of the classifier than the one in eq. (2.12), to be trained

with values of the parameters that are unrelated with the morphing basis points. The parametrization employed in the Quadratic Classifier is arguably more convenient, as it is simpler, universal and bounded to $f \in (0,1)$ interval owing to the positivity of eq. (2.10). Importantly, also the condition $\widehat{r}(x,0) \equiv 1$ is built-in in the Quadratic Classifier. However this could be enforced in the morphing parameterization as well by selecting $c = 0$ as one of the basis points. If this is done, we do not expect [4] a degradation of the performances if employing the morphing-based parametrization rather then ours. Indeed, we believe that the key of the success of the Quadratic Classifier that we observe stems from the appropriate choice of the values of c used for training, and not from the specific parametrization we employ. The non-optimal performances of the morphing strategy observed in Ref. [68] (on a different process than the one we study) are probably to be attributed to a non-optimal choice. Further investigations on this aspect are beyond the scope of this study.

2.2 Fully leptonic ZW

Consider ZW production at the LHC with leptonic decays, namely $Z \to \ell^+\ell^-$ and $W \to \ell\nu$, where $\ell = e, \mu$. As explained in the Introduction, this is arguably the simplest process, of established EFT relevance, where a multivariate approach is justified and potentially improves the sensitivity. We focus on the high-energy tail of the process, with a selection cut on the transverse momentum of the Z-boson, $p_{T,\mathrm{Z}} > 300$ GeV, because of two independent reasons. First, because at high energy we can approximate the differential cross section analytically and define a realistic enough Toy problem to assess the optimality of the method. Second, because at high-energy the statistics is sufficiently limited (less than 5×10^3 expected events at the HL-LHC, including both W charges) to expect systematic uncertainties not to play a dominant role. The reach we will estimate in the $p_{T,\mathrm{Z}} > 300$ GeV region, on purely statistical bases, should thus be nearly realistic.

The high-energy regime, in spite of the relatively limited statistics, is the most relevant one to probe those EFT operators that induce energy-growing corrections to the SM amplitudes. There are only two CP-preserving and flavor-universal operators in the ZW channel that induce quadratically energy-growing terms, namely [5]

$$\mathcal{O}_{\varphi q}^{(3)} = G_{\varphi q}^{(3)} \left(\overline{Q}_L \sigma^a \gamma^\mu Q_L\right) \left(i H^\dagger \overleftrightarrow{D}_\mu H\right), \qquad \mathcal{O}_W = G_W \varepsilon_{abc} W_\mu^{a\,\nu} W_\nu^{b\,\rho} W_\rho^{c\,\mu}. \tag{2.15}$$

We thus focus on these operators in our analysis. The BSM relevance of these operators is well established in the literature. Searching for $\mathcal{O}_{\varphi q}^{(3)}$ in di-boson final states is among the most promising probes of Higgs compositeness at the HL-LHC [75]. The operator $\mathcal{O}_W$ is generated with loop-suppressed coefficient in perturbative BSM models and in composite

[4] Provided that the possibility of having f outside the $(0,1)$ interval is not a problem when training, for instance, with the cross-entropy loss function.

[5] We use the definition $H^\dagger \overleftrightarrow{D}_\mu H = H^\dagger D_\mu H - (D_\mu H)^\dagger H$.

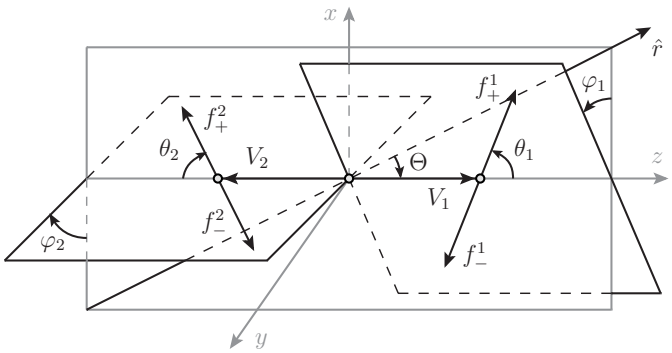


Figure 2.1: The kinematical variables in the "special' coordinate frame [76].

Higgs models of SILH [11] type. However it does not experience loop suppression for strongly-coupled BSM and is even potentially enhanced by the strong coupling in particular scenarios [85].

Both $\mathcal{O}_{\varphi q}^{(3)}$ and $\mathcal{O}_W$ contribute to the ZW production amplitudes with quadratically energy-growing terms of order $G \cdot s$, where s is the center-of-mass energy squared of the diboson system. However the way in which this energy growth manifests itself in the cross section is rather different for the two operators (see e.g. Refs. [76, 75]). The $\mathcal{O}_{\varphi q}^{(3)}$ operator mainly contributes to the "00" helicity amplitude, in which the gauge bosons are longitudinally polarized. The SM amplitude in this channel is sizable and has a constant behavior with s at high energy. As a consequence, a sizable quadratically-growing interference term between the SM and the BSM amplitudes is present in the cross section. This happens even at the "inclusive" level, i.e. when only the hard scattering variables describing ZW production (and not the decay ones) are measured.

On the contrary, the $\mathcal{O}_W$ operator induces quadratically-growing contributions only in the transverse polarization channels with equal helicity for the two gauge bosons (namely, $++$ and $--$). In the SM this channel is very suppressed at high energy, since its amplitude decreases as m_W^2/s. Therefore in inclusive observables the interference between $\mathcal{O}_W$ and the SM does not grow with the energy and is very small. In order to access (or "resurrect" [76]) the interference, which is the dominant new physics contribution since the Wilson coefficient of the operator is small, the vector bosons decay variables must be measured. We thus expect that the sensitivity to $\mathcal{O}_W$ will benefit more from an unbinned multivariate analysis technique than the one on $\mathcal{O}_{\varphi q}^{(3)}$.

The relevant kinematical variables that characterize the four-leptons final state are defined as in Ref. [76] and depicted in Figure 2.1, where V_1 is identified with the Z and V_2 with the W boson. The figure displays the kinematics in the rest frame of the ZW system, obtained from the lab frame by a boost along the direction of motion (denoted as $\hat{r}$ in the figure) of the ZW pair, followed by a suitable rotation that places the Z along

the positive z axis and $\hat{r}$ on the x-z plane with positive x component. The "inclusive" variables associated with ZW production are the center-of-mass energy squared s and $\Theta \in [0, \pi]$, which is defined as the angle between $\hat{r}$ and the Z-boson momentum. The decay kinematics is described by the polar and azimuthal decay angles $\theta_{1,2}$ and $\varphi_{1,2}$. The latter angles are in the rest frame of each boson and they are defined as those of the final fermion or anti-fermion with helicity $+1/2$ (e.g. the ℓ^+ in the case of a W^+ and the $\overline{\nu}$ for a W^-), denoted as $f_+^{1,2}$ in the figure.[6] The remaining variables that are needed to characterize the four leptons completely are weakly sensitive to the presence of the EFT operators and can be ignored, with the exception of the total transverse momentum of the ZW system, $p_{T,\mathrm{ZW}}$, which is a useful discriminant at NLO [75].

The variables described above are useful for the theoretical calculation of the cross section, but they cannot be used for our analysis because they are not experimentally accessible. The "measured" variables we employ are defined as follows. First, since we do not measure the neutrino (longitudinal) momentum, this needs to be reconstructed by imposing the on-shell condition for the W. The reconstructed neutrino momentum, rather than the true one, is used to define the kinematical variables and in particular s and Θ. Moreover, since we do not measure the helicity of the fermions but only their charge, the decay angles of the Z boson, denoted as θ_Z and φ_Z, are defined in terms of the charge-plus lepton rather than of the helicity plus lepton. Depending on the (unobserved) leptons helicities these angles are either equal to θ_1 and φ_1, or to $\pi - \theta_1$ and $\varphi_1 + \pi$, respectively. The W decay angles, defined in terms of the lepton or the reconstructed neutrino depending on the charge of the W as previously explained, are denoted as θ_W and φ_W. In summary, the variables we employ in the analysis are

$$\{s,\, \Theta,\, \theta_W,\, \varphi_W,\, \theta_Z,\, \varphi_Z,\, p_{T,\mathrm{ZW}}\}\,, \tag{2.16}$$

where of course $p_{T,\mathrm{ZW}}$ is non-vanishing only at NLO.

The on-shell condition for the W boson has no real solution if the W-boson transverse mass is larger that the W pole mass m_W. The neutrino is reconstructed in this case by assuming that the neutrino rapidity is equal to the one of the lepton. If instead the transverse mass is smaller than m_W, the condition has two distinct real solutions, each of which produces a different reconstructed kinematics. For our analysis we picked one of the two solutions at random on an event-by-event basis, while for the analysis of the actual data it would be arguably more convenient to duplicate the kinematical variables vector by including both solutions. Nothing changes in the discussion that follows if this second option is adopted.

[6] The correct definition of φ_2 appears in version four of Ref. [76].

2.2.1 Analytic approximation

At the tree-level order, and based on the narrow-width approximation for the decays, it is easy to approximate the cross section analytically in the high-energy regime. The crucial simplification is that the reconstructed 3-momentum of the W boson (with any of the two solutions for the neutrino) becomes exact when the W is boosted, so that the reconstructed Θ and s variables approach the "true" ones of Figure 2.1. Notice that Θ is the angle between the Z and the direction of motion of the ZW system in the lab frame, which corresponds at tree-level to the direction of motion of the most energetic incoming parton. In the kinematical region we are interested in, the (valence) quark is more energetic than the anti-quark in more than 80% of the events. Therefore we can identify Θ as the angle between the Z and the u quark or the d quark in the $u\overline{d} \to ZW^+$ and $d\overline{u} \to ZW^-$ processes, respectively.

With these identifications, the non-vanishing on-shell helicity amplitudes $\mathcal{M}_{h_Z h_W}$ for the hard scattering process $u\overline{d} \to ZW^+$, at leading order in the high-energy expansion, read

$$\mathcal{M}_{00} = -\frac{g^2 \sin\Theta}{2\sqrt{2}} - \sqrt{2} G^{(3)}_{\varphi q} s \sin\Theta\,, \qquad \mathcal{M}_{++} = \mathcal{M}_{--} = \frac{3 g c_{\rm w} G_W s \sin\Theta}{\sqrt{2}}\,, \tag{2.17}$$
$$\mathcal{M}_{-+} = -\frac{g^2(s^2_{\rm w} - 3\,c^2_{\rm w}\cos\Theta)}{3\sqrt{2}c_{\rm w}} \cot\frac{\Theta}{2}\,, \qquad \mathcal{M}_{+-} = \frac{g^2(s^2_{\rm w} - 3c^2_{\rm w}\cos\Theta)}{3\sqrt{2}c_{\rm w}} \tan\frac{\Theta}{2}\,,$$

where g is the $\mathrm{SU}(2)_L$ coupling, $c_{\rm w}$ and $s_{\rm w}$ are the cosine and the sine of the Weak angle. An overall factor equal to the cosine of the Cabibbo angle has not been reported for shortness. The amplitudes for the $d\overline{u} \to ZW^-$ process can be obtained from the ones above with the formal substitutions $\Theta \to -\Theta$ and $s^2_{\rm w} \to -s^2_{\rm w}$. The amplitudes are non-vanishing only for left-chirality initial quarks. Notice that the above formulas depend on the conventions in the definition of the wave-function of the external particles, and that these conventions must match the ones employed in the decay amplitude for the consistency of the final results. The wave-function reported in Ref. [86] are employed.

Let us now turn to the vector bosons decays. The decay amplitudes assume a very simple form in terms of the $\theta = \theta_{1,2}$ and $\varphi = \varphi_{1,2}$ variables, namely

$$\mathcal{A}_h = -\sqrt{2} g_V m_V e^{ih\varphi} d_h(\theta)\,, \tag{2.18}$$

where h is the helicity of the decaying vector boson ($V = V_{1,2} = Z, W$) and $d_h(\theta)$ are the Wigner d-functions. The overall coupling factor g_V depends on the nature of the boson and, in the case of the Z, on the electric charge of the helicity-plus fermion it decays to. Specifically, $g_W = g/\sqrt{2}$ for the W, $g_Z = g_L = -g(1 - 2\,s^2_{\rm w})/2c_{\rm w}$ if the Z decays to an helicity-plus ℓ^+ and $g_Z = g_R = g\,s^2_{\rm w}/c_{\rm w}$ if the Z decays to an helicity-plus ℓ^-. The two options for the helicity (which are physically distinct) correspond to two terms in the cross section. In the first one the Z decay amplitude is evaluated with the g_L coupling, with $\theta = \theta_1 = \theta_Z$ and $\varphi = \varphi_1 = \varphi_Z$. In the second one we employ g_R, $\theta = \theta_1 = \pi - \theta_Z$

and $\varphi = \varphi_1 = \varphi_Z + \pi$. There is no helicity ambiguity in the W-boson decay angles. However the reconstruction of the azimuthal decay angle is exact in the high-energy limit only up to a twofold ambiguity [76]. Namely the reconstructed φ_W approaches φ_1 on one of the two solutions for the neutrino (and we do not know which one), and $\pi - \varphi_1$ on the other. Since we are selecting one solution at random, we should average the cross section over the two possibilities $\varphi = \varphi_2 = \varphi_W$ and $\varphi = \varphi_2 = \pi - \varphi_W$ for the W azimuthal angle. The polar angle is instead $\theta = \theta_2 = \theta_W$ in both cases.

Production and decay are conveniently combined using the density matrix notation. We define the hard density matrix

$$d\rho^{\text{hard}}_{h_Z h_W h'_Z h'_W} = \frac{1}{24\, s} \mathcal{M}_{h_Z h_W} (\mathcal{M}_{h'_Z h'_W})^* \, d\Phi_{\text{ZW}} \,, \tag{2.19}$$

where $d\Phi_{\text{ZW}}$ is the two-body phase space and the factor $1/24s$ takes care of the flux and of the averages over the colors and the helicities of the initial quarks. The decay processes are instead encoded into decay density matrices. The one for the Z-boson, including the sum over the $\ell^\pm$ helicities as previously explained, reads

$$d\rho^{Z}_{h_Z h'_Z} = \frac{1}{2 m_Z \Gamma_Z} \left[\mathcal{A}_{h_Z} \mathcal{A}^*_{h'_Z} \Big|_{g_L, \theta_Z, \varphi_Z} + \mathcal{A}_{h_Z} \mathcal{A}^*_{h'_Z} \Big|_{g_R, \pi - \theta_Z, \varphi_Z + \pi} \right] d\Phi_{\ell^+ \ell^-} \,, \tag{2.20}$$

where Γ_Z is the Z decay width. For the W, since we average on the neutrino reconstruction ambiguity, we have

$$d\rho^{W}_{h_W h'_W} = \frac{1}{2 m_W \Gamma_W} \frac{1}{2} \left[\mathcal{A}_{h_W} \mathcal{A}^*_{h'_W} \Big|_{\frac{g}{\sqrt{2}}, \theta_W, \varphi_W} + \mathcal{A}_{h_W} \mathcal{A}^*_{h'_W} \Big|_{\frac{g}{\sqrt{2}}, \theta_W, \pi - \varphi_W} \right] d\Phi_{\ell\nu} \,. \tag{2.21}$$

The complete partonic differential cross section is finally simply given by

$$d\hat{\sigma} = 4 \sum d\rho^{\text{hard}}_{h_Z h_W h'_W h'_Z} \, d\rho^{Z}_{h_Z h'_Z} \, d\rho^{W}_{h_W h'_W} \,, \tag{2.22}$$

where the sum is performed on the four helicity indices and the factor of 4 takes into account the decay channels into electrons and muons.

2.2.2 Monte Carlo Generators

For our analysis we use three Monte Carlo generators, of increasing accuracy.

The first one is the Toy generator that implements the analytic approximation of the cross section in eq. (2.22), with the hard amplitudes expanded up to order $G \cdot s$ in the EFT contribution and up to order s^0 in the SM term, as in eq. (2.17). This implies, in particular, that in the Toy Monte Carlo all the mixed transverse/longitudinal helicity channels vanish exactly, that only the $\pm\mp$ and 00 channels are retained in the SM and

that new physics is just in the 00 and $\pm\pm$ channels for $\mathcal{O}_{\varphi q}^{(3)}$ and $\mathcal{O}_W$, respectively. The Toy Monte Carlo employs a simple fit to the ($u\bar{d}$ or $d\bar{u}$) parton luminosities obtained from the nCTEQ15 [87] PDF set (implemented through the `ManeParse` [88] Mathematica package). The variable s is sampled according to the parton luminosity, while all the other variables are sampled uniformly. The cut $p_{T,Z} = \sqrt{s}/2 \sin\Theta > 300$ GeV is implemented at generation level. Since the analytical distribution is extremely fast to evaluate, this basic approach is sufficient to obtain accurate Monte Carlo integrals and large unweighted event samples in a very short time.

The second generator is MadGraph [89] at LO, with the EFT operators implemented in the UFO model of Ref. [90]. We simulate the $2 \to 4$ process $pp \to \mu^+\mu^- e\,\nu_e$, with the Z and the W decaying to opposite flavor leptons for a simpler reconstruction, and multiply the resulting cross section by 4. The cut on $p_{T,\mathrm{Z}}$, defined as the sum of the μ^+ and μ^- momenta, is imposed at generation level, as well as the cuts

$$m_{T,e\nu} \leq 90\,\mathrm{GeV}\,, \qquad 70\,\mathrm{GeV} \leq m_{\mu\mu} \leq 110\,\mathrm{GeV}\,, \tag{2.23}$$

on the transverse mass of the virtual W and the invariant mass of the virtual Z. These are needed to suppress non-resonant contributions to the production of the 4 leptons. Standard acceptance cuts on the charged leptons are also applied. The unweighted events obtained with MadGraph are further processed to compute the kinematical variables in eq. (2.16) after neutrino reconstruction, as detailed at the beginning of this section.

The MadGraph LO generator is slightly more accurate than the Toy one. It contains all the ZW helicity amplitudes and no high-energy approximation. Furthermore, it describes non-resonant topologies and off-shell vector bosons production, which affects the reconstruction of the neutrino and in turn the reconstruction of the Z and W decay variables [76]. Nevertheless on single-variable distributions the Toy Monte Carlo and the LO one agree reasonably well, at around 10%.

The third and most refined generator is MadGraph at NLO in QCD, interfaced with Pythia 8.244 [91, 92] for QCD parton showering. The complete $2 \to 4$ process is generated like at LO, but no cuts could be applied at generation level apart from default acceptance cuts on the leptons and the lower cut on $m_{\mu\mu}$ in eq. (2.23). At NLO, the cut on $p_{T,\mathrm{Z}}$ needs to be replaced with the cut $p_{T,\mathrm{V}} > 300$ GeV, with $p_{T,\mathrm{V}} = \min[p_{T,\mathrm{Z}}, p_{T,\mathrm{W}}]$. This cut suppresses soft or collinear vector boson emission processes, which are insensitive to the EFT. In order to populate the $p_{T,\mathrm{V}} > 300$ GeV tail of the distribution with sufficient statistics, events were generated with a bias. The bias function was equal to one for $p_{T,\mathrm{V}}$ above 290 GeV, and equal to $(p_{T,\mathrm{V}}/290\,\mathrm{GeV})^4$ below. The momenta of the charged leptons and the transverse momentum of the neutrino in the generated events were read with MadAnalysis [93] and the kinematical variables in eq. (2.16) reconstructed like at LO. The cut $p_{T,\mathrm{V}} > 300$ GeV and the remaining cuts in eq. (2.23) were imposed on the reconstructed events. The total cut efficiency on the Monte Carlo data, thanks to

the bias, was large enough (around 17%) to allow for an accurate prediction of the cross section and for the generation of large enough event samples.

Even if ours is an electroweak process, it is known that NLO QCD corrections can in principle affect significantly the sensitivity to the EFT operators. Relevant effects are related with the tree-level zero [94] in the transverse amplitude, which is lifted at NLO, and with the appearance of same-helicity transverse high-energy amplitudes due to real NLO radiation [95]. All these effects are properly modeled by the MadGraph NLO generator.

2.3 Optimality on Toy data

Our goal is to reconstruct the EFT-over-SM cross section ratio $r(x, c)$ as accurately as possible using the methods introduced in Section 2.1. Since r is known analytically for the Toy problem, a simple qualitative assessment of the performances could be obtained by a point-by-point comparison (see Figures 2.7 and 2.8) of $r(x, c)$ with its approximation $\widehat{r}(x, c)$ provided by the trained Neural Network. However a point-by-point comparison is not quantitatively relevant, since the level of accuracy that is needed for $\widehat{r}(x, c)$ can be vastly different in different regions of the phase-space, depending on the volume of expected data and on the discriminating power of each region (i.e. on how much r is different from one).

The final aim of the entire construction is to obtain an accurate modeling of the extended log-likelihood ratio in eq. (2.2), to be eventually employed in the actual statistical analysis. A quantitative measure of the r reconstruction performances is thus best defined in terms of the performances of the final analysis that employs $\widehat{r}$, instead of r, in the likelihood ratio. Among all possible statistical analyses that could be carried out, frequentist tests to the EFT hypothesis $H_0(c)$ (regarded as a simple hypothesis for each given value of c), against the SM one, H_1, are considered for the illustration of the performances.

Four alternative test statistic variables are employed. One is the standard Poisson binned likelihood ratio (see below). The others are unbinned and take the form

$$t_c(\mathcal{D}) = \mathrm{N}(\mathrm{X}|H_0) - \mathrm{N}(\mathrm{X}|H_1) - \sum_{i=1}^{\mathcal{N}} \tau_c(x_i)\,, \tag{2.24}$$

where $\tau_c(x)$ is either equal to the exact $\log[r(x, c)]$ or to $\log[\widehat{r}(x, c)]$, as reconstructed either with the Standard Classifier or with the Quadratic Classifier described in Section 2.1.1 and 2.1.2, respectively. In each case the probability distributions of t in the two hypotheses are computed with toy experiments (or with the simpler strategy of Section 2.4.1), and used to estimate the expected (median) exclusion reach on c at 95% Confidence Level if the SM hypothesis is true. In formulas, the 95% reaches ($c_{2\sigma}$) we quote in what follows

are solutions to the implicit equation

$$p(t_{\text{med}}(c_{2\sigma}); c_{2\sigma}) = 0.05, \text{ with } t_{\text{med}}(c) = \text{Median}\left[t_c(\mathcal{D})|H_1\right] , \tag{2.25}$$

where the p-value is defined as

$$p(t_c; c) = \int_{t_c}^{\infty} dt'_c \, \text{pdf}(t'_c|H_0(c)) \, . \tag{2.26}$$

The two Wilson coefficients $c = G_{\varphi q}^{(3)}$ and $c = G_W$ are considered separately. Therefore the results that follow are single-operator expected exclusion reaches.

Summarizing, the four methodologies we employ are

i) *Matrix Element (ME)*

In this case we set $\tau_c(x) = \log[r(x, c)]$ in eq. (2.24), with r computed analytically using eq. (2.22). Therefore t coincides with the log-likelihood ratio λ in eq. (2.2), which in turn is the optimal discriminant between H_0 and H_1 due to the Neyman–Pearson lemma [70]. Namely, a straightforward application of the lemma guarantees that by employing $t = \lambda$ as test statistic we will obtain the optimal (smallest) $c_{2\sigma}$ reach, better than the one we could have obtained using any other variable. The Matrix Element Method is thus optimal in this case, and the optimality of the other methods can be assessed by comparing their $c_{2\sigma}$ reach with the one of the Matrix Element.

ii) *Standard Classifier (SC)*

The second method consists in setting $\tau_c(x) = \log[\widehat{r}(x, c)]$ in eq. (2.24), with $\widehat{r}$ reconstructed by the Standard Classifier as in Section 2.1.1. Notice that a separate training is needed to reconstruct $\widehat{r}(x, c)$ for each value of the Wilson Coefficient. Therefore computing $c_{2\sigma}$, as defined in eq. (2.25), requires scanning over c, performing first the Neural Network training and next the calculation of the distributions of t by toy experiments. For the Quadratic Classifier (and for the Matrix Element Method), the first step is not needed. The details on the Neural Network architecture and training, and of its optimization, will be discussed in Section 2.5.

iii) *Quadratic Classifier (QC)*

The third approach is to employ $\widehat{r}(x, c)$ as reconstructed by the Quadratic Classifier of Section 2.1.2. Implementation details are again postponed to Section 2.5, however it is worth anticipating that the key for a successful reconstruction is to train using values for the Wilson coefficients that are significantly larger than the actual reach.

Specifically, we used

$$\begin{aligned} G_{\varphi q}^{(3)} \quad &: \quad \{\pm 50, \pm 20, \pm 5\} \times 10^{-2}\ \text{TeV}^{-2}\,, \\ G_W \quad &: \quad \{\pm 20, \pm 10, \pm 5\} \times 10^{-2}\ \text{TeV}^{-2}\,. \end{aligned} \tag{2.27}$$

These values have been selected as those that induce order one departures from the SM cross section in the low, medium and high regions of the $p_{T,\text{Z}}$ distribution. If willing to compute cross-section in each $p_{T,\text{Z}}$ region by quadratic interpolation, using the values selected with this criterion can be shown to maximize the accuracy on the reconstruction of the linear term, while still allowing for a good determination of the quadratic term. We expect this choice to be optimal for the Quadratic Classifier training as well. Also notice that the total number of training Monte Carlo events is the same one (6M, see Section 2.5) employed for each of the separate trainings performed on the Standard Classifier.

iv) *Binned Analysis (BA)*

Finally, in order to quantify the potential gain of the unbinned strategy, we also perform a binned analysis. The test statistic in this case is provided by the sum over the bins of the log-ratio of the SM over EFT Poisson likelihoods, with the expected countings as a function of the Wilson coefficients computed from Monte Carlo simulations. The test statistic distributions, and in turn the reach by eq. (2.25), are computed with toy experiments like for the other methods and no asymptotic formulas are employed.

For both $G_{\varphi q}^{(3)}$ and G_W we considered 3 bins in $p_{T,\text{Z}}$, with the following boundaries

$$p_{T,\text{Z}}[\text{GeV}] \ : \quad \{300, 500, 750, 1200\}\ \text{GeV}\,. \tag{2.28}$$

The $p_{T,\text{Z}}$ variable is an extremely important discriminant because it is sensitive to the energy growth induced by the EFT. The three bins are selected based on the studies in Refs. [76, 75], and a narrower binning has been checked not to improve the sensitivity significantly. A cut $\cos\Theta < 0.5$ is imposed in the analysis targeting $G_{\varphi q}^{(3)}$, because this helps [75] in isolating the longitudinal helicity channel thanks to the amplitude zero in the transverse SM amplitudes. For $\mathcal{O}_W$, no $\cos\Theta$ cut is performed, and each $p_{T,\text{Z}}$ is split in two bins, for $\cos 2\,\varphi_W$ larger and smaller than zero. This is sufficient to partially capture the leading EFT/SM interference term as discussed in Ref. [76].

Most likely the binned analysis could be improved by considering more (and/or better) variables and a narrower binning. However it should be noticed that the simple strategies described above already result from an optimization, targeted to the specific operators at hand, and that the reach we obtain is consistent with the sensitivity projections available in the literature.

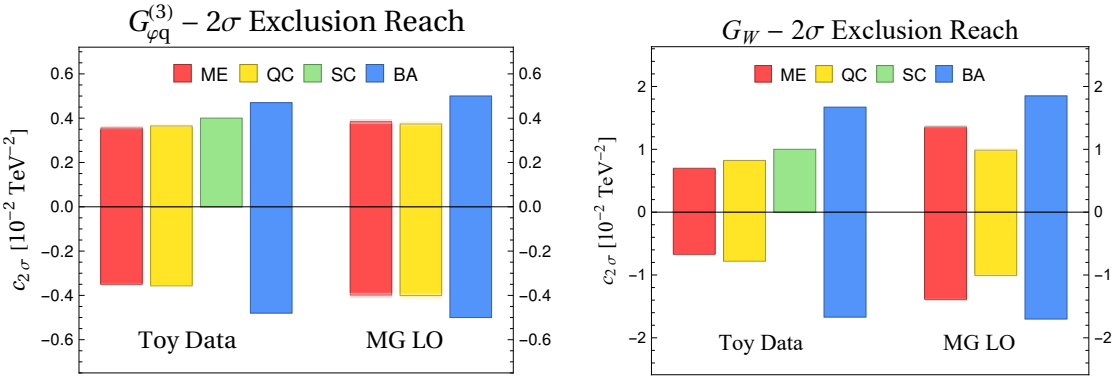


Figure 2.2: Expected exclusion reach on $G^{(3)}_{\varphi q}$ (left) and on G_W (right). The results are also reported in Table 2.1. Light-color stacked bars represent the errors.

2.3.1 Results

The results of the four methods are shown in Figure 2.2 (see also Table 2.1), together with the ones obtained with the MadGraph LO description of the ZW process, to be discussed in Section 2.3.2. The 2σ sensitivities reported in the figure are obtained by interpolating the median p-value as a function of the Wilson Coefficient c, in the vicinity of the reach, and computing $c_{2\sigma}$ by solving eq. (2.25). Further details on this procedure, and the associated error, are given in Section 2.4.1.

The figure reveals a number of interesting aspects. First, by comparing the Matrix Element reach with the one of the Binned Analysis we can quantify the potential gain in sensitivity offered by a multivariate strategy. The improvement is moderate (around 30%) for $G^{(3)}_{\varphi q}$, but it is more than a factor of 2 (of 2.4) in the case of the G_W operator coefficient. The different behavior of the two operators was expected on physical grounds, as discussed in details below. The figure also shows that the Quadratic Classifier is nearly optimal. More precisely, the reach is identical to the one of the Matrix Element for $G^{(3)}_{\varphi q}$, and $< 20\%$ worse for G_W. We will see in Section 2.5 that the residual gap for G_W can be eliminated with more training points than the ones used to produce Figure 2.2. Suboptimal performances are shown in the figure in order to outline more clearly, in Section 2.5, that our method is systematically improvable as long as larger and larger Monte Carlo samples are available.

Finally, we see in the figure that the Standard Classifier is slightly less sensitive than the Quadratic one, but still its performances are not far from optimality. This is reassuring in light of possible applications of Statistical Learning methodologies to different problems, where the dependence of the distribution ratio on the new physics parameters is not known and the Quadratic Classifier approach cannot be adopted. On the other hand, the Standard Classifier method is rather demanding. First, because it requires separate trainings on a grid of values of c, out of which the reach should be

extracted by interpolation. In turn, this requires a much larger number of training points than the Quadratic Classifier, since at each point of the grid we use as many training points as those the Quadratic Classifier needs in total for its training. Second, because we observed hyperparameters optimization depends on the specific value of c that is selected for training. Because of these technical difficulties, we only report sensitivity estimates for the positive Wilson coefficients reach. Furthermore these estimates (see Table 2.1) are based on the p-value obtained at a given point of the c grid without interpolation. For the same reason, we did not try to apply the Standard Classifier methodology to the LO and to the NLO data and we focus on the Quadratic Classifier in what follows.

Let us discuss now the physical origin of the different behaviors observed for the $\mathcal{O}_{\varphi q}^{(3)}$ and for the $\mathcal{O}_W$ operator. The point is that the new physics effects due to $\mathcal{O}_{\varphi q}^{(3)}$ have very distinctive features which can be easily isolated even with a simple binned analysis with few variables. Indeed $\mathcal{O}_{\varphi q}^{(3)}$ (see eq. (2.17)) only contributes to the 00 polarization amplitude, which is non-vanishing in the SM as well and proportional to $\sin\Theta$. The squared 00 amplitude thus contributes to the cross section with a sizable interference term, which is peaked in the central scattering region $\cos\Theta \sim 0$. The other helicity channels play the role of background, and are peaked instead in the forward region. They are actually almost zero (at LO) at $\cos\Theta \simeq 0$. Therefore a binned analysis targeting central scattering (this is why we imposed the cut $\cos\Theta < 0.5$) is sufficient to isolate the effects of $\mathcal{O}_{\varphi q}^{(3)}$ at the interference level and thus to probe $G_{\varphi q}^{(3)}$ accurately. By including the decay variables as in the multivariate analysis we gain sensitivity to new terms in the cross section, namely to the interference between the 00 and the transverse amplitudes, however these new terms are comparable with those that are probed already in the Binned Analysis and thus they improve the reach only slightly.

The situation is very different for the $\mathcal{O}_W$ operator. It contributes to the $++$ and $--$ helicity channels, that are highly suppressed in the SM and set exactly to zero in the Toy version of the problem we are studying here. The $p_{T,\mathrm{Z}}$ (and Θ) distribution depends only at the quartic level on G_W, i.e. through the square of the BSM amplitude, because the interference between different helicity channels cancels out if we integrate the cross section in eq. (2.22) over the ZW azimuthal decay angles. Our Binned Analysis is sensitive to the interference term through the binning in φ_W, however this is not enough to approach the optimal reach because all the other decay variables (and Θ as well) do possess some discriminating power, from which we can benefit only through a multivariate analysis. More specifically, one can readily see by direct calculation that the dependence on all our kinematical variables of the G_W interference contribution to the differential cross section is different from the SM term. By integrating on any of this variables we partially lose sensitivity to this difference, and this is why the multivariate analysis performs much better than the binned one.

2.3.2 MadGraph Leading Order

The analyses performed for the Toy dataset can be easily replicated for the MadGraph LO Monte Carlo description of the process, obtaining the results shown in Figure 2.2.

The most noticeable difference with what was found with the Toy Monte Carlo is the strong degradation of the Matrix Element reach, and the fact that it gets weaker than the one of the Quadratic Classifier. As usual, the effect is more pronounced for the $\mathcal{O}_W$ operator. This is not mathematically inconsistent because the analytic ratio $r(x, c)$ we employ for the Matrix Element test statistic is not equal anymore to the ratio of the true distributions according to which the data are generated. Therefore it is not supposed to give optimal performances. On the other hand the observed degradation is quantitatively surprising for G_W, especially in light of the fact that the MadGraph LO Monte Carlo distributions seem quite similar to the ones of the Toy data at a superficial look. The degradation is not due to the high-energy approximation in the ZW production amplitude, indeed the results we are reporting are obtained with the exact tree-level helicity amplitudes, which are employed in eq. (2.22) in place of the ones in eq. (2.17). It is due to the other approximations we performed in the calculation of the cross section, namely to the assumption that the initial quark is always more energetic than the anti-quark, which allows us to interpret Θ as the angle between the quark and the Z, and to the one of a perfect reconstruction (up to the ambiguity) of the neutrino momentum. We verified that this is the case by repeating the Matrix Element analysis using the true neutrino momentum and the actual direction of motion of the quark in the Monte Carlo events. In this case the reach on G_W gets closer to the one obtained with the Toy data.

The degradation of the Matrix Element reach should be contrasted with the relative stability of the Quadratic Classifier method. Notice that the method is applied on the MadGraph LO data in the exact same way as on the Toy data, namely the architecture is the same, as well as the number of training point and the values of the Wilson coefficients in eq. (2.27) used for training. The computational complexity of the distribution ratio reconstruction is thus identical in the two cases, in spite of the fact that the MadGraph LO Monte Carlo offers a slight more complete (or "complex") description of the data. The total computational cost is somewhat higher in the MadGraph LO case, but just because the process of Monte Carlo events generation is in itself more costly. Similar considerations hold at NLO, where the cost of event generation substantially increases.

2.4 The reach at Next-to-Leading Order

Including NLO QCD corrections is in general essential for an accurate modeling of the LHC data. Therefore it is imperative to check if and to what extend the findings of the

previous section are confirmed with the MadGraph NLO Monte Carlo description of the process, introduced in Section 2.2.2. As far as the reconstruction of $\widehat{r}(x,c)$ is concerned, using MadGraph NLO does not pose any conceptual or technical difficulty, provided of course the (positive and negative) Monte Carlo weights are properly included in the loss function as explained in Section 2.1. Computing the distribution of the test statistic variable that we obtain after the reconstruction (or of the one we employ with the Matrix Element method, for which the exact same issue is encountered) is instead slightly more complicated than with the Toy and MadGraph LO data. This point is discussed in the following section, while the illustration of the results is postponed to Section 2.4.2.

2.4.1 Estimating the test statistics distributions

As soon as $\tau_c(x)$ is known, either as an analytic function in the case of the Matrix Element or as a (trained) Neural Network in the case of the Quadratic Classifier, the test statistic $t_c(\mathcal{D})$, as defined in eq. (2.24), is fully specified. Namely we can concretely evaluate it on any dataset $\mathcal{D} = \{x_i\}$, consisting of $\mathcal{N}$ repeated instances of the variable x, for each given value of c. However in order to perform the hypothesis test, and eventually to estimate the reach $c_{2\sigma}$, we also need to estimate the probability distribution of $t_c(\mathcal{D})$ under the H_0 and under the H_1 hypotheses. This is the problematic step at NLO, after which the evaluation of $c_{2\sigma}$ proceeds in the exact same way as for the Toy and for the LO data. Specifically, once we are given with

$$\text{pdf}(t_c|H_0(c)) \;\text{ and }\; \text{pdf}(t_c|H_1)\,, \tag{2.29}$$

we obtain the p-value as a function of t_c and c as in eq. (2.26) from the former, while from the latter we compute the median value of t_c and in turn

$$p_{\text{med}}(c) \equiv p(t_{\text{med}}(c);c)\,, \tag{2.30}$$

as a function of c. After scanning over c and interpolating $p_{\text{med}}(c)$ in the vicinity of the reach (actually we interpolate the logarithm of $p_{\text{med}}(c)$, using three points in c and quadratic interpolation), we can solve the equation $p_{\text{med}}(c_{2\sigma}) = 0.05$ and obtain the reach as defined in eq. (2.25). Given the error on $p_{\text{med}}(c)$ at the three points used for the interpolation, the error on the estimate of $c_{2\sigma}$ is obtained by error propagation.

It is conceptually trivial (but numerically demanding) to estimate the distributions if artificial instances of the dataset $\mathcal{D}$ (aka "toy" datasets) are available. In this case one can simply evaluate $t_c(\mathcal{D})$ on many toy datasets following the $H_0(c)$ and the H_1 hypotheses and estimate the distributions. More precisely, one just needs the empirical cumulative in $H_0(c)$ and the median of t_c in H_1. Toy datasets are readily obtained from unweighted Monte Carlo samples by throwing $\mathcal{N}$ random instances of x from the sample, with $\mathcal{N}$ itself thrown Poissonianly around the total expected number of events. This is impossible at NLO because the events are necessarily weighted, therefore they are not a

		Toy Data	LO	NLO
$G_{\varphi q}^{(3)}$	ME	$[-0.350(6), 0.356(8)]$	$[-0.399(13), 0.384(12)]$	$[-0.55(4), 0.464(14)]$
	SC	$\gtrsim 0.4\,(p = 0.077(5))$	—	—
	QC	$[-0.357(6), 0.365(8)]$	$[-0.401(12), 0.374(10)]$	$[-0.426(22), 0.401(21)]$
	BA	$[-0.48, 0.47]$	$[-0.50, 0.50]$	$[-0.58, 0.55]$
G_W	ME	$[-0.673(14), 0.697(11)]$	$[-1.390(21), 1.357(22)]$	$[-1.51(7), 1.93(14)]$
	SC	$\lesssim 1\,(p = 0.038(3))$	—	—
	QC	$[-0.781(13), 0.822(13)]$	$[-1.007(27), 0.987(26)]$	$[-0.99(4), 1.08(12)]$
	BA	$[-1.67, 1.67]$	$[-1.70, 1.85]$	$[-1.63, 1.98]$

Table 2.1: Bounds on the $G_{\varphi q}^{(3)}$ and G_W coefficients obtained for the Toy, LO and NLO datasets. The rows correspond to the Matrix Element (ME), Standard Classifier (SC), Quadratic Classifier (QC) and Binned Analysis (BA) approach. Notice that the errors on the Binned Analysis bounds are negligible. The results are given in $10^{-2}\,\text{TeV}^{-2}$ units.

sampling of the underlying distribution of the variable x. As emphasized in Section 2.1, NLO Monte Carlo data can only be used to compute expectation values of observables $O(x)$ as in eq. (2.5). For instance we can compute the cross section in any region of the X space, and the mean or the higher order moments of the variable of interest, $\tau_c(x)$.

This suggests two options to estimate the distributions of the test statistic at NLO. The first one is to compute the distribution of $\tau_c(x)$ by means of a (weighted) histogram with many and very narrow bins. By knowing the cross section of each bin in τ_c, we know how many events are expected to fall in that bin and generate toy datasets for τ_c accordingly. This procedure is quite demanding, and it relies on a careful choice of the τ_c binning, which can only be performed on a case-by-case basis. It is still useful to validate the strategy we actually adopt, described below.

The second option is to approximate the distribution of t_c in a "nearly Gaussian" form, based on the Central Limit theorem. Namely we notice that t_c is in a trivial linear relation (see eq. (2.24)) with the variable

$$\mathcal{T}_c(\mathcal{D}) \equiv \frac{1}{\mathrm{N}} \sum_{i=1}^{\mathcal{N}} \tau_c(x_i)\,, \tag{2.31}$$

where $\mathcal{N}$ is Poisson-distributed with expected N, with $\mathrm{N} = \mathrm{N}(\mathrm{X}|H)$ and $H = H_0$ or $H = H_1$. The x_i's are independent and sampled according to $\mathrm{pdf}(x|H)$. The cumulant-generating function of $\mathcal{T}_c$ (which is a so-called "compound" Poisson variable [96]) is

readily computed

$$\mathrm{K}_{\mathcal{T}_c}(\xi) \equiv \log\left\{\mathrm{E}\left[e^{\xi\mathcal{T}_c}\middle|\ H\right]\right\} = \mathrm{N}\,\mathrm{E}\left[e^{\frac{\xi}{\mathrm{N}}\tau_c}\middle|\ H\right] - \mathrm{N}\,, \tag{2.32}$$

by first taking the expectation on the x_i's conditional to $\mathcal{N}$ and next averaging over the Poisson distribution of $\mathcal{N}$. Therefore the cumulants of $\mathcal{T}_c$,

$$\kappa^n_{\mathcal{T}_c} \equiv \left.\frac{d^n \mathrm{K}_{\mathcal{T}_c}(\xi)}{d\xi^n}\right|_{\xi=0} = \mathrm{N}^{1-n}\mathrm{E}\left[\,\tau_c^n\middle|\ H\right]\,, \tag{2.33}$$

are increasingly suppressed with N for larger and larger $n > 1$. Since N is of the order of several thousands in our case, neglecting all cumulants apart from the first and the second one, i.e. adopting a Gaussian distribution for $\mathcal{T}_c$, might be a good approximation.

Actually it turns out that in order to model properly the 5% tail of the distribution, which we need to probe for the exclusion limit, non-Gaussianity effects can be relevant. These are included by using a skew-normal distribution for $\mathcal{T}_c$, which contains one more adjustable parameter than the Gaussian to model the skewness. The mean, standard deviation and skewness of $\mathcal{T}_c$ are immediately obtained from eq. (2.33)

$$\mu(\mathcal{T}_c) = \langle\tau_c\rangle\,, \qquad \sigma(\mathcal{T}_c) = \frac{1}{\sqrt{\mathrm{N}}}\sqrt{\langle\tau_c^2\rangle} \qquad \mu_3(\mathcal{T}_c) = \frac{1}{\sqrt{\mathrm{N}}}\frac{\langle\tau_c^3\rangle}{\langle\tau_c^2\rangle^{3/2}}\,, \tag{2.34}$$

where $\langle\cdot\rangle$ is used to denote expectation for brevity. By computing the expectation values of τ_c, τ_c^2 and τ_c^3 using the Monte Carlo data, we thus find the parameters of the skew-normal distribution for $\mathcal{T}_c$ and in turn the distribution of t_c. We finally obtain the median p-value from the definition in eq. (2.30). The errors on the expectation values of τ_c are estimated from the fluctuations in the means on subsets of the entire Monte Carlo sample. These errors are propagated to the p-value and eventually to the $c_{2\sigma}$ estimated reach as previously explained. Accurate results (see Table 2.1) are obtained with relatively small Monte Carlo samples. Namely, 500k event were used at NLO, 1M at LO and 3M for the Toy data.

We cross-checked the above procedure in multiple ways. First, it reproduces within errors the LO and Toy p-values obtained with the toy experiments. Second, we validated it against the approach based on τ_c binning on NLO data, as previously mentioned. We also verified that including the skewness changes the results only slightly, with respect to those obtained in the Gaussian limit. Further improving the modeling of the non-Gaussiantiy with more complex distributions than the skew-normal, with more adjustable parameters in order to fit higher order moments of $\mathcal{T}_c$, is therefore not expected to affect the results.

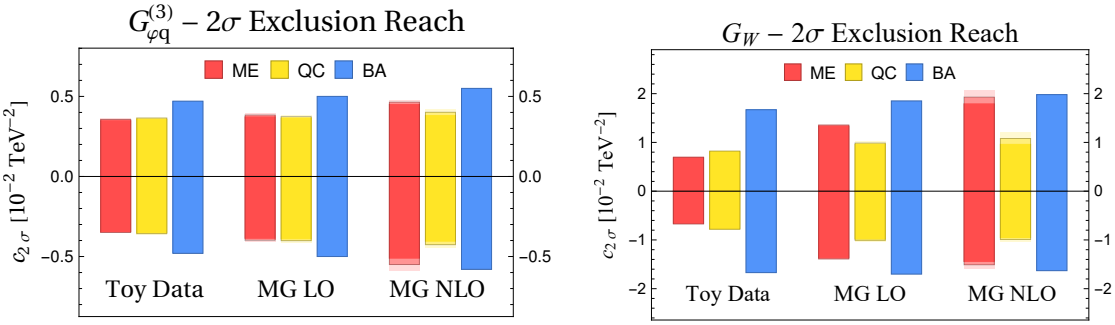


Figure 2.3: Expected exclusion reach on $G_{\varphi q}^{(3)}$ (left) and on G_W (right) with the various methodologies described in the text. The results are also reported in Table 2.1.

2.4.2 Results

Our results with the MadGraph NLO Monte Carlo are reported in Figure 2.3 and in Table 2.1. They essentially confirm the trend we already observed in the transition from the Toy to the MadGraph LO data. The Matrix Element keeps losing sensitivity because the analytic distribution ratio is now even more faraway from the actual distribution ratio since it does not include NLO QCD effects. The reach of the Binned Analysis deteriorates less, so that it becomes comparable to the one of the Matrix Element. The Quadratic Classifier reach is remarkably stable. Actually it slightly improves with respect to the LO one for G_W. This is probably due to the appearance of same-helicity SM transverse amplitudes (see Section 2.2.2) and of the corresponding interference term for the $\mathcal{O}_W$ operators.

Notice few minor differences in the implementation of the Quadratic Classifier and of the Binned Analysis at NLO. The Quadratic Classifier now also employs the variable $p_{T,\mathrm{ZW}}$, as discussed in Section 2.2. The Binned Analysis for $G_{\varphi q}^{(3)}$ employs $p_{T,\mathrm{ZW}}$ as well, through a cut $p_{T,\mathrm{ZW}}/p_{T,\mathrm{V}} < 0.5$. This improves the reach [75] because it helps recovering (partially) the background suppression due to the zero of the transverse amplitudes in the central region.

2.5 Neural Network implementation and validation

The strategies described in Section 2.1 were implemented in `Pytorch` [97] and run on NVIDIA GeForce GTX 1070 graphics card. Fully connected feedforward deep Neural Networks were employed, acting on the features vector

$$x = \{s,\, \Theta,\, \theta_W,\, \theta_Z,\, p_{T,\mathrm{ZW}},\, p_{T,\mathrm{Z}},\, \sin\varphi_W,\, \cos\varphi_W,\, \sin\varphi_Z,\, \cos\varphi_Z\}\,, \tag{2.35}$$

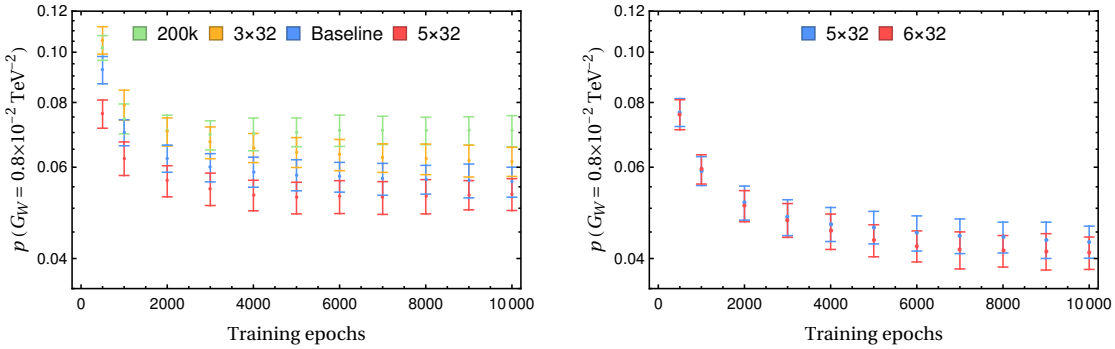


Figure 2.4: Evolution of the p-value for different architectures and training sample sizes. On the left plot we compare the baseline setup with the baseline architecture Network trained with 200k points per value of c (for a total of 2.4M points), and with the baseline number of training points (500k, times 12) on architectures with one less ("3×32") and one more ("5×32") hidden layer. On the right plot, a similar analysis is performed, but with 3M points per value of c.

for a total of 10 features. Each feature is standardized with a linear transformation to have zero mean and unit variance on the training sample. For the Quadratic Classifier training, the Wilson coefficient employed in the parametrization (2.12) were scaled to have unit variance on the training sample. Employing the redundant variables (i.e., $p_{T,\mathrm{Z}}$, and the cosines and sines of $\varphi_{W,Z}$) is helpful for the performances, especially the angular ones, which enforce the periodicity of the azimuthal angular variables. The "baseline" results presented in Figures 2.2, 2.3 and in Table 2.1 were all obtained with the features vector above and employing a total of 6 million training Monte Carlo points for each of the two Wilson coefficients. Training was always performed with a single batch (which was found to perform better in all cases), even if in practice the gradients calculation was split in mini-batches of 100k points in order to avoid saturating the memory of the GPU. Apart from these common aspects, the optimization of the Neural Network design and of the training strategy is rather different for the Quadratic and for the Standard Classifier methods. They are thus discussed separately in what follows.

2.5.1 The Quadratic Classifier

For the Quadratic Classifier, best performances were obtained with ReLU activation functions and with the Adam `Pytorch` optimizer. The initial learning rate (set to 10^{-3}) does not strongly affect the performances. Other attempts, with Sigmoid activation and/or with SGD optimizer, produced longer execution time and worse performances. The baseline architecture for the two Neural Networks n_α and n_β in eq. (2.12) consists of 4 hidden layers with 32 neurons, namely the architecture $\{10, 32, 32, 32, 32, 1\}$, including the input and the output layers. Weight Clipping was implemented as a bound on the L_1 norm of the weights in each layer, but found not to play a significant role. The total

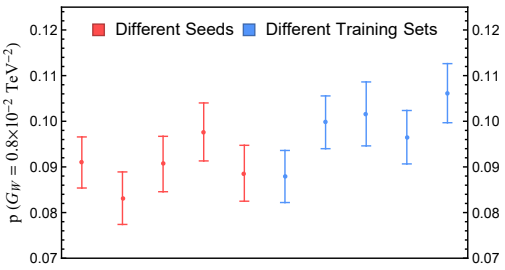


Figure 2.5: Results of 5 different trainings of the same architecture (Baseline architecture trained with 2.4M points) using: the same training data but different initialization seeds (red points) and the same initialization but different training data samples (blue points).

training time, for 10^4 training epochs, is around 5 hours for the baseline architecture and with the baseline number (6 million) of training points.

The Neural Network architecture was selected based on plots like those in Figure 2.4. The left panel shows the evolution with the number of training epochs of the median p-value (see eq. (2.30)) on Toy data for $c = G_W = 0.8 \times 10^{-2}\,\text{TeV}^{-2}$, with the baseline and with larger and smaller Networks. We see that adding or removing one hidden layer to the baseline architecture does not change the performances significantly. The plot also shows that 10^4 epochs are sufficient for the convergence and that no overfitting occurs. The degradation of the performances with less training point is also illustrated in the plot. Of course, the p-value is evaluated using independent Monte Carlo samples, not employed for training. The errors on the p-value are estimated from the error on the skew-normal distribution parameters as explained in Section 2.4.1. In the baseline configuration we used 500k EFT Monte Carlo training points for each of the 6 values of G_W in eq. (2.27), plus 500k for each associated SM sample. Each sample consists instead of 3M points in the extended configuration employed on the right panel of Figure 2.4, for a total of 36M. The same value of $G_W = 0.8 \times 10^{-2}\,\text{TeV}^{-2}$ is employed. The baseline architecture becomes insufficient, and best results are obtained with the 6 hidden layers of 32 neurons each.

The figure also demonstrates that the method is systematically improvable towards optimality. The value of G_W considered in the figure was not within the 95% CL reach with the baseline setup, while it becomes visible with the extended configuration. All the reaches reported in Table 2.1 would expectedly improve with the extended configuration. The G_W reach on Toy data becomes $[-0.732(9), 0.764(14)]\,10^{-2}\,\text{TeV}^{-2}$, which is now only less than 10% worse than the optimal Matrix Element reach. Training takes around 30 hours with the extended configuration, while generating and processing the required training points with MadGraph NLO (which is the most demanding generator) would take around 10 days on a 32-cores workstation. We could thus try to improve also the NLO reach even with limited computing resources.

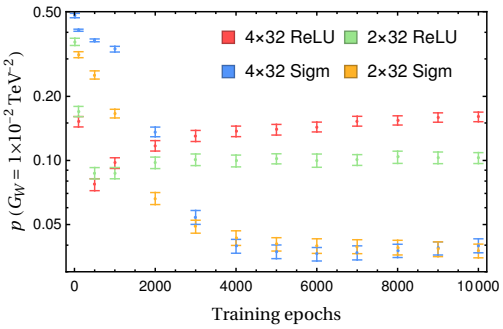


Figure 2.6: The p-value evolution during training for the Standard Classifier using different architectures and activation functions. The value $G_W = 1 \times 10^{-2}\,\mathrm{TeV}^{-2}$ is employed.

For the reproducibility of our results we also study how the performances depend on the Neural Network initialization and on the statistical fluctuations of the Monte Carlo training sample. This analysis is performed in a reduced setup, with a total of 2.4 million training point, and for $G_W = 0.8 \times 10^{-2}\,\mathrm{TeV}^{-2}$. We see in Figure 2.5 that the p-value fluctuates by varying the random seed used for training at a level comparable with the error on its determination. Similar results are observed by employing different independent Monte Carlo training samples. Notice that these fluctuations should not be interpreted as additional contributions to the error on the p-value. Each individual Neural Network obtained from each individual training defines a valid test statistic variable, on which we are allowed to base our statistical analysis. Since the fluctuations are comparable to the p-value estimate errors, our sensitivity projections were obtained by randomly selecting one of the seed/training set configuration.

2.5.2 The Standard Classifier

Hyperparameters optimization is rather different for the Standard Classifier. We see in Figure 2.6 that Networks with ReLU activation like those we employed for the Quadratic Classifier displays overfitting, and Sigmoid activations need to be employed. The results in Figures 2.2 and in Table 2.1 were obtained with 2 hidden layers with 32 neurons each and Sigmoid activation. The figure shows that increasing the complexity does not improve the performances.

This different behavior of the Standard Classifier compared with the Quadratic one is probably due to the fact that training is performed on small Wilson coefficient EFT data, whose underlying distribution is very similar to the one of the SM data sample. Therefore there is not much genuine difference between the two training sets, and the Network is sensitive to statistical fluctuations in the training samples. The Quadratic Classifier instead is trained with large values of the Wilson coefficients. The optimizer

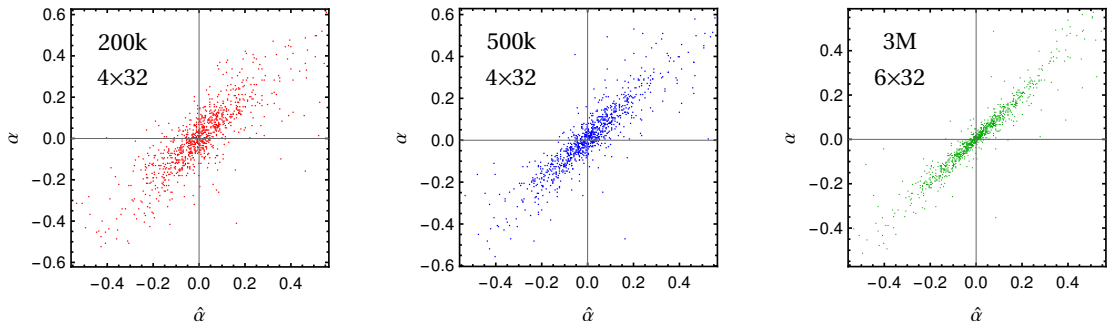


Figure 2.7: Comparison between the reconstructed ($\widehat{\alpha}$) and true (α) linear term of the distribution ratio for the G_W operator.

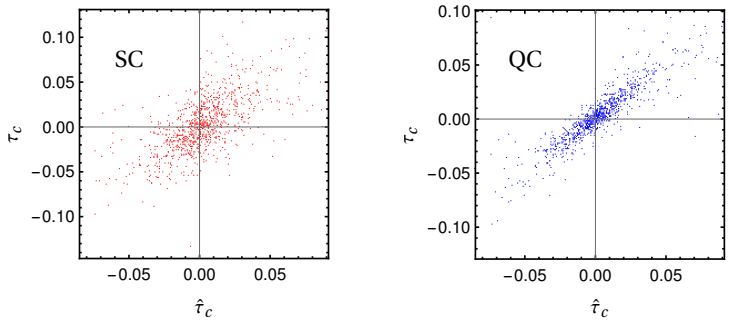


Figure 2.8: Comparison between the reconstructed ($\widehat{\tau}_c$) and true (τ_c) distribution log-ratio for $G_W = 1 \times 10^{-2}\,\mathrm{TeV}^{-2}$. The Standard Classifier and the Quadratic one are considered in the left and right panel of the figure, respectively.

thus drives the Neural Networks towards the deep minimum that corresponds to a proper modeling of the distribution ratio, which is more stable against statistical fluctuations of the training samples.

2.5.3 Validation

An important question is how to validate as "satisfactory" the outcome of the hyperparameters optimization described above. This is straightforward for the Toy version of the problem, because we have to our disposal a rigorous notion of statistical optimality, through the Neyman–Pearson lemma, and we do have direct access to the true distribution ratio through which the data are generated. Therefore we know that we can stop optimization as soon as the reach of the Neural Network becomes sufficiently close to the one of the Matrix Element method. We can also rely on a more naive validation test, based on comparing point-by-point the distribution ratio learned by the Neural Network with the true one, which is known analytically. For instance in Figure 2.7 we compare the true linear term $\alpha(x)$ in eq. (2.10) (for the $\mathcal{O}_W$ operator) with its estimator $\widehat{\alpha}(x) \equiv \widehat{\mathrm{n}}_\alpha(x)$ provided by the trained Neural Network. The baseline architecture is employed, with

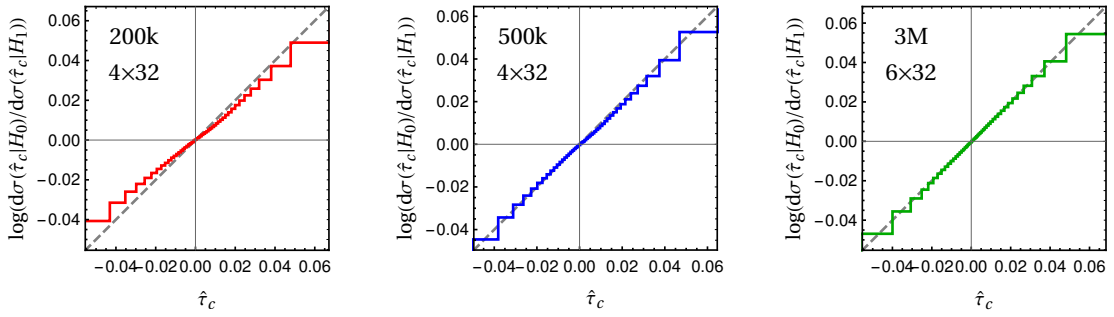


Figure 2.9: Distribution log-ratio for $\widehat{\tau}_c$, for $c = G_W = 0.8 \times 10^{-2}\,\text{TeV}^{-2}$. The accurate determination displayed in the plots is obtained by the reweighting of a single Toy SM Monte Carlo sample. The same approach, based on reweighting, could have been adopted to asses the quality of the distribution ratio reconstruction on MadGraph Monte Carlo data, using MadWeight.

increasing number of training points. While it is impossible to extract quantitative information, a qualitative comparison between the three scatter plots confirms that more training points improve the quality of the reconstruction. We also show, in Figure 2.8, the correlation between the true and the reconstructed ratios (for $G_W = 1 \times 10^{-2}\,\text{TeV}^{-2}$, which corresponds to the Standard Classifier 95% reach) obtained with the Quadratic and with the Standard Classifier. The reconstruction obtained with the Quadratic Classifier is more accurate as expected.

Validation is of course less easy if, as it is always the case on real problems, the true distribution ratio is not known. One option is to proceed like we did in the present work. Namely to identify a Toy version of the problem that is sufficiently close to the real one and for which the distribution ratio is known. Since it is unlikely that the true distribution is much harder to learn than the Toy distribution, and since we can establish optimality on the Toy data using a certain architecture and training dataset size, we can argue heuristically that the same configuration will be optimal also with a more refined Monte Carlo description.

Finally, one can monitor heuristically how accurately the distribution ratio is reconstructed, as follows. The true distribution log-ratio $\tau_c(x) = \log r(x, c)$, seen as a statistical variable for each fixed value of c, obeys, by definition, the equation

$$\frac{d\sigma_0}{d\tau_c} = e^{\tau_c} \frac{d\sigma_1}{d\tau_c}\,. \tag{2.36}$$

Therefore if we computed the distribution of τ_c (if it was known) in the EFT hypothesis $H_0(c)$ and in the SM hypothesis H_1, and take the log-ratio, the result would be a straight line as a function of τ_c. By computing the same distributions for the reconstructed distribution log-ratio $\widehat{\tau}_c = \log \widehat{r}(x, c)$, we can thus get an indication of how closely $\widehat{r}(x, c)$ approximates $r(x, c)$. While no quantitative information can be extracted from these

plots, they clearly illustrate the improvement achieved by enlarging the size of the training sample and the Neural Network architecture, as Figure 2.9 shows.

2.6 Conclusions and outlook

We studied the potential gain in sensitivity of EFT searches at the LHC from multivariate analysis techniques. The results reported in Figure 2.3 show that a considerable improvement is possible, especially for operators (like $\mathcal{O}_W$) with a complex interference pattern that is difficult to capture with a Binned Analysis.

Multivariate analyses based on Statistical Learning techniques are particularly promising, and should be considered as an alternative to the more standard (though not yet employed for EFT LHC searches) Matrix Element method. The advantage is eminently practical, because the Matrix Element method is optimal in principle, as much as the Statistical Learning approach. However the Matrix Element method needs to be designed case-by-case, and re-designed for each new effect one is willing to add for a more accurate modeling of the distribution ratio. It already required some effort to compute the approximate distribution in Section 2.2.1, which in turn provides the simplest modeling of the distribution ratio to be employed in the Matrix Element approach, and we saw that this modeling is inadequate to describe the LO and even less adequate at NLO. In order to improve the modeling in the case at hand one should model the neutrino reconstruction more accurately, for instance by performing the integral over the neutrino momentum point-by-point in the space of the observed kinematical variables. The integral on the radiation should be also performed if willing to add QCD NLO effects. The predictions should be further refined including transfer functions for the detector effects, if the method has to be employed on real data.

The situation is radically different with the Statistical Learning approach. We saw that the exact same computational effort is required to reconstruct the distribution ratio at the Toy level, at LO and at NLO. Furthermore the accuracy of the reconstruction can be systematically improved using more training points and bigger Networks. The limiting factor is not reconstructing the distribution ratio by the Neural Network training. That step takes a quite small fraction of the computing time. The most time-consuming part of the procedure is the generation of the Monte Carlo training data, which becomes increasingly demanding as the sophistication of the Monte Carlo code increases. Even if we are still far from the limit for our analysis, it would be worth investigating improvements on this aspect based on Monte Carlo reweighting techniques.

It should be emphasized that Machine Learning methodologies are useful for EFT studies not only in view of the possible application to the analysis of the real data. After the conceptual and technical framework is in place, it is very easy to run the Machine Learning algorithm on the specific EFT problem at hand, and to get a feeling of the

potential improvement of the reach compared with other methods. For instance our results show that the Binned Analysis we employed is inadequate for G_W, and that even for $G^{(3)}_{\varphi q}$ it could be improved. Furthermore they provide a target for the sensitivity such improvements should attain. Similarly, the results outline the importance of neutrino reconstruction modeling and of NLO QCD corrections being implemented in the Matrix Element method, if one is willing to adopt that strategy.

When it comes to the direct applicability of the method to the data, of the ZW process for instance, two additional steps are needed. The first one is to further improve the level of detail of the simulation. Detector effects could be added very easily with DELPHES [98]. However the reliability of the DELPHES description of the detectors should be cross-checked with a complete simulation by the experimental collaborations, and the DELPHES simulation replaced with a full detector simulation, which is much more demanding, if needed.

The second aspect is to include systematic uncertainties of theoretical and experimental origin. It should be stressed that this is not more problematic in the Machine Learning framework than it is in the Matrix Element or any other multivariate approach. In particular it should be noticed that one has full control on the choice of the input variables that are given to the Neural Network and from which the sensitivity emerges. For instance in our case these would be the kinematical variables of the high-level reconstructed leptons, better if including photon recombination, in order to reduce the sensitivity to detector effects and showering, which might not be modeled accurately enough. Similarly if jets were used in the final state, high level IR-safe observables would be employed to be insensitive to hadronization, exactly like one would do for the Matrix Element method. It should also be stressed, as explained in the Introduction, that our method can be employed also in the presence of reducible backgrounds that must be extracted from the data because no reliable Monte Carlo generator is available.

The simplest strategy to deal with uncertainties is to merely quantify their impact on the sensitivity, using as discriminating variable the distribution ratio reconstructed from the nominal Monte Carlo generator that does not incorporate uncertainties. This is suboptimal, but sufficient to obtain conservative (i.e. correct) results, and to identify the irrelevant sources of uncertainties. For better results one can include the uncertainties in the likelihood (i.e. in the reconstructed distribution ratio) in the form of nuisance parameters. This is perfectly compatible with the Machine Learning approach, and already implemented in MADMINER [69] through morphing. Actually the Quadratic Classifier we employ in this work could be useful also for this task. We will return to this point at the end of the Section. While conceptually straightforward, it is quantitatively important to assess the impact of uncertainties on the sensitivities we obtained in Figure 2.3 on purely statistical grounds. This is left to future work.

One interesting technical element of the present work is the Quadratic Classifier, in-

troduced in Section 2.1.2. We have found that it performs better than the Standard Classifier, as expected since it is designed to be sensitive to the small departures from the SM due to the EFT by exploiting the exact knowledge of the (quadratic) functional dependence of the distribution ratio on the Wilson coefficients. Furthermore it is computationally much more convenient and thus feasible also when several EFT operators are considered simultaneously and the scan over the Wilson coefficients becomes unfeasible. The Quadratic Classifier has been found to be nearly optimal, with a rigorous notion of optimality based on the Neyman–Pearson lemma.

We described in the body of the chapter the connection between the Quadratic Classifier and other techniques based on Statistical Learning available in the literature, but we did not yet discuss the relation with the most sophisticated such techniques, namely the ones that exploit "hidden" information from the Monte Carlo simulator [67]. The basic idea is that the simulator does contain the analytic information on the underlying distribution, and so it does contain a representation of the EFT/SM distribution ratio in terms of latent variables. One can incorporate this information in the loss function, so that the machine does not need to learn the likelihood ratio from scratch, but only the distortions of the likelihood ratio due to the transition between the latent and the true variables. The Quadratic Classifier trick is orthogonal to this interesting idea, and it could be straightforwardly implemented in the simulator-assisted methods by modifying the loss function in close analogy with eq. (2.13). The advantages of parametrization in that context could be the same we observed here.

On the other hand, simulator-assisted methods have also potential limitations, in two respects. First, because there is a clear benefit from exploiting the latent-space distribution ratio if the latter is similar to the one in the space of observables, but this is not necessarily the case. For instance in ZW we saw that a proper modeling of the neutrino reconstruction is crucial for the performances, and this is not captured by the latent-variables ratio that involves the true neutrino momentum. This can be a problem for the validation of the approach, due to the fact that any additional effect we include in the simulation, which further distorts the observed ratio, might be more and more difficult for the machine to learn. For instance a simulator-assisted method should be trivially optimal on the Toy data, where the latent space coincides with the observed space and thus the likelihood ratio employed in training coincides with the true one and the machine has nothing to learn. However this does not mean that it will work on the LO data (using the appropriate LO latent-variables ratio) because now the machine has the non-trivial task to integrate out the neutrino. Instead for our method, that learns the distribution ratio using no information from the Monte Carlo apart from the event sample itself, it is arguably equally difficult to model the distribution ratio on the Toy, on the LO and on the NLO data. Therefore the optimality on Toy data, which we can establish rigorously because we know the exact distribution ratio, heuristically indicates that the algorithm is optimal at LO and NLO as well. The second problem of simulator-assisted method is that the required information on the latent-space distribution ratio might not be made

available by the Monte Carlo code. In light of this, it is reassuring to have an alternative method that does not rely on latent-space information, that is feasible and optimal, at least in the case at hand.

Finally, it should be noticed that the parametrization trick is not specific of the EFT and it could be applied to any situation where the functional dependence of the distribution on the parameters is either exactly or approximately known. One should just replace the quadratic dependence of eq. (2.12) on c with the appropriate (polynomial or not) functional form. This could be useful to include the effect of nuisance parameters in the likelihood. Nuisance parameters effects on the distribution can be normally modeled linearly (or with an exponential, to avoid negative distributions) to good approximation because their effects are small. However if they are too small (but still potentially competitive with the EFT ones) it could be difficult for the machine to learn them using simulations where the nuisances are varied within their one-sigma interval. If the analytic dependence on the nuisance parameters is incorporated in the classifier, we could ameliorate the situation by training with larger values of the parameters like we did in this study to reconstruct the small EFT effects. Exploring this direction is left to future work.

2.A The general Quadratic Classifier

Any quadratic-order real polynomial of $n-1$ variables c_i, $i = 1, \dots n-1$, with arbitrary constant, linear and quadratic terms, can be written as a quadratic form in the n-dimensional variable

$$v(c) = (1,\, c_1,\, \dots,\, c_{n-1})^T\,. \tag{2.37}$$

Namely, we write the polynomial as

$$P(c) = v^T(c) A\, v(c)\,, \tag{2.38}$$

with A a generic n-dimensional real symmetric square matrix.

If $P(c)$ is non-negative for any value of c, it is easy to show that the matrix A must be positive semi-definite. Being real, symmetric and positive semi-definite, it is possible to use the Cholesky decomposition for A, and write it as

$$A = L^T L\,, \tag{2.39}$$

where L is a upper-triangular (i.e., $L_{ij} = 0$ for $j < i$) real matrix. Therefore the most general positive quadratic order polynomial reads

$$P(c) = v^T(c) L^T L\, v(c) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} L_{ij} v_j(c) \right)^2 = \sum_{i=1}^{n} \left(L_{i1} + \sum_{j=2}^{n} L_{ij} c_{j-1} \right)^2\,, \tag{2.40}$$

which is manifestly non-negative because it is the sum of square terms. Moreover for $c = 0$, since $L_{i1} = L_{11}\delta_{i1}$, we have $P(0) = L_{11}^2$. The Cholesky decomposition is unique up to sign flips of the rows of L. Rather than resolving this ambiguity, for instance by choosing the diagonal entries of L to be positive, we adopt eq. (2.40) without further constraints as the most general (though redundant) parametrization of $P(c)$.

The EFT differential cross section is a positive quadratic polynomial in the Wilson Coefficient c_i at each phase-space point x, and it reduces to the SM cross section for $c = 0$. It must therefore take the form

$$d\sigma_0(x;c) = d\sigma_1(x) \sum_{i=1}^{n} \left[\delta_{i1} + \sum_{j=2}^{n} \lambda(x)_{ij} c_{j-1}\right]^2 , \tag{2.41}$$

with $\lambda(x)$ an upper-triangular matrix of real functions. If only one c parameter is present (i.e., $n = 2$), this reduces to eq. (2.10) with the identifications

$$\lambda(x)_{12} = \alpha(x) \qquad \lambda(x)_{22} = \beta(x)\,. \tag{2.42}$$

The Quadratic Classifier that generalizes eq. (2.12) is thus defined as

$$f(x,c) \equiv \frac{1}{1 + \sum\limits_{i=1}^{n} \left[\delta_{i1} + \sum\limits_{j=2}^{n} \mathrm{n}(x)_{ij} c_{j-1}\right]^2} , \tag{2.43}$$

in terms of an upper-triangular matrix $\mathrm{n}(x)$ of real-output Neural Networks.

2.B Minimization of the parametrized loss

In the Large Sample limit, the loss function in eq. (2.13) becomes

$$L[\mathrm{n}(\cdot)] \overset{\mathrm{LS}}{=} \sum_{c\in\mathcal{C}} \left\{ \int d\sigma_0(x;c)[f(x,c)]^2 + \int d\sigma_1(x)[1 - f(x,c)]^2 \right\} , \tag{2.44}$$

with the Quadratic Classifier f defined in eq. (2.43). By simple algebraic manipulations, this can be rewritten as

$$L[\mathrm{n}(\cdot)] \overset{\mathrm{LS}}{=} \sum_{c\in\mathcal{C}} \left\{ \int \frac{d\sigma_1(x) d\sigma_0(x;c)}{d\sigma_1(x) + d\sigma_0(x;c)} + \int [d\sigma_1(x) + d\sigma_0(x;c)] \left[f(x,c) - \frac{1}{1 + r(x,c)} \right]^2 \right\} , \tag{2.45}$$

with $r(x,c) = d\sigma_0(x;c)/d\sigma_1(x)$. The first integral is independent of f and thus it is irrelevant for the minimization of the loss. The second one is the integral of a non-negative

function of x which attains its global minimum (i.e., it vanishes) if and only if

$$f(x,c) = f_{\min}(x,c) = \frac{1}{1+r(x,c)}\,, \qquad \forall\, c \in \mathcal{C}\,. \tag{2.46}$$

By using eq. (2.41), and comparing with eq. (2.43), we immediately conclude that the configuration $\mathrm{n}(x)_{ij} = \lambda(x)_{ij}$ is a global minimum of the loss and that this minimum is unique provided the set $\mathcal{C}$ contains at least two distinct non-vanishing elements. More precisely, this holds only up to sign ambiguities, associated with those of the Cholesky decomposition. However this is irrelevant because the ambiguity cancels out in f, and in turn it cancels out in the reconstructed distribution ratio $\widehat{r}(x,c) = 1/\widehat{f}(x,c) - 1$.

We have shown that the Quadratic Classifier reconstructs the distribution ratio exactly (in the Large Sample limit and for infinitely complex Neural Network) at the global minimum of the loss, and that this minimum is unique. Notice however that we could not show that the Large Sample limit loss does not possess additional local minimums, as it is instead readily proven for the standard classifier of Section 2.1.1 by variational calculus.

3 EW radiation at a High Energy Muon Collider

The perspective of a future muon collider with high energy and high luminosity [99], whose feasibility is being investigated by the International Muon Collider Collaboration [100], has triggered a growing interest in the physics potential of lepton colliders with a center of mass energy of 10 TeV or more [101–140]. Such a Very High Energy Lepton Collider (VHEL) could greatly advance the post-LHC knowledge of fundamental physics [99] by directly searching for new heavy particles (see e.g. [101–104]), and by precise measurements of ElectroWeak (EW) scale processes exploiting the high luminosity of virtual vector bosons pairs (see e.g. [104–106]). By these two search modes, the VHEL reach on new physics is generally comparable to that of the most ambitious future collider projects, in the corresponding domains of investigation. In particular it is comparable to the combined reach of the -ee and -hh stages of the FCC program. The sensitivity is slightly weaker or slightly stronger depending on the specific new physics target and, obviously, on the assumed VHEL energy and luminosity.

At a VHEL, however, there also exists a third strategy of investigation [99, 105], based on hard processes with energy scale comparable to the collider energy $E_{\rm cm} \sim 10$ TeV. As the indirect effects of new heavy particles are enhanced at high energy, these processes are extremely sensitive probes of new physics. With the target integrated luminosity of 10 ab^{-1}, $2 \to 2$ scattering processes at $E_{\rm cm} = 10$ TeV can be generically measured with percent or few-percent precision. Such measurements are therefore sensitive to putative new physics at a scale $\Lambda \sim 100$ TeV when its effects, relative to the SM cross-section, scale like $(E_{\rm cm}/\Lambda)^2$. In an Effective Field Theory (EFT) description of new physics, this corresponds to an enhanced sensitivity to those dimension-6 operators that contribute to the $2 \to 2$ processes with a quadratically growing-with-energy term. The VHEL sensitivity to new physics through this kind of "high-energy" probes vastly and generically exceeds the potential of any other future project that is currently under consideration [105]. In particular it exceeds the sensitivity of precision measurements of EW-scale (~ 100 GeV) processes at future Higgs factories, where new physics at $\Lambda \sim 100$ TeV produces effects at the unobservable level of one part per million. It also exceeds the potential sensitivity of

a 100 TeV proton collider, because the effective luminosity for partonic collisions above 10 TeV is significantly lower than that of the VHEL. The possibility of probing new physics at the 100 TeV scale, and in particular of probing new physics that is either flavor-universal or endowed with a flavor protection mechanism, is thus a unique feature of the VHEL that deserves an extensive investigation.[1]

The above mentioned high-energy strategy exploits simple $2 \to 2$ processes with extremely low or negligible background, whose target accuracy is statistically limited to 1%. At a superficial look, it might thus seem that its implementation will not pose any specific challenge, neither as concerns the feasibility of the measurements, nor as concerns the theoretical predictions that are needed for their interpretation. However the situation is slightly more complex, both experimentally and theoretically, owing to the phenomenon of EW radiation [141–161], which becomes prominent at 10 TeV or above. This happens because of the existence of a large separation between the hard scale E of the process and the vector boson mass scale $m_{\rm W}$, which acts as an IR cutoff. As the hard scale is increased, large logarithms $\log E^2/m_{\rm W}^2$ (squared) enhance both virtual corrections and real emission, up to the point where fixed-order perturbation theory becomes insufficient and resummation is needed. The experimental implications of the copious emission of real EW massive vector bosons should be assessed. It particular it should be studied if and how it impacts the accuracy of the relevant cross-section measurements. In this chapter we investigate the implications of EW radiation on the theoretical predictions and, assuming purely statistical experimental errors, on the VHEL sensitivity to new physics.

EW radiation obviously displays some similarities with QCD radiation, but also remarkable differences. First, EW scattering processes violate the KLN theorem assumptions [162, 163] because the initial state particles are not EW singlets. Therefore no cancellation takes place between virtual and real contributions, not even in "fully-inclusive" cross-sections [141, 142]. Moreover the observables that are fully inclusive in the sense of Ref. [141] are insufficient to characterize new physics because they require summing over the "color" of the hard final-state particles. In the EW context, color sum means, for instance, including transversely-polarized W and Z bosons and photons (or, longitudinal W, Z and Higgs) in the same observable, while we need to keep them separate for a comprehensive exploration of new physics. Unlike QCD (and QED), EW radiation effects cannot be eliminated or systematically mitigated with a judicious choice of the observables. They unavoidably play an important role in the predictions.

The second peculiarity of EW radiation is that the IR cutoff scale is physical, and the theory is weakly-coupled at the IR scale. It should thus be possible to characterize the radiation fully by first-principle suitably resummed perturbative calculations. Unlike

[1] Hard processes are also useful to investigate flavor non-universal effects, as we will see in Section 3.2.3 for Top Compositeness. See also Ref. [135] for a study of new physics potentially responsible for the $g-2$ muon anomaly.

QCD and QED, one can consider observables with an arbitrary degree of radiation exclusiveness, among which "exclusive" scattering cross-sections with a fixed number (2, in $2 \to 2$ processes) of hard partons in the final state and no extra massive vector bosons.[2] Fully-inclusive cross-sections can be also considered, however they are not sufficiently informative on new physics as previously mentioned. In this chapter we employ "semi-inclusive" final states, consisting of 2 hard partons with fixed EW color and flavor carrying a large fraction of the total available energy $E_{\rm cm}$ and accompanied by an arbitrary number of massive vectors, photons and other light particles[3]. Our resummed predictions for semi-inclusive observables at the double logarithm (DL) accuracy are obtained by extending the IR Evolution Equation (IREE) treatment of EW radiation developed in Ref. [143]. Similarly, we employ the IREE to compute the more standard exclusive cross-sections. Single-logarithmic terms turn out to be relevant, and they are included at fixed one-loop order in the exclusive cross-sections using the results of Ref. [165].

Finally, there is an interplay between EW radiation and short-distance physics that has no analog in QED and QCD [105]. Based on a simplistic fixed order intuition, this interplay can be exemplified by noticing that the emission of a soft W from one initial lepton changes the total charge of the initial state of the hard $2 \to 2$ scattering process. By allowing for the charged W emission one thus obtains a scattering cross-section that is sensitive to charged current new physics interactions, while the exclusive process where no radiation is allowed is only sensitive to neutral currents.[4] The combined measurement of the two types of cross-section thus enables a more effective and complete exploration of new physics. In reality the situation is slightly more complex, because the neutral and the charged current hard amplitudes both contribute to the resummed expression of the neutral exclusive and of the charged and neutral semi-inclusive cross-sections. However, since they appear in different combinations in the different cross-sections, the conclusion is unchanged.

At a more quantitative level, we will see that the VHEL energy sits in an interestingly "intermediate" regime for EW radiation. The energy is on one hand large enough for the radiation effects to be important and require resummation. Accurate resummation techniques, more accurate than the one considered in the present chapter, will thus be needed to fully exploit the potential of the machine. On the other hand, the energy is not yet in the asymptotic regime where the Sudakov exponentials entail a strong suppression

[2] In order to cope with QED and QCD radiation, the observable must still be inclusive over the emission of photons and other light particles. The cross-section we define as "exclusive" coincides with the "semi-inclusive" cross-section of Ref. [143]. Correspondingly, the "semi-inclusive" cross-section we will readily introduce was not considered in Ref. [143]. See Section 3.1 for details.

[3] A similar observable is discussed in [164] to show the impact of weak gauge boson emission at LHC. The final state they consider is somehow intermediate between the "fully-inclsuive" of [141] and the "semi-inclusive" we study in this chapter

[4] More precisely, the charged and neutral current process depend on different linear combinations of the Wilson coefficients of the operators parametrizing new physics.

of the non-emission probability. Therefore in this regime the exclusive cross-sections are still large, and comparable with the semi-inclusive (and fully-inclusive) ones. Observables with a different degree of radiation inclusiveness can thus be measured with comparable statistical accuracy and combined, potentially boosting, as previously explained, the indirect sensitivity to heavy new physics.

The rest of the chapter is organized as follows. We start, in Section 3.1, by reviewing the IREE approach to the DL resummation of exclusive cross-sections, extending it to semi-inclusive ones. In this section we also present the calculation of the exclusive and semi-inclusive di-fermion and di-boson production processes, to be employed in Section 3.2 to estimate the sensitivity to dimension-6 EFT current-current contact interactions of muon colliders with a center of mass energy ranging from 3 to 30 TeV. The estimates include experimental reconstruction efficiencies at the level expected for the CLIC detector at 3 TeV, which we extract from Ref. [166]. Our results do not depend on the nature of the colliding leptons. In particular they do not include the effect of Beamstrahlung, which is expectedly small at muon colliders but large at e^+e^- colliders of the same energy. Up to this caveat, our results thus also apply to hypothetical linear e^+e^- collider based on plasma wake field acceleration [167].[5] In Section 3.2 we also study the impact of the EFT sensitivity projections on concrete scenarios for Beyond-the-SM (BSM) physics such as Higgs and Top Compositeness and a minimal Z' extension of the EW sector. Our conclusions and an outlook on future work are reported in Section 3.3. The first two appendices contain technicalities about the the derivation of the IREE. The third one collects our sensitivity projections on EFT operators relevant for the Top Compositeness scenario, while the fourth appendix extensively summarizes the BSM sensitivity of muon colliders of different center of mass energies.

3.1 Theoretical predictions

Several approaches have been considered in the literature for the resummation of the effects of EW radiation. Double logarithm (DL) contributions, of the form $(\alpha \log^2 E^2/m_{\rm W}^2)^n$ with arbitrary n, have been resummed in fully-inclusive and exclusive cross-sections, using respectively Asymptotic Dynamics [141, 142] and IREE [143, 144]. In Soft-Collinear Effective Theory (SCET) the expansion is already organized in exponential form. In that case the resummation of leading logarithms (LL)[6] as well as its extension to next-to-leading (NLL) logarithm [150] has been studied. The study of EW parton distribution or fragmentation functions [153–157] is obviously connected, but not directly relevant for very high energy processes, where the main role is played by the emission of EW radiation that is both collinear and soft. Notice however that in some framework [153–157] soft effects are partially or completely included in parton distributions and fragmentation

[5] Notice that Beamstrahlung potentially entails a strong depletion of the high-energy luminosity peak, which is the part of the luminosity spectrum that is relevant for the high-energy probes.

[6] These include but do not coincide with the pure DL, as explained, for instance, in Ref. [150].

functions.

In this chapter we employ DL predictions based on the robust diagrammatic methodology of the IREE [143], which we review and further develop in Section 3.1.1. We also supplement our computations by the available virtual single logarithms (SL) at 1-loop [165, 168]. Based on these results, we present in Sections 3.1.2 and 3.1.3 our theoretical predictions for the di-fermion and di-boson production processes at the VHEL. While it will emerge that single logarithms are potentially relevant, a systematic consideration of these effects goes beyond the scope of the present work. Our finding that electroweak radiation can be used to our own advantage in the exploration of new physics, strongly motivates the systematic inclusion of subleading effects. A first simple step would be to include in our predictions the single logarithms from real emissions at tree level. A complete treatment including resummation would likely best be achieved by using SCET.

3.1.1 IR Evolution Equations

The basic idea of the IREE is to introduce an unphysical IR regulator λ with dimension (energy)^2 in the calculation of the observables and to write down a differential equation for the evolution with λ of the result. Denoting by "E^2" the hardness of the observable under consideration, the choice $\lambda \sim E^2$ eliminates all the IR enhancements and makes fixed-order perturbation theory well-behaved. For $\lambda \sim E^2$, the Born level computation therefore offers a reliable approximation, which can be used as the initial condition for the evolution equation to lower λ. The physical predictions are obtained from the solution of the IREE in the limit $\lambda \to 0$.

In order to define the IR regulator, consider the 4-momenta k_i of the external hard particles that participate in the scattering process. They will correspond in our setup to the 4 legs of a central energetic $2 \to 2$ processes. With the exception of the masses $k_i^2 \ll E^2$, all the Lorentz invariants constructed with the k_i's are therefore large and of order E^2. Given a pair ij of external hard particles and given a radiation particle with 4-momentum q we define its hardness relative to the ij pair as

$$\mathfrak{h}^{ij}(q) \equiv 2 \left| \frac{(k_i \cdot q)(k_j \cdot q)}{(k_i \cdot k_j)} \right| . \tag{3.1}$$

The IR regulator is provided by the lower bound λ

$$\mathfrak{h}(q) \equiv \min_{i \neq j} \mathfrak{h}^{ij}(q) > \lambda \, . \tag{3.2}$$

on the "absolute" hardness $\mathfrak{h}$ of the radiation. Notice that the bound is imposed on the 4-momentum of each individual radiation particle, either virtual or real. Specifically, eq. (3.2) applies to the off-shell loop momenta describing virtual radiation, as well as to the on-shell momenta of real radiation particles in the final state of the process. The

specific definition of the radiation hardness in eq. (3.1) is convenient because it reflects the structure of the denominators that appear in the calculation of the IR-enhanced diagrams, as we will readily see. At this stage, it is enough to remark that the lower cut on $\mathfrak{h}(q)$ in eq. (3.2) is a valid IR regulator as it eliminates both the regions where q is soft and those where it is collinear to one of the hard partons.

The main peculiarity of the IREE formalism applied to EW radiation stems from the presence of the physical scale $m_{\rm W} \sim 100$ GeV associated to the masses of the EW bosons. We will see that $m_{\rm W}$ acts as a threshold that separates two different regimes, $\lambda \gg m_{\rm W}^2$ and $\lambda \ll m_{\rm W}^2$. In the former regime, the cut on the radiation hardness in eq. (3.2) is so strong that the mass of the radiation particles can be safely neglected and the IREE computed in the unbroken $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ EW gauge theory. Starting from the initial condition at $\lambda \sim E^2$, the evolution is thus performed with the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ evolution kernel down to $\lambda \sim m_{\rm W}^2$. At $\lambda \ll m_{\rm W}^2$, the massive vector bosons do not contribute to the evolution and the kernel is purely determined by the unbroken $\mathrm{U}(1)_Q$ group of electromagnetism.

Amplitude evolution

We start, following Ref. [143], from the IREE for the scattering amplitude with purely hard external quanta and with regulator λ on the internal lines. While the discussion holds for an arbitrary number of external legs, we focus for definiteness on $2 \to 2$ amplitudes, which we indicate by

$$\mathcal{M}^{\alpha}_{\lambda} = \mathcal{M}_{\lambda}\Big[\mathrm{p}_1(k_1,\alpha_1)\,\mathrm{p}_2(k_2,\alpha_2) \to \mathrm{p}_3(k_3,\alpha_3)\,\mathrm{p}_4(k_4,\alpha_4)\Big]\,, \tag{3.3}$$

where α_i denotes the gauge group index of the external state p_i, which is taken to transform in an irreducible representation of the group. The amplitude is labeled by the 4 indices $\alpha = \alpha_1\alpha_2\alpha_3\alpha_4$, and it is IR-regulated according to eq. (3.2). Since no real radiation is involved, the cut acts only on the momenta of virtual vector bosons in loop diagrams. We aim at writing down the IREE for $\mathcal{M}^{\alpha}_{\lambda}$ and to solve it given the initial condition

$$\mathcal{M}^{\alpha}_{E^2} = \mathcal{B}^{\alpha} \equiv \text{Born Amplitude}\,. \tag{3.4}$$

As we explained, for $\lambda \gg m_{\rm W}^2$ the effects of EW symmetry breaking (EWSB) can be ignored, and $\mathcal{M}^{\alpha}_{\lambda}$ equals the (IR-regulated) amplitude of the unbroken EW gauge theory. More precisely, EWSB gives negligible relative corrections of order $m_{\rm W}/\sqrt{\lambda}$ (or powers thereof) to all those amplitudes that are not forbidden by the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ exact symmetry of the unbroken theory. The other amplitudes are sensitive to EWSB at the leading order and therefore they cannot be studied in the unbroken theory.[7] However

[7] For instance the amplitude with 3 transversely- and one longitudinally-polarized W bosons is suppressed by $m_{\rm W}/E$ already at the Born level, owing to the fact that it is impossible to form an $\mathrm{SU}(2)_L$

their contribution to the physical scattering process is negligible and they can be safely excluded from the discussion[8] . Similarly, for the allowed processes, up to negligible power corrections of order $m_{\rm W}/E$, the amplitude $\mathcal{M}^{\alpha}_{\lambda}$ is an $\mathrm{SU}(2)_L{\times}\mathrm{U}(1)_Y$ invariant tensor satisfying the charge conservation equation

$$(G^A_{1^c})^{\alpha}_{\ \beta}\mathcal{M}^{\beta}_{\lambda} + (G^A_{2^c})^{\alpha}_{\ \beta}\mathcal{M}^{\beta}_{\lambda} + (G^A_{3})^{\alpha}_{\ \beta}\mathcal{M}^{\beta}_{\lambda} + (G^A_{4})^{\alpha}_{\ \beta}\mathcal{M}^{\beta}_{\lambda} \overset{\lambda\gg m_{\rm W}^2}{=} 0\,, \qquad \forall\, A,\, \alpha\,. \tag{3.5}$$

In the equation, G^A_i denotes the generators associated with the representation of each hard particle "i" under the EW group, acting only on the corresponding index "α_i" of the amplitude tensor. For instance

$$(G^A_3)^{\alpha}_{\ \beta} = \delta^{\alpha_1}_{\beta_1}\delta^{\alpha_2}_{\beta_2}(G^A_3)^{\alpha_3}_{\ \beta_3}\delta^{\alpha_4}_{\beta_4}\,. \tag{3.6}$$

Notice that, in our notation, (α_3, α_4) run in the representations of the outgoing states, while (α_1, α_2) run in the conjugate representation of the incoming particles. Consequently in eq. (3.5), $G_{1^c} = -G_1^*$ and $G_{2^c} = -G_2^*$.

The IREE is obtained by computing the variation of the amplitude under an infinitesimal variation $\lambda \to \lambda + \delta\lambda$ of the IR cutoff in eq. (3.2). This computation dramatically simplifies in the leading DL approximation as one can infer by inspecting diagrams involving a number n of soft/collinear virtual vector bosons. Indeed the maximal logarithm power arises from the region where momenta are hierarchically separated $E^2 \gg \mathfrak{h}(q_1) \gg \mathfrak{h}(q_2) \gg \cdots \gg \mathfrak{h}(q_n)$ with the softer legs dressing the subdiagrams involving the harder legs, as shown in the left panel of Figure 3.1. In this configuration only the outermost virtual vector can reach a virtuality $\mathfrak{h}(q_n) \sim \lambda$, the inner ones being much harder in the dominant region of integration. The effect on $\mathcal{M}^{\alpha}_{\lambda}$ of the variation of λ is then computed by considering the variation of the endpoint of the integral over the momentum of such outermost vector. More precisely we have that $-\delta\mathcal{M}^{\alpha}_{\lambda} \equiv \mathcal{M}^{\alpha}_{\lambda} - \mathcal{M}^{\alpha}_{\lambda+\delta\lambda}$ equals the integral over the outermost loop momentum in the infinitesimal strip

$$\delta\sigma = \{q :\, \mathfrak{h}(q) \in [\lambda,\, \lambda + \delta\lambda]\}\,. \tag{3.7}$$

The contribution to the variation from the vector that connects a given pair of hard external legs can be depicted like on the right panel of Figure 3.1. The vector boson is represented with a double line to indicate that its momentum q must be integrated only over the strip $\delta\sigma$.

As we already said the leading contribution comes from the integration region where q is soft (and also collinear), in which the vector boson emission is described by the eikonal

singlet with one doublet (i.e., the representation of longitudinal W's owing to the Equivalence Theorem) and three triplets.

[8] Regarding power suppressed amplitudes, in [169] it has been found that they can receive positive Sudakov enhancements. However due to the power-like suppression we don't expect this to affect our discussion.

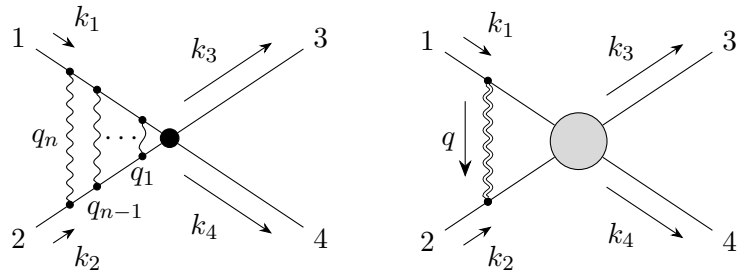
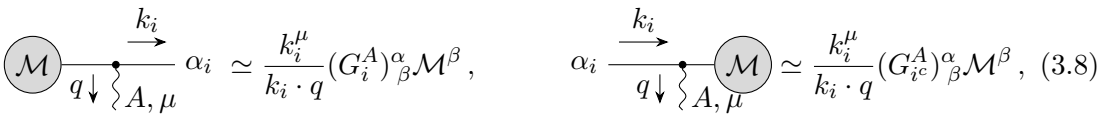


Figure 3.1: Left panel: the leading Sudakov diagrams topology. Right panel: a diagrammatic representation of the contributions to the amplitude variation that are logarithmically enhanced. The momentum q is integrated over the infinitesimal strip $\delta\sigma$ (3.7).

formula



with G_i the group generator acting on particle "i" as in eq. (3.6). In line with our conventions, as explained above, the generators of the charge-conjugate representation G_{i^c} appear in the eikonal formula for vector boson emission from an incoming particle. For brevity, we have included the gauge coupling constants in the definition of the generators G_i. In terms of the canonical $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ generators we then have

$$G_i^{1,2,3} = g\, T_i^{1,2,3}\,, \qquad G_i^Y = g'\, Y_i\,. \tag{3.9}$$

The integration over the soft q momentum factorizes with respect to the integral over the harder lines, represented as a blob in Figure 3.1. Indeed in the eikonal ($q \to 0$) limit the virtual vector boson momentum can be neglected and the sub-amplitude blob evaluated on the momenta k_i of the external legs before the virtual vector boson emission/absorption. Therefore the blob gives us back the original amplitude, with one less loop but this is immaterial as $\mathcal{M}_\lambda^\alpha$ is the all-loops amplitude. We can thus express the amplitude variation in terms of the amplitude itself, eventually obtaining an evolution equation. In covariant gauges, the leading DL contributions only arise from virtual vectors connecting two distinct external lines. Therefore, we have

$$\delta\mathcal{M}_\lambda^\alpha = \frac{-i}{(2\pi)^4} \sum_{j<i} \int_{\delta\sigma} d^4q\, \frac{1}{q^2 + i\epsilon} \frac{k_i \cdot k_j}{(q \cdot k_i)(q \cdot k_j)} \Big[\sum_A G_i^A \cdot G_j^A \Big]^\alpha{}_\beta\, \mathcal{M}_\lambda^\beta\,, \tag{3.10}$$

where the sum extends over the unordered ij pairs of distinct external legs and it is understood that the conjugate generators must be employed for the incoming legs $i, j = 1, 2$, due to eq. (3.8).

The evaluation of the integral in eq. (3.10) is quite straightforward, and it is reported in Appendix 3.A for completeness. This gives

$$\delta\mathcal{M}_\lambda^\alpha = -\frac{1}{8\pi^2}\frac{\delta\lambda}{\lambda}\,\log\frac{E^2}{\lambda}\,\frac{1}{2}\sum_{A,\,i}(G_i^A)^\alpha{}_\beta\sum_{j\neq i}(G_j^A)^\beta{}_\gamma\,\mathcal{M}_\lambda^\gamma\,, \tag{3.11}$$

up to non logarithmically enhanced terms. Notice that in the equation we traded the sum over unordered ij pairs for an ordered sum times $1/2$. The sum over $j\neq i$ can be performed using charge conservation according to eq. (3.5), giving

$$\begin{aligned}\delta\mathcal{M}_\lambda^\alpha &= \frac{1}{16\pi^2}\frac{\delta\lambda}{\lambda}\,\log\frac{E^2}{\lambda}\,\sum_i\Big[\sum_A G_i^A G_i^A\Big]^\alpha{}_\beta\,\mathcal{M}_\lambda^\beta \\ &= \frac{1}{16\pi^2}\frac{\delta\lambda}{\lambda}\,\log\frac{E^2}{\lambda}\,\sum_i\Big[g^2c_i + g'^2y_i^2\Big]\,\mathcal{M}_\lambda^\alpha\,,\end{aligned} \tag{3.12}$$

where for any given external particle with weak isospin spin t_i and hypercharge y_i, the coefficients $c_i = t_i(t_i+1)$ and y_i^2 are nothing but the Casimirs of respectively $\mathrm{SU}(2)_L$ and $\mathrm{U}(1)_Y$. We thus recovered the familiar result that, in DL accuracy, IR effects are universal for each individual external particle and purely determined by the Casimir of the corresponding gauge group representation.

We finally obtain an IREE

$$\frac{d\mathcal{M}_\lambda^\alpha}{d\log^2(E^2/\lambda)} = -\frac{1}{2}\mathcal{K}\,\mathcal{M}_\lambda^\alpha\,,\quad \text{where}\ \ \mathcal{K}\overset{\lambda\gg m_{\rm W}^2}{=}\frac{1}{16\pi^2}\sum_i\Big[g^2c_i + g'^2y_i^2\Big]\,, \tag{3.13}$$

with, since the Casimir operators are proportional to the identity, an evolution kernel $\mathcal{K}$ that is a mere multiplicative constant. Solving eq. (3.13) starting from the initial condition (3.4) gives the amplitude evaluated with an IR cutoff scale $\lambda = m_{\rm W}^2$

$$\mathcal{M}^\alpha_{m_{\rm W}^2} = \exp\left[-\sum_i\frac{g^2c_i + g'^2y_i^2}{32\pi^2}\,\log^2(E^2/m_{\rm W}^2)\right]\mathcal{B}^\alpha\,. \tag{3.14}$$

In order to continue the amplitude evolution to lower λ, we should now consider the regime $\lambda \ll m_{\rm W}^2$, write the corresponding IREE and solve them using eq. (3.14) as initial condition. This is straightforward, because we have seen that all that matters for the derivation of the IREE are the loop integrals in a strip where the virtual radiation hardness is infinitesimally close to the cutoff λ as in eq. (3.7). In this region, a logarithmic enhancement of the amplitude variation only originates from photon exchange diagrams.[9] The IREE evolution kernel is thus immediately obtained by specifying the previous

[9]The calculation of the loop integral in Appendix 3.A shows explicitly that no enhancement emerges from the exchange of vectors with mass m_V much larger than λ.

formulae to the $\mathrm{U}(1)_Q$ gauge group of QED

$$\mathcal{K} \overset{\lambda \ll m_{\mathrm{W}}^2}{=} \frac{1}{16\pi^2} \sum_i \left[e^2 q_i^2 \right] . \tag{3.15}$$

Notice that in order to derive the IREE in this regime, only conservation of electric charge must be employed. The conservation of the full $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ charges of eq. (3.5) is not valid for $\lambda \ll m_{\mathrm{W}}^2$, where the effects of electroweak symmetry breaking are important.

Solving eq. (3.15) produces the regular QED Sudakov factors, which go to zero in the physical limit $\lambda \to 0$ where the IR regulator is removed. Therefore the amplitude $\mathcal{M}_0^\alpha$ vanishes, and so does the cross-section of the corresponding fully-exclusive scattering process, in which no extra radiation is present in the final state. More inclusive observables need to be considered for a non-vanishing result. One possibility is to allow for the presence of real photon radiation up to an upper threshold of order m_{W}^2 on the hardness $\mathfrak{h}$. This defines a cross-section that we denote as *exclusive* because it indeed excludes the radiation of massive EW bosons. In fact, it is easy to check that $\mathfrak{h}(q) > m^2$ for the emission of a real radiation quantum with $q^2 = m^2$. An upper cut $\mathfrak{h}(q) < m_{\mathrm{W}}^2$ then excludes the presence of massive EW bosons in the final state, but allows for (sufficiently) soft photons. Ref. [143] considered this same observable (but calling it "semi-inclusive") showing that it stops evolving with λ below m_{W}^2, due to the cancellation of real and virtual IR effects in QED. Cross-sections that are exclusive according to our definition can thus be computed at the DL accuracy by just squaring the $\lambda = m_{\mathrm{W}}^2$ amplitude (3.14). At the end of the next section we will re-derive the result of Ref. [143] for exclusive cross-sections by a slightly different methodology, which is also suited for the calculation of the other type of cross-sections we are interested in.

Density matrix evolution

It is possible to extend the IREE methodology to more complex quantities than the hard Feynman amplitude. Specifically, we consider the hard "density matrix" [10]

$$\mathcal{D}_\lambda^{\alpha\bar{\alpha}} \equiv \mathcal{M}_\lambda^\alpha (\mathcal{M}_\lambda^{\bar{\alpha}})^* + \sum_{N=1}^{\infty} \int \mathrm{dPh}_{N,\lambda}^{\mathcal{H}} \sum_{\rho_1 \cdots \rho_N} \mathcal{M}_\lambda^{\alpha;\rho} (\mathcal{M}_\lambda^{\bar{\alpha};\rho})^* , \tag{3.16}$$

which incorporates the emission of an arbitrary number N of radiation particles, with gauge group indices denoted as $\rho = \rho_1 \ldots \rho_N$. In the equation, $\mathcal{M}_\lambda^\alpha$ is the hard amplitude with no extra emissions as in the previous section, while $\mathcal{M}_\lambda^{\alpha;\rho}$ is the amplitude for the production of the 2 hard particles plus the radiation. The virtual radiation particles exchanged in the Feynman diagrams for the amplitude are subject to the IR hardness cutoff λ as in eq. (3.2). The phase-space volume element $\mathrm{dPh}_{N,\lambda}^{\mathcal{H}} = \prod_{k=1}^N \mathrm{dPh}_{k,\lambda}^{\mathcal{H}}$ for the emission of real radiation is also constrained by eq. (3.2). The $^{\mathcal{H}}$ superscript refers to

[10] The same object was dubbed "overlap matrix" in Ref. [142].

the possible presence of an upper cutoff on the radiation hardness $\mathfrak{h}(q) < \mathcal{H}$. In what follows we will first consider processes we define as *semi-inclusive*, for which $\mathcal{H} \sim E^2$. For these processes the upper radiation cut is effectively absent, and plays no role in the discussion. The exclusive processes defined in the previous section instead simply correspond to $\mathcal{H} = m_{\rm W}^2$.

It should be noted that eq. (3.16) formally violates the conservation of the total energy and momentum, because in the radiation terms we are employing the same hard 4-momenta that obey energy and momentum conservation in the absence of radiation. It is understood that this makes sense only in the presence of an upper cutoff on the total energy and momentum of the radiation, say a one tenth of E. In this way, the radiation plays a minor role in the total balance of energy and momentum conservation or, equivalently, the hard 4-momenta can be readjusted to balance the radiation emission up to small corrections in the corresponding Feynman amplitudes. In practice, the cutoff allows us to factorize the total phase-space into that for radiation, on one hand, and that for the hard $2 \to 2$ process on the other, with the latter also including the delta function of 4-momentum conservation. The density matrix (3.16) can thus be related to the physical scattering cross-section.

An upper cut $E_{rad} < E/10$ on the total radiation energy and momentum does not affect the predictions at the double logarithm accuracy. Indeed a simple modification of the real radiation integral (see the discussion around eq. (3.18) computation in Appendix 3.A shows that the effect of this cut on the q momentum of the radiated particle merely entails reduction of the double logarithm from $\log^2 E^2/\lambda$ to $\log^2 E_{rad}^2/\lambda$. The difference is then of order $\log E^2/\lambda \times \log E^2/E_{rad}^2$ and falls into the same class as single logarithms as long as E/E_{rad} is not too small, with $1/10$ qualifying.

The hard density matrix (3.16) is a simple generalization of the scattering cross-section in which the conjugated amplitude indices $\bar\alpha$ are not equal to the indices α of the non-conjugated amplitude. It is a useful generalization because it obeys charge conservation equations similar to eq. (3.5). Namely, in the regime $\lambda \gg m_{\rm W}^2$, we have

$$\sum_{i=1^c,2^c,3,4} \left[(G_i^A)^{\alpha}_{\ \beta} \mathcal{D}_\lambda^{\beta\bar\alpha} + (G_{i^c}^A)^{\bar\alpha}_{\ \bar\beta} \mathcal{D}_\lambda^{\alpha\bar\beta} \right] \overset{\lambda \gg m_{\rm W}^2}{=} 0\,, \qquad \forall\, A,\, \alpha,\, \bar\alpha\,, \tag{3.17}$$

where the obvious relations $[1^c]^c \equiv 1$ $[2^c]^c \equiv 2$ should be understood. That way the generators acting on the indices $\bar\beta$ of the complex conjugated amplitude are those of the corresponding charge conjugated representation. Eq. (3.17) holds only for $\lambda \gg m_{\rm W}^2$, because in this regime both the virtual and the real emissions are nearly insensitive to EWSB effects as previously explained. For $\lambda \ll m_{\rm W}^2$, only the electric charge generator is conserved.

The IREE can be obtained like in the previous section by computing the variation of $\mathcal{D}_\lambda$ under $\lambda \to \lambda + \delta\lambda$, taking now also into account also the effect of the IR cutoff on real

emission. The contribution of virtual loop momentum integrals is thus accompanied by that of integrals over the momentum of real radiation. All integrals have to be performed over the infinitesimal strip $\delta\sigma$ defined in eq. (3.7). Logarithmically enhanced terms only arise from the exchange of virtual or real gauge bosons between different external legs ($i \neq j$), like in Figure 3.2. The effects and the corresponding diagrams can be divided into two classes. The first, in the left panel of Figure 3.2, is given by *primary* radiation diagrams where vector bosons are exchanged between the hard legs. The second, in the right panel, is given by *secondary* radiation diagrams where vector bosons connect to at least one real radiation leg.

We will first consider the effects of primary radiation. The virtual radiation integral gives the result already mentioned in eq. (3.11), and, as we show in Appendix 3.A, the result is exactly the same for the real radiation integral. The total variation from primary radiation is then

$$\delta\mathcal{D}_\lambda^{\alpha\bar{\alpha}} = -\frac{1}{16\pi^2}\frac{\delta\lambda}{\lambda}\log\frac{E^2}{\lambda}\sum_{i=1^c,2^c,3,4}\sum_A\Bigg[(G_i^A)^\alpha{}_\beta\sum_{j\neq i}\Big[(G_j^A)^\beta{}_\gamma\mathcal{D}_\lambda^{\gamma\bar{\alpha}}+(G_{j^c}^A)^{\bar{\alpha}}{}_{\bar{\beta}}\mathcal{D}_\lambda^{\beta\bar{\beta}}\Big] \\ +(G_{i^c}^A)^{\bar{\alpha}}{}_{\bar{\beta}}\sum_{j\neq i}\Big[(G_j^A)^\alpha{}_\beta\mathcal{D}_\lambda^{\beta\bar{\beta}}+(G_{j^c}^A)^{\bar{\beta}}{}_{\bar{\gamma}}\mathcal{D}_\lambda^{\alpha\bar{\gamma}}\Big]\Bigg]. \tag{3.18}$$

The argument of the first sum, over the four external legs, collects the contributions of all the radiation emitted from the leg "i" of the amplitude and of the conjugated amplitude. A factor 1/2 is included to avoid double-counting. Notice that both virtual and real radiation connecting one leg with itself is excluded from the sum, because, as we already mentioned, no enhancement arises from those diagrams.

We can now proceed as in the previous section, and use the charge conservation in eq. (3.17) to perform the sum over j in eq. (3.18). We find the IREE

$$\frac{d\mathcal{D}_\lambda^{\alpha\bar{\alpha}}}{d\log^2(E^2/\lambda)} = -\mathcal{K}_{\beta\bar{\beta}}^{\alpha\bar{\alpha}}\,\mathcal{D}_\lambda^{\beta\bar{\beta}}\,, \tag{3.19}$$

with an evolution kernel that is the direct sum of universal terms for each external leg

$$\begin{aligned}\mathcal{K}_{\beta\bar{\beta}}^{\alpha\bar{\alpha}} &\overset{\lambda\gg m_{\rm W}^2}{=} \frac{1}{32\pi^2}\sum_i\left[\Big[\sum_A G_i^A G_i^A\Big]^\alpha{}_\beta\delta_{\bar{\beta}}^{\bar{\alpha}}+\delta_\beta^\alpha\Big[\sum_A G_{i^c}^A G_{i^c}^A\Big]^{\bar{\alpha}}{}_{\bar{\beta}}+2\sum_A(G_i^A)^\alpha{}_\beta(G_{i^c}^A)^{\bar{\alpha}}{}_{\bar{\beta}}\right] \\ &= \frac{g^2}{16\pi^2}\sum_i\Big[c_i\,\delta_{\beta_i}^{\alpha_i}\delta_{\bar{\beta}_i}^{\bar{\alpha}_i}+\sum_{A=1,2,3}(T_i^A)^{\alpha_i}{}_{\beta_i}(T_{i^c}^A)^{\bar{\alpha}_i}{}_{\bar{\beta}_i}\Big]\Big[\prod_{j\neq i}\delta_{\beta_j}^{\alpha_j}\delta_{\bar{\beta}_j}^{\bar{\alpha}_j}\Big] \\ &= \frac{g^2}{16\pi^2}\sum_i\Big[K_i\Big]_{\beta_i\bar{\beta}_i}^{\alpha_i\bar{\alpha}_i}\Big[\prod_{j\neq i}\delta_{\beta_j}^{\alpha_j}\delta_{\bar{\beta}_j}^{\bar{\alpha}_j}\Big]\,.\end{aligned} \tag{3.20}$$

The kernel contains one term, provided by the $\mathrm{SU}(2)_L$ Casimir $c_i = t_i(t_i+1)$, which

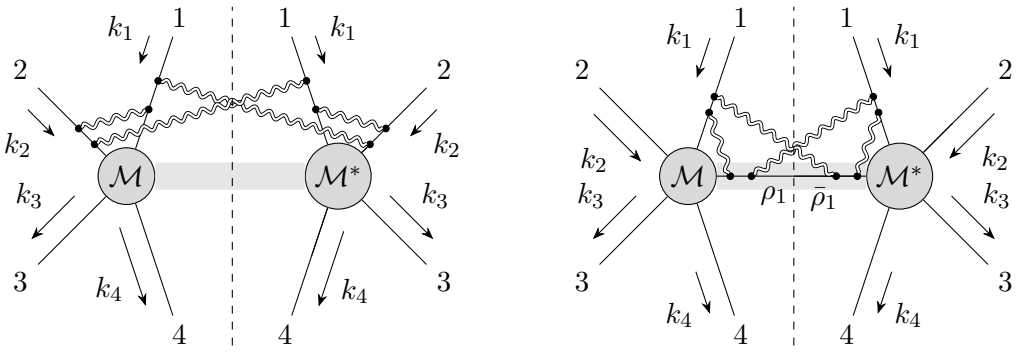


Figure 3.2: Diagrammatic representation of the contributions to the density matrix variation from primary (left panel) and secondary (right panel) radiation. The vector bosons are represented as double lines to indicate that their momenta have to be integrated over the infinitesimal strip (3.7).

is proportional to the identity in the color indices of the density matrix tensor, plus a non-diagonal term constructed with the $\mathrm{SU}(2)_L$ group generators matrices T_i^A of the external legs. Notice that the contribution of the $\mathrm{U}(1)_Y$ hypercharge generator cancels.

There is one peculiarity of eq. (3.20) that is worth emphasizing. The semi-inclusive cross-sections we are interested in are the diagonal entries ($\alpha = \bar{\alpha}$) of the density matrix, with no sum performed over the gauge group index α of the scattering particles.[11] However one can also consider *inclusive* cross-sections, where the sum over the gauge index α_i is performed for one or several external legs. By setting $\bar{\alpha}_i = \alpha_i$ and summing over α_i, the $\mathrm{SU}(2)_L$ generators in eq. (3.20) recombine to form the Casimir operator, and the contribution to the evolution kernel from leg "i" cancels. We thus find that, at DL accuracy, the cancellation between real and virtual IR effects in inclusive cross-sections occurs on a leg-by-leg basis. Namely, the effects of soft/collinear emissions associated to each individual leg cancel in the cross-section (and in the entire density matrix) for processes that are inclusive over the color of the corresponding particle. This result is stronger than the KLN theorem, which foresees a cancellation only when summing over the color of all legs. The reason for the added strength is that we are here considering radiation that is both soft and collinear. Notice however that fully inclusive observables of practical relevance can only involve summation on the color of the final state particles. This retains the IR effects associated with the colliding particles in the initial state (e.g., two left-handed leptons $\ell_L^+\ell_L^-$) which are not $\mathrm{SU}(2)_L$ singlets. The resulting non-cancellation of IR effects in "fully-inclusive" cross-sections, coincides with the result of Ref. [142].

So far we have ignored the secondary radiation diagrams, depicted in the right panel

[11] This is true only in a basis where the gauge indices α_i label the on-shell SM particles, while for the calculation of di-boson cross-sections we work in a different basis. See Section 3.1.3 and Appendix 3.B for details.

of Figure 3.2. We show now that their contribution vanishes, giving full justification to eq. (3.20). Secondary radiation diagrams correspond to the effect of the λ cutoff variation on virtual or real vector bosons attached to one of the intermediate "ρ" particles in the definition of the density matrix (3.16). Clearly these effects are potentially enhanced only if the intermediate particle is relatively hard, such that a significant separation is present between the IR cutoff λ and the scalar product between the intermediate particle and the external leg momenta. We thus start considering vector bosons attached to the hardest intermediate particle, with gauge index "ρ_1" as in the figure. The density matrix is inclusive over the color of the intermediate particle. However we can momentarily define an "extended" density matrix $\mathcal{D}_\lambda^{\alpha\,;\rho_1\,\bar{\alpha}\,;\bar{\rho}_1}$ with labels ρ_1 and $\bar{\rho}_1$ for the gauge indices of the amplitude and of the conjugate amplitude, as in the figure. The actual density matrix is eventually obtained by first setting $\rho_1 = \bar{\rho}_1$ and then summing. The effect on the extended density matrix variation of all the radiation emitted from ρ_1 and $\bar{\rho}_1$ can be written in a form similar to eq. (3.18) and then simplified using the analog of eq. (3.17) for the extended density matrix. The resulting contribution to the evolution kernel from the intermediate ρ_1 leg is the analog of that from the hard external legs in eq. (3.20). But this contribution cancels out in the evolution of the actual density matrix, which is inclusive over the ρ_1 leg, because of the previously explained leg-by-leg cancellation mechanism. The argument can of course be repeated for the diagrams involving the second hardest intermediate particle, showing, as anticipated, that all the secondary radiation diagrams can be ignored in the calculation of the evolution kernel.

It is straightforward to adapt the previous results to the regime $\lambda \ll m_{\rm W}^2$, in which only the exchange of photons contributes to the evolution, as discussed in the previous section. By specifying eq. (3.20) to the Abelian $\mathrm{U}(1)_Q$ group we immediately find that the kernel vanishes, owing to the well-known cancellation between real and virtual IR effects in QED. For the calculation of the physical ($\lambda \to 0$) density matrix, and in turn of the semi-inclusive cross-section, we thus only need to solve the IREE with the $\lambda \gg m_{\rm W}^2$ kernel (3.20), down to $\lambda = m_{\rm W}^2$.

For $\lambda = E^2$ the hard density matrix (3.16) is well-approximated by its tree-level expression, which serves as the initial condition for the evolution

$$\mathcal{D}_{E^2}^{\alpha\bar{\alpha}} = \mathcal{B}^{\alpha}(\mathcal{B}^{\bar{\alpha}})^* \,. \tag{3.21}$$

The kernel is the direct sum of tensors, denoted as K_i in eq. (3.20), each acting on the pair $\alpha_i, \bar{\alpha}_i$ associated to the i-th external particle. Therefore the solution of the IREE reads

$$\mathcal{D}_{\rm si}^{\alpha\bar{\alpha}} \equiv \mathcal{D}_{m_{\rm W}^2}^{\alpha\bar{\alpha}} = \left\{ \prod_i \left[\exp\left[-\frac{g^2}{16\pi^2} K_i \log^2(E^2/m_{\rm W}^2) \right] \right]_{\beta_i\bar{\beta}_i}^{\alpha_i\bar{\alpha}_i} \right\} \mathcal{B}^{\beta}(\mathcal{B}^{\bar{\beta}})^* \,, \tag{3.22}$$

where the "$_{\rm si}$" subscript denotes the density matrix of the semi-inclusive process, with no

upper cut on the real radiation hardness. The explicit form of the K_i exponentials in the above equation is reported in eq.s (3.83) and (3.88) for external legs in the doublet and triplet $\mathrm{SU}(2)_L$ representations. Applications of eq. (3.22) to specific processes are shown in Sections 3.1.2 and 3.1.3.

We have defined the density matrix (3.16) allowing for the presence of an upper cutoff $\mathcal{H}$ on the real radiation, but this played no role in the previous discussion because this cutoff is effectively absent ($\mathcal{H} \sim E^2$) in our definition of semi-inclusive processes. In exclusive processes we instead set $\mathcal{H} = m_{\rm W}^2$, namely we veto real radiation particles with hardness above $m_{\rm W}^2$. Obviously, for $\lambda \gg m_{\rm W}^2$ this upper cut is in contradiction with the IR cutoff in eq. (3.2) on the radiation phase-space. Therefore in the density matrix for the exclusive process no real radiation is present and in the $\lambda \gg m_{\rm W}^2$ regime the result simply equals the square of the hard amplitude in eq. (3.3). The evolution up to $\lambda = m_{\rm W}^2$ can thus be obtained from the hard amplitude evolution (3.14) we obtained in the previous section, or easily re-derived by dropping the terms in eq. (3.18) (namely, the second and the third) that are due to real radiation. The contribution of real radiation is instead restored for $\lambda \ll m_{\rm W}^2$ and the evolution stops due to the cancellation between virtual and real QED radiation as previously explained. The physical ($\lambda \to 0$) density matrix for exclusive processes can thus be written in a simple closed form as

$$\mathcal{D}_{\rm ex}^{\alpha\bar{\alpha}} = \exp\left[-\sum_i \frac{g^2 c_i + g'^2 y_i^2}{16\pi^2}\, \log^2(E^2/m_{\rm W}^2)\right] \mathcal{B}^{\alpha}(\mathcal{B}^{\bar{\alpha}})^* \,. \tag{3.23}$$

In Sections 3.1.2 and 3.1.3 we employ this formula to compute exclusive di-fermion and di-boson production cross-sections, and discuss the need of supplementing it with fixed-order single-logarithmic terms, from Ref. [165, 168].

Before concluding this section it is worth commenting on the experimental definition of the semi-inclusive and exclusive processes, and on the perspectives for their actual experimental detectability. The semi-inclusive process is characterized by two central (specifically, emitted from 30 to 150 degrees from the beam line) energetic particles of specific EW color and flavor. In particular we will require them to carry a total center of mass energy above 85% of the VHEL $E_{\rm cm}$, enforcing in this way the upper cut on the total radiation 4-momentum required for the definition of the hard density matrix as discussed below eq. (3.16). The two particles can be accompanied by the radiation of EW bosons, photons, or any other soft particle.

Notice that in our calculation at the DL order we could ignore all the effects of collinear (rather than soft-collinear) radiation, which emerge at the single logarithm. On the other hand, the single logarithms associated with low-virtuality (below $m_{\rm W}$) photon splittings are much larger than $\log E^2/m_{\rm W}^2$. In particular, the emission of real photons that are energetic but collinear to a light charged hard particle (e.g., an electron or a muon) with mass m_ℓ produces terms proportional to $\log E^2/m_\ell^2$. By the KLN theorem these terms

will be canceled by the corresponding virtual contributions, but only in suitably-defined observables that recombine the emitted photons in the experimental definition of the hard particle 4-momentum. With a lower threshold of order $m_{\rm W}$ on the energy of the photons to be recombined, the net effect on our prediction should be of the order of a single EW logarithm $\log E^2/m_{\rm W}^2$. A more detailed assessment of this aspect, and of the possible interplay between the QED and the EW bosons collinear emissions, requires the inclusion of single logarithms and goes beyond the scope of the present work. Similar considerations hold for the collinear emission of QCD gluons to be collected into jets, in the case of colored final states.

Up to the caveats outlined above, there are good perspectives for the actual direct experimental detectability of semi-inclusive cross-sections. The situation is arguably more problematic for the exclusive cross-section. In exclusive final states, we require the presence of the two hard particles defined as above, plus the absence of any massive vector boson (since $\mathfrak{h}(q) > q^2 = m^2$, as discussed at the end of the previous section), or photons above the hardness upper threshold $m_{\rm W}^2$. However, it is experimentally impossible to impose this radiation veto strictly because the limited coverage of the detector in the forward and backward regions will not allow to tag EW bosons or photons that are collinear to the beam. Furthermore our definition of the exclusive cross-section is problematic in the case of QCD-colored final states. Indeed if the upper cut $\mathfrak{h}(q) < m_{\rm W}^2$ had to be imposed also on gluon radiation, QCD effects should be included in the exclusive density matrix evolution (but not in the semi-inclusive one, where they cancel because of color inclusivity), resulting in a large QCD Sudakov suppression factor in eq. (3.23). This factor is as small as $\exp[-\alpha_s/(4\pi)(8/3)\log^2 E_{\rm cm}^2/m_{\rm W}^2] \sim 0.03$ for di-quark final states at the highest VHEL energy $E_{\rm cm} = 30$ TeV, entailing a strong suppression of the cross-section. Avoiding this suppression requires a definition of the exclusive cross-section with a higher threshold on the QCD radiation. We will further comment in the Conclusions on the limitations of the exclusive cross-section definition employed in this chapter.

3.1.2 Di-fermion production

The first process we investigate is the production of a highly energetic pair of fermions

$$\ell^+(k_1)\,\ell^-(k_2) \to \bar{f}(k_3)\,g(k_4) + X\,, \tag{3.24}$$

where f and g can be one of the six quarks, a lepton $\ell' \neq \ell$ or a neutrino $\nu_{\ell'}$. We do not discuss explicitly the final states with the same leptonic flavor as the initial state, $\ell' = \ell$, but these processes will be employed for the muon collider sensitivity projections in Section 3.2. As previously discussed, the final state is characterized (both for exclusive and semi-inclusive processes) by an invariant mass for the $(\bar{f}, g)$ pair that is almost equal to the center of mass energy $E_{\rm cm}$ of the colliding leptons and by central scattering angle

$\theta_* \in [30°, 150°]$. Here θ_* is the angle between the incoming ℓ^+ and the final anti-fermion $\bar{f}$ in the lab frame. Notice that θ_* almost coincides with the scattering angle in the center of mass frame of the hard process, because of the tight cut on the invariant mass of the $(\bar{f}, g)$ pair.

In order to resum the DL it is convenient to organize the calculation of the cross-section in terms of amplitudes and density matrices whose external legs are canonical irreducible representations of the EW group. This is trivial to achieve for the di-fermion process because the helicity eigenstates of quarks and leptons in the massless limit do indeed transform as canonical representations (doublets and singlets, with specific hypercharge), reported for completeness in Appendix 3.B. Furthermore, since we restrict our attention to inelastic processes $\ell' \neq \ell$, the only sizable helicity amplitudes are those with the same chirality χ_I (χ_O) for the two incoming (outgoing) fermions, corresponding to helicities $\bar{\psi}_{+1/2}\psi_{-1/2}$ for $\chi = L$ and $\bar{\psi}_{-1/2}\psi_{+1/2}$ for $\chi = R$. The dominance of such amplitudes holds in the SM because of the vector-like structure of gauge interaction, and it will be preserved by the 4-fermions new physics contact interaction operators we will study in Section 3.2. We thus have to deal with four polarized cross-sections for each di-fermion production process, labeled by $\chi_I\chi_O = LL, LR, RL, RR$. Each such cross-section will be obtained from the diagonal $\alpha = \bar{\alpha}$ entries of the density matrices of Section 3.1.1, times the appropriate phase-space factors.

Exclusive processes

Exclusive cross-sections are readily obtained from eq. (3.23), and take the form

$$\frac{d\sigma_{\rm ex}}{d\cos\theta_*} = e^{\rm DL}\frac{d\sigma_B}{d\cos\theta_*}\,, \tag{3.25}$$

in terms of the corresponding Born-level differential cross-sections. The Double Log exponent DL is of order $g^2/16\pi^2\log^2(E^2_{\rm cm}/m^2_{\rm W})$, which ranges from 0.14 at $E_{\rm cm} = 3$ TeV up to 0.25 (0.38) for $E_{\rm cm} = 10(30)$ TeV, times the sum of the four SU(2) Casimir of the external legs. For LL chirality processes this factor is as large as $4 \times 1/2(1/2+1) = 3$, showing that DL resummation is mandatory at VHEL energies $E_{\rm cm} \geq 10$ TeV, at least for this chirality. Double logs are still considerable for LR and RL chirality, while they get smaller in the RR configuration because $g'^2 \sim g^2/4$. Resummation might instead not be necessary for $E_{\rm cm} = 3$ TeV. However it will still be needed to include the effects of radiation at fixed order since we aim, eventually, at theoretical predictions with percent-level accuracy.

The DL Sudakov exponents in eq. (3.25) are listed in Table 3.1. The processes are labeled taking into account that electric charge conservation enforces $g = f$ in eq. (3.24), since a charge mismatch cannot be compensated by the emission of charged W bosons, which is forbidden in exclusive processes. The table also reports single logarithm (SL)

	3 TeV			10 TeV			30 TeV		
	DL	$e^{\mathrm{DL}}-1$	$\mathrm{SL}(\frac{\pi}{2})$	DL	$e^{\mathrm{DL}}-1$	$\mathrm{SL}(\frac{\pi}{2})$	DL	$e^{\mathrm{DL}}-1$	$\mathrm{SL}(\frac{\pi}{2})$
$\ell_L \to \ell'_L$	-0.46	-0.37	0.25	-0.82	-0.56	0.33	-1.23	-0.71	0.41
$\ell_L \to q_L$	-0.44	-0.36	0.25	-0.78	-0.54	0.34	-1.18	-0.69	0.42
$\ell_L \to e_R$	-0.32	-0.27	0.13	-0.56	-0.43	0.17	-0.85	-0.57	0.21
$\ell_L \to u_R$	-0.27	-0.24	0.11	-0.48	-0.38	0.15	-0.72	-0.51	0.18
$\ell_L \to d_R$	-0.24	-0.21	0.10	-0.43	-0.35	0.13	-0.64	-0.47	0.16
$\ell_R \to \ell'_L$	-0.32	-0.27	0.13	-0.56	-0.43	0.17	-0.85	-0.57	0.21
$\ell_R \to q_L$	-0.30	-0.26	0.12	-0.53	-0.41	0.16	-0.79	-0.55	0.21
$\ell_R \to \ell'_R$	-0.17	-0.16	0.07	-0.30	-0.26	0.09	-0.46	-0.37	0.12
$\ell_R \to u_R$	-0.12	-0.12	0.05	-0.22	-0.20	0.07	-0.33	-0.28	0.08
$\ell_R \to d_R$	-0.09	-0.09	0.04	-0.17	-0.16	0.05	-0.25	-0.22	0.06

Table 3.1: Double and single logarithmic corrections to the exclusive processes $\ell^+\ell^- \to \bar{f}f$. The single-logarithmic corrections are evaluated at $\theta_* = \pi/2$.

contributions computed at the fixed one loop order, which we extract from Ref.s [165].[12] Specifically, we employ the general formulae of Ref.s [165] to compute the 1-loop log-enhanced cross-section, we subtract the corresponding DL and normalize to the Born cross-section. We also subtract the single logarithms from the Renormalization Group evolution, because we decided to compute the Born amplitude with the EW couplings at the hard scale E_{cm}.[13] Notice that the threshold for photon recombination into the hard final state particles matters at the single-logarithmic order. Here we assume a scale of recombination of order m_{W}, for which the SL terms can be easily obtained by adding a fictitious photon mass $m_\gamma = m_{\mathrm{W}}$ to the calculations of Ref.s [165, 168]. The SL terms obtained in this way can be used for "improved" theoretical predictions

$$\frac{d\sigma^{\mathrm{SL_1}}_{\mathrm{ex}}}{d\cos\theta_*} = e^{\mathrm{DL}}(1+\mathrm{SL}(\theta_*))\frac{d\sigma_B}{d\cos\theta_*}\,, \tag{3.26}$$

that include single logarithms at fixed 1-loop order. We see in Table 3.1 that the SL contributions are relatively large. It is unclear whether they require resummation or if including them at fixed order (definitely higher than 1-loop, if we target 1% accuracy) is sufficient.

Notice that, unlike double logarithms, the single logarithm contributions are not proportional to the Born-level amplitude of the same scattering process. Namely the amplitudes of the neutral-current processes in Table 3.1 receive SL corrections that are proportional to Born charged-current amplitudes. Therefore it should be kept in mind the SL terms in

[12] Two loops NLL results for four-fermion processes are also available in [170, 171].

[13] The calculation is similar to the one performed by two of us in Ref. [172]. We refer the reader to Section 2.3 of [172] for additional details, concerning in particular the inclusion of non-log-enhanced angular-dependent terms.

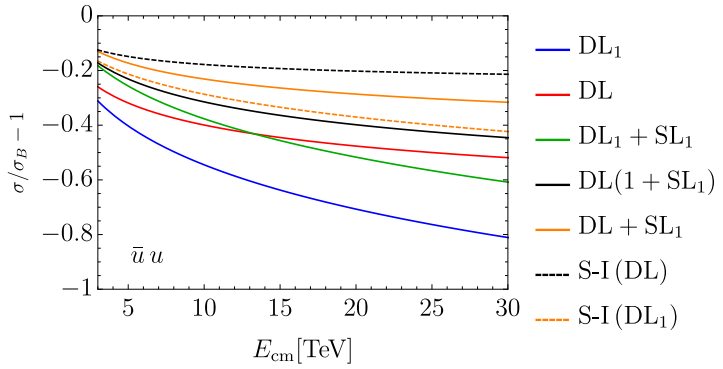


Figure 3.3: Impact of radiative corrections on the production of two up quarks at the VHEL. The solid lines represent different predictions for the exclusive cross-section. The dashed lines are double-logarithm semi-inclusive cross-sections resummed (in black) or at one loop (in orange).

eq. (3.26), which we normalized to the Born cross-section of the process, depend on the ratio between charged and neutral current Born amplitudes. We evaluated the amplitude ratio within the SM to produce the results in Table 3.1. However the amplitude ratio depends on the new physics contact interactions we consider in Section 3.2, entailing a dependence of the SL terms on the new physics parameters. This is not the case for the double logarithms, which are completely universal and insensitive to short-distance physics. The single logarithms also carry a non-trivial dependence on the scattering angle θ_*, as explicitly indicated in eq. (3.26). In Table 3.1 they are evaluated at central angle $\theta_* = \pi/2$, where they are always positive. They can become negative, and typically increase in magnitude, in the forward and backward scattering regions, which we however exclude with the central cut $\theta_* \in [30°, 150°]$. Finally, notice that the SL terms are affected by the sizable mass of the top quark, which we do include in the $t\bar{t}$ production process.

The impact of EW radiation effects on the total (unpolarized) cross-section in the central region, relative to the Born, is displayed in Figure 3.3 as a function of $E_{\rm cm}$. The production of two light up-type quarks is considered for illustration, but the results for the other final states are similar. The blue line is the one-loop DL prediction without exponentiation, while in red we report the resummed DL prediction in eq. (3.25). The green line (labeled DL$_1$+SL$_1$) represents the fixed-order one loop DL plus SL, while in black we report the SL-improved prediction in eq. (3.26). The dashed lines are semi-inclusive cross-sections computed below. We notice a significant cancellation between double and single logarithmic terms. However this cancellation is not expected to be structural and to survive at higher orders in perturbation theory.

We do not try to assign theoretical uncertainties to our predictions. However an upper bound can be obtained by considering the orange line in the figure, in which the resummed

DL are combined additively with the SL (i.e., as $e^{\rm DL} + {\rm SL}$), rather than multiplicatively. An alternative estimate of the uncertainties could be obtained by varying the scale of the EW couplings employed for the evaluation of the radiation terms DL and SL. Varying this scale from $m_{\rm W}$ (which we employ for our predictions) to $E_{\rm cm}$, the relative change of the radiation effects is rather small, typically at the 10% level or less.

Semi-inclusive processes

The semi-inclusive cross-sections are the diagonal $\alpha = \bar{\alpha}$ entries of the semi-inclusive density matrix in eq. (3.22), with the appropriate K_i exponential factors for each external leg. The factors only depend on the $\mathrm{SU}(2)_L$ quantum numbers of the legs and not of their hypercharge (and QCD color). They are provided by eq. (3.83) for L-chirality external external legs (which transform as doublets or conjugate-doublets) and they are trivial for the R-handed singlets. Notice that eq. (3.83) (and the same is true for the triplet exponential factor (3.88)) does not mix diagonal with off-diagonal entries of the density matrix. Namely if we set $\alpha = \bar{\alpha}$ we obtain a tensor that is diagonal in β and $\bar{\beta}$. Therefore the DL resummed cross-sections, collected in a vector $d\vec{\sigma}_{\rm si}$, are linear combinations of the Born cross-sections $d\vec{\sigma}_B$. We express this relation as

$$\frac{d\vec{\sigma}_{\rm si}}{d\cos\theta_*} = e^{\rm DL} \cdot \frac{d\vec{\sigma}_B}{d\cos\theta_*}\,, \tag{3.27}$$

where the Double-Logarithm terms "DL" are now matrices connecting the Born cross-sections of different processes unlike for exclusive processes (3.25).

For an explicit illustration of the semi-inclusive cross-section calculation, and of the main features of the result, we consider the RL-chirality production processes. In this case, we have

$$d\vec{\sigma}_{\rm si} = \begin{pmatrix} d\sigma_{\rm si}(\ell^+_R\ell^-_R \to \bar{u}_L u_L) \\ d\sigma_{\rm si}(\ell^+_R\ell^-_R \to \bar{u}_L d_L) \\ d\sigma_{\rm si}(\ell^+_R\ell^-_R \to \bar{d}_L u_L) \\ d\sigma_{\rm si}(\ell^+_R\ell^-_R \to \bar{d}_L d_L) \end{pmatrix}, \quad d\vec{\sigma}_B = \begin{pmatrix} d\sigma_B(\ell^+_R\ell^-_R \to \bar{u}_L u_L) \\ 0 \\ 0 \\ d\sigma_B(\ell^+_R\ell^-_R \to \bar{d}_L d_L) \end{pmatrix} = d\sigma_B(\ell^+_R\ell^-_R \to \bar{q}_L q_L) \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \tag{3.28}$$

where "u" and "d" denote here the up and down components of a L-handed fermion doublet. The exponentiated DL matrix reads

$$e^{\rm DL} = \frac{1}{4}\, e^{-\mathcal{L}} \begin{pmatrix} 4\cosh^2(\mathcal{L}/2) & 2\sinh(\mathcal{L}) & 2\sinh(\mathcal{L}) & 4\sinh^2(\mathcal{L}/2) \\ 2\sinh(\mathcal{L}) & 4\cosh^2(\mathcal{L}/2) & 4\sinh^2(\mathcal{L}/2) & 2\sinh(\mathcal{L}) \\ 2\sinh(\mathcal{L}) & 4\sinh^2(\mathcal{L}/2) & 4\cosh^2(\mathcal{L}/2) & 2\sinh(\mathcal{L}) \\ 4\sinh^2(\mathcal{L}/2) & 2\sinh(\mathcal{L}) & 2\sinh(\mathcal{L}) & 4\cosh^2(\mathcal{L}/2) \end{pmatrix}, \tag{3.29}$$

where $\mathcal{L} = g^2/16\pi^2 \log^2(E^2_{cm}/m^2_{\rm W})$.

We see that DL effects induce a non-vanishing cross-section for charged processes with $g \neq f$ in eq. (3.24), such as $\bar{u}_L d_L$ and $\bar{d}_L u_L$ production. Clearly this stems from the emission of real soft W-bosons, which is allowed in the semi-inclusive final state. Such charged cross-sections are proportional to the Born cross-section for the corresponding neutral ($\bar{u}_L u_L$ or $\bar{d}_L d_L$) processes, and they are not drastically smaller than those because the double-logarithm is sizable at VHEL energies. Therefore they can be measured bringing additional sensitivity to the charge-preserving Born amplitudes and to the corresponding short-distance new physics effects. The interplay with short-distance physics is even more interesting for the LL-chirality process. In that case, $\vec{\sigma}_{\rm si}$ is a 16-dimensional vector that contains 4 observable ($\ell^+\ell^-$-initiated) processes with final states $\bar{u}_L u_L$, $\bar{u}_L d_L$, $\bar{d}_L u_L$ and $\bar{d}_L d_L$. DL is a 16×16 matrix that relates the observable processes to 16 Born amplitudes, among which those (like, e.g., $\bar{\nu}_\ell \ell^- \to \bar{u}_L d_L$) that are sensitive to new charged current interactions. We can thus probe the latter interactions even with the neutral $\ell^+\ell^-$ VHEL collisions.

The black dashed lines in Figure 3.3 quantify the impact of the EW radiation effects on the neutral semi-inclusive cross-sections relative to the Born predictions. The effects are smaller than for exclusive cross-sections, as qualitatively expected owing to the partial cancellation between virtual and real radiation. While this suggests that resummation might play a less relevant role in semi-inclusive predictions, we point out that one-loop double logarithms are insufficient for accurate predictions. This is shown in the purple dashed line in the figure, which is obtained by truncating at the one-loop order the exponentiated DL matrix. It would be interesting to study the impact of single logarithms on the predictions. This could be achieved by combining the single radiative logs from Ref.s [165, 168] with the factorized formulas for real emissions in Ref.s [158, 86] (which however would have to be extended to include also the soft radiation region), but is left to future work.

As a final technical note, we remark that the DL matrix is negative semi-defined with a single vanishing eigenvector that corresponds to the "fully-inclusive" cross-section, further averaged over the $\mathrm{SU}(2)_L$ color of the initial states. Specifically the vanishing eigenvector of eq. (3.29) is $(1,1,1,1)^t$, which corresponds to the sum of the cross-sections over the $\mathrm{SU}(2)_L$ gauge indices of the final states. Therefore in this case the double logarithmic effects cancel on the "fully-inclusive" cross-section, in accordance with the KLN theorem since the right-handed initial leptons are $\mathrm{SU}(2)_L$ singlets. Clearly this does not happen for the LL-chirality processes (nor for LR-chirality) and the average over leptons and neutrinos in the initial states would be necessary for the cancellation. The vanishing eigenvalue controls the behavior of the DL exponential at asymptotically high energies.

In the case of eq. (3.29), we have

$$e^{\rm DL} \xrightarrow{E_{\rm cm}\to\infty} \frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \tag{3.30}$$

and all the semi-inclusive cross-section listed in eq (3.28) become equal. Notice however this only holds at asymptotic energies, way above the VHEL energies. Cross-sections equality becomes a reasonable (better than order-one) approximation only for if $g^2/16\pi^2 \log^2(E_{\rm cm}^2/m_{\rm W}^2)$ is as large as ~ 1.5, i.e. $E_{\rm cm} \gtrsim 10000$ TeV.

3.1.3 Di-boson production

We now turn to the production of two energetic vector or Higgs bosons. We are interested in reactions that are not power-like suppressed at high energy, therefore we restrict our attention to "longitudinal" processes entailing the production of zero-helicity W and Z bosons and Higgs, and to "transverse" di-boson processes where the W and the Z (or, the photon) have ± 1 helicities. Indeed the "mixed" longitudinal/transverse production processes are suppressed by $m_{\rm W}/E_{\rm cm}$ at the amplitude level, as readily understood (see e.g. [161, 86]) by combining the Goldstone Boson Equivalence Theorem with the selection rules associated with the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ SM group.

The new physics interactions we consider in Section 3.2 only affect longitudinal di-boson production cross-sections, which thus play the role of the signal in our analysis. We nevertheless also need the transverse cross-sections for an estimate of the background. We discuss the calculation of the (exclusive and semi-inclusive) cross-sections for the two type of processes in turn.

Longitudinal di-boson

We consider the production, out of $\ell^+\ell^-$, of one of the following hard final states

$$W_0^+ W_0^- \,, \qquad Z_0 h \,, \qquad W_0^\pm Z_0 \,, \qquad W_0^\pm h \,, \tag{3.31}$$

where the subscript "$_0$" refers to the helicity of the massive vectors, and "h" denotes the physical Higgs particle. Obviously only the first two final states can be produced in an exclusive process, while the latter ones require the emission of at least one charged W and therefore they only occur at the semi-inclusive level. Notice that the ones listed above are the only hard final states with longitudinal bosons and Higgs that can be produced by soft EW bosons radiation out of sizable Born-level $2 \to 2$ cross-sections. Therefore they are the only longitudinal di-boson processes that can be considered for

precise VHEL measurements in the high-energy regime.

At energies much above $m_{\rm W}$, the adequate description of longitudinally-polarized massive vectors is provided by the charged and neutral Goldstone boson scalars $\pi^\pm$ and π_0 (see Appendix 3.B). Together with the Higgs, they form a canonical $\mathrm{SU}(2)_L\times\mathrm{U}(1)_Y$ doublet H with 1/2 hypercharge, reported in eq. (3.78). We thus need to consider amplitudes and density matrices associated with the hard processes

$$\begin{aligned}\ell^+_{-1/2}(k_1)\ell^-_{+1/2}(k_1) &\to \bar{H}(k_3,\,\alpha_3^{\bar{\rm d}})\,H(k_4,\,\alpha_4^{\rm d})\,,\\ \bar{\ell}_{+1/2}(k_1,\,\alpha_1^{\bar{\rm d}})\ell_{-1/2}(k_2,\,\alpha_2^{\rm d}) &\to \bar{H}(k_3,\,\alpha_3^{\bar{\rm d}})\,H(k_4,\,\alpha_4^{\rm d})\,,\end{aligned}\tag{3.32}$$

for, respectively, L-handed and R-handed production.[14] For the gauge group indices we employ the same notation as in eq. (3.3), supplemented by the superscripts $^{\rm d}$ ($^{\bar{\rm d}}$) to indicate that the indices belong to the doublet (conjugate-doublet) representation. With a slight abuse of notation we are denoting as $\ell_{-1/2}=(\nu_{\ell,-1/2},\ell^-_{-1/2})^t$ the lepton doublet with $-1/2$ helicity and with $\ell_{+1/2}$ the conjugate-doublet with helicity $+1/2$. Notice that final states with two H or two $\bar{H}$ need not to be included because they are power-like suppressed at high energy by hypercharge conservation.

The relevant density matrices are obtained as a straightforward application of the results in Section 3.1.1. The need for employing H and $\bar{H}$ as external states does not pose any additional difficulty (relative to the di-fermion processes) in the evaluation of exclusive cross-sections. That is because the double logs are mere multiplicative factors in front of the Born-level density matrix (3.23). Therefore the exclusive cross-sections still take the form of eq. (3.25) and are proportional to the corresponding Born-level predictions. For the semi-inclusive cross-section, we can proceed as for di-fermions in the determination of the K_i exponential factors, using in particular eq. (3.83) which also holds in the present case because H and $\bar{H}$ are doublets. However in order to apply eq. (3.22) we must first express the $\mathcal{D}^{\alpha\bar{\alpha}}_{\rm si}$ density matrix, which is written in the isospin basis (H and $\bar{H}$), in the physical basis of the charge and CP eigenstates h, $Z_0=\pi_0$ and $W_0^\pm=\pi^\pm$. This is achieved by simply inverting eq. (3.78). The final result can again be expressed in terms of the Born-level cross-sections in the form of eq. (3.27).

The results display the same qualitative features as di-fermions. In particular we observe the same interplay between short-distance physics affecting the neutral- and the charged-current Born amplitudes, which we investigate in Section 3.2 in details. Also at the quantitative level, the relative impact of radiation radiation is similar, as expected because $\mathrm{SU}(2)_L$ doublets are involved also in these processes. This is shown in the left panel of Figure 3.4, where we show the exclusive and semi-inclusive cross-section predictions for $W_0^+W_0^-$. The different predictions are obtained as explained in the previous section for the di-fermion processes. Notice in particular the exclusive predictions that include

[14] The production from opposite-chirality leptons is negligible, both in the SM and in the presence of the new contact interactions we investigate in the following section.

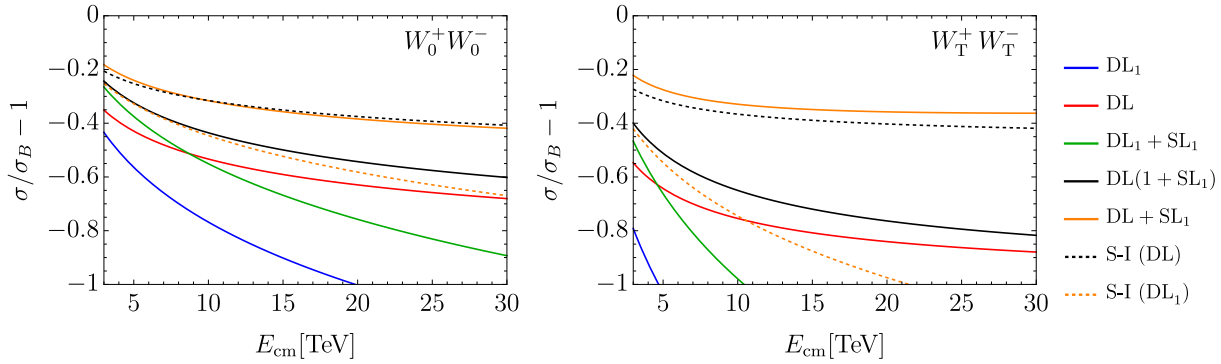


Figure 3.4: Same as Figure 3.3, but for di-boson production. As explained in Section 3.2.2, the cross-sections for W^+W^- production are integrated in the angular region $\theta_* \in [67°, 150°]$.

one-loop single logarithms as in eq. (3.26). We employ these predictions for exclusive cross-section in the phenomenological studies of Section 3.1.1.

Transverse di-boson

Vector bosons (W, Z, or γ) with transverse helicity $\mathrm{T} = \pm 1$ have zero hypercharge and they decompose as a real triplet plus a singlet under the SM $\mathrm{SU}(2)_L$, as in eq. (3.79). Therefore three non-power-suppressed hard processes have to be considered for L-handed production

$$\begin{aligned} \bar{\mu}_{+1/2}(k_1, \alpha_1^{\bar{\mathrm{d}}})\ell_{-1/2}(k_2, \alpha_2^{\mathrm{d}}) &\to B(k_3)\, B(k_4)\,, \\ \bar{\mu}_{+1/2}(k_1, \alpha_1^{\bar{\mathrm{d}}})\ell_{-1/2}(k_2, \alpha_2^{\mathrm{d}}) &\to W(k_3, \alpha_3^{\mathrm{t}})\, B(k_4)\,, \\ \bar{\mu}_{+1/2}(k_1, \alpha_1^{\bar{\mathrm{d}}})\ell_{-1/2}(k_2, \alpha_2^{\mathrm{d}}) &\to W(k_3, \alpha_3^{\mathrm{t}})\, W(k_4, \alpha_4^{\mathrm{t}})\,, \end{aligned} \tag{3.33}$$

while only one is relevant for the production initiated by R-handed leptons [15]

$$\ell^+_{-1/2}(k_1)\ell^-_{+1/2}(k_2) \to B(k_3)\, B(k_4)\,, \tag{3.34}$$

The "t" superscript in eq. (3.33) refers to the triplet nature of the W indices.

Unlike for di-fermion and longitudinal di-boson, the transverse di-boson cross-sections for L-handed initial leptons are linear combinations of several distinct density matrices with different $\mathrm{SU}(2)_L$ quantum numbers. Therefore the exclusive cross-sections are not proportional, unlike in eq. (3.25), to the corresponding Born cross-sections in general. For instance in the $\gamma\gamma$ cross-section the contribution from the WW final state experiences a stronger Sudakov suppression (3.23) than one from the BB (or WB) final state, owing to the higher $\mathrm{SU}(2)_L$ Casimir of the final states.

[15] The Born process $\ell^+_{-1/2}(k_1)\ell^-_{+1/2}(k_2) \to W(k_3, \alpha_3^{\mathrm{t}})\, W(k_4, \alpha_4^{\mathrm{t}})$ is power-suppressed in the SM.

The evaluation of the semi-inclusive cross-sections proceeds as for the longitudinal di-bosons. Namely we derive the cross-sections for the physical states by inverting eq. (3.79) and we compute the double-logarithm exponentials using eq. (3.88) on the $SU(2)$ triplet subspace. Clearly the need of combining different density matrices complicates the calculation, but it does not introduce any novel conceptual aspect. At the quantitative level instead, the situation is significantly different than for di-fermions and longitudinal di-bosons. As shown on the right panel of Figure 3.4, EW radiation effects are much larger due to the larger Casimir $c_{\rm t} = 2$ of the triplet representation. A sufficiently accurate modeling of these effects will probably require resummation even at the lowest VHEL energy $E_{\rm cm} = 3$ TeV.

The figure reports the cross-section of the $W_{\rm T}^+ W_{\rm T}^-$ final state. This final state, together with $W_{\rm T}^+ Z_{\rm T}$, is the only transverse di-boson process we will consider in Section 3.2 (as a background to the corresponding longitudinal processes). Notice however that there are many other transverse di-boson processes (namely ZZ, $Z\gamma$, $\gamma\gamma$, and $W\gamma$) that can be measured at the VHEL. These processes probe heavy new physics in the EW sector. In particular, as shown in Refs. [136, 137, 140], they are sensitive (together with di-fermions) to minimal Dark Matter in large-multiplets. The large effects of EW radiation might have a strong impact on these studies.

3.2 Sensitivity projections

As described in the Introduction, we target effects from short-distance new physics that grow quadratically with the collision energy, to be probed in $\ell^+\ell^-$ collisions at the highest available energy $E = E_{\rm cm}$. In this section we consider the dimension-6 EFT operators listed in Table 3.2, and we estimate the sensitivity of muon colliders of energies $E_{\rm cm} = 3$, 10, 14 or 30 TeV to their Wilson coefficients. We assume a baseline integrated luminosity [99]

$$\widehat{\mathcal{L}} = 10 \text{ ab}^{-1} \left(\frac{E_{\rm CM}}{10 \text{ TeV}} \right)^2 . \tag{3.35}$$

Semi-quantitative comments on the impact of a reduced luminosity target are postponed to the Conclusions. We base our projections on statistically-dominated measurements of exclusive and semi-inclusive cross-sections for the processes listed in Table 3.3. In the table, for each process we label with a check mark the operators that produce a quadratically growing-with-energy correction to the SM cross-section.

The target EFT operators are selected to represent generic manifestations, at energies much below the new physics scale, of the BSM scenarios we investigate in Section 3.2.3. These are Composite Higgs, Composite Top and a minimal Z' model, which we select as concrete examples of new physics in the Higgs, Top and EW-gauge sectors. Among the many operators that emerge in these scenarios, we focused our attention on those that display energy growth in $2 \to 2$ scattering processes at the muon collider. We will see in

Section 3.2.3 that other operators offer a weaker sensitivity to the same BSM scenarios.

The phenomenological analysis of the various processes listed in Table 3.3 is described in Sections 3.2.1 and 3.2.2, focusing respectively on the effects of the "W&Y" and of the "Di-boson" operators of Table 3.2. In an attempt to mimic realistic experimental results, we include reconstruction (and, in some case, mistag) efficiencies at a level that is comparable with the CLIC detector performances, which we extract, whenever possible, from Refs. [166, 173]. Table 3.3 displays surprisingly low efficiencies for certain processes (e.g., $t\bar{t}$), entailing a considerable degradation of the measurement uncertainty. In Sections 3.2.1 and 3.2.2 we also present our results for the sensitivity of muon colliders to the corresponding set of operators, with the main aim of outlining the impact of the EW radiation effects on the analysis. The operators in the last class, dubbed "3$^{\rm rd}$ family" in Table 3.2, are not discussed explicitly but the sensitivity projection results are reported in Appendix 3.C. The relevant final states, tt, bb and tb are discussed in Section 3.2.1.

3.2.1 W&Y operators

The first two operators we consider are those associated with the W and Y parameters of LEP EW precision tests [32], namely O_{2W} and O_{2B} defined as in Table 3.2. These operators arise in the so-called *universal* scenarios [32, 31], that is new physics that couples dominantly to the bosonic sector of the SM. Employing O_{2W} and O_{2B} is convenient in the low-energy context of the LEP experiment, however for our purpose it is better to trade them for the current-current operators O'_{2W} and O'_{2B} (see again Table 3.2), using the SM equations of motion. In doing so, we neglect the contribution to the O'_W and O'_B operators, which are expected to have no impact on the sensitivity. In what follows we parameterize the O'_{2W} and O'_{2B} operator coefficients

$$G'_{2W} = -\frac{g^2 \mathrm{W}}{2m_{\mathrm{W}}^2}\,, \qquad G'_{2B} = -\frac{g'^2 \mathrm{Y}}{2m_{\mathrm{W}}^2}\,, \tag{3.36}$$

in terms of the dimensionless parameters W and Y.

The relevant scattering processes, listed in Table 3.3, are the production of two energetic fermions in the central region of the detector. Specifically, as explained at the end of Section 3.1, we have in mind the two hard particles whose invariant mass is higher than around 85% of the total collider $E_{\rm cm}$, and a scattering angle $\theta_* \in [30^\circ, 150^\circ]$. We assume perfect detector sensitivity to massive W and Z bosons of arbitrary low 3-momentum, enabling the measurement of exclusive scattering cross-sections where the emission of massive vectors (and of photons with hardness above m_{W}^2) is vetoed. The exclusive cross-section measurements are combined with the semi-inclusive cross-sections, where the emission of an arbitrary number (including zero) of massive vectors or hard photons is allowed.

	SILH basis	Warsaw-like basis
W&Y	$O_{2W} = (D_\mu W^{\mu\nu,a})^2$ $O_{2B} = (\partial_\mu B^{\mu\nu})^2$	$O'_{2W} = J_L^{a,\mu} J_{L,\mu}^a \qquad J_L^{a,\mu} = \frac{1}{2}\sum_f \bar{f}\gamma^\mu\sigma^a f$ $O'_{2B} = J_Y^\mu J_{Y,\mu} \qquad J_Y^\mu = \sum_f Y_f \bar{f}\gamma^\mu f$
Di-boson	$O_W = \frac{ig}{2}(H^\dagger\sigma^a \overleftrightarrow{D}_\mu H) D^\nu W^a_{\mu\nu}$ $O_B = \frac{ig'}{2}(H^\dagger \overleftrightarrow{D}_\mu H)\partial^\nu B_{\mu\nu}$	$O'_W = \frac{g^2}{4}(H^\dagger i\overleftrightarrow{D}_\mu \sigma^a H)(\bar{L}_L\gamma^\mu\sigma^a L_L)$ $O'_B = -\frac{g'^2}{4}(H^\dagger i\overleftrightarrow{D}_\mu H)(\bar{L}_L\gamma^\mu L_L)$ $-\frac{g'^2}{2}(H^\dagger i\overleftrightarrow{D}_\mu H)(\bar{l}_R\gamma^\mu l_R)$
3rd family	$O_{qD}^{(3)} = (\bar{q}\gamma^\mu\sigma^a q)\left(D^\nu W^a_{\mu\nu}\right)$ $O_{qD}^{(1)} = (\bar{q}\gamma^\mu q)(\partial^\nu B_{\mu\nu})$ $O_{tD} = (\bar{t}\gamma^\mu t)(\partial^\nu B_{\mu\nu})$	$O'^{(3)}_{qD} = (\bar{q}\gamma^\mu\sigma^a q)\, J^a_{L,\mu}$ $O'^{(1)}_{qD} = (\bar{q}\gamma^\mu\sigma^a q)\, J_{Y,\mu}$ $O'_{tD} = (\bar{t}\gamma^\mu\sigma^a t)\, J_{Y,\mu}$

Table 3.2: The operators under consideration in their "SILH" [11] form and, after using the equations of motion, expressed as a linear combination of Warsaw [51] operators. Y_f is the hypercharge of the fermionic field f. In the operators involving the 3rd family the fields t and q denote respectively the right-handed and left-handed top quark.

Process	N (Ex)	N (S-I)	Eff.	O'_{2W}	O'_{2B}	O'_W	O'_B	$O'^{(3)}_{qD}$	$O'^{(1)}_{qD}$	O'_{uD}
e^+e^-	6794	9088	100%	✓	✓					
$e\nu_e$	—	2305	100%	✓	✓					
$\mu^+\mu^-$	206402	254388	100%	✓	✓					
$\mu\nu_\mu$	—	93010	100%	✓	✓					
$\tau^+\tau^-$	6794	9088	25%	✓	✓					
$\tau\nu_\tau$	—	2305	50%	✓	✓					
jj (Nt)	19205	25725	100%	✓	✓					
jj (Ch)	—	5653	100%	✓	✓					
$c\bar{c}$	9656	12775	25%	✓	✓					
cj	—	5653	50%	✓	✓					
$b\bar{b}$	4573	6273	64%	✓	✓			✓	✓	
$t\bar{t}$	9771	11891	5%	✓	✓			✓	✓	✓
bt	—	5713	57%	✓	✓			✓	✓	✓
Z_0h	680	858	26%			✓	✓			
$W_0^+W_0^-$	1200	1456	44%			✓	✓			
$W_T^+W_T^-$	2775	5027	44%							
$W^\pm h$	—	506	19%			✓	✓			
$W_0^\pm Z_0$	—	399	23%			✓	✓			
$W_T^\pm Z_T$	—	2345	23%							

Table 3.3: The exclusive and semi-inclusive processes employed for the sensitivity projections. The operators that give a growing-with-energy contribution to each operator are labeled with a check mark. The expected number of events (before efficiencies) is for $E_{\rm cm} = 10$ TeV with the integrated luminosity (3.35).

For each inclusive and semi-inclusive final state, we employ cross-section measurements in 10 equally-spaced bins of $\cos\theta_*$ in the range $[-\sqrt{3}/2, \sqrt{3}/2]$. In processes (e.g., jj, or $b\bar{b}$) where the two final states are effectively indistinguishable, $\cos\theta_*$ is defined to be positive and 5 bins are employed. We assume cross-section measurements with purely statistical uncertainties, which we estimate based on the number of events that are expected in the SM.

Of course in order to combine the exclusive and semi-inclusive cross-sections for the same (neutral) hard final state we must take into account that the exclusive events are also counted in the measurement of the semi-inclusive cross-section. It is thus convenient to consider a cross-section *with radiation*, defined as the difference between the semi-inclusive and the exclusive cross-sections

$$\sigma_{\rm rad} \equiv \sigma_{\rm si} - \sigma_{\rm ex} \,. \tag{3.37}$$

The measurement of $\sigma_{\rm rad}$ can be combined with the one of $\sigma_{\rm ex}$ since they are statistically independent. For charged hard final states there is instead only one type of cross-section, which necessarily involves EW radiation emission by charge conservation. We will refer to the charged cross-section as "semi-inclusive" or "with radiation" interchangeably.

We now discuss the di-fermion processes individually.

- $\boldsymbol{e^+e^-}$, $\boldsymbol{\mu^+\mu^-}$ **and** $\boldsymbol{\tau^+\tau^-}$**:** We assume 100% reconstruction efficiency for muon and electrons, and an efficiency of 50% [166] for each τ lepton. Notice that the cross-section for muons is around 30 times larger than for the other leptons. This is mostly due to the t-channel enhancement of the elastic $\mu^+\mu^-$ scattering.
- $\boldsymbol{c\bar{c}}$ **and** $\boldsymbol{b\bar{b}}$**:** We assume 50% and 80% efficiency for tagging respectively charm and bottom quark jets [166]. We ignore the mis-tag of light jets, as well as c/b misidentification. No information on the charge of the tagged quark is employed.
- $\boldsymbol{jj}$**:** We consider the production of two light quarks u, d or s, which we suppose to be reconstructed as jets with 100% efficiency. In Table 3.3 we report separately the production of a neutral (Nt) and of a charged (Ch) quark/anti-quark pair, but the two processes are collected into a single 2-jets final state. We also include the contribution from mistagged b and c quarks.
- $\boldsymbol{t\bar{t}}$**:** Based on Ref.s [173, 174], we estimate as 5% the total efficiency for the reconstruction of the $t\bar{t}$ pair. This (somewhat low) efficiency estimate only includes the semi-leptonic $t\bar{t}$ final states, in which the charge of the tagged top quarks can be measured.
- $\boldsymbol{tb}$ **and** $\boldsymbol{cj}$**:** We use 50% and 80% tag efficiency for the charm and the bottom, respectively, and $\sqrt{0.05} = 20\%$ efficiency for the top. The charge of the top quark is assumed to be reconstructed.
- $\boldsymbol{e\nu_e}$, $\boldsymbol{\mu\nu_\mu}$ **and** $\boldsymbol{\tau\nu_\tau}$**:** The efficiency is 100% for muons and electrons, and 50% for

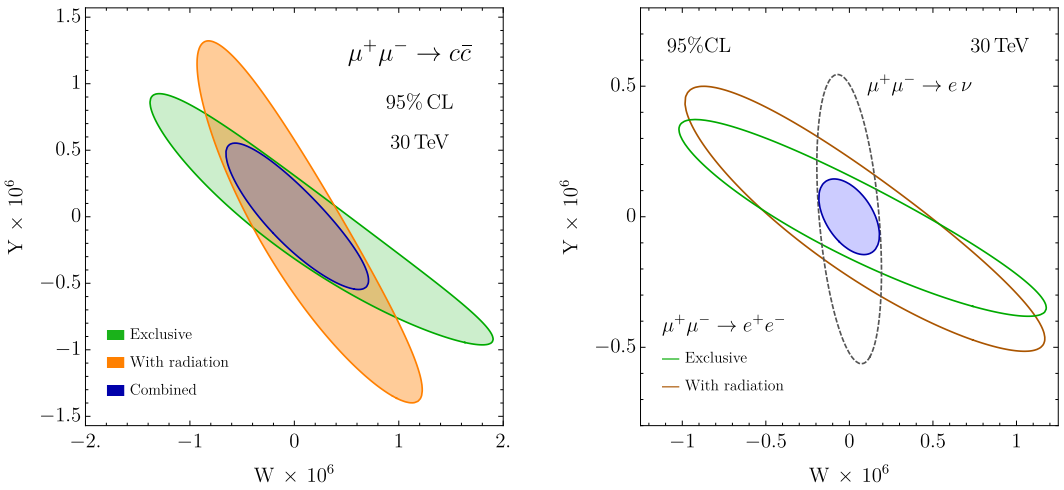


Figure 3.5: 95% CL sensitivities to the W and Y parameters of the 30 TeV muon collider. Exclusive and "with radiation" (i.e., semi-inclusive minus exclusive) cross-section measurements of the $c\bar{c}$ process are considered in the left panel. The right panel shows the impact of e^+e^- (exclusive and "with radiation") and $e\nu$ (that only exists at the semi-inclusive level) final states.

the τ. It should be noted that, because of the invisible neutrino, the hard scattering region of this final state can not be selected with a cut on the invariant mass of the two particles. The selection will have instead to be performed on the energy and the transverse momentum of the observed charged lepton.

The different dependence on W and Y of the neutral- and charged-current Born amplitudes entails (see Section 3.1.2) a different dependence on these parameters of the exclusive and semi-inclusive cross-sections. The statistical combination of the two types of cross-sections can thus increase the sensitivity, as illustrated in Figure 3.5. The left panel displays the 95% CL sensitivity of $c\bar{c}$ production to W and Y, comparing the impact of the exclusive cross-section (in green) to that (in orange) of the cross-section with radiation. The two measurements probe different regions of the W and Y parameter space, and a significant sensitivity gain is observed in their combination (in blue). The green and blue lines on the right panel of Figure 3.5 display a similar complementarity pattern for the e^+e^- final state. There also appears an even stronger complementarity with the measurement of the $e\nu$ cross-section, reported as a gray dashed line. The emergence of the $e\nu$ process, as well as the other charged final states in Table 3.3, is entirely due to EW radiation. Nevertheless its (semi-inclusive) cross-section is large, because EW radiation is indeed a prominent phenomenon at $E_{\rm cm} \simeq 10$ TeV. Furthermore the cross-section displays a peculiar dependence on new physics, producing a sensitivity contour that is different from that of the e^+e^- measurements. The statistical combination of the three measurements (in blue) improves the sensitivity significantly.

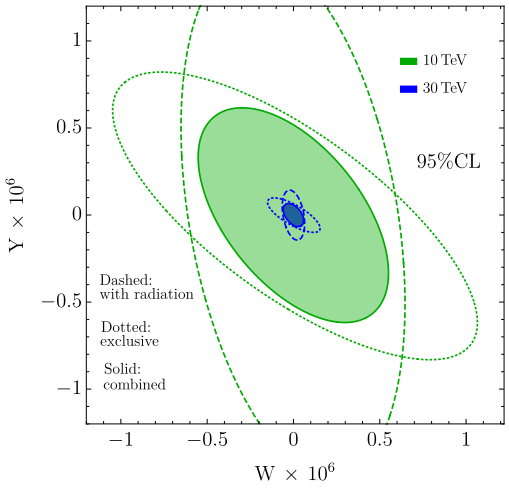


Figure 3.6: 95% CL sensitivities to the W and Y at the 10 and 30 TeV muon collider.

The final results of our analysis including all channels are summarized in Figure 3.6 and in Table 3.4. The figure displays the sensitivity contours of exclusive measurements as dotted lines, and the combined impact of charged and of neutral "with radiation" cross-sections, in dashed. The combination of all measurements is also shown. The table reports the results for 3, 10, 14 and 30 TeV, comparing the sensitivity of exclusive cross-sections alone with the total combination.

At the High-Luminosity LHC (HL-LHC), it will be possible to probe the W and Y parameters at the level of $4 \cdot 10^{-5}$ and $8 \cdot 10^{-5}$, respectively, at 95% CL [172, 175, 38]. Table 3.4 shows that the 3 TeV muon collider would improve by one order of magnitude or more, and the sensitivity improves quadratically with the muon collider energy. Among the other future collider projects [36], CLIC at 3 TeV has the best sensitivity, of around $4 \cdot 10^{-6}$ for both parameters [166]. This is of course comparable with the 3 TeV muon collider sensitivity, and a factor 10 worst than that of the muon collider at 10 TeV. The comparison with FCC-hh projections is even more favorable to the muon collider.

3.2.2 Diboson operators

The setup for this analysis is similar to that of Ref. [105]. Namely we consider the SILH operators O_W and O_B, we convert them into the current-current interactions O'_W and O'_B as in Table 3.2, and we study their effect on the production of high-energy vector bosons and Higgs. Notice that, by the equivalence theorem, O'_W and O'_B only significantly affect the production of longitudinally polarized vector bosons. We are therefore here studying the production of high-energy longitudinally vector bosons and Higgs, with the production of transversely polarized vector bosons playing merely the role of background. Since the effects are quadratically enhanced by the energy, such high-energy di-boson processes are by far the best probe of these operators at the muon collider [105].

	Exclusive-only [95% CL]			Combined [95% CL]		
	W$\times 10^7$	Y$\times 10^7$	$\rho_{\mathrm{W,Y}}$	W$\times 10^7$	Y$\times 10^7$	$\rho_{\mathrm{W,Y}}$
3 TeV	$[-53, 53]$	$[-48, 48]$	-0.72	$[-41, 41]$	$[-46, 46]$	-0.60
10 TeV	$[-5.71, 5.71]$	$[-4.47, 4.47]$	-0.74	$[-3.71, 3.71]$	$[-4.16, 4.16]$	-0.54
14 TeV	$[-3.11, 3.11]$	$[-2.31, 2.31]$	-0.74	$[-1.90, 1.90]$	$[-2.13, 2.13]$	-0.52
30 TeV	$[-0.80, 0.80]$	$[-0.52, 0.52]$	-0.75	$[-0.42, 0.42]$	$[-0.47, 0.47]$	-0.48

Table 3.4: Single-operator 95% CL reach and correlation for the W&Y parameters at different muon collider energies including only exclusive cross-sections and combining all measurements. Since the likelihood is dominated by the linear terms in the new physics parameters, the single parameter reach plus the correlation characterizes our results completely.

We thus consider, among those in Table 3.3, the following final states

- $\boldsymbol{Zh}$**:** Following Ref. [105], we consider an efficiency of 26% for tagging the two hard and central final state particles, with a selection that reduces the background to a manageable level. Notice that this final state is dominated by the longitudinal helicity channel $Z_0 h$.
- $\boldsymbol{W^+W^-}$**:** Again like in [105], we assume a 44% efficiency for the detection of the two W bosons in the semi-leptonic decay channel, where the charge of the W's can be reconstructed. Transverse WW production plays here the role of background.
- $\boldsymbol{Wh}$**:** We consider an efficiency of 19%, having in mind the leptonic W decay, and $h \to b\bar{b}$. Like for Zh, there is no relevant background from transverse production.
- $\boldsymbol{WZ}$**:** We apply an efficiency of 23%, which corresponds to the leptonic W and the hadronic Z decay. The background from transverse WZ production is considerable, and is taken into account.

In our analysis we do not consider the possibility of employing the decay angles of the bosons to extract information on their polarization. Therefore the transverse di-bosons processes $W_{\mathrm{T}}^+ W_{\mathrm{T}}^-$ and $W_{\mathrm{T}} Z_{\mathrm{T}}$ are effectively irreducible backgrounds to the corresponding longitudinal processes, and the scattering angle θ_* is the only discriminating variable. An increased lower cut on θ_* benefits the sensitivity, as it suppresses the t-channel enhancement of the transverse background processes. After optimization we find, like in Ref. [105], that a good signal sensitivity is obtained by the measurement of fiducial WW and WZ cross-sections in the range

$$\theta_* \in [67^\circ, 150^\circ]\,. \tag{3.38}$$

The possibility of binning θ_* has been considered, but found not to improve the sensitivity. Our analysis will thus be only based on the measurement of the fiducial WW and WZ

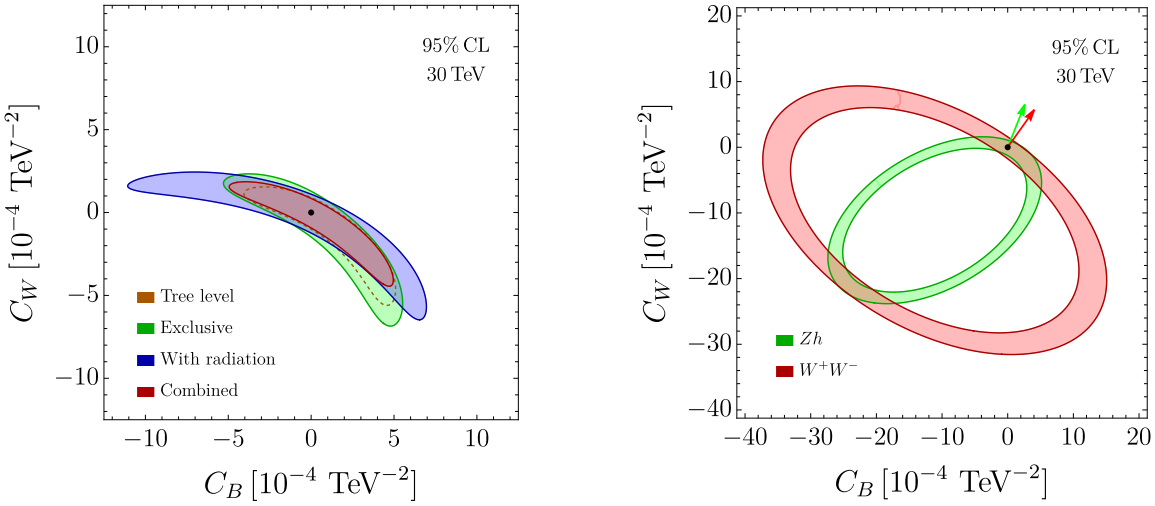


Figure 3.7: Left: 95% sensitivity contours in the (C_B, C_W) plane at the 30 TeV muon collider. A second allowed region, not shown in the figure, can be eliminated by other measurements [105]. Right: Zh and WW likelihood contours at tree-level. Notice that the ellipses for WW and Zh are tangent in two points, one being the SM, the other being the point where the amplitudes have the same magnitude as in the SM but opposite sign.

cross-sections in the above region, and of the Zh and Wh cross-sections for $\theta_* \in [30°, 150°]$. As in the previous section, both exclusive and semi-inclusive cross-sections will be employed for the neutral processes WW and Zh, plus the semi-inclusive charged cross-sections for Wh and WZ.

The results of our analysis are reported in Table 3.5 and on the left panel of Figure 3.7, in terms of the dimensionful coefficients (C_B and C_W) of the O'_B and O'_W operators of Table 3.2. Our finding are quantitatively similar to the ones of Ref. [105]. We can thus refer to that article for the (very favorable) assessment of the muon collider sensitivity to C_B and C_W in comparison with current knowledge and with other future colliders. We devote the rest of this section to discuss the approximate flat direction of the likelihood in the (C_B, C_W) plane, which we observe in Figure 3.7 (left panel).

The flat direction entails a strong degradation of the marginalized sensitivity, as in Table 3.5. Furthermore this degradation brings the marginalized C_B and C_W limits to large values, in a region where the likelihood is considerably affected by the contributions to the cross-sections of the terms that are quadratic in the new physics parameters. In theories like Composite Higgs where $C_{B,W} \sim 1/m_*^2$, this fact implies that the marginalized limits correspond to a new physics scale m_* not much above the collider energy. In fact, looking at Table 3.5 we notice that the 30 TeV C_B reach corresponds to $m_* = 43$ TeV. Thus, if new physics happened to sit along the flat direction in Figure 3.7, diboson processes would fail to extend the muon collider sensitivity well above the direct mass-reach. We do not have reasons to expect new physics to lie in that direction. Actually in

	Single Operator (Exclusive-only) [95% CL]			
	$C_B[10^{-4}\ \text{TeV}^{-2}]$		$C_W[10^{-4}\ \text{TeV}^{-2}]$	
	Linear	Quadratic	Linear	Quadratic
3 TeV	$[-170, 170]$	$[-189, 157]$	$[-77.4, 77.4]$	$[-81, 74.4]$
10 TeV	$[-15.3, 15.3]$	$[-17, 14.2]$	$[-8.18, 8.18]$	$[-8.62, 7.82]$
14 TeV	$[-7.86, 7.86]$	$[-8.69, 7.25]$	$[-4.40, 4.40]$	$[-4.65, 4.20]$
30 TeV	$[-1.73, 1.73]$	$[-1.92, 1.6]$	$[-1.1, 1.1]$	$[-1.16, 1.04]$

	Single Operator (Combined) [95% CL]			
	$C_B[10^{-4}\ \text{TeV}^{-2}]$		$C_W[10^{-4}\ \text{TeV}^{-2}]$	
	Linear	Quadratic	Linear	Quadratic
3 TeV	$[-153, 153]$	$[-169, 142]$	$[-65.8, 65.8]$	$[-68.2, 63.6]$
10 TeV	$[-12.8, 12.8]$	$[-13.9, 11.9]$	$[-6.14, 6.14]$	$[-6.37, 5.93]$
14 TeV	$[-6.40, 6.40]$	$[-6.95, 5.99]$	$[-3.17, 3.17]$	$[-3.29, 3.06]$
30 TeV	$[-1.34, 1.34]$	$[-1.44, 1.25]$	$[-0.71, 0.71]$	$[-0.737, 0.686]$

	Marginalized (Exclusive-only) [95% CL]			
	$C_B[10^{-4}\ \text{TeV}^{-2}]$		$C_W[10^{-4}\ \text{TeV}^{-2}]$	
	Linear	Quadratic	Linear	Quadratic
3 TeV	$[-478, 478]$	$[-352, 596]$	$[-217, 217]$	$[-583, 125]$
10 TeV	$[-53.2, 53.2]$	$[-35.2, 50]$	$[-28.4, 28.4]$	$[-53.5, 14.2]$
14 TeV	$[-29.4, 29.4]$	$[-18.6, 25]$	$[-16.5, 16.5]$	$[-27.5, 7.82]$
30 TeV	$[-7.98, 7.98]$	$[-4.45, 5.19]$	$[-5.04, 5.04]$	$[-6.16, 2.05]$

	Marginalized (Combined) [95% CL]			
	$C_B[10^{-4}\ \text{TeV}^{-2}]$		$C_W[10^{-4}\ \text{TeV}^{-2}]$	
	Linear	Quadratic	Linear	Quadratic
3 TeV	$[-442, 442]$	$[-341, 535]$	$[-189, 189]$	$[-426, 115]$
10 TeV	$[-44, 44]$	$[-33.4, 43.4]$	$[-21.1, 21.1]$	$[-35.1, 12.3]$
14 TeV	$[-23.1, 23.1]$	$[-17.6, 21.6]$	$[-11.4, 11.4]$	$[-17.6, 6.6]$
30 TeV	$[-5.24, 5.24]$	$[-4.12, 4.43]$	$[-2.79, 2.79]$	$[-3.70, 1.62]$

Table 3.5: Single operator and marginalized 95% reach on C_B and C_W, at different muon collider energies. The sensitivity of exclusive cross-section measurements alone is shown separately from the combination of all the measurements. The significant degradation of the marginalized bounds relative to the single-operators ones, and the strong sensitivity to the quadratic terms at the marginalized level, is due to the approximately flat direction displayed in Figure 3.7

Process	σ_B/σ_0
$(\mu\mu)_L \to Zh$	$\left[t_{\rm w}^2(1+E_{\rm cm}^2 C_B) - (1+E_{\rm cm}^2 C_W)\right]^2$
$(\mu\mu)_L \to WW$	$\left[t_{\rm w}^2(1+E_{\rm cm}^2 C_B) + (1+E_{\rm cm}^2 C_W)\right]^2$
$(\mu\mu)_R \to WW/Zh$	$\left[2t_{\rm w}^2(1+E_{\rm cm}^2 C_B)\right]^2$
$(\mu\nu)_L \to WZ/Wh$	$\left[\sqrt{2}(1+E_{\rm cm}^2 C_W)\right]^2$

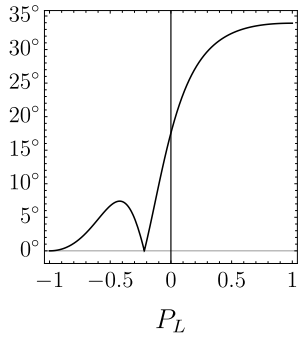


Figure 3.8: Left: Born-level cross-sections, with $t_{\rm w}$ the tangent of the Weinberg angle, normalized to a common σ_0 (whose expression is irrelevant). Right: the angle between the ZH and WW cross-section gradients as a function of the beam polarization fraction.

certain Composite Higgs models one expects it to lie in the nearly orthogonal direction $C_B = C_W$ [105]. However the presence of the flat direction is an obstruction to the broad exploration of new physics and to the characterization of a putative discovery. It is thus worth explaining its origin and discussing strategies to eliminate it.

The origin of the flat direction in the tree-level sensitivity contour (showed dashed, on the left panel of Figure 3.7) is readily understood analytically, by considering the gradients "∇" of the Born-level cross-sections in the (C_B, C_W) plane, around the SM point $(0,0)$. Using the results for WW and Zh shown in Figure 3.8 and rescaling the gradients by the common factor $2\,E_{\rm cm}^2\sigma_0$ one finds

$$\nabla_L^{Zh} = (1-t_{\rm w}^2)\{-t_{\rm w}^2, +1\}\,, \qquad \nabla_L^{WW} = (1+t_{\rm w}^2)\{+t_{\rm w}^2, +1\}\,, \qquad \nabla_R^{Zh} = \nabla_R^{WW} = 4\,t_{\rm w}^4\{1,0\}\,, \tag{3.39}$$

where sup- and sub-scripts refer respectively to the final states and to the chirality of the incoming fermions. Notice that the Zh and WW gradients for right-handed initial states are perfectly aligned, so that this contribution to the cross sections has a flat direction (orthogonal to the gradient). The degeneracy can only be lifted by the left-handed contribution to the cross sections. However, given the small value of $t_{\rm w}^2 \simeq 0.3$, the gradients ∇_L^{Zh} and ∇_L^{WW} also form a relatively small angle, $\sim 30°$. They are thus not very effective in lifting the flat direction when considering the total $(L+R)$ contribution to the WW and Zh cross-section. Indeed, the angle between $\nabla_L^{Zh} + \nabla_R^{Zh}$ and $\nabla_L^{WW} + \nabla_R^{WW}$ is in the end only $\sim 17°$ and thus the flat directions of the two cross-section measurements essentially coincide, as the right panel of Figure 3.7 shows. The combined likelihood is consequently also flat, in the same direction.

As evident in eq. (3.39), the L-gradients form a large angle with the R-gradient. Therefore, if one could use polarized beams, the degeneracy would be eliminated by measuring the contribution of each chirality. Considering a polarization fraction $-P_L$ for the muon, and $+P_L$ for the anti-muon beam, the cross-section gradients read (we indicate by ∇_R the

identical ∇_R^{Zh}, ∇_R^{WW})

$$\nabla_{P_L}^{Zh} = \frac{(1+P_L)^2}{4}\nabla_L^{Zh} + \frac{(1-P_L)^2}{4}\nabla_R\,, \qquad \nabla_{P_L}^{WW} = \frac{(1+P_L)^2}{4}\nabla_L^{WW} + \frac{(1-P_L)^2}{4}\nabla_R\,. \tag{3.40}$$

The angle between the two gradients steeply increases for positive P_L, as indicated by the plot in the right panel of Figure 3.8. Correspondingly, even a modest amount of polarization has a considerable impact on the sensitivity. The left panel of Figure 3.9 displays our sensitivity projections in a scheme where the VHEL integrated luminosity is equally split between positive and negative $P_L = \pm 30\%$. The likelihood contour (in green) corresponding to $P_L = +30\%$ is significantly smaller than that (in blue) for $P_L = -30\%$, owing to the lifting of the flat direction achieved for positive P_L. On the other hand, the measurements at $P_L = -30\%$ probe a direction complementary to that probed at $P_L = +30\%$. The combination of the two measurements thus benefits the sensitivity. The impact of beam polarization was emphasized already in Ref. [105]. Here we confirm that result, using more accurate predictions and including the entire set of exclusive and semi-inclusive cross-section measurements previously described.

Up to this point, we discussed the flat direction in the un-polarized likelihood (left panel of Figure 3.7) by employing the tree-level cross-sections. When considering also EW radiation, the predictions are significantly affected, but the flat direction is not fully eliminated. For the exclusive Zh and WW cross-sections this is easily understood, since virtual radiation suppresses the L-processes more than the R ones, owing to the larger Sudakov for incoming left-handed muons. The exclusive Zh and WW cross-sections gradients are thus even more aligned than the gradients of the corresponding tree-level predictions. The semi-inclusive cross-sections for Zh and WW production are also quite aligned, among them and with the exclusive cross-sections. This was expected because the partial cancellation between real and virtual logarithms make semi-inclusive cross-sections not vastly different from the tree-level ones.

On the contrary, the measurement of the charged processes WZ and Wh could have been expected to eliminate or mitigate the flat direction, because they are strongly sensitive to the Born cross-section of the charged scattering $(\mu\nu)_L \to WZ/Wh$ (see the left panel of Figure 3.8). The associated gradient

$$\nabla_L^{\rm ch} = 2\{0, +1\}\,, \tag{3.41}$$

points in a different direction than ∇_L^{Zh}, ∇_L^{WW} and ∇_R. Therefore the gradient of $\sigma^{Wh/WZ}$ could in principle point in a direction that is completely different from that of the (nearly parallel) gradients of the Zh and WW cross-sections. However, by expanding at the first order in $\mathcal{L} = g^2/16\pi^2 \log^2(E_{\rm cm}^2/m_{\rm W}^2)$, the unpolarized (longitudinal) WZ and

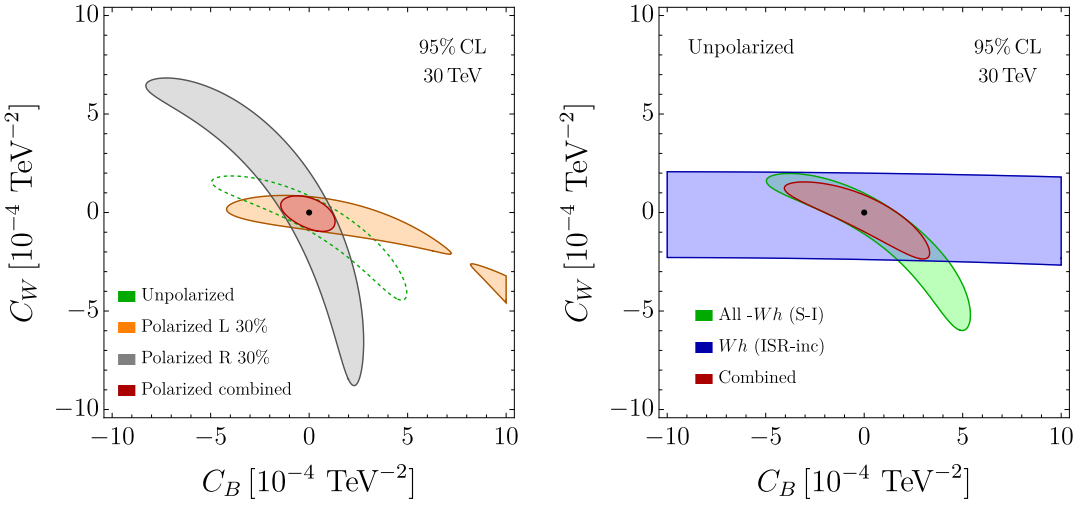


Figure 3.9: Left Panel: 95% CL contours in for $P_L = \pm 30\%$ beam polarization. Right Panel: the impact of the ISR-inclusive Wh cross-section measurement.

Wh cross-sections are approximately equal and read

$$\sigma^{WZ} \simeq \sigma^{Wh} \simeq \frac{1}{4}\mathcal{L} \cdot (\sigma_B^{Zh} + \sigma_B^{WW} + \sigma_B^{\mathrm{ch}})\,, \tag{3.42}$$

where σ_B^{ch} is the charged Born cross-section reported on the left panel of Figure 3.8 (times 1/4, from the polarization average) and $\sigma_B^{Zh,WW}$ are the Born cross-sections of the neutral processes. Therefore, the charged cross-section gradient ∇_L^{ch} must compete with the (nearly parallel) gradients of Zh and WW, and its size happens to be insufficient to produce a large misalignment angle between the $\sigma^{Wh/WZ}$ and $\sigma^{Zh/WW}$ gradients.

The situation would be improved, if we could tailor an observable in which the σ_B^{Zh} and σ_B^{WW} contributions in eq. (3.42) are eliminated or reduced. Notice, for that purpose, that the Zh and WW terms in eq. (3.42) can be interpreted as due to one hard $\mu\mu$ neutral-current scattering, followed by the radiation of one charged W boson from the final legs of the hard process. The W is thus preferentially collinear to the final states. The σ_B^{ch} term comes instead from the radiation of a W from the initial state, collinear to the beam axis, followed by a hard $\mu\nu$ scattering.[16] This suggests to consider alternative Wh and WZ cross-sections that exclude final state radiation (FSR) while being inclusive on initial state radiation (ISR). FSR consists of soft radiation collinear to the hard particles in the final state, which is precisely the source of the first two terms in eq. (3.42). Excluding FSR, the resulting "ISR-inclusive" cross-sections are expected be roughly

$$\sigma^{WZ/Wh}_{\mathrm{ISR-inc}} \simeq \frac{1}{4}\mathcal{L}\,\sigma_B^{\mathrm{ch}}\,. \tag{3.43}$$

[16]This interpretation would straightforwardly correspond to Feynman diagrams in a physical gauge, where DL's are associated to emissions from individual legs. We already remarked that in covariant gauges instead they arise from the interference between emission from strictly different legs.

This observable should thus be mostly sensitive to C_W, and its measurement should produce a nearly horizontal band in the (C_B, C_W) plane, thus eliminating the flat direction.

Unfortunately we are unable to produce resummed predictions for the ISR-inclusive cross-sections with the IREE methodology. We can however illustrate the impact of such measurements using tree-level `MadGraph` [89] predictions with the SMEFT@NLO model [176], focusing in particular on the Wh channel. Specifically, we simulate the process

$$\mu^+\mu^- \to W^+W^-h\,, \tag{3.44}$$

at $E_{\rm cm} = 30$ TeV, with the following selection cuts. First, we identify as "hard" the W boson that forms, together with the Higgs, the pair with the highest invariant mass. Secondly we ask this mass to be above $0.85 \cdot E_{\rm cm} = 25.5$ TeV and the hard W and h to be within the central region $\theta_* \in [30^\circ, 150^\circ]$. These selections enforce the occurrence of a hard scattering, and correspond to our definition of a "semi-inclusive" process. We further restrict to the "ISR-inclusive" region by asking the other ("soft") W to be parallel to the beam, in a cone of 20°. Since the emission of at least one soft W is required for Wh production, the latter cut effectively corresponds to a veto on central EW radiation.[17]

The above estimate of the ISR-inclusive cross-section produces the blue band on the right panel of Figure 3.9. As expected, the band is nearly horizontal. In the figure we also display, in green, the 95% CL contour of the likelihood including all the measurements discussed in the present section, apart from the measurement of the semi-inclusive Wh cross-section which is correlated with the ISR-inclusive measurement. The combination of the two contours, shown in red, strongly mitigates the flat direction issue. Notice however that our tree-level estimate of the ISR-inclusive cross-section could be subject to large errors, and resummed predictions should be employed for a conclusive assessment of the sensitivity gain.

3.2.3 BSM sensitivity

Composite Higgs. As a first concrete scenario of new physics we consider Composite Higgs [177, 6, 4]. In this scenario, the Higgs is a composite Pseudo-Nambu-Goldstone boson emerging from some strong dynamics at a scale m_*. In principle the underlying dynamics could arise from gauge interactions, like in QCD. However the only concrete realistic constructions, accounting for the origin of both the fermion masses and the scale m_* itself, have been obtained in the context of warped compactifications. In these constructions, compositeness occurs in a holographic sense. Within the Composite Higgs scenario, the size of the Wilson coefficients in the resulting low energy EFT, can be estimated, under simple but robust dynamical assumptions, in terms of the mass scale

[17]The attempt made in Ref. [105] to exploit the WWh process did not impose the crucial angular cut that defines the ISR-inclusive region.

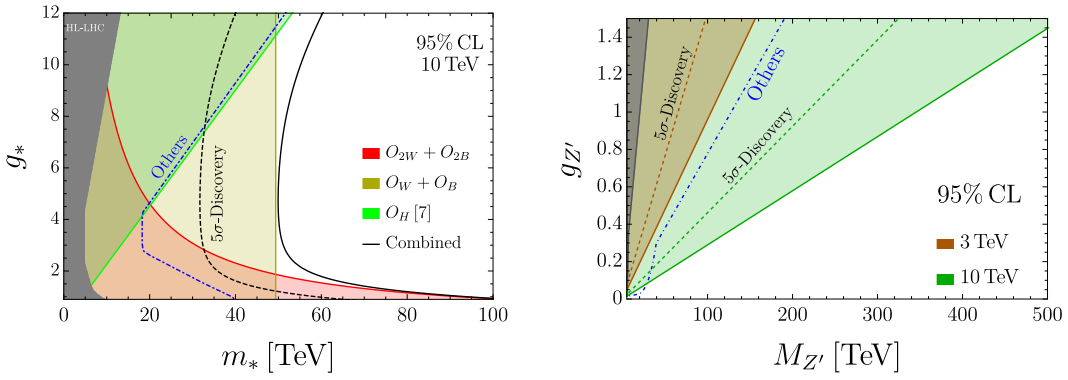


Figure 3.10: Left Panel: 95% exclusion reach on the Composite Higgs coupling-mass parameter space. The reach for O_H is taken from [105]. Right Panel: sensitivity projections for a Y-universal Z' model. The gray band and the blue dash-dot line represent respectively the region probed by the HL-LHC program and the sensitivity projections for all other future collider projects [37].

m_* and overall coupling strength g_* of the underlying strong dynamics [11]. Furthermore, simple considerations suggest $g_* \lesssim 4\pi$, while the existence of $\mathcal{O}(1)$ couplings within the SM implies $g_* \gtrsim 1$. The Composite Higgs power-counting rules predict the Wilson coefficients of the operators in the left column of Table 3.2 to scale as

$$C_{2W} = -\frac{c_{2W}}{2}\frac{g^2}{g_*^2 m_*^2}, \quad C_{2B} = -\frac{c_{2B}}{2}\frac{g'^2}{g_*^2 m_*^2}, \quad C_W = c_W \frac{1}{m_*^2}, \quad C_B = c_B \frac{1}{m_*^2}, \tag{3.45}$$

where the dimensionless coefficients c_{2W}, c_{2B}, c_W, c_B are expected to be of order 1. Even though it does not affect the processes studied in this chapter, an important role is also played by

$$O_H = \partial_\mu (H^\dagger H)\partial^\mu (H^\dagger H), \qquad C_H = \frac{c_H}{2}\frac{g_*^2}{m_*^2}\,. \tag{3.46}$$

In our sensitivity projections we will report the corresponding bounds, as obtained in [105] by studying the process $\mu^+\mu^- \to hh\nu\nu$ at tree level. Other probes of C_H at the muon collider, from Higgs coupling measurements, are superior or competitive at the lower energy muon colliders [105], but they are not considered in the sensitivity plots.

Using the above scalings, and setting all the c coefficients to 1, we can translate the bounds of Section 3.2 for a 10 TeV muon collider into sensitivity estimates in the plane $(m_*,\ g_*)$, as in Figure 3.10. In the same plot we also report the HL-LHC sensitivity projections, the envelope of the 95% CL sensitivity contours of all the future collider projects that have been considered for the 2020 update of the European Strategy for Particle Physics [37]. The advantage of the muon collider is evident. Results at muon colliders with different energies, with an integrated luminosity scaling as in eq. (3.35), are reported in Appendix 3.D.

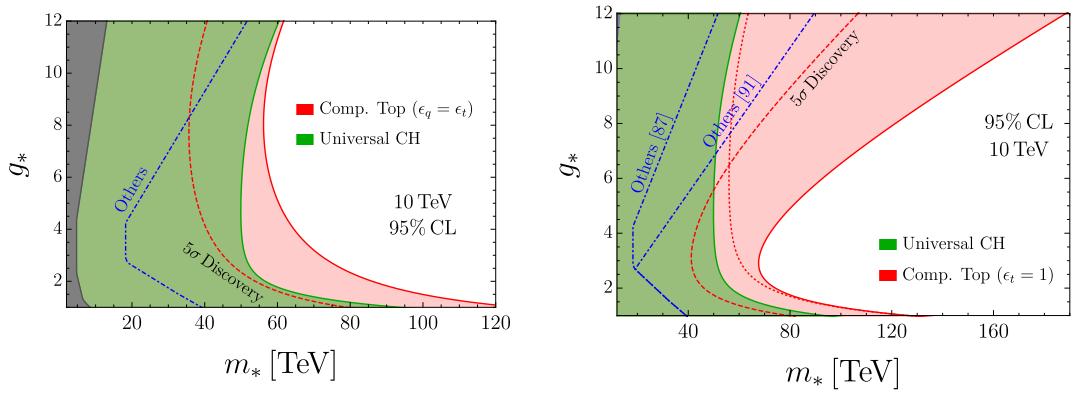


Figure 3.11: 95% exclusion reach for the two partial compositeness scenarios under consideration. The green shapes represent the combined bound from the flavor universal measurements, while the red contours also includes the di-top and di-bottom constraints. The projected sensitivity of other future collider projects and the gray band of HL-LHC are taken from Ref. [37]. The right panel ($\epsilon_t = 1$) also includes the stronger CLIC sensitivity estimated in Ref. [178].

Composite top. The results for purely bosonic operators we just discussed apply robustly to basically all composite Higgs scenarios. Operators involving fermions are more sensitive to the assumptions on the flavor dynamics, but one convenient option is offered by the mechanism of partial compositeness [10], under which the elementary fermions mix linearly with heavy partners in the strong sector. Due to its large Yukawa coupling, the top quark is expected to have a large mixing with its partners and therefore precise measurements involving the third family represent an appealing opportunity to probe new physics.

At the muon collider the most relevant effects are expected in $t\bar{t}$ and $b\bar{b}$ production.[18] Indeed the dimension-6 operators in the last block of Table 3.2 gives rise to contributions that grow with $E_{\rm cm}$ and which can be exploited at the large energy of the muon collider. In a model-independent approach one can parametrize the "amount of compositeness" of respectively the 3$^{\rm rd}$ quark family left-handed doublet and right-handed up-type singlet by ϵ_q and ϵ_t. These quantities range from 0 to 1. Given the universal coupling strength g_* of the strong sector the resulting top Yukawa coupling scale as [11]

$$y_t \sim \epsilon_q \epsilon_t g_* \,. \tag{3.47}$$

The relevant Wilson coefficients are then expected to scale as (see [174] for a short review)

$$C_{qD}^{(3)} = c_{qD}^{(3)} \frac{g\,\epsilon_q^2}{m_*^2}\,, \qquad C_{qD}^{(1)} = c_{qD}^{(1)} \frac{g'\epsilon_q^2}{m_*^2}\,, \qquad C_{tD} = c_{tD} \frac{g'\epsilon_t^2}{m_*^2}\,, \tag{3.48}$$

where the c_i are, as usual, expected to be order 1 coefficient. For concreteness we focus

[18] See [179] for a similar analysis for CLIC.

on two benchmark scenarios, where we fix ϵ_t and ϵ_q and leave g_* and m_* free. In the first scenario, the right-handed top quark is assumed to be fully composite, corresponding to $\epsilon_t = 1$ and $\epsilon_q = y_t/g_*$. In the second, the two top chiralities are assumed equally composite, that is $\epsilon_q = \epsilon_t = \sqrt{y_t/g_*}$.

Notice that the contribution of the operator

$$O_{tt} \equiv \frac{1}{2}(\bar{t}_R\gamma^\mu t_R)(\bar{t}_R\gamma_\mu t_R)\,, \tag{3.49}$$

to the Wilson coefficients of the O_{tD}, through Renormalization Group (RG) evolution, is not negligible in the scenario of total right-handed top quark compositeness [178]. Using the power-counting estimate

$$C_{tt} = \epsilon_t^4 \frac{g_*^2}{m_*^2} c_{tt}\,, \tag{3.50}$$

we obtain a correction [178] to the C_{tD} coefficient at a scale $\mu = E_{\rm cm}$

$$C_{tD}(\mu) = C_{tD}(m_*) + C_{tt}(m_*)\frac{32}{9}\frac{g'}{16\pi^2}\log\left(\frac{m_*^2}{\mu^2}\right) = \epsilon_t^2\frac{g'}{m_*^2}\left(c_{tD} + c_{tt}\frac{32}{9}\frac{\epsilon_t^2 g_*^2}{16\pi^2}\log\left(\frac{m_*^2}{\mu^2}\right)\right)\,. \tag{3.51}$$

This correction is sizable if $\epsilon_t \sim 1$, especially for large g_*, because the sensitivity of the muon collider extends to a scale m_* that is significantly larger than $E_{\rm cm}$.

There are in principle three more operators O_{Ht}, $O_{Hq}^{(1)}$ and $O_{Hq}^{(3)}$ (defined as in the "Warsaw" basis [51]) that mix significantly with those in eq. (3.48) through RG evolution. However, the their effects can only be important in the case where $\epsilon_q \sim 1$, which we do not contemplate in our analysis. We will therefore neglect the RG effects of the latter three operators and consider only those of O_{tt}.

Our results are summarized in Figure 3.11, where we report the projected exclusion reach in the g_* and m_* plane in the two scenarios under consideration for $E_{\rm cm} = 10$ TeV. Additional results can be found in Appendix 3.D. Starting from the scenario of equal compositeness (left panel) we notice that at g_* the additional hypothesis of top compositeness extends the muon collider potential to probe the scale of Higgs compositeness m_*. The effect is even stronger for fully composite t_R (right panel), which shows that di-top measurements can cover up to $m_* \sim 150$ TeV for $g_* \gtrsim 8$. We should point out, however, that this result depends on the exact $\mathcal{O}(1)$ value of the c_{tt}, c_{tD} coefficients in eq. (3.51). This dependence is illustrated in Figure 3.16, where we set $c_{tD} = 1$ and we vary the value of c_{tt}.

Finally we remark that a detailed analysis of the composite Higgs scenario with partial compositeness would require specific hypotheses on the flavor dynamics and a detailed inspection of the flavor observables. Depending on those hypotheses, principally the maximality or minimality of the underling flavor symmetry, the resulting flavor constraints

on the new physics scale m_* can vary dramatically. While a comprehensive analysis clearly exceeds the purposes of this work, a perspective can be gained by considering available studies. As shown in Ref. [46], under the strongest assumptions, that is for a symmetry structure offering the best protection from unwanted effects, flavour and CP observables could start exploring the range $m_* = O(10)$ TeV in the next decade or so, given the availability of better measurements and assuming better theoretical predictions. This is roughly the same range explored by a 3 TeV muon collider. Moreover the $m_* \sim 50$ TeV reach of a 10 TeV muon collider vastly surpasses any conceivable improvement of flavour constraints, and competes with the more stringent flavour bounds obtained by making more generic assumptions on the flavor dynamics. Notice also that the present lepton flavor universality anomalies in B-decays, at least the seemingly more prominent ones in neutral currents, suggest a new physics scale in the ~ 30 TeV range, which could be explored both directly and indirectly by the muon collider.

Y-Universal Z' model The Y-universal Z' model represents a simple extension of the SM, featuring an additional heavy gauge boson, of mass $M_{Z'}$, on top of the SM particles.[19] In this benchmark scenario the new vector charges are aligned with the SM hypercharge with coupling $g_{Z'}$. Requiring the width of the Z' to not exceed $0.3 M_{Z'}$ sets the perturbative limit on the coupling to be $g_{Z'} \lesssim 1.5$. At energies below $M_{Z'}$, integrating out the Z' only generates the O'_{2B} operator of Table 3.2. The Wilson coefficient of the operators corresponds, by eq. (3.36), to

$$\mathrm{Y} = \left(\frac{g_{Z'}\, m_W}{g'\, M_{Z'}} \right)^2 . \tag{3.52}$$

The sensitivity projections are reported in the right panel of Figure 3.10. The orange and green regions are the ones probed by muon colliders at 3 and 10 TeV energy, respectively. The gray band represents the expected exclusion reach from HL-LHC, while the blue line indicates the combined sensitivity from other future collider projects (dominantly FCC, and the 3 TeV stage of CLIC). The 3 TeV muon collider sensitivity is obviously similar to the one of CLIC. A 10 TeV machine would greatly improve this result probing up to 500 TeV for large (but still perturbative) coupling. The dashed lines represent the discovery reach, showing that already at 3 TeV there are vast opportunities for indirect discovery, well above the region that the HL-LHC might exclude. Results at higher muon collider energies are reported in Figure 3.14.

[19] See [37] for details. Concrete BSM scenarios featuring additional Z's can be found, for instance, in Ref. [180].

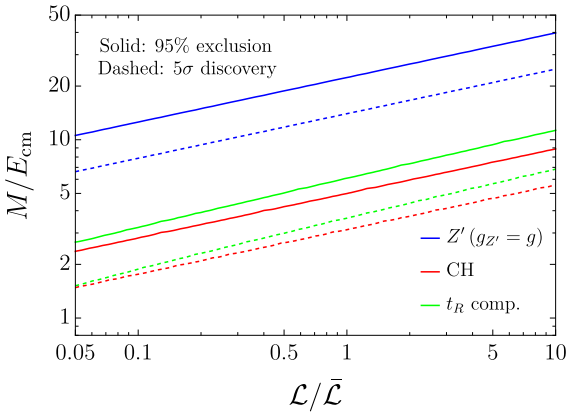


Figure 3.12: Reach on the new physics scale at 2σ (continuous) and at 5σ (dashed), relative to the collider energy, as a function of the integrated luminosity normalized to eq. (3.35). The red lines are for Universal manifestations of Higgs compositeness, while the green ones include the effect of Top compositeness in the $\epsilon_t = 1$, $\epsilon_q = y_t/g_*$ scenario. The blue lines are for the Y-universal Z' for a fixed coupling $g_{Z'} = g$.

3.3 Conclusions and outlook

We have studied the interplay between two classes of novel phenomena, which can be observed at future lepton colliders with very large center of mass energy. The first class consists of hard scattering processes induced by new physics at around 100 TeV. The second class consists of the long-distance effects of EW radiation. Both phenomena play a relevant role at lepton colliders with $E_{\rm cm} \sim 10$ TeV. In particular they are relevant at muon colliders, which are the main target of the present work.

The interplay manifests itself in two ways. The first one is simply that EW radiation effects on the SM predictions are large (see Sections 3.1.2 and 3.1.3) and require to be included and resummed with high accuracy in order to isolate the putative BSM contribution to the measurements. EW radiation thus plays for muon colliders a similar role as QCD radiation for the LHC, with the difference (discussed in the Introduction) that its effects can not be mitigated by the choice of suitable (inclusive) observables. Therefore they are actually even more important for muon colliders than QCD is for the LHC.

The second and possibly more interesting aspect of the interplay is given by the influence on the pattern of EW radiation operated by the presence of new physics in the hard scattering amplitude. This makes observables that require or that exclude the presence of radiation display a different dependence on the new physics parameters, and the sensitivity profits from their combined measurements as illustrated in Sections 3.2.1 and 3.2.2, and in Appendix 3.C.

Our sensitivity projections rely on putative measurements of exclusive and semi-inclusive

cross-sections. Both classes of processes are characterized by the occurrence of a hard scattering, with two particles in the final state carrying almost all the available energy. The emission of additional EW bosons and hard photons is vetoed in the exclusive case and allowed in the semi-inclusive one. We computed the resummed semi-inclusive cross-sections in double logarithm (DL) approximation by extending the IREE methodology [143], as described in Section 3.1.1. The exclusive cross-sections were computed at DL, but also including single logarithms at 1-loop, which we found to be sizable.

The studies performed in this chapter should be improved and extended in many directions. Better predictions will be definitely needed in order to approach the percent-level accuracy target that is needed to fully exploit the statistical precision potential of a muon collider. Moreover, given the magnitude of the radiation effects we observed, it is possible that more accurate predictions will considerably affect some of our sensitivity projections. A first step in that direction, which we leave for future work, is the inclusion of single logarithms at fixed leading order in the semi-inclusive predictions. That could be achieved by combining one loop virtual logarithms with a factorized treatment of real emission. That calculation, would possibly also help clarify the connection between soft-collinear effects (studied in this chapter) and the PDF/Fragmentation Function treatment of EW radiation. It is not unconceivable that the same approach could be extended at the two loop order. In parallel, the impact of resummation beyond DL should be assessed. The SCET methodology currently offers the most promising approaches.

Another priority is to investigate further classes of cross-sections, sensitive to different EW radiation patterns. Our results indicate that investigation should be done on the basis of the structure of short distance new physics. At the end of Section 3.2.2 we took one step in that direction, showing that the approximately flat direction in the (C_B, C_W) likelihood contours is mitigated, in the absence of polarized beams, by considering an "ISR-inclusive" cross-section. This third cross-section type is inclusive on radiation collinear to the beam, but vetoes centrally emitted radiation. We could not compute the ISR-inclusive cross-section at DL with our IREE methodology and limited ourself to a tree-level estimate. A straightforward direction for progress would be to perform that calculation and verify if and how it impacts our findings.

The definition and study of cross-sections should be also based on experimental considerations. We already pointed out that exclusive cross-sections are problematic in that regard. Indeed imposing the radiation veto requires experimental sensitivity to EW radiation that is emitted in all directions, including the forward and backward regions along the beam line. The angular coverage of the muon collider detector is still to be quantified, however we expect that it will be insufficient for the measurement of exclusive cross-sections.

In view of the above, it is important to emphasize that our sensitivity projections have been verified to not change radically when exclusive cross-section measurements are not

available. This conclusion is not in contradiction with (and cannot be inferred from) our sensitivity plots, where (see e.g. Figure 3.6) we observe a strong complementarity between "exclusive" observables and observables "with radiation". Indeed for neutral processes the latter observables consists of the difference between the semi-inclusive and the exclusive cross-sections. Therefore the impact of eliminating the exclusive measurements can not be estimated by suppressing the "exclusive" cross section measurement in the computation of the likelihood. The proper estimate is obtained by employing the semi-inclusive neutral processes without subtraction, combined with the charged measurements, produces a combined reach that is not much inferior to the one that exploits the exclusive measurements. In essence, the main sensitivity gain due to radiation stems from the emergence of the charged processes and from their complementarity with the neutral ones. The complementarity between neutral measurements with different degrees of radiation inclusiveness (e.g., exclusive versus semi-inclusive) plays instead a less relevant role in our results. This same qualitative behavior can be observed in the comparison between the neutral and charged lepton production processes on the right panel of Figure 3.5.

On the other hand, the complementarity between exclusive and semi-inclusive measurements exists and can benefit the sensitivity as we illustrated on the left panel of Figure 3.5. It plays a marginal role in the combined fit to the limited number of EFT operators we considered in this chapter. It could however be relevant in a more global exploration and characterization of putative new physics. One way to recover sensitivity, if exclusive measurements were indeed unavailable, could be to exploit the ISR-inclusive cross-sections. These are easier to measure because they do not require sensitivity to radiation in the forward and backward regions. This aspect should be investigated.

Our phenomenological results strengthen and extend previous estimates of the muon collider sensitivity to heavy new physics. We have a considered a variety of BSM scenarios with new physics coupled to the SM with electroweak strength. We found that a $E_{\rm cm} = 10$ TeV muon collider can probe up to a scale ranging from 50 to 200 TeV. The reach improves linearly with $E_{\rm cm}$. These figures are significantly above the potential (direct and indirect) sensitivity of other future collider projects, and above the direct sensitivity reach of the muon collider itself, which is obviously bounded by the collider energy.

The indirect sensitivity to scales that are well above the direct reach is a great addition to the physics case of a muon collider, whose relevance would not be diminished, but on the contrary augmented, by the occurrence of a direct discovery. Indeed, direct hints for new particles observed at the muon collider will turn into a full-fledged discovery of new physics only after unveiling the underlying theoretical description of their dynamics. The possibility of probing this dynamics well above the particle's mass will play a decisive role in this context. Furthermore, the direct manifestation of new physics might be hard to detect. Perhaps, indirect probes will provide the first hint of its existence, to be eventually confirmed by targeted direct searches. Finally, indirect searches for BSM

phenomena based on precise measurements guarantee a sound output of the project. The connection with the phenomenon of EW radiation, which is interesting per se, adds scientific value to the program.

Before concluding, we discuss the impact of the integrated luminosity on our results. We employed the baseline luminosity in eq. (3.35), which corresponds to 90 ab^{-1} for a 30 TeV muon collider. Since the possibility of reducing the 90 ab^{-1} target is under discussion, it is worth assessing the impact of a lower integrated luminosity on our conclusions. An important aspect is associated with the actual experimental feasibility of the relevant measurements. While a conclusive assessment will require dedicated studies, the expected number of events in Table 3.3 [20] and the corresponding efficiencies show that, with a factor 10 reduction in luminosity, some of the processes we employed would be left with a handful of observed events, possibly preventing the corresponding measurements. If the reduction in luminosity were less extreme, the sensitivity to the scale of new physics would simply deteriorate as the fourth root of the luminosity, as shown in Figure 3.12 for some of the BSM scenarios we studied in Section 3.2.3. The figure displays the exclusion and discovery reach on the new physics scale normalized to the collider energy $E_{\rm cm} = 30$ TeV. This is the right figure of merit, since the goal is to extend the muon collider sensitivity above the direct reach. The reduction by a factor of $10^{1/4} = 1.8$ due to a factor 10 luminosity reduction partially undermines this goal, especially for what concerns the generic manifestations of the Composite Higgs scenario.

3.A Radiation integrals

The contribution of virtual radiation to the amplitude variation in eq. (3.10) is proportional to the integral

$$I \equiv \frac{-i}{(2\pi)^4} \int_{\delta\sigma} d^4q \, \frac{1}{q^2 - m_V^2 + i\epsilon} \frac{(k_i \cdot k_j)}{(q \cdot k_i)(q \cdot k_j)} \,, \tag{3.53}$$

where we included a mass $m_V \sim m_{\rm W}$ for the virtual vector in order to verify explicitly that the integral is log-enhanced only in the $\lambda \gg m_{\rm W}^2$ regime, where the IR cutoff is much above the EW scale. We now proceed (following [181]) to the evaluation of I assuming, for simplicity, exactly massless hard 4-momenta $k_i^2 = k_j^2 = 0$. The integral is Lorentz-invariant, therefore it can only depend on the scalar product $(k_i \cdot k_j)$, that we set to

$$(k_i \cdot k_j) = \frac{1}{2}(k_i + k_j)^2 = \frac{1}{2}E^2 \,, \tag{3.54}$$

in what follows.

The calculation is conveniently performed in Sudakov coordinates [182]. Namely we

[20] The table is for $E_{\rm cm} = 10$ TeV, however with the scaling in eq. (3.35) the results depend weakly on the energy.

parametrize the loop momentum q as

$$q = uk_i + vk_j + q_\perp^1 \zeta_1 + q_\perp^2 \zeta_2 \,, \tag{3.55}$$

where $(\zeta_1)^2 = (\zeta_2)^2 = -1$ and $\zeta_{1,2} \cdot k_i = \zeta_{1,2} \cdot k_j = 0$. In these coordinates

$$q^2 = u\, v\, E^2 - |q_\perp|^2 \,, \tag{3.56}$$

and the infinitesimal strip $\delta\sigma$ (3.7) that defines the integration region is expressed as

$$|u\, v| E^2 \in [\lambda, \lambda + \delta\lambda] \,. \tag{3.57}$$

After the change of variables, the integral reads

$$I = \frac{i}{(2\pi)^4} \int_{\delta\sigma} \frac{du dv}{uv} \int \frac{d^2 q_\perp}{|q_\perp|^2 - uvE^2 + m_V^2 - i\epsilon} \,. \tag{3.58}$$

The $d^2 q_\perp$ integral must be performed up to an upper cutoff that justifies the usage of the Eikonal approximation formula in eq. (3.8) for the gauge boson vertices. In particular we notice that the actual denominators of the virtual legs in the diagram are not $(k_{i,j} \cdot q)$ as in the Eikonal formula, but rather $(k_{i,j} \cdot q) \pm q^2/2$, with plus or minus for incoming and outgoing particles, respectively. Neglecting $q^2/2$ is justified only if $|q_\perp|^2/2$ is not as large as to compete with the minimum among $|k_i \cdot q| = |u|\, E^2/2$ and $|k_j \cdot q| = |v|\, E^2/2$. Therefore the $q_\perp$ integral should be cutoff at

$$|q_\perp|^2 < \Lambda^2 \,, \qquad \text{with} \qquad \Lambda^2 = E^2 \min\Big[|u|, |v|\Big] \,. \tag{3.59}$$

Similarly, the term proportional to uv in q^2 should not be large compared to $(k_{i,j} \cdot q)$, therefore the u and v integrals are also bounded, in the region

$$|u| < 1 \,, \qquad |v| < 1 \,, \tag{3.60}$$

supplemented by eq. (3.57).

The integration boundaries of u and v are invariant under $u \to -u$ and under $v \to -v$ reflections. We can thus perform the integral (3.58) in the first quadrant of the (u, v) plane, provided we duly symmetrize the integrand. Furthermore we notice that first-quadrant integration region (restricted to $u < 1$ and $v < 1$, owing to eq. (3.60)) is conveniently described by the coordinates τ and y, defined by

$$u = \sqrt{\tau} e^y \,, \qquad v = \sqrt{\tau} e^{-y} \,. \tag{3.61}$$

Indeed in these coordinates the condition (3.57) merely sets the value of τ to

$$\tau = \frac{\lambda}{E^2}\,, \qquad \text{with} \qquad \delta\tau = \frac{\delta\lambda}{E^2}\,, \tag{3.62}$$

while the upper bound (3.59) on the $|q_\perp|$ integral reads

$$\Lambda^2 = E^2\sqrt{\tau}e^{-|y|} = E\sqrt{\lambda}\,e^{-|y|}\,. \tag{3.63}$$

We can thus express the integral as

$$I = \frac{2\,i}{(2\pi)^3}\frac{\delta\lambda}{\lambda}\int_0^{y_{\rm M}} dy \int_0^{E\sqrt{\lambda}e^{-y}} d|q_\perp|^2 \left[\frac{1}{|q_\perp|^2 - \lambda + m_V^2 - i\epsilon} - \frac{1}{|q_\perp|^2 + \lambda + m_V^2 - i\epsilon}\right]. \tag{3.64}$$

where, since Λ^2 is symmetric under $y \to -y$, we could take the y integral from 0 to

$$y_{\rm M} = -\frac{1}{2}\log\tau = \log\frac{E}{\sqrt{\lambda}}\,, \tag{3.65}$$

and multiply by a factor 2.

The $d|q_\perp|^2$ integral in eq. (3.64) is readily performed. It is convenient to separate the terms that emerge from the upper $|q_\perp|^2$ integration extreme from the one of the lower extreme $|q_\perp|^2 = 0$, obtaining two contributions to I, $I_{\rm U}$ and $I_{\rm L}$. We will readily see that fhe former contribution is suppressed, therefore

$$I \simeq I_{\rm L} = \frac{1}{8\pi^2}\frac{\delta\lambda}{\lambda}\log\frac{E^2}{\lambda}\frac{i}{\pi}\log\left[\frac{\lambda + m_V^2 - i\,\epsilon}{-\lambda + m_V^2 - i\,\epsilon}\right]. \tag{3.66}$$

The logarithm gives quite different results in the two regimes $\lambda \gg m_{\rm W}^2 \sim m_V^2$ and $\lambda \ll m_{\rm W}^2 \sim m_V^2$. In the second one, the argument of the logarithm has positive real part, almost equal to to 1 up to m_V^2/λ power-corrections. In the first regime, the argument has negative real part and the log equals $+i\,\pi$, plus λ/m_V^2 corrections. Namely

$$I \overset{\lambda\gg m_{\rm W}^2}{=} -\frac{1}{8\pi^2}\frac{\delta\lambda}{\lambda}\log\frac{E^2}{\lambda}\Big[1+\mathcal{O}(m_V^2/\lambda)\Big]\,, \qquad I \overset{\lambda\ll m_{\rm W}^2}{=} -\frac{1}{8\pi^2}\frac{\delta\lambda}{\lambda}\log\frac{E^2}{\lambda}\cdot\mathcal{O}(\lambda/m_V^2)\,, \tag{3.67}$$

from which we recover eq. (3.11). More precisely, notice that the integral for each pair of external legs ij is controlled by the specific scale $E^2 = 2(k_i \cdot k_j)$. In eq. (3.11) we set all these scales equal up to corrections that are not log enhanced, but of order one. This in turn corresponds to order $1/\log$ corrections to the evolution kernel and to single-logarithm corrections to the solutions of the IREE.

The contribution to I from the upper $|q_\perp|^2$ integration extreme is suppressed. To see

this it is convenient to employ the integration variable $\rho = \exp(y - y_{\rm M})$, obtaining

$$I_{\rm U} = \frac{2\,i}{(2\pi)^3}\frac{\delta\lambda}{\lambda}\int_{\sqrt{\lambda}/E}^{1} d\rho\,\frac{1}{\rho}\log\left[\frac{1-\rho(1-m_V^2/\lambda)}{1+\rho(1+m_V^2/\lambda)}\right], \tag{3.68}$$

where we could drop the $-i\epsilon$ because the argument of the logarithm has positive real part in the entire range of integration. The ρ integral is finite for $\sqrt{\lambda}/E \to 0$, therefore it does not produce log-enhanced contributions. In particular the integral gives $-\pi^2/4$ for $\lambda \gg m_{\rm W}^2 \sim m_V^2$ and it is power-suppressed in the opposite regime.

We now compute the contribution of real radiation to the density matrix variation, which we employ in eq. (3.18). The relevant integral that controls the contribution from a real radiation diagram like those in Figure 3.2, reads

$$I_{\rm R} \equiv -\frac{1}{(2\pi)^3}\int_{\delta\sigma}\frac{d^3q}{2q^0}\frac{(k_i\cdot k_j)}{(k_i\cdot q)(k_j\cdot q)} = -\frac{1}{(2\pi)^3}\int_{\delta\sigma} d^4q\,\delta(q^2-m_V^2)\,\theta(q^0)\frac{(k_i\cdot k_j)}{(k_i\cdot q)(k_j\cdot q)}\,, \tag{3.69}$$

where we employed the eikonal formula in eq. (3.8), but ignored the "G" factors that are included separately in eq. (3.18). Notice the presence of a minus sign, which is due to three factors of -1. The first minus is due to the fact that applying the eikonal formula to the conjugate amplitude gives the complex conjugate of the generator matrix "G_i" of the corresponding leg, while eq. (3.18) is expressed in terms of the generators $G_{i^c} = -G_i^*$ of the conjugate representation. The second minus sign emerges from the sum over the polarizations of the real vector boson, which gives $-\eta_{\mu\nu}$. The third minus is because the contribution to the variation is minus the integral over the strip $\delta\sigma$, since the λ cutoff is a lower bound on the hardness.

In Sudakov coordinates (3.55), and setting $(k_i\cdot k_j) = E^2/2$, the integral becomes

$$I_{\rm R} = -\frac{1}{(2\pi)^2}\int_{\delta\sigma}\frac{du dv}{uv}\,\theta(u+v)\int d|q_\perp|^2\delta\left(uvE^2 - |q_\perp|^2 - m_V^2\right). \tag{3.70}$$

The integration extremes of all the variables are like those of the virtual integral, including the condition (3.57) that defines the infinitesimal integration strip $\delta\sigma$. The delta function in eq. (3.70) has support only if $u\,v > 0$, and the theta function further restricts the integral to the first quadrant of the (u, v) plane. We can thus employ the τ and y coordinates in eq. (3.61) and readily obtain

$$I_{\rm R} = -\frac{1}{8\pi^2}\frac{\delta\lambda}{\lambda}\log\frac{E^2}{\lambda}\cdot\theta(\lambda - m_V^2)\,. \tag{3.71}$$

Evidently, the theta function condition is not satisfied for $\lambda \ll m_{\rm W}^2 \sim m_V^2$, therefore the

Bosons	$\mathrm{SU}(2)_L$	$\mathcal{Y}$
W	t	0
B	s	0
H	d	$+1/2$
$\bar{H}$	$\bar{\mathrm{d}}$	$-1/2$

Leptons	$\mathrm{SU}(2)_L$	$\mathcal{Y}$
$(\nu_{\ell,-1/2}\,, \ell^-_{-1/2})$	d	$-1/2$
$(\bar{\nu}_{\ell,+1/2},\, \ell^+_{+1/2})^t$	$\bar{\mathrm{d}}$	$-1/2$
$\ell^-_{+1/2}$	s	-1
$\ell^+_{-1/2}$	s	$+1$

Quarks	$\mathrm{SU}(2)_L$	$\mathcal{Y}$
$(u_{-1/2},\, d_{-1/2})^t$	d	$+1/6$
$(\bar{u}_{+1/2},\, \bar{d}_{+1/2})^t$	$\bar{\mathrm{d}}$	$-1/6$
$u_{+1/2}$	s	$+2/3$
$\bar{u}_{-1/2}$	s	$-2/3$
$d_{+1/2}$	s	$-1/3$
$\bar{d}_{-1/2}$	s	$+1/3$

Table 3.6: $\mathrm{SU}(2)_L{\times}\mathrm{U}(1)_Y$ quantum numbers of the SM particles in the high-energy regime.

integral vanishes in this regime. In the other regime

$$I_{\mathrm{R}} \overset{\lambda\gg m_{\mathrm{w}}^2}{=} -\frac{1}{8\pi^2}\frac{\delta\lambda}{\lambda}\log\frac{E^2}{\lambda}\,, \tag{3.72}$$

which is equal to the virtual radiation integral, as anticipated in the main text. Notice that the pre-factor of eq. (3.18) contains an additional 1/2, due to the fact that the real radiation contribution is effectively counted twice in the equation by the two terms proportional to $(G_i^A)(G_{j^c}^A)$ and to $(G_{i^c}^A)(G_j^A)$, which are equal after summing over i and j.

3.B High-energy EW multiplets

The EW symmetry is effectively unbroken at energies much above the EW scale. Therefore in this regime it is convenient to describe the SM particles in terms of representations of the (linearly-realized) SM group $\mathrm{SU}(2)_L{\times}\mathrm{U}(1)_Y$, with generators

$$\mathcal{T}^A = \{\mathcal{T}^a, \mathcal{Y}\}\,, \qquad a = 1,2,3\,. \tag{3.73}$$

The generators act on the single particle states as

$$\mathcal{T}^A|\mathrm{p}(k,\alpha)\rangle = |\mathrm{p}(k,\beta)\rangle(T_{\mathrm{r}}^A)^\beta_\alpha\,, \tag{3.74}$$

with generator matrices $T_{\rm r}^A$ that define the representation "r" of the particle multiplet. The field $\Phi_{\rm r}^\alpha$ that interpolates the multiplet from the vacuum, namely

$$\langle 0|\Phi_{\rm r}^\alpha(0)|{\rm p}(k,\beta)\rangle \propto \delta_\beta^\alpha\,, \tag{3.75}$$

transforms with the same generator matrix, i.e.

$$\left[\Phi_{\rm r}^\alpha(x), \mathcal{T}^A\right] = (T_{\rm r}^A)^\alpha_\beta \Phi_{\rm r}^\alpha(x)\,. \tag{3.76}$$

The $\mathrm{SU}(2)_L$ representations of the SM particles and the corresponding $\mathrm{U}(1)_Y$ charges are listed in Table 3.6.

All the fermionic particles with helicity $-1/2$ transforms as doublets (i.e., r = d), the anti-particles with helicity $+1/2$ transform in the conjugate (r = $\bar{\rm d}$) of the doublet representation, while all the others are singlets. Obviously this is true only in the high energy limit where the fermions are effectively massless and the helicity corresponds the chirality of the corresponding interpolating fields. The doublet representation matrices are the standard

$$T_{\rm d}^a = \left\{ \begin{pmatrix} 0 & +1/2 \\ +1/2 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i/2 \\ +i/2 & 0 \end{pmatrix}, \begin{pmatrix} +1/2 & 0 \\ 0 & -1/2 \end{pmatrix} \right\}, \qquad T_{\bar{\rm d}}^a = -(T_{\rm d}^a)^* = -(T_{\rm d}^a)^t\,. \tag{3.77}$$

The EW multiplets of bosonic particles are less well-known, but equally straightforward to work out employing the standard Goldstone Boson Equivalence Theorem, or better its stronger formulation in Ref.s [86, 183]. The point is that for massive $W^\pm$ and Z vector bosons with 0 (longitudinal) helicity one can employ interpolating fields that are a specific combination of the regular gauge fields $W_\mu^\pm$ and Z_μ and of the Goldstone boson scalar fields $\pi^\pm$ and π_Z. The longitudinal states are thus a linear combination of the quantum fluctuation modes associated to this two different type of fields, and the scattering amplitudes with external longitudinal states are a linear combination of Feynman diagrams where the external states are represented either as gauge fields or as Goldstone fields. This is a convenient formalism in the high energy limit because the polarization vector associated with gauge external lines vanishes as $m_{\rm W}/E$ (unlike the regular longitudinal polarization, that diverges as $E/m_{\rm W}$), and only the Goldstone diagrams survive. In essence this means that the Goldstones, and not the gauge, are the adequate interpolating fields for the longitudinal particles at high energy. Therefore the EW quantum numbers of the longitudinal particles are the ones of the Goldstones, and not of the gauge fields. The Goldstone bosons, together with the Higgs, form a doublet with $+1/2$ hypercharge, and the corresponding conjugate doublet

$$H = \left(\pi^+\,,\ \frac{h + i\pi_Z}{\sqrt{2}}\right)^t, \qquad \bar{H} = \left(\pi^-\,,\ \frac{h - i\pi_Z}{\sqrt{2}}\right)^t. \tag{3.78}$$

Vector bosons with transverse helicity $\mathrm{T} = \pm 1$ are instead well-described by gauge fields even at high energy. Therefore they decompose into a triplet plus a singlet EW multiplet, which are readily obtained by undoing the Weinberg rotation

$$W^a = \left(\frac{W^+_\mathrm{T} + W^-_\mathrm{T}}{\sqrt{2}} \,,\, i\frac{W^+_\mathrm{T} - W^-_\mathrm{T}}{\sqrt{2}} \,,\, c_\mathrm{w} Z_\mathrm{T} + s_\mathrm{w} \gamma_\mathrm{T} \right)^t , \qquad B = -s_\mathrm{w} Z_\mathrm{T} + c_\mathrm{w} \gamma_\mathrm{T} \,. \tag{3.79}$$

Explicitly, the triplet generator matrices are

$$T^a_\mathrm{t} = \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} , \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} , \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\} . \tag{3.80}$$

We now proceed to the evaluation of the K_i exponentials in the explicit formula for the semi-inclusive density matrix (3.22), for external legs in the doublet (or anti-doublet) or in the triplet canonical representations defined by eq. (3.77) and (3.80). The K tensors in eq. (3.20) are

$$\Big[K_\mathrm{d}\Big]^{\alpha\bar\alpha}_{\beta\bar\beta} = \Big[K_{\bar\mathrm{d}}\Big]^{\alpha\bar\alpha}_{\beta\bar\beta} = c_\mathrm{d}\, \delta^\alpha_\beta \delta^{\bar\alpha}_{\bar\beta} + \sum_a (T^a_\mathrm{d})^\alpha{}_\beta (T^a_{\bar\mathrm{d}})^{\bar\alpha}{}_{\bar\beta} = \delta^\alpha_\beta \delta^{\bar\alpha}_{\bar\beta} - \frac{1}{2} \delta^{\alpha\bar\alpha} \delta_{\beta\bar\beta} \,, \tag{3.81}$$

$$\Big[K_\mathrm{t}\Big]^{\alpha\bar\alpha}_{\beta\bar\beta} = c_\mathrm{t}\, \delta^\alpha_\beta \delta^{\bar\alpha}_{\bar\beta} + \sum_a (T^a_\mathrm{t})^\alpha{}_\beta (T^a_\mathrm{t})^{\bar\alpha}{}_{\bar\beta} = 2\, \delta^\alpha_\beta \delta^{\bar\alpha}_{\bar\beta} + \delta^\alpha_{\bar\beta} \delta^{\bar\alpha}_\beta - \delta^{\alpha\bar\alpha} \delta_{\beta\bar\beta} \,. \tag{3.82}$$

The exponential of eq. (3.81) is trivial and we readily obtain

$$\left[\exp\Big(-\mathcal{L}\, K_\mathrm{d}\Big) \right]^{\alpha\bar\alpha}_{\beta\bar\beta} = e^{-\mathcal{L}} \delta^\alpha_\beta \delta^{\bar\alpha}_{\bar\beta} + e^{-\mathcal{L}/2} \sinh(\mathcal{L}/2) \delta^{\alpha\bar\alpha} \delta_{\beta\bar\beta} \,, \tag{3.83}$$

where we defined

$$\mathcal{L} = \frac{g^2}{16\pi^2} \log^2(E^2/m^2_\mathrm{w}) \,. \tag{3.84}$$

The exponential of eq. (3.82) is slightly more involved. First of all notice that the three tensor structures in the right-hand-side of eq. (3.82) commute so we can exponentiate all of them independently. In particular we get

$$\exp\Big(-2\mathcal{L} \delta^\alpha_\beta \delta^{\bar\alpha}_{\bar\beta}\Big) = e^{-2\mathcal{L}} \delta^\alpha_\beta \delta^{\bar\alpha}_{\bar\beta} \tag{3.85}$$

$$\exp\Big(-\mathcal{L} \delta^\alpha_{\bar\beta} \delta^{\bar\alpha}_\beta\Big) = \cosh(\mathcal{L}) \delta^\alpha_\beta \delta^{\bar\alpha}_{\bar\beta} - \sinh(\mathcal{L}) \delta^\alpha_{\bar\beta} \delta^{\bar\alpha}_\beta \,, \tag{3.86}$$

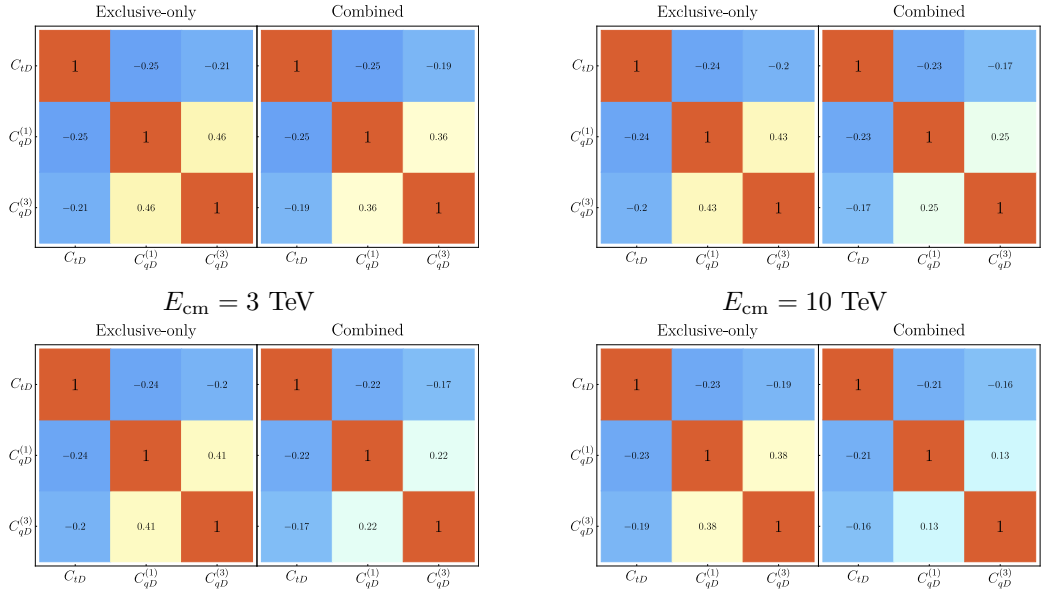
$$\exp\Big(\mathcal{L} \delta^{\alpha\bar\alpha} \delta_{\beta\bar\beta}\Big) = \delta^\alpha_\beta \delta^{\bar\alpha}_{\bar\beta} + \frac{1}{3}\Big(e^{3\mathcal{L}} - 1\Big) \delta^{\alpha\bar\alpha} \delta_{\beta\bar\beta} \,. \tag{3.87}$$

Taking the product of the factors just found, we get

$$\left[\exp\left(-\mathcal{L}\,K_{\rm t}\right)\right]^{\alpha\bar{\alpha}}_{\beta\bar{\beta}} = e^{-2\mathcal{L}}\left(\cosh(\mathcal{L})\delta^{\alpha}_{\beta}\delta^{\bar{\alpha}}_{\bar{\beta}} - \sinh(\mathcal{L})\delta^{\alpha}_{\bar{\beta}}\delta^{\bar{\alpha}}_{\beta}\right) + \frac{2}{3}e^{-\frac{3}{2}\mathcal{L}}\sinh\left(3\mathcal{L}/2\right)\delta^{\alpha\bar{\alpha}}\delta_{\beta\bar{\beta}}\,. \tag{3.88}$$

3.C 3rd family operators

The sensitivity to the 3rd family operators in Table 3.2 are summarized in this section. In Figure 3.13 we report the two-dimensional contours in the $(C_{tD},\, C^{(3)}_{qD})$ and $(C^{(1)}_{qd},\, C^{(3)}_{qD})$ planes, with the third operator set to zero. We notice that the "with radiation" cross-section measurements (see the main text) is mostly effective to probe $C^{(3)}_{qD}$ producing a significant sensitivity improvement on the combined bound in this direction. The effect is milder in the orthogonal directions. The likelihood is dominated by the linear term in the new physics parameters so all our result can be expressed in terms of the single operator reaches (at 95% CL) of and the correlations matrices in Table 3.7. In the table we report the sensitivity of exclusive cross-section measurements alone, and the combination of all the measurements.



$E_{\text{cm}} = 3$ TeV $E_{\text{cm}} = 10$ TeV

$E_{\text{cm}} = 14$ TeV $E_{\text{cm}} = 30$ TeV

	Exclusive-only [95% CL]			Combined [95% CL]		
	C_{tD}	$C_{qD}^{(1)}$	$C_{qD}^{(3)}$	C_{tD}	$C_{qD}^{(1)}$	$C_{qD}^{(3)}$
3 TeV	$[-24.4, 24.4]$	$[-9.47, 9.47]$	$[-6.68, 6.68]$	$[-23.1, 23.1]$	$[-8.59, 8.59]$	$[-5.45, 5.45]$
10 TeV	$[-2.24, 2.24]$	$[-0.97, 0.97]$	$[-0.71, 0.71]$	$[-2.04, 2.04]$	$[-0.81, 0.81]$	$[-0.49, 0.49]$
14 TeV	$[-1.15, 1.15]$	$[-0.52, 0.52]$	$[-0.38, 0.38]$	$[-1.03, 1.03]$	$[-0.42, 0.42]$	$[-0.25, 0.25]$
30 TeV	$[-0.26, 0.26]$	$[-0.13, 0.13]$	$[-0.10, 0.10]$	$[-0.22, 0.22]$	$[-0.10, 0.10]$	$[-0.05, 0.05]$

Table 3.7: Single-operator 95% CL reach and correlation matrices for the Wlison coefficient $C_{qD}^{(3)}$, $C_{qD}^{(1)}$ and $C_{qD}^{(3)}$ of the operators of Table 3.2 at different collider energies. All results include exclusive cross-sections or combined measurements. The Wilson coefficient are expressed in $10^{-4}\,\text{TeV}^{-2}$. Since the likelihood is dominated by the linear terms in the new physics parameters, the single parameter reach and the correlation characterize our results completely.

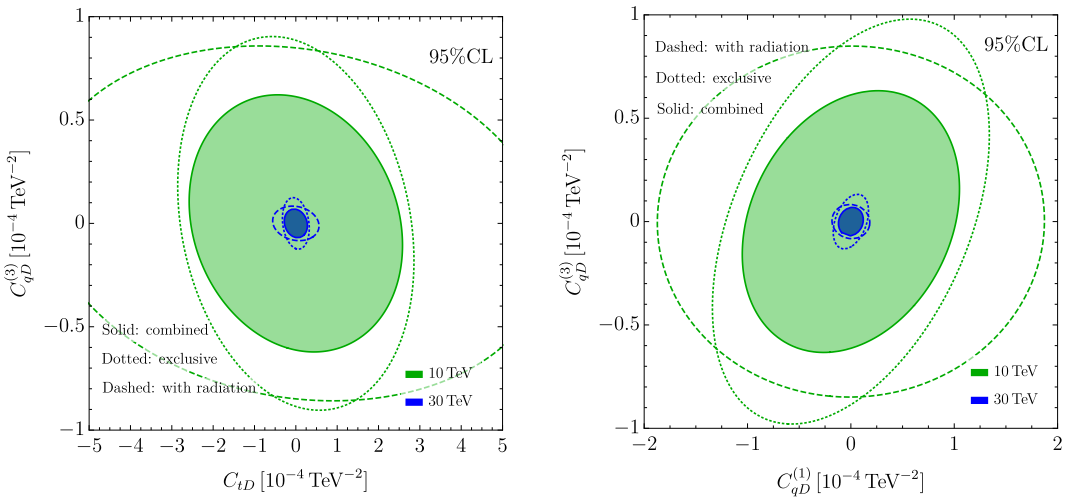


Figure 3.13: 95% CL contours in the $(C_{tD}, C^{(3)}_{qD})$ (left) and $(C^{(1)}_{qd}, C^{(3)}_{qD})$ (right) planes at the 10 and 30 TeV muon colliders.

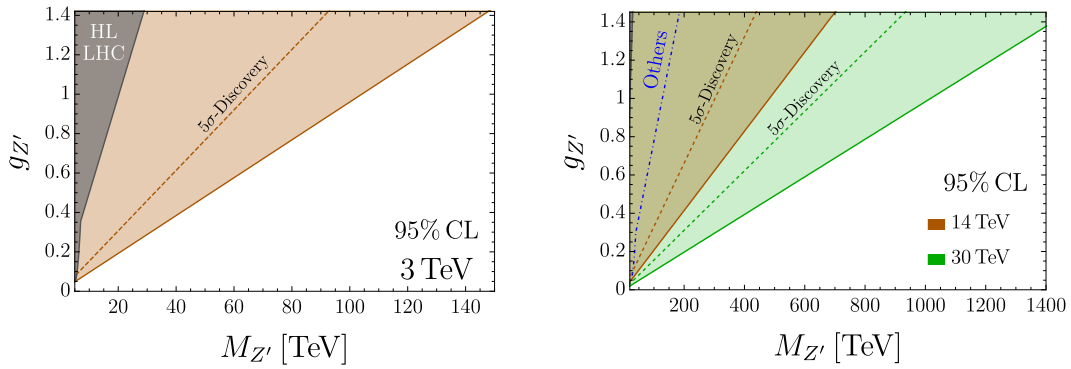


Figure 3.14: The same as the right panel of Figure 3.10 for various collider energies.

3.D Summary plots

In this appendix we collect additional results skipped in the main text. In particular in Figure 3.14 we report the sensitivity projections for the Y-universal Z' model, in the $(g_{Z'}, M_{Z'})$ plane for the different collider energies. In Figure 3.15 we collect the sensitivity projections for the composite Higgs model in the (m_*, g_*) plane for $E_{\rm cm} = 3$, 14, 30 TeV. Projections including composite top measurements can be found in Figure 3.17. Finally, Figure 3.16 shows the dependence of the bound on the value of the c_{tt} coefficients, as explained in the main text.

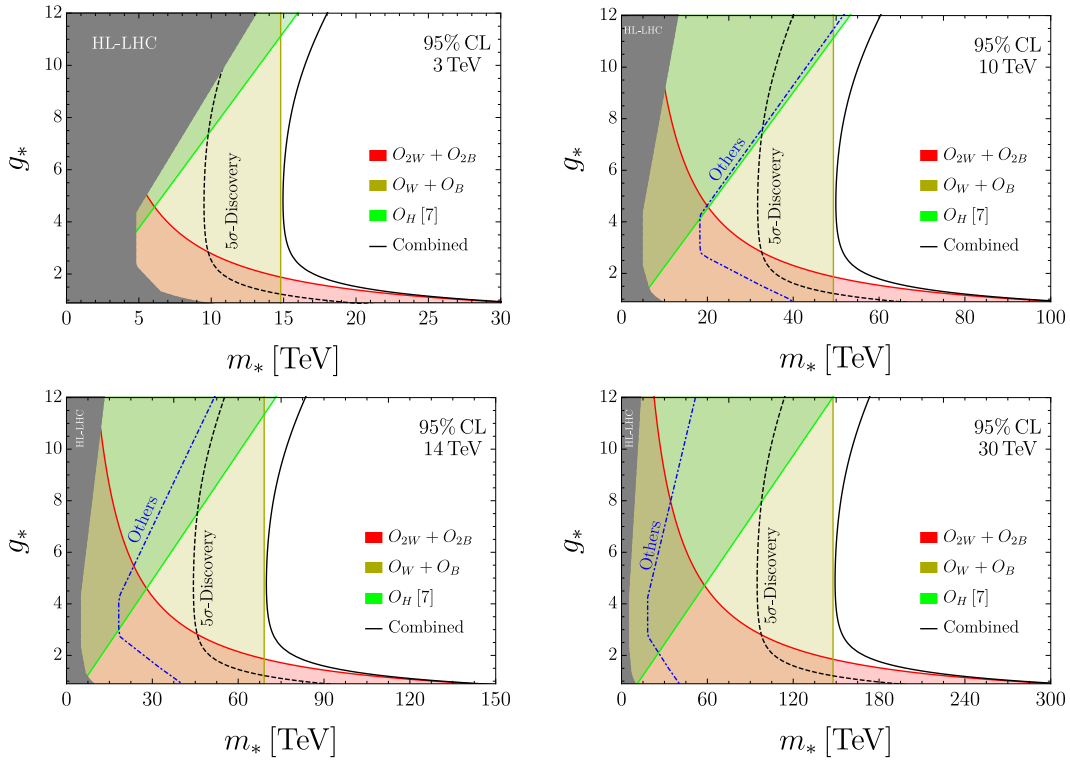


Figure 3.15: The same as the left panel of Figure 3.10 for various collider energies.

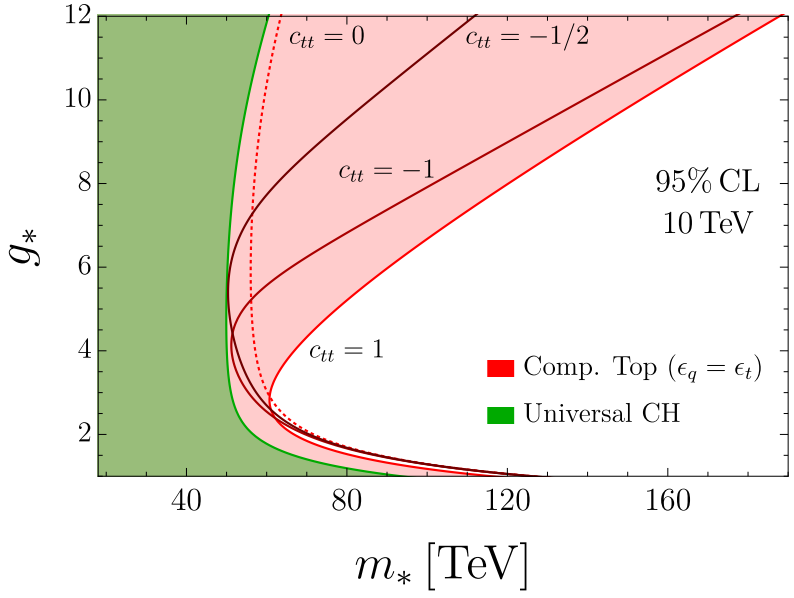


Figure 3.16: The same as the right panel of Figure 3.11 for different values of c_{tt} to show the model dependence of the result.

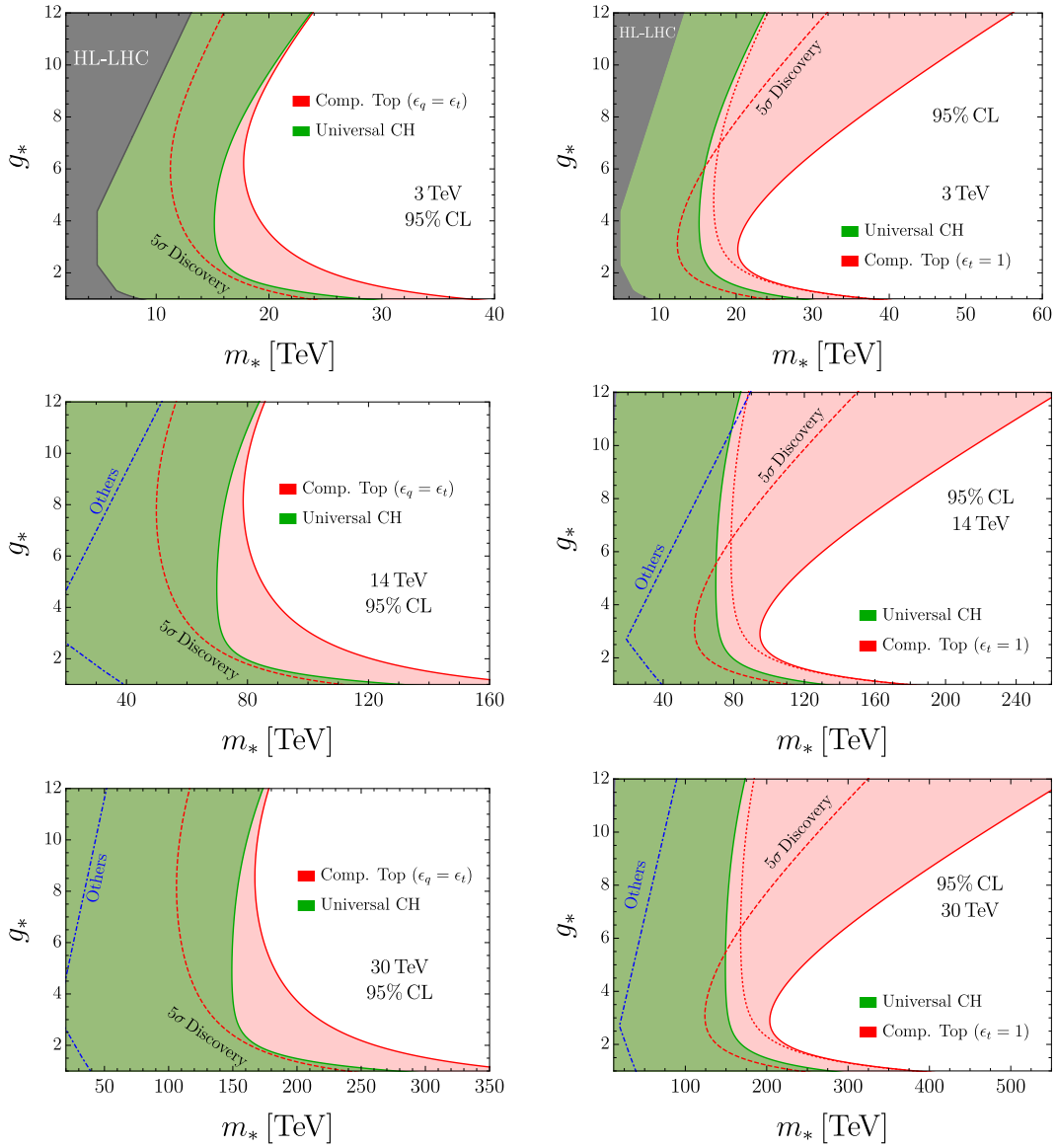


Figure 3.17: The same as Figure 3.11 for various collider energies. The blue line on the equally-composite (left panels) projections are taken from [37], while on the right-handed composite top scenario are taken from [178].

4 Tetraquarks in a Non-Relativistic Effective Field Theory

The thesis was so far focused on the search for physics Beyond the Standard Model, in particular at energies much higher than the electroweak scale. In the examples presented in the previous chapters, it is clear that an Effective Field Theory extension of the Standard Model is a powerful way of parametrizing the low-energy effects of an unknown underlying microscopic theory in a model independent way.

However, Effective Field Theories are extremely useful also in cases where the microscopic theory is well known but cannot be used to make accurate predictions. In this chapter we will see an application of such techiques for the study of the $X(3872)$ exotic resonance, a QCD tetraquark candidate whose fundamental nature is not yet understood due to the non-perturbative nature of the theory. We will see that the seemingly accidental tuning between the mass of this objects and the $D - D^*$ meson threshold will allow us to describe its dynamics in terms of a non-relativistic EFT. Understanding this theory will hopefully lead to a general parametrization for these objects in terms of Lagrangian couplings that might be used to describe the experimental data on such resonances and to understand their nature.

In sec. 4.1 we will give a short review about exotic QCD resonances and in particular the $X(3872)$. In sec. 4.2 we present the two main interpretations of such states as bound states of two mesons or as compact QCD tetraquarks, showing hot recent studies seem to prefer the latter possibility to the former. In sec 4.3 we study whether the exchange of pions could mimic some of the results found in the previous section. Finally in sec. 4.4 we recap the main results and outlook future research directions on the topic.

4.1 Exotic resonances and the X(3872)

It is known since [184] that while QCD is written in terms of colored objects, quarks $q \sim \mathbf{3} \in SU(3)$, anti-quarks $\bar{q} \sim \bar{\mathbf{3}} \in SU(3)$ and gluons $G \sim \mathbf{8} \in SU(3)$, at low energies the physical states of the theory are color singlets, bound states of these elementary

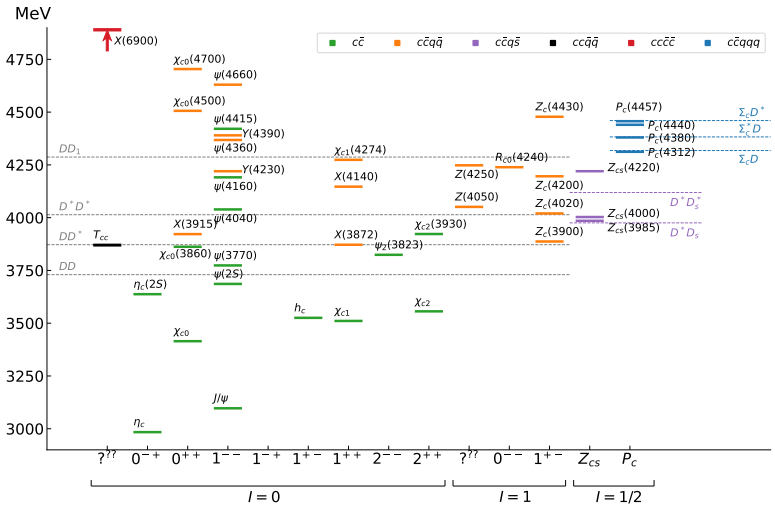


Figure 4.1: Summary of ordinary charmonia and the exotic $XYZP$ tetraquark and pentaquark states. Figure taken from [187]

objects. The standard classification of such bound states is based on their constituents and traditionally goes as

$$\text{Meson} = \bar{q}_i q^i\,, \qquad \text{Baryion} = \epsilon_{ijk} q^i q^j q^k\,, \qquad \text{Anti-baryion} = \epsilon^{ijk} \bar{q}_i \bar{q}_j \bar{q}_k\,.$$

However, the theory itself does not prevent other colorless combinations. Example of these "exotic" states are the glueball $G^a G^a$, the tetraquark $\bar{q}\bar{q}qq$ and the pentaquark $qqqq\bar{q}$. As it stands today, there seem to be evidence of the existence of several tetraquark and pentaquark states (see [185, 186] for recent reviews on the topic). These evidences are considerably stronger for the so-called hidden charm $c\bar{c}$ and hidden beauty $b\bar{b}$ states. The reason is twofold. First of all, for states containing only light quarks it is difficult to properly define the concept of an exotic state, since at strong coupling there is no strong separation of scales and it is impossible to properly define what is the difference between a $q\bar{q}$ state as opposed to a $qq\bar{q}\bar{q}$ state. Furthermore, for the charmonia and bottomonia $Q\bar{Q}$ states it is possible to predict the spectrum with high accuracy with potential models since the quark masses are large enough to allow for a non-relativistic Coulomb-like description of the bound state in terms of their constituent quarks. This has made possible to identify several new states that do not agree with the predicted spectrum. The first of such unexpected resonances was the $X(3872)$, observed for the first time in 2003 by Belle [188], in the decay

$$B^0 \to K^0 X \to K^0 J/\psi \rho\,. \tag{4.1}$$

The decay into a J/ψ signals the presence of a c and $\bar{c}$ quarks in the X particle, however this state does not fit into the charmonium predicted spectrum. The smoking gun that confirmed the existence of multi-quark states was however the $Z^+(4430)$ charged

charmonium discovered by Belle in 2007 [189] in the decay

$$B^0 \to K^- Z^+ \to K^- J/\psi \rho^+ \,. \tag{4.2}$$

It is indeed clear that the state Z has to contain both a $c\bar{c}$ pair and some light quarks in order to have it charged.

Since then several other exotic states have been observed. The naming convention, although not universally established, is as follows

- X states, e.g. $X(3872)$, for neutral states observed in decays into ψ and pions and positive parity.
- Y states, e.g. $Y(4260)$, for neutral states observed in e^+e^- annihilations with initial state radiation, therefore with negative parity.
- Z states, e.g. $Z(4430)$, for isospin triplet states, thus both charged and neutral, and positive parity
- P states, for pentaquark candidates.

A list of such states in the charmonium sector can be found in Fig. 4.1, taken from [187]. The most recent at the time of writing is the doubly charmed $T_{cc}^+(3875)$ state [190, 191] that shares some of the properties that we will discuss in the chapter regarding the $X(3872)$ state.

Understanding the nature of these unconventional mesons is one of the greatest present challenges in hadronic physics. Indeed, while it is now established that these states contain four valence quarks, their binding mechanism is far less clear: are these resonances a pure QCD tetraquark state bound by color interaction or are they bound states of two color-neutral mesons? So far there has been no direct evidence for either these hypotheses, meaning maybe that a satisfactory description of these objects should include both possibilities. In the next section we will review the main physical properties and observational consequences of the two models applied to the $X(3872)$ resonance. But before that it is necessary to briefly list what is experimentally known about this state.

As already mentioned, the $X(3872)$ was first discovered in the $J/\psi\pi^+\pi^-$ invariant mass spectrum from a B meson decay, by the Belle experiment and has since then been confirmed by multiple experiments (BaBar [192], CDF [193], D∅ [194], CMS [195], LHCb [196]) from both B meson decays and in prompt production from pp or $p\bar{p}$ collisions. The state is confirmed to be neutral and have $J^{PC} = 1^{++}$ quantum numbers to a high precision. One of the peculiarity of this hadron that makes it so puzzling is that its mass is extremely fine-tuned to the $D^0\bar{D}^{0*}$ threshold. The current value for the difference

from threshold is still zero within the uncertainties, but seems to be around [197]

$$\epsilon_X \equiv |m_X - m_{D^0} - m_{D^{0*}}| = -0.07 \pm 0.12\,\text{MeV}\,. \tag{4.3}$$

This tuning is what prompted the interpretation of this state as a molecule between the D^0 and D^{0*} mesons. The main decay channel is indeed in $D^0\bar{D}^0\pi^0$ with a branching ratio of [26]

$$\text{Br}[X \to D^0\bar{D}^0\pi^0] = 0.5 \pm 0.2\,. \tag{4.4}$$

The X can also decay to $J/\psi\,\rho$ and $J/\psi\,\omega$ with roughly the same rate

$$\text{Br}[X \to J/\psi\,\rho] \approx \text{Br}[X \to J/\psi\,\omega] \approx 0.05\,, \tag{4.5}$$

pointing to a large violation of isospin. The possibility of decaying in a $I = 1$ state strongly disfavors the interpretation of the X as a pure charmonium or a $\bar{c}cG$ hybrid state. The total width of the X has not yet been measured exactly and varies quite a lot based on the parametrization used for the lineshape. The current value seems to be around [198]

$$\Gamma_X \approx 0.1 - 1\,\text{MeV}\,. \tag{4.6}$$

The measurement of the width is still challenging and requires to understand the most general dynamics underlying this state. In the molecule picture one would expect the width to be at least as large as the width of the shortest lived component, in this case the D^{0*}, that is expected around 85 KeV. The total width is thus a fundamental parameter that has to be measured precisely for this state in order to better study the nature of this state.

4.2 Hadronic Molecule or Tetraquark?

4.2.1 Hadronic Molecule

Given the extreme closeness to the D^0D^{*0} threshold of the $X(3872)$, a possible description of this state that has been proposed is that of an hadronic molecule [199–201]. Similar to how the deuterium is a bound state between a proton and a neutron, the X could be a state made of the two D mesons bound by a potential slightly tuned in order to give a particularly small binding energy. Given the measured value of the charge conjugation of this state, $C = +1$, the molecule must have the structure

$$X = \frac{D^0\bar{D}^{0*} + D^{0*}\bar{D}^0}{\sqrt{2}}\,. \tag{4.7}$$

To form the molecule we assume an attractive potential between the two mesons given by the exchange of lighter mesons. In this hypothesis the mass difference with the threshold

must be assumed to be negative and it corresponds to the molecule binding energy B

$$\epsilon_X = B \approx -100\,\mathrm{KeV}\,. \tag{4.8}$$

Such a small binding energy compared to the masses of the molecule constituents leads to several important consequences. First of all, the typical momentum of the constituents is very non-relativistic. Indeed we have

$$k \sim \sqrt{2\mu_r B} \approx 14\,\mathrm{MeV} \ll m_{D^0}\,, \tag{4.9}$$

where $\mu_r = 976\,\mathrm{MeV}$ is the reduced mass of the D^0 and D^{0*}. This momentum scale is also smaller than the one coming from the interaction range between the two mesons, that can be estimated to be $\lesssim m_\pi$ from the Yukawa potential of a single pion exchange. As we will see in the next section however, this estimate is not correct due to an accidental fine tuning of the D^{0*} mass and the $D^0 - \pi^0$ threshold. Indeed it turns out that

$$\delta \equiv m_{D^{0*}} - m_{D^0} - m_{\pi^0} \approx 7\,\mathrm{MeV}\,. \tag{4.10}$$

We will see in section 4.3 that this leads to a potential with a longer range compared to the naive estimate. Still for the following arguments we can ignore this fact, since we will see that it does not lead to any large deviations from what we will now find.

In the following, for simplicity of notation, we will refer to the D^0 and the D^{0*} simply as D and D^*.

Quantum Mechanical approach

The scale separation between the momenta k and the mass m_π, or equivalently between the size of the bound state and the maximum potential range, leads to an effective description of the system in terms of long distance parameters that are insensitive to the details of the potential. For example, just by traditional quantum mechanical arguments we could proceed as follows [202]. We consider the wavefunction of the $C = +1$ combination of the $DD*$ system in the center of mass frame. In the non-relativistic limit, this system is predominantly in s-wave, so we can neglect the dependence on the angular momentum. The radial s-wave Schrödinger equation for the radial wavefunction $\psi(r) = \chi(r)/r$ of the DD^* system in the center of mass frame reads

$$\chi(r)'' + 2\mu_r[E - U(r)]\chi(r) = 0\,. \tag{4.11}$$

As justified above, we assume that $U(r)$ has a range of the order of m_π^{-1} and it is tuned in such a way to allow a single bound state with energy $B \ll U$. In this hypothesis the equation can be separated in two regions, for $r < \bar{r}$ and $r > \bar{r}$, with $\bar{r} \sim m_\pi^{-1}$ and we will consider low momenta $k\bar{r} \ll 1$. When this happens, in the inner region we can neglect E

compared to U

$$\chi(r)'' - 2\mu_r U(r)\chi(r) = 0 \qquad r < \bar{r}\,, \tag{4.12}$$

while in the outer region we can neglect $U(r)$

$$\chi(r)'' + 2\mu_r E\chi(r) = 0 \qquad r > \bar{r}\,. \tag{4.13}$$

This shows that for long wavelengths the potential only acts as a boundary condition at $r = \bar{r} \approx 0$ for the free Schrödinger equation. Furthermore a boundary condition is energy independent since it is determined by the first equation and thus can be computed in the special case $E = -B$. This leads to a universal prediction for the wavefunction of the bound state

$$\chi = e^{-r\sqrt{2\mu_r B}}\,, \tag{4.14}$$

and for the scattering state for $E = k^2/2\mu_r > 0$

$$\chi = \sin(kr + \delta_s)\,, \tag{4.15}$$

where δ_s is the phase shift

$$k\cot\delta_s = -\sqrt{2\mu_r B}\,. \tag{4.16}$$

We see that in this approximation the whole dynamics is described by the single parameter B. This last result can also be rewritten in terms of the $D - D^*$scattering amplitude

$$f_s = \frac{1}{k\cot\delta_s - ik} = \frac{1}{-\sqrt{2\mu_r B} - ik}\,, \tag{4.17}$$

from which we see that the ϵ parameter is related to the scattering length a_s by

$$a_s = \frac{1}{\sqrt{2\mu_r B}} \approx 14\,\text{fm} \gg m_\pi^{-1}\,. \tag{4.18}$$

The small k expansion of the phase shift, known as the effective range expansion

$$k\cot\delta = -\frac{1}{a_S} + \frac{1}{2}r_0 k^2 + O(k^4)\,, \tag{4.19}$$

is in this case exact. The effective range r_0 arises from the finer details of the potential that we have neglected and is thus expected to be of the order of m_π^{-1}.

Non-relativistic EFT

The same system can be equivalently studied in a more modern Effective Field Theory language. This theory will contain two interpolating fields for the scalar mesons D, $\bar{D}$ and two vector fields $\mathbf{D}$ and $\bar{\mathbf{D}}$ for the spin-1 mesons and will be built by performing two expansions: an expansion in low velocities $v \ll 1$, that will give us the non-relativistic limit, and an expansion in low momenta compared to the scale m_π. The most general

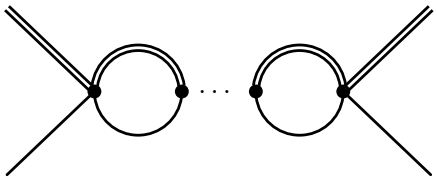


Figure 4.2: The class of diagrams that contribute to the $D^0\bar{D}^{*0}$ scattering amplitude in the non-relativistic EFT. The single line represents a D meson and the double line represents a D^*.

Lagrangian reads

$$\begin{aligned}\mathcal{L} =& D^\dagger\left(i\partial_t+\frac{\nabla^2}{2m_D}\right)D+\bar{D}^\dagger\left(i\partial_t+\frac{\nabla^2}{2m_D}\right)\bar{D}\\ &+\mathbf{D}^\dagger\left(i\partial_t+\frac{\nabla^2}{2m_{D^*}}\right)\mathbf{D}+\bar{\mathbf{D}}^\dagger\left(i\partial_t+\frac{\nabla^2}{2m_{D^*}}\right)\bar{\mathbf{D}}\\ &+\frac{\lambda}{2}\left(D\bar{\mathbf{D}}+\mathbf{D}\bar{D}\right)^\dagger\left(D\bar{\mathbf{D}}+\mathbf{D}\bar{D}\right)+\mathcal{O}\left(\frac{\nabla^2}{m_\pi^2}\right)\end{aligned} \tag{4.20}$$

This Lagrangian deserves several comments. First of all, the meson fields become complex in the non-relativistic limit. This is because in this limit due to the conservation of the numbers of particles, each operator contains only either creation or annihilation operators, and not both. This conservation law becomes in the effective theory a $U(1)$ global symmetry for each of the fields in the Lagrangian. The coupling λ represents the short distance interaction between the $C=+$ combination of mesons. We neglect other combinations of couplings since they do not contribute to the X bounds state.

The conservation of the number of particles at each vertex makes it such that the only diagrams that contribute to the $DD^*\to DD^*$ scattering amplitude are the ones in Fig. 4.2, given by a series of bubbles. This simple structure makes it possible to resum the whole series of diagrams in order to obtain the exact amplitude non-perturbatively in λ.

The resummed amplitude in the center of mass frame is

$$\mathcal{A}_s=-\frac{\lambda}{1+i\lambda L(E)}\,, \tag{4.21}$$

where

$$L(E)=-i\frac{\mu_r}{\pi^2}\Lambda+\frac{i\mu_r}{2\pi}\sqrt{-2E\mu_r-i\varepsilon} \tag{4.22}$$

is the linearly divergent bubble regularized with a hard cutoff $\Lambda\sim m_\pi$ and $\varepsilon>0$ tells us the correct sheet for the square root. Renormalizing the amplitude, imposing that at

$E = 0$ one has $\mathcal{A} = -\lambda_R$ gives the relation

$$\frac{1}{\lambda_R} = \frac{1}{\lambda} + \frac{\mu_r}{\pi^2}\Lambda \tag{4.23}$$

and the renormalized amplitude

$$\mathcal{A}_s = -\frac{\lambda_R}{1 - \frac{\lambda_R \mu_r}{2\pi}\sqrt{-2E\mu_r - i\varepsilon}}\,. \tag{4.24}$$

When $E = k^2/2\mu_r$ this amplitude becomes

$$\mathcal{A}_s = \frac{2\pi}{\mu_r}\left(-\frac{2\pi}{\mu_r \lambda_R} - ik\right)^{-1} = \frac{2\pi}{\mu_r} f_s\,, \tag{4.25}$$

where f_s is the same as eq. (4.17). From this result we immediately read the scattering length

$$a_s = \frac{\lambda_R \mu_r}{2\pi}\,, \tag{4.26}$$

while the effective radius r_0 is zero. A contribution to r_0 would come from higher derivative operators, such as

$$O_6 = \frac{\lambda_6}{16 m_\pi^2}\left(D\bar{\mathbf{D}} + \mathbf{D}\bar{D}\right)^\dagger \left(D \overleftrightarrow{\nabla}^2 \bar{\mathbf{D}} + \mathbf{D} \overleftrightarrow{\nabla}^2 \bar{D}\right)\,, \tag{4.27}$$

where $\overleftrightarrow{\nabla} = \overleftarrow{\nabla} - \overrightarrow{\nabla}$ and is thus expected to be $r_0 \sim m_\pi^{-1}$.

The amplitude (4.24) has a pole at the binding energy of the bound state $E = -B$ with

$$B = \frac{2\pi^2}{\lambda_R^2 \mu_r^3} = \frac{1}{2\mu_r a_s^2}\,, \tag{4.28}$$

that agrees with what found in eq. (4.18) using the simple Quantum Mechanics arguments.

Finally, from eq. (4.23) we can see what are the conditions necessary to obtain an anomalously large scattering length or equivalently a small binding energy. We can indeed rewrite the equation in terms of the physical scattering length coming from λ_R and a "bare" scattering length coming from λ as

$$\frac{1}{a_s} \sim \frac{1}{a_{\text{bare}}} + \frac{2}{\pi} m_\pi\,, \tag{4.29}$$

where we have also used the fact that $\Lambda \sim m_\pi$. By dimensional analysis we would expect a natural size for the bare scattering length of $a_{\text{bare}} \sim m_\pi^{-1}$. To obtain a scattering length that is larger of a factor ~ 10 than m_π^{-1}, as it happens in the real case, there must be an accidental cancellation of one part in ten between the two terms in eq. (4.29).

4.2.2 Compact tetraquark

The second interpretation of the $X(3872)$ that we want to consider is the possibility that this state is a four quarks state tightly bound by QCD interactions [203, 204]. In this section we will first review some of the problems that the molecule model faces when compared to experimental data and then move on to see how the tetraquark hypothesis can address some of those.

Problems of the Hadronic Molecule model

We have seen that it is possible to generate a shallow bound state between the two D and D^* mesons thanks to a moderate tuning and we have computed the consequences of such hypothesis, in particular regarding the structure of the $D - D^*$ scattering amplitude in the presence of such bound state. The most important prediction that we found is that this amplitude has a large scattering length fixed by the X binding energy by eq. (4.18) while it has a small $\sim m_\pi^{-1}$ effective range.

The main reason that motivated the work of this chapter is indeed related to the measured value of r_0. It was recently found [198] that the latest experimental fit to the resonance lineshape seem to give a value for scattering length and effective range of

$$a_s^{\text{LHCb}} \sim 28 \text{ fm}, \quad r_0^{\text{LHCb}} \sim -5.3 \text{ fm}\,. \tag{4.30}$$

This result is particularly interesting because of its sign and size. From general quantum mechanics arguments [202, 205] it can be proven that for a fully attractive potential r_0 is expected to be positive . Furthermore as we have seen, we would expect its size to be of the order of $m_\pi^{-1} \approx +1.4$ fm. This is indeed what happens for the deuteron, a loosely bound state of a proton and a neutron for which we also have a large scattering length $a_s^{\text{d}} \approx +5.4$ fm but a natural effective range $r_0^{\text{d}} \approx +1.7$ fm. Explaining the size and sign of r_0 in the molecular model would require a second numerical accident, in addition to the B tuning, that enhances the size of this coefficient. However, as we will see in the next section, a more natural explanation would be the mixing with a compact state such as a charmonium or a tetraquark.

Another important observable that could help discern the nature of the $X(3872)$ would be the partial width in J/ψ and a ρ or ω meson. Intuitively the, the molecular model would predict a large width for the decay into $D^0\bar{D}^{*0}$ or $D\bar{D}\pi$ due to the molecular dissociation or the decay of its $\bar{D}^*$ component. The decay into a compact charmonium should be extremely suppressed due the large difference in size between the molecule and the charmonium. The probability that the charm and anti-charm quarks lie in a region

of size of ~ 1 fm is roughly given by

$$P \sim \left(\frac{1\text{ fm}}{a_s}\right)^3 \sim 10^{-3}\,. \tag{4.31}$$

Using this estimate, and computing the available phase space for the decays one can estimate [185]

$$\Gamma(X \to J/\psi\rho) \sim \Gamma(X \to J/\psi\omega) \sim 5\,\text{KeV}\,. \tag{4.32}$$

This result is compatible with the experimental measurement if the total width of the X is of the order of $\lesssim 100\,\text{KeV}$. A more precise measurement of Γ_X is thus needed to assess this.

Another hint that the molecular picture for the $X(3872)$ cannot be the full story, comes from the observation of this resonance in prompt hadronic collisions. In particular, comparing the production at high transverse momentum with those of deuterium or other light nuclei, shows that the X is produced at much higher rates than what one would expect for a loosely bound state [206, 207, 185] while it is comparable to the rates of production of a standard charmonium. The intuitive picture behind this simple: we have seen that the relative momenta of the D and D^* inside the molecule is small from equation (4.9). It then is clear then that the probability of creating the two hadrons at high transverse momentum ($p_\perp \gtrsim$ GeV) but low relative momentum should be much smaller compared to creating a charmonium of similar mass with the same $p_\perp$. The experimental measure for such cross sections shows instead

$$\sigma(p\bar{p} \to \psi(2S) + \dots; p_\perp > 5\,\text{GeV})_{\text{prompt}} = 67\text{ nb} \tag{4.33}$$

$$\sigma(p\bar{p} \to X(3872) + \dots; p_\perp > 5\,\text{GeV})_{\text{prompt}} = 30 - 70\text{ nb}\,. \tag{4.34}$$

In the molecular picture, this last cross section has an upper bound given by the production cross section of a D and D^* with a relative momentum of $\sim 14\,\text{KeV}$ of (4.9)

$$\begin{aligned}\sigma(p\bar{p} \to X(3872) + \dots) &\approx \left|\int_{\mathcal{R}} \psi(p)\,\langle DD^* + \dots |p\bar{p}\rangle\right|^2 d^3p \\ &\lesssim \left|\int_{\mathcal{R}} \langle DD^* + \dots |p\bar{p}\rangle\right|^2 d^3p\,,\end{aligned} \tag{4.35}$$

where $\psi(p)$ is the $X(3872)$ molecule wave function in momentum space and $\mathcal{R}$ is a ball in momentum space with radius k from (4.9). This upper bound was estimated by Monte Carlo simulations and found to be ~ 300 times smaller than what experimentally measured. This result depends on the region of integration $\mathcal{R}$ and has been criticized, for example in [208]. In [207] however the prompt production of the X was compared with the experimental data for several light nuclei, that are obviously hadronic molecules, and was still shown to be much higher than expected.

On the same line, the prompt production of the X and the $\psi(2S)$ charmonium were

compared in events with high multiplicity. In this kind of events final state interactions become important and give very different predictions for the molecule and tetraquark scenarios. Since the many final state collisions tend to lower the average relative momentum, the rate of production of the molecule compared to the charmonium increases with the number of particles in the final state. On the other hand for the tetraquark picture the rate decreases. The results from [209] applied to the experimental results of LHCb [210] show a strong preference with the tetraquark picture.

Finally, another problem of the molecular interpretation that is worth mentioning is the fact that even though it has not yet been measured precisely, the total width of the $X(3872)$ seems to be larger or at best comparable to its binding energy. However, since the binding energy and width are inversely related to the time it takes to respectively form and destroy the bound state, this would lead to the paradoxical conclusion that the X is a state that decays before it even has the chance to form. More accurate measurements of these two parameters are however needed in order to confirm this.

Weinberg compositeness criterion

The question on how to distinguish a composite state from a compact one by looking at long distance observables was already answered by Weinberg in 1965 [211]. His result is now known as the Weinberg compositeness criterion and showed how it is possible to determine whether the deuterium is an elementary particle or a proton-neutron bound state from scattering experiments. In this context, the words "elementary" and "composite" depend on the length scales that can be probed by the experiment. In this case what we mean is that there exists an hierarchy of length scales set by the size of the states l and the collision momentum k of the type

$$l_n \approx l_p \ll l_d \lesssim 1/k\,. \tag{4.36}$$

In this limit the proton and neutron can be considered as point-like objects, while their bound state cannot. The fact that the size of the deuteron is larger than the typical nuclear size is due to another accidental tuning that, similarly to the $X(3872)$ in the previous sector, makes the binding energy particularly small ($B \approx 2.22\,\text{MeV}$).

The starting point of Weinberg's argument is a Hamiltonian H composed of a free part H_0 and a potential term V that describes all the states with the same quantum numbers of the deuteron. H_0 will then contain a continuum of two particles proton-neutron states $|n\,p(\mathbf{k})\rangle$, where $\mathbf{k}$ is the relative momentum, and one (or more) discrete single particle states $|\mathfrak{d}\rangle$ that correspond to a possible elementary deuteron with a size comparable to the nucleons. The full Hamiltonian H contains instead the physical deuteron state $|d\rangle$ that has energy

$$H\,|d\rangle = -B\,|d\rangle \tag{4.37}$$

with respect to the $p-n$ threshold. The existence of this state comes from the potential V that contains both interactions between proton and neutron, that could generate a bound state in the absence of $|\mathfrak{d}\rangle$, and interaction between the elementary deuteron and the continuum. The physical deuteron $|d\rangle$ can then be written as a mixture of the two H_0 eigenstates, weighted by an overlap parameter Z defined as

$$Z \equiv |\langle \mathfrak{d}|d\rangle|^2 . \tag{4.38}$$

This leads to the decomposition

$$|d\rangle = \sqrt{Z}\,|\mathfrak{d}\rangle + \int \frac{d^3\mathbf{k}}{(2\pi)^3} C(\mathbf{k})\,|n\,p(\mathbf{k})\rangle \ , \tag{4.39}$$

where

$$\int \frac{d^3\mathbf{k}}{(2\pi)^3} |C(\mathbf{k})|^2 = 1 - Z . \tag{4.40}$$

The overlap Z satisfies $0 \leq Z \leq 1$, where the case $Z = 0$ corresponds to the case in which the physical deuterum is a pure bound state, while the case $Z = 1$ to the extreme limit in which the theory is free, there are no interactions with the $|np\rangle$ continuum and the physical deuteron coincides with the bare one.

Weinberg then computes the elastic proton-neutron scattering amplitude in the small scattering energy limit in which the potential can be treated as approximately local. Explicitly he defines an effective coupling between the continuum and the physical state $|d\rangle$ as

$$g \equiv |\langle np|V|d\rangle|_{E=0} , \tag{4.41}$$

valid until $\sim 2\mu_r E/m_\pi^2 \ll 1$. This effective coupling can be rewritten using eq. (4.37) and (4.39) as

$$g^2 = \frac{2\pi\sqrt{2\mu B}}{\mu^2}(1-Z) . \tag{4.42}$$

From this, Weinberg solves the Lippmann-Schwinger equation to compute the scattering phase-shift as a function of these parameters. We will not repeat the steps of his derivations here since, as we will show in the next section, his results can be more easily computed with a non-relativistic EFT language. The final result of the computation gives for the scattering length and the effective range

$$\begin{aligned} a_s &= \frac{2(1-Z)}{2-Z}\frac{1}{\sqrt{2\mu_r B}} + O(m_\pi^{-1}) \\ r_0 &= -\frac{Z}{1-Z}\frac{1}{\sqrt{2\mu_r B}} + O(m_\pi^{-1}) . \end{aligned} \tag{4.43}$$

We can see that in the limit $Z \to 0$, that corresponds to the case in which the deuteron is a pure bound state, we find the results of the previous section and in particular the fact that r_0 is zero, up to finite potential range corrections. When $0 < Z \lesssim 1$ instead,

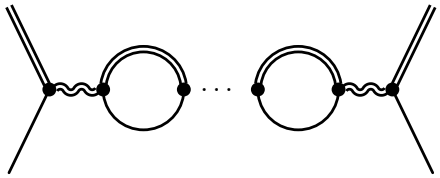


Figure 4.3: The class of diagrams that contribute to the $D^0\bar{D}^{*0}$ scattering amplitude in the non-relativistic EFT with the compact tetraquark. The double wiggly line represents the **X** field.

we see that, up to m_π^{-1} corrections, the effective range can be large for a shallow bound state and, most importantly, negative.

By measuring scattering length and effective range one can determine the parameter Z and thus measure if there exists a compact component in the bound state. If one applies this criterion to the deuterium, where $r_0 \approx +1.75$ fm $\approx m_\pi^{-1}$ one can conclude that Z has to be small and thus the deuterium is simply a proton-neutron bound state. In the case of the $X(3872)$ instead r_0 is measured to be negative and much larger than m_π^{-1} from eq. (4.30) and this leads to the conclusion that Z has to be significantly different from zero, pointing out to the existence of a compact component in the X. Indeed substituting the measured numbers from eq. (4.30) in eq. (4.43) one finds

$$Z \sim 0.14\,. \tag{4.44}$$

EFT with the tetraquark

In this section we will extend the results of sec. 4.2.1 by showing how to extend the non-relativistic EFT to include a compact tetraquark components. We will see that this EFT reproduces Weinberg's results [212].

To extend the theory we need to add a vector field **X** that will represent the elementary component of the $X(3872)$ and a cubic coupling between this object and the $C = +$ combination of the meson fields. We will assume the bare mass of the X to be tuned to be close to the $D^0\bar{D}^{*0}$ threshold. The cubic coupling clearly violates the conservation of the number of particles that is however a consequence of the non-relativistic limit. Such a violation is however allowed as long as the mass difference between these states is small, as it is in our hypotheses.

To be more specific, the non-relativistic limit of a QFT is taken by performing field redefinitions of the type

$$\phi(t,\mathbf{x}) = \frac{1}{\sqrt{2m}}\left(e^{-imt}\psi(t,\mathbf{x}) + e^{imt}\psi^\dagger(t,\mathbf{x})\right)\,. \tag{4.45}$$

In order to allow for the $X \leftrightarrow DD^*$ transitions, we can perform a redefinition like

$$\mathbf{X}_{\text{rel}} = \frac{1}{\sqrt{2(m_D + m_{D^*})}} \left(e^{-i(m_D+m_{D^*})t} \mathbf{X} + e^{i(m_D+m_{D^*})t} \mathbf{X}^\dagger \right) . \tag{4.46}$$

The difference between the mass of the $\mathbf{X}$ and the threshold $m_D + m_{D^*}$ then becomes an additional term in the kinetic part of the Hamiltonian that gives it a non-zero rest energy. The full Lagrangian of the theory then is

$$\begin{aligned}
\mathcal{L} =& D^\dagger \left(i\partial_t + \frac{\nabla^2}{2m_D} \right) D + \bar{D}^\dagger \left(i\partial_t + \frac{\nabla^2}{2m_D} \right) \bar{D} \\
&+ \mathbf{D}^\dagger \left(i\partial_t + \frac{\nabla^2}{2m_{D^*}} \right) \mathbf{D} + \bar{\mathbf{D}}^\dagger \left(i\partial_t + \frac{\nabla^2}{2m_{D^*}} \right) \bar{\mathbf{D}} \\
&+ \mathbf{X}^\dagger \left(i\partial_t + \frac{\nabla^2}{2(m_D + m_{D^*})} - \epsilon \right) \mathbf{X} \\
&+ \frac{\lambda}{2} \left(D\bar{\mathbf{D}} + \mathbf{D}\bar{D} \right)^\dagger \left(D\bar{\mathbf{D}} + \mathbf{D}\bar{D} \right) \\
&+ g \left(D\bar{\mathbf{D}} + \mathbf{D}\bar{D} \right)^\dagger \mathbf{X} + h.c. ,
\end{aligned} \tag{4.47}$$

where the parameter

$$\epsilon = m_X - m_{D^*} - m_D , \tag{4.48}$$

represents the "bare" distance from the threshold.

Also for this theory it is possible to compute the $D^0 \bar{D}^{*0}$ scattering amplitude non-perturbatively in both λ and g. The cubic interactions adds new diagrams where each of the λ vertices of Fig. 4.2 can be replaced by X propagators. An example of such diagrams can be seen in Fig. 4.3. The simplest way to see how to compute it is by solving the diagrammatic equation of Fig. 4.4. The solution is easy since all the loops factorize as simple products. The resummed renormalized amplitude obtained is

$$\mathcal{A}_s = -\frac{g^2 + (E-\epsilon)\lambda}{E - \epsilon - \frac{\mu_r}{2\pi}[g^2 + (E-\epsilon)\lambda]\sqrt{-2\mu_r E - i\varepsilon}} . \tag{4.49}$$

From this we can extract the scattering length and effective range by expanding at low momenta for $E = k^2/2\mu_r$ to find

$$a_s = \frac{\mu_r}{2\pi} \left(\lambda - \frac{g^2}{\epsilon} \right) , \tag{4.50}$$

$$r_0 = -\frac{2\pi g^2}{(g^2 - \epsilon\lambda)^2 \mu_r^2} = -\frac{g^2}{2\pi a_s^2 \epsilon^2} . \tag{4.51}$$

We can see that in the limit $g \to 0$ we recover the result of the molecule model and, in particular, $r_0 \to 0$. When $g \neq 0$ instead we find that r_0 is negative as expected from Weinberg's criterion.

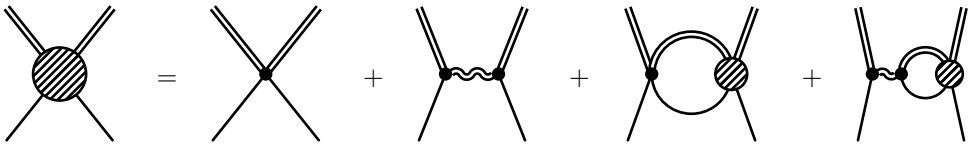


Figure 4.4: Diagrammatic equation for the full scattering amplitude (grey blob) of the NR-EFT containing the D (single line) and D^* (double line) mesons, and the compact tetraquark $\mathbf{X}$ (double wiggly line).

These results are the same as Weinberg's original paper. The matching is done more easily in the limit $\lambda = 0$. In this limit the amplitude looks like

$$\mathcal{A}_s = -\frac{g^2}{E - \epsilon_0 - g^2\left[-\frac{\mu_r}{\pi^2}\Lambda + \frac{\mu_r}{2\pi}\sqrt{-2\mu_r E - i\varepsilon}\right]}\,, \tag{4.52}$$

where we have reinstated the cutoff dependence in order to show more clearly the origin of the tuning and we have thus introduced the parameter ϵ_0 to distinguish it from the renormalized ϵ in the previous equations.

In Weinberg's language ϵ_0 is the energy of the elementary X state in the free theory. The amplitude has a pole at $E = -B$ that corresponds to the mass of the physical X state in the interacting theory and is obtained by the equation

$$-\epsilon \equiv -\epsilon_0 + g^2\frac{\mu_r}{\pi^2}\Lambda = B + g^2\frac{\mu_r}{2\pi}\sqrt{2\mu_r B}\,. \tag{4.53}$$

In this case we see that in order to have a state close to threshold, that is a small B, we need a tuning in the value of the bare ϵ_0, a parameter that is determined entirely by short distance interactions mediated by gluons.

Weinberg's effective coupling (4.41) can be obtained by taking the residue of the bound state pole in the scattering amplitude

$$\mathrm{Res}[\mathcal{A}_s, E = -B] \equiv -g_{\text{eff}}^2 = -\frac{g^2}{1 + g^2\mu_r^2/(2\pi\sqrt{2\mu B})}\,. \tag{4.54}$$

The proportionality constant between g^2 and g_{eff}^2 is the wavefunction renormalization constant Z between the elementary and physical X states

$$Zg^2 = g_{\text{eff}}^2\,. \tag{4.55}$$

Indeed eliminating g in favor of Z from (4.54) gives

$$g^2 = \frac{2\pi\sqrt{2\mu_r B}}{\mu_r^2}(1 - Z)\,, \tag{4.56}$$

exactly like in (4.42).

The final step to match with Weinberg is then to rewrite eq. (4.50) in terms of B and Z using (4.53) and (4.42). This gives

$$a_s = \frac{2(1-Z)}{2-Z}\frac{1}{\sqrt{2\mu_r B}} \tag{4.57}$$

$$r_0 = -\frac{Z}{1-Z}\frac{1}{\sqrt{2\mu_r B}} \tag{4.58}$$

exactly like eq. (4.43) and Weinberg's original result.

4.3 The role of the pion on the effective range

As we have already anticipated earlier, there is a second accidental tuning in the $D^0\bar{D}^{*0}$ system. The mass of the D^{*0} is indeed close to the D^0 π threshold as shown in eq. (4.10). This leads to an interaction whose range is longer than what naively expected and the pion should thus be included in the long distance EFT [213–216]. One might then wonder whether this longer range interaction might lead to violations of the arguments given before and could give a large correction to the effective radius. In this section we address this problem using a non-relativistic quantum mechanical treatment based on the distorted wave Born approximation. An expression for the effective range for the $X(3872)$, including the effect of pions, has already appeared in [216], where a non-relativistic effective theory for the D^{*0}, the D^0 and pion has been used. The approach we use here considerably simplifies the problem, avoiding the calculation of several loop Feynman diagrams. Despite the long range interactions, the pion is sufficiently weakly coupled not to generate a large effective radius for a $D^0\bar{D}^{*0}$ molecule. Moreover, it introduces a new feature in the effective range expansion, i.e. that the latter can be complex.

4.3.1 The $D^0\bar{D}^{*0}$ potential

The non-relativistic $D^0\bar{D}^{*0}$ potential is related to the Fourier transform of the tree level $D^0\bar{D}^{*0} \to D^0\bar{D}^{*0}$ scattering amplitude. At lowest order in the pion coupling, the diagrams contributing to the process are those in Figure 4.5. Because of the accidents described above, there is little phase space available to all the particles involved, which are then characterized by small velocities. As we have already seen, the problem can be treated in a non-relativistic effective field theory [213–216]. The theory including the pion is given by adding to eq. (4.20) the interaction terms

$$\mathcal{L}_{\text{pion}} = \left(\frac{g}{\sqrt{2}f_\pi}\right)\frac{1}{2m_\pi}\left(D\mathbf{D}^\dagger\cdot\nabla\pi + \bar{D}^\dagger\bar{\mathbf{D}}\cdot\nabla\pi^\dagger\right) + h.c.\,. \tag{4.59}$$

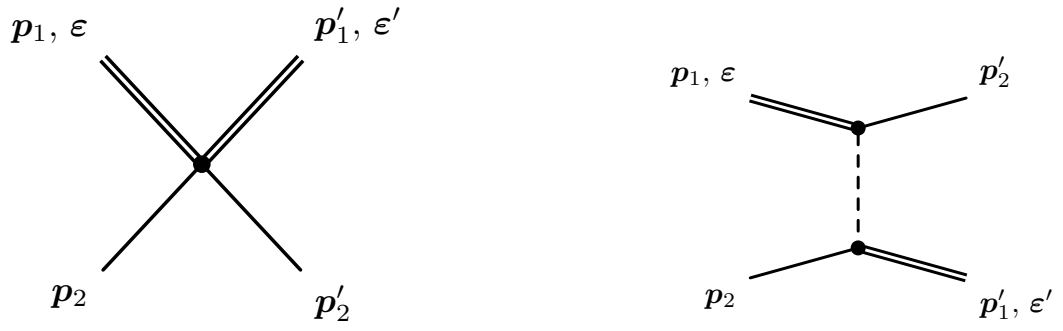


Figure 4.5: Contact (left) and pion-mediated (right) tree level $D^0\bar{D}^{*0}$ interactions. The single line represents a D^0, the double line a D^{*0} and the dashed one a neutral pion. Finally, $\boldsymbol{\varepsilon}$ and $\boldsymbol{\varepsilon}'$ are the polarization vectors of the incoming and outgoing D^{*0}.

At tree level we then find the following amplitudes

$$\mathcal{A}_a = -\lambda\,, \tag{4.60a}$$

$$\mathcal{A}_b = -\frac{g_\pi^2}{2f_\pi^2}\frac{\boldsymbol{\varepsilon}\cdot(\boldsymbol{p}_1-\boldsymbol{p}_2')\,\boldsymbol{\varepsilon}'\cdot(\boldsymbol{p}_1-\boldsymbol{p}_2')}{\mu^2-(\boldsymbol{p}_1-\boldsymbol{p}_2')^2}\,, \tag{4.60b}$$

where $\mu^2 \equiv 2m_\pi\delta$, with $\delta \equiv m_{D^*} - m_D - m_\pi$. As mentioned, $\mu^2 \ll m_\pi^2$. The coupling in the second line can be extracted from the known $D^{*\pm}$ decay width, and it is given by $g_\pi^2/(4m_\pi f_\pi^2) \simeq 3.7\times 10^{-8}$ MeV^{-3} [215]. Moreover, to keep our expressions simple, we are working at lowest order in $m_\pi \ll m_D$.[1] To be consistent with power counting, at this order one should also include the higher derivative correction to the contact $D^0\bar{D}^{*0}$ interaction as in eq. (4.27), as explained, for example, in [213, 214]. Nonetheless, we expect that such corrections give a contribution to the effective range of order m_π^{-1}. Here we are interested in tracking possible contributions substantially larger than that, and we will therefore ignore these corrections. All the short distance contributions are parametrized with the simplest contact interaction.

Denoting the initial positions and momenta of the D^{*0} and of the D^0 as $(\boldsymbol{x}_1,\boldsymbol{p}_1)$ and $(\boldsymbol{x}_2,\boldsymbol{p}_2)$ respectively, and the final ones as $(\boldsymbol{x}_1',\boldsymbol{p}_1')$ and $(\boldsymbol{x}_2',\boldsymbol{p}_2')$, the position space matrix element for the potential operator, $U \equiv \langle \boldsymbol{x}_1',\boldsymbol{x}_2'|V|\boldsymbol{x}_1,\boldsymbol{x}_2\rangle$, is obtained from the Fourier transform of the non-relativistic amplitude. Specifically, given the amplitude, one has

$$U = -\int \frac{d^3p_1}{(2\pi)^3}\frac{d^3p_2}{(2\pi)^3}\frac{d^3p_1'}{(2\pi)^3}\frac{d^3p_2'}{(2\pi)^3}\,\mathcal{M}\,e^{i\left(\boldsymbol{p}_1\cdot\boldsymbol{x}_1+\boldsymbol{p}_2\cdot\boldsymbol{x}_2-\boldsymbol{p}_1'\cdot\boldsymbol{x}_1'-\boldsymbol{p}_2'\cdot\boldsymbol{x}_2'\right)}\,(2\pi)^3\delta\Big(\boldsymbol{p}_1+\boldsymbol{p}_2+\boldsymbol{p}_1'+\boldsymbol{p}_2'\Big)\,. \tag{4.61}$$

[1] We checked that including corrections to this approximation [215] changes our results by a factor smaller than 10%.

For the contact interaction, $\mathcal{A}_a$, one gets

$$U_a = -\lambda\, \boldsymbol{\varepsilon}\cdot\boldsymbol{\varepsilon}'\delta^{(3)}(\boldsymbol{r})\delta^{(3)}\Big(\boldsymbol{x}_1-\boldsymbol{x}_1'\Big)\delta^{(3)}\Big(\boldsymbol{x}_2-\boldsymbol{x}_2'\Big)\,, \tag{4.62}$$

where $\boldsymbol{r}\equiv\boldsymbol{x}_1-\boldsymbol{x}_2$ is the relative distance, and we are working in a basis where the polarization vectors are constant. The meaning of the δ-functions is clear: the positions of the D^{*0} right before and right after the collision coincides, and the same is true for the D^0. Moreover, the interaction happens at short distances.

The Fourier transform of $\mathcal{A}_b$, instead, returns (see also [217])

$$\begin{aligned} U_b = {}& \frac{g_\pi^2}{8\pi f_\pi^2}\delta^{(3)}\Big(\boldsymbol{x}_1-\boldsymbol{x}_2'\Big)\delta^{(3)}\Big(\boldsymbol{x}_2-\boldsymbol{x}_1'\Big) \\ & \times\left\{\left[\boldsymbol{\varepsilon}\cdot\hat{\boldsymbol{r}}\,\boldsymbol{\varepsilon}'\cdot\hat{\boldsymbol{r}}\left(\frac{3}{r^3}+\frac{3i\mu}{r^2}-\frac{\mu^2}{r}\right)-\boldsymbol{\varepsilon}\cdot\boldsymbol{\varepsilon}'\left(\frac{1}{r^3}+\frac{i\mu}{r^2}\right)\right]e^{-i\mu r}+\frac{4\pi}{3}\delta^{(3)}(\boldsymbol{r})\right\} \qquad (4.63a) \\ \xrightarrow{\text{S-wave}} {}& -\frac{g^2}{8\pi f_\pi^2}\delta^{(3)}\Big(\boldsymbol{x}_1-\boldsymbol{x}_2'\Big)\delta^{(3)}\Big(\boldsymbol{x}_2-\boldsymbol{x}_1'\Big)\frac{1}{3}\boldsymbol{\varepsilon}\cdot\boldsymbol{\varepsilon}'\left(\mu^2\frac{e^{-i\mu r}}{r}+4\pi\delta^{(3)}(\boldsymbol{r})\right)\,, \qquad (4.63b) \end{aligned}$$

where in the last line we projected on S-wave as appropriate for a loosely bound molecule. The width of the D^{*0} appearing in μ clearly serves as an IR regulator.

Some comments regarding the above matrix element are now in order. First of all, we notice that the δ-functions dictate that the D^0 and D^{*0} exchange their positions right after the collision, in agreement with the intuitive picture arising from the diagram (b) in Figure 4.5. Secondly, this potential is complex and has infinite range. This reflects the fact that, since the D^{*0} decays into $D^0\pi^0$, the intermediate pion in this process can be real, and hence propagate at arbitrarily large distances. For this reason, as we show in the next section, the contribution to the effective range arising from these interactions is now complex as well.

4.3.2 The scattering amplitude and the effective range

We compute the $D^0\bar{D}^{*0}$ scattering amplitude at lowest order in the pion interaction using the so-called distorted wave Born approximation [218]. In order to do that, one separates the potential into a short and a long distance contribution, $V_{DD^*}=V_s+V_\ell$, where the latter is smaller than the former. In particular, from Eqs. (4.62) and (4.63), they are given by

$$V_s=-\left(\lambda+\frac{g_\pi^2}{6f_\pi^2}\right)\delta^{(3)}(\boldsymbol{r})\,\delta_{jj'}\,,\ V_\ell=-\alpha\frac{e^{-i\mu r}}{r}\delta_{jj'}\,, \tag{4.64}$$

where j and j' are the initial and final spins of the D^{*0}, and we defined $\alpha \equiv g_\pi^2 \mu^2/(24\pi f_\pi^2)$ for simplicity. The scattering amplitude can then be computed perturbatively in V_ℓ,

$$f_{DD^*} = f_s - \frac{\mu_r}{2\pi}\langle \psi_s^- | V_\ell | \psi_s^+ \rangle + \mathcal{O}\left(V_\ell^2\right) , \tag{4.65}$$

where μ_r is the $D^0\bar{D}^{*0}$ reduced mass, and f_s and $\psi_s^\pm$ are the amplitude and in/out states arising from the short distance potential alone. They can be computed exactly and, in S-wave, they are [219, 220]

$$f_s = \frac{1}{-1/a_s - ik} , \tag{4.66a}$$

$$\psi_s^\pm(r) = \frac{\sin(kr)}{kr} + \frac{1}{-1/a_s \mp ik}\frac{e^{\pm ikr}}{r} = e^{\pm i\delta_s}\frac{\sin(kr+\delta_s)}{kr} , \tag{4.66b}$$

where k is the relative $D^0\bar{D}^{*0}$ momentum. Morever, δ_s is the short distance scattering phase, related to the scattering length, a_s, by $k\cot\delta_s = -1/a_s$. The latter can be related to the couplings in Eq. (4.64). However, it needs to be renormalized due to UV divergences [219, 221] and, consequently, it is not computable within the long distance theory: it must be taken from data. These results of course agree with what we found in the previous section, with the difference that now we have included a short distance contribution coming from the pion together with the usual λ.

With this at hand, using the distorted Born approximation we find that the leading order correction to the scattering amplitude is

$$\langle \psi_s^- | V_\ell | \psi_s^+ \rangle = -\alpha\, e^{2i\delta_s} \int d^3r \frac{e^{-i\mu r}}{r}\frac{\sin^2(kr+\delta_s)}{(kr)^2} , \tag{4.67}$$

where the spin of the D^{*0} plays no role since it is trivially conserved before and after the collision. Once again, the amplitude above features a UV divergence, encoding the unknown short distance physics. We choose to regularize it using a cutoff at $r = \bar{r}$. At leading order in $\bar{r} \to 0$ one gets

$$\begin{aligned}\langle \psi_s^- | V_\ell | \psi_s^+ \rangle = \frac{\pi\alpha}{k^2}\Bigg\{ &-\left(1-e^{2i\delta_s}\right)^2 (i\pi + \gamma_E) - \log\left[\bar{r}(\mu+2k)\right] \\ &+ 2e^{2i\delta_s}\log(\bar{r}\mu) - e^{4i\delta_s}\log\left[\bar{r}(\mu-2k)\right]\Bigg\} + \mathcal{O}(\bar{r}) ,\end{aligned} \tag{4.68}$$

where γ_E is the Euler-Mascheroni constant. For small relative momenta, the inverse amplitude reads

$$f_{DD^*}^{-1} = -\frac{1}{a_R} - ik + \frac{2\alpha\mu_r}{\mu^2}\left(\frac{2}{a_R^2\mu^2} - 1 - \frac{8i}{3a_R\mu}\right)k^2 + \mathcal{O}\left(k^4\right) , \tag{4.69}$$

where the $\mathcal{O}\left(k^2\right)$ term determines the effective range and a_R is the renormalized scattering length, which again reabsorbs the UV divergence in a measurable quantity. The expression in Eq. (4.69) reproduces exactly the result obtained in [216], up to an unspecified constant, which the authors called F_2. Unfortunately, the origin of such a term is not clear in [216]. The scattering length can be determined from the distance of the $X(3872)$ from the $D^0\bar{D}^{*0}$ threshold as $a_R = 1/\sqrt{2\mu_r B}$. Considering the current experimental bound, for a molecular state one has 0 keV $\lesssim B \lesssim$ 120 keV, from which one obtains the following bounds on the real and imaginary parts of the effective range due to long range pion exchange:

$$-0.20 \text{ fm} \lesssim \text{Re}\, r_0 \lesssim -0.15 \text{ fm}\,, \tag{4.70a}$$

$$-0.19 \text{ fm} \lesssim \text{Im}\, r_0 \lesssim 0 \text{ fm}\,. \tag{4.70b}$$

As one can see, the pion is too weakly coupled to generate a large, negative effective range in a purely molecular scenario. Nonetheless, r_0 can now have an imaginary part, albeit small.

Note that this computation misses the contributions from the D^* self-energy diagrams from pion corrections. How to properly include such corrections in this formalism is still part of our ongoing work.

4.4 Partial conclusions and Outlook

We have seen that the two description of the $X(3872)$ as a DD^* mesonic molecule or as a compact tetraquark give very different predictions. There are many experimental hints that point to a compact component of the X. We have seen in particular that the effective range can be a direct measure of the mixing between the compact and molecular component of the system. In the case of the $X(3872)$, such parameter is measured to be large and negative. According to Weinberg's criterion, r_0 is expected to be $\mathcal{O}(m_\pi^{-1})$ for a pure bound state, while for a compact object close to a threshold it should be negative, with magnitude that can be larger than m_π^{-1}.

We have then seen that the interaction due to pion exchanges, that have a longer range than naively expected due to the mass tuning between the D^* and the $D\pi$ threshold, while giving a negative contribution to r_0, do not effect the validity of the argument and cannot generate corrections to the effective range larger than m_π^{-1}. The same arguments could be applied to other exotic states close to meson threshold. For example the recently discovered $T_{cc}^+(3875)$ was also measured to have a negative effective range (see Fig. 4.6).

Both the $X(3872)$ and the $T_{cc}^+(3875)$ are exciting examples of fine-tuning. Understanding their nature might tell us something about QCD and maybe teach us something about fine-tunings in general. From the theoretical side, a complete understanding of these

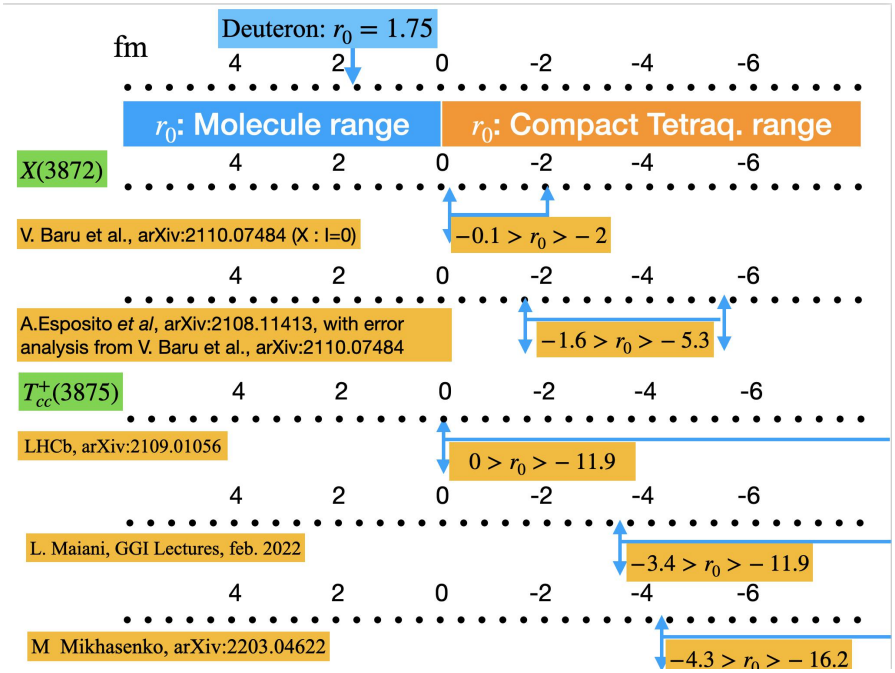


Figure 4.6: The value of the effective range for the $X(3872)$ and the $T_{cc}^+(3875)$ from the experimental lineshape from various studies. Plot taken from [186]

exotic resonances would require the understanding of what are the relevant parameters to describe them. As mentioned, it is unclear whether the lineshape currently employed by LHCb is the most general one, able to accommodate for the presence of compact states and for additional thresholds. It is thus important to construct a general Effective Theory that describes this states in order to identify which and how many parameters are needed to fully capture the nature of these states.

From the experimental side, the smoking gun that might confirm the compact tetraquark nature of this state would be the detection of its charged partners $X^{\pm}(3872)$ that are predicted from isospin symmetry. The approximate $SU(3)_f$ flavor symmetry among the three lighter quarks indeed leads to a prediction of the existence of flavor multiplets for the tetraquarks exactly like what happens for the conventional hadrons [222]. This pattern is not expected in molecular models. The color singlet mesons would be bound together by forces mediated by the exchange of lighter color singlets. Due to the difference in range between the exchange of pions and heavier mesons, different isospin/strangeness combinations might not bind at all. Some of the exotic resonances that have been observed seem indeed to fill in a flavor multiplet. For example in the $J^{PC} = 1^{++}$ family there have been observed two other resonances: the $X(4140)$, containing also a $s\bar{s}$ pair and the $Z_{cs}(4003)$ whose mass is at the exact midpoint between the $X(4140)$ and the $X(3872)$ as predicted by the quark model. If the charged partners of the $X(3872)$ were observed this would confirm the validity of the tetraquark picture of these resonances.

4.A The $B \to K J/\psi\rho$ amplitude

We now briefly show that, up to some overall short distance couplings, near the $X(3872)$ pole the $B \to K J/\psi\rho$ amplitude coincides with the $D^0\bar{D}^{*0} \to D^0\bar{D}^{*0}$ one. The Hilbert space is schematically made of the $|B\rangle$, $|K\rangle$, $|DD^*\rangle$ and $|J/\psi\rho\rangle$ states. The relevant Hamiltonian can be written as

$$H = H_0 + V_{DD^*} + V_{BK} + V_{J/\psi\rho}\,, \tag{4.71}$$

where the only non-vanishing matrix elements are given by $\langle DD^*|V_{DD^*}|DD^*\rangle$, $\langle DD^*K|V_{BK}|B\rangle$ and $\langle J/\psi\rho|V_{J/\psi\rho}|DD^*\rangle$. Using the Lipmann-Schwinger equation, the scattering matrix can be computed perturbatively in V_{BK} and $V_{J/\psi\rho}$,

$$\begin{aligned} T &= V + V(E - H + i\epsilon)^{-1}V \\ &\supset V_{J/\psi\rho}(E - H_{DD^*} + i\epsilon)^{-1}V_{BK}\,, \end{aligned} \tag{4.72}$$

with $H_{DD^*} \equiv H_0 + V_{DD^*}$. It then follows that

$$\begin{aligned} T(B \to K J/\psi\rho) = T(D^0\bar{D}^{*0} \to D^0\bar{D}^{*0}) \\ \times \langle J/\psi\rho|V_{J/\psi\rho}V_{DD^*}^{-1}|DD^*\rangle \\ \times \langle DD^*K|V_{DD^*}^{-1}V_{BK}|B\rangle\,, \end{aligned} \tag{4.73}$$

where we used the fact that, near the pole of the $X(3872)$, one has $T(D^0\bar{D}^{*0} \to D^0\bar{D}^{*0}) \simeq \langle DD^*|V_{DD^*}(E - H_{DD^*} + i\epsilon)^{-1}V_{DD^*}|DD^*\rangle$. The matrix elements $\langle J/\psi\rho|V_{J/\psi\rho}V_{DD^*}^{-1}|DD^*\rangle$ and $\langle DD^*K|V_{DD^*}^{-1}V_{BK}|B\rangle$ are dominated by short distance contributions and, therefore, the two amplitudes in Eq. (4.73) are related by a simple constant rescaling.

5 EW Baryogenesis and Symmetry non-restoration

The existence of a non-vanishing baryon number density in the universe is a fact of life and a mystery of physics. According to our present understanding of the early universe, such density cannot be simply accounted for by initial conditions, as any original density would have been completely diluted by inflation. Dynamics during the Big Bang era must be responsible for the observed ratio of baryon number over entropy $\eta_b \equiv n_b/s \simeq 10^{-10}$. The conditions such dynamics should satisfy were famously spelled out by Sakharov a long time ago. The first condition, concerning the existence of baryon number violating interactions, is satisfied by the Standard Model (SM). This happens in a remarkable way: on the one hand, at low temperature, baryon number emerges as an accidental symmetry whose violation is negligibly small; on the other hand, at temperatures above the weak scale, fast baryon number violation occurs through sphaleron processes. The other two conditions, concerning CP-violation and departure from thermodynamic equilibrium, are however not satisfied by the SM. The reasons for that are more quantitative than structural. Indeed the SM is endowed with CP violation in the CKM matrix, but that turns out to be insufficient, in view of the small value of quark masses and mixings. The SM, given electroweak symmetry breaking, could also, in principle, experience an epoch of departure from equilibrium when transiting from a high temperature symmetric phase to a low temperature non-symmetric phase. However, given its parameters, in particular given the Higgs mass, such transition is known to be a smooth second order crossover. Baryogenesis thus requires new physics.

A potentially testable option is having baryogenesis occurring at the electroweak phase transition [223][224][225]. A realistic model of electroweak (EW) scale baryogenesis should involve new states so as to provide new sources of CP violation and a first order electroweak phase transition. The new states should range from ~ 100 GeV to several hundreds GeV, with the upper range attainable only in the presence of somewhat strong couplings (see for instance [226]). While this is interesting in view of a full test of this scenario, it also implies, given the lack of clear evidence so far of new physics in direct and indirect searches, that the models are already under pressure. In particular one

major source of pressure comes from the need for new CP-violating phases: on the one hand these are directly constrained by the searches for EDMs of elementary particles, on the other their presence is typically associated to new sources of flavor violation, with the corresponding strong bounds. To be more quantitative, a new physics sector characterized by a mass scale Λ and CP-odd phases $\phi_{\rm CP}$ is expected to generate an electron EDM either at 1-loop (when direct couplings of order g to the SM fermions are present), or 2-loop (when mainly couplings g to the bosonic sector of the SM exist). Taking g of order the EW gauge coupling for illustration, in these two cases we estimate:

$$\frac{|d_e|}{e} = \begin{cases} \sin\phi_{\rm CP}\frac{g^2}{16\pi^2}\frac{m_e}{\Lambda^2} \sim 10^{-29}\text{cm} \times \sin\phi_{\rm CP}\left(\frac{50\ \text{TeV}}{\Lambda}\right)^2 & 1-\text{loop} \\ \sin\phi_{\rm CP}\left(\frac{g^2}{16\pi^2}\right)^2\frac{m_e}{\Lambda^2} \sim 10^{-29}\text{cm} \times \sin\phi_{\rm CP}\left(\frac{2.5\ \text{TeV}}{\Lambda}\right)^2 & 2-\text{loop (Zee}-\text{Barr)} \end{cases} \tag{5.1}$$

The recent improved bound on the electron EDM $|d_e| < 1.1 \times 10^{-29} e$ cm [227] then suggests that Λ should be well above the weak scale if $\sin\phi_{\rm CP} \sim 1$. Scenarios with suppressed CP-phases may allow the new physics to stay in the hundreds of GeV, but in those cases achieving a sizable baryon asymmetry is rather difficult (see, e.g., [228]). Even ignoring possible associated flavor-changing effects we must conclude that realizing conventional scenarios of EW baryogenesis is at best extremely challenging. Clever flavor model-building provides one way to limit this pressure, nonchalance provides another. The goal of this chapter is to study an alternative scenario, where all constraints from flavor and CP violation are structurally eliminated.

Our scenario is described as follows. Up to some high scale $\Lambda \gg m_W$, say $\Lambda = 100 - 1000$ TeV, the only sources of flavor and CP violation are the SM Yukawa couplings. One should think of Λ as the scale of flavor, at which new sizable sources of CP and flavor violation become active, without significantly affecting low energy observables. In view of that, Λ seems the natural scale where to realize baryogenesis. Our idea is then simply that the electroweak phase transition happens at $T \sim \Lambda$ rather than at $T \sim m_W$ and that moreover such a transition is first order with consequent departure from thermodynamic equilibrium and generation of a baryon asymmetry. In order for such asymmetry to be kept unsuppressed until the present day, the EW symmetry must remain broken at all temperatures below $\sim \Lambda$, so as to avoid wash-out from sphaleron processes. In the SM, thermal effects, dominated by the top quark, are known to restore the EW symmetry at a temperature $T_c \simeq 160$ GeV (see e.g. [229] for a recent calculation). In order to realize this scenario there should therefore exist new degrees of freedom in the 100 GeV range and coupled to the Higgs so as to prevent electroweak symmetry restoration at temperatures above the weak scale and below Λ. The simplest option for such states is given by a set of N SM neutral scalar fields S_i bilinearly coupled to the Higgs. As we shall discuss in detail, there exists a minimal value of N where the Higgs vacuum expectation value (VEV) at finite temperature is big enough to suppress sphaleron processes at $T < \Lambda$ while preserving perturbativity. For reasons that emerge combining the structural consistency of the model with the significant impact of the top Yukawa on the Higgs potential, this

minimal N turns out to be quite large, $N \gtrsim 100$. Such large numbers, however, do not imply complexity in the structure of the model: the scalars S_i could fit in a single multiplet of a global or local symmetry and the Lagrangian could be described by a handful of couplings. Indeed we shall also make a simple remark showing there exists a scaling region for the couplings and for N, where finite temperature effects are large while zero temperature ones, like collider production rates or corrections to electroweak precision quantities, decouple with inverse powers of N.

The main goal of this chapter is to study in detail the dynamics at temperatures below Λ. The S_i will have a mass in the 100 GeV range, so we must ensure they serve our purpose (symmetry non-restoration) compatibly with all theoretical and phenomenological constraints. In particular one issue concerns their problematic relic density, which we shall eliminate by introducing additional interactions which can in principle be rather weak.

For what concerns the EW phase transition dynamics at the scale Λ, there is great freedom as there are no significant constraints on model-building at scales $\Lambda \sim 100 - 1000$ TeV from low energy phenomenology. That was indeed the goal of the whole construction. In view of that we shall limit ourselves to a simple sketch of possible scenarios at the scale Λ.

This chapter is organized as follows. In Section 5.1.1 we identify a sufficient (and largely model-independent) condition to ensure the primordial $B + L$ asymmetry is not washed-out by sphaleron processes at $T < \Lambda$. Our $O(N)$ model is introduced in Section 5.1.2, where its domain of validity as a perturbative effective field theory (EFT) is established. Especially important for the reminder of the chapter is how the perturbativity and stability conditions scale with $N \gg 1$. We will see that large-N sectors have the ability to qualitatively impact the finite-temperature dynamics while remaining essentially invisible around the vacuum. In Section 5.1.3 we take a first look at the finite T behavior of our model and roughly assess how large N needs to be in order to ensure EW symmetry non-restoration. The need for a refined analysis is emphasized in Section 5.1.4. Here we carefully assess the perturbative regime of the finite T version of our scenario and discuss some computational subtleties that characterize our class of large-N models. A detailed discussion of the effective potential at finite T is presented in Section 5.1.5. A large-N technique is implemented to reliably take into account the most important contributions. A numerical analysis is then used to identify the allowed parameter space of the model. The main phenomenological constraints on the minimal $O(N)$ scenario are presented in Section 5.2. All direct and indirect signatures of S_i at colliders are suppressed by powers of N and may be hard to see. Simultaneously, the minimal model predicts N-enhanced signals at direct dark matter experiments and is therefore incompatible with current data. We thus discuss various possible extensions that eliminate the problem. A subset of these extensions also features a realistic dark matter candidate. In particular in subsection 5.2.3 we focus on a scenario where the $O(N)$ symmetry is partially gauged and where dark

matter can plausibly be made of bound states. This is a novel dark matter scenario that crucially combines compositeness and large-N. In view of its intricate phenomenology, subsection 5.2.3 somewhat blew out of proportions: in spite of its interest this section may be skipped in a first reading. In Section 5.3 we sketch possible variants of our scenario. These include models with different symmetry structure but a comparable or even larger number of degrees of freedom than in the model discussed in detail here. However we also point out that a large N is not strictly necessary to our program if we introduce additional EW-charged scalars. In Section 5.4 we show that there exists no structural obstruction to the construction of complete models for EW baryogenesis. We illustrate this by briefly discussing two possible scenarios of new physics at the flavor scale Λ, one strongly- and one weakly-coupled. We present our conclusions in Section 5.5.

5.1 The low energy sector

In this section, assuming baryogenesis takes place via a first order electroweak phase transition at $T \sim \Lambda \gg m_W$ (see Section 5.4 for some models for the phase transition), we shall study the conditions on the effective theory at $E \ll \Lambda$ for the baryon asymmetry to survive. The crucial request concerns the suppression of sphaleron wash-out. After quantifying this request in the next subsection, we will introduce a specific model and study it in detail.

5.1.1 Avoiding sphaleron washout

Our basic assumption is that the baryon asymmetry is generated by the anomalous electroweak baryon number violation during the electroweak phase transition at $T \sim \Lambda$. More precisely, this implies the asymmetry arises along the direction $B + L$, which is anomalously broken by the electroweak interactions, while it vanishes in the exactly conserved direction $B - L$. Because of that we must ensure that the anomalous $B + L$-violating processes, the so-called sphalerons, remain inactive at $T < \Lambda$, as they would otherwise wash-out this very same combination. [1]

The rate of variation of the baryon number density normalized to the entropy density $\Gamma_b(T) \equiv -\frac{d}{dt} \ln \eta_b$ crucially depends on the weak coupling $\alpha_w = g^2/4\pi$, on the temperature T, and on the temperature-dependent Higgs vacuum expectation value (VEV) $h(T)$. [2] One can distinguish three qualitatively different regimes, which respectively correspond

[1] This issue of course does not arise in other scenarios of high scale baryogenesis, such as leptogenesis, where the asymmetry is generated along the non-anomalous $B - L$ direction.

[2] Throughout the chapter we adopt the convention $h(0) \equiv v = 246$ GeV, for which $m_W^2 = g^2 h^2/4$ and $m_t^2 = y_t^2 h^2/2$.

to high, intermediate and low temperature:

$$\begin{aligned} \text{high T}: \quad & T \gtrsim \frac{gh(T)}{\alpha_w} = \frac{4\pi}{g} h(T) \\ \text{inter. T}: \quad & \frac{g}{4\pi} h(T) \lesssim T \lesssim \frac{4\pi}{g} h(T) \\ \text{low T}: \quad & T \lesssim \frac{g}{4\pi} h(T). \end{aligned} \tag{5.2}$$

These ranges are essentially determined by the T-dependent vector boson mass $m_W \sim gh(T)$, and by the energy of the sphaleron configuration

$$E_{\rm sph}(T) \equiv B \frac{gh(T)}{\alpha_w}, \tag{5.3}$$

where $B = B(\lambda_H/g^2, g'^2/g^2)$ is a number of order unity. According to the zero-temperature analysis of refs. [230][231][232], $B \sim 1.9$.

In the high temperature regime the EW-symmetry is effectively unbroken: the sphaleron barrier is overcome by thermal fluctuations, the rate is $\Gamma_{\rm b} \sim 20\alpha_w^5 T \gg H$ [233][229] and any asymmetry would be washed-out in a fraction of a Hubble time. On the other hand, the low temperature regime approaches the zero temperature result, where $B+L$-violating effects are $\propto e^{-2\pi/\alpha_w}$ and can be safely neglected under all relevant circumstances. Finally, in the intermediate regime one finds

$$\Gamma_b(T) \sim 0.1\, T \left(\frac{gh}{T}\right)^7 \left(\frac{4\pi}{\alpha_w}\right)^3 e^{-E_{\rm sph}/T}, \tag{5.4}$$

where to estimate the prefactor we used the results of [234] with $\kappa \sim 1$ and $\omega_- \sim gh$. The reliability of this estimate of the actual sphaleron rate can be assessed observing that (5.4) fits the numerical results of [229] quite well provided $B \to B_{\rm fit} \simeq 2.2 - 2.3$. [3] We thus see that finite T effects modify the rate (5.4) appreciably. In what follows we will still use (5.4), but will adopt the conservative value $B = 1.8 - 2.1$.

As we will see shortly, in order to preserve the baryon asymmetry generated at the phase transition, it is sufficient to ensure that, when the universe was in the range of temperatures $m_W \lesssim T \lesssim \Lambda$, the rate in eq.(5.4) was small enough. One can easily see that the constraint on $h(T)/T$ is strongest at the latest times, that is the lowest T's in the range of interest. Indeed the preservation of the baryon asymmetry essentially gives the constraint $\int \Gamma_b dt \lesssim 1$, which by using Hubble law can be conveniently written as the dT integral

$$\int \Gamma_b(T) \frac{M_P}{\sqrt{g_*}} \frac{dT}{T^3} = \int F(gh(T)/T) \frac{M_P}{\sqrt{g_*}T} \frac{dT}{T} \lesssim 1 \tag{5.5}$$

where $F(x) \propto x^7 e^{-Bx/\alpha_w}$ — implicitly defined by eqs.(5.4) and (5.5) — goes monotonically

[3] We thank M. Shaposhnikov for suggesting this fit.

to 0 as $x \to 0$. The strongest constraint on $h(T)/T$ is dominated by the region where T is of order the weak scale and reads

$$\frac{h(T)}{T} \gtrsim 1.2 - 1.5, \tag{5.6}$$

with the upper (lower) value corresponding to $B \simeq 1.8\,(2.1)$. As promised at the beginning of the paragraph, it is now apparent that requiring (5.6) for all $T < \Lambda$ automatically prevents us from entering the high T regime, where washout is unsuppressed. We therefore conclude that eq. (5.6) represents a sufficient condition to avoid washout. Note also that in the models we shall study $h(T)/T$ is minimized at around the weak scale, and therefore the argument leading to the numerical values shown in eq. (5.6) is justified.

5.1.2 A model

In order to guarantee the condition on $h(T)/T$ discussed in the previous section we must suitably modify the Higgs dynamics in the energy range between m_W and Λ. A simple model realizing that is constructed by adding to the SM a multiplet of scalars S_i transforming as the fundamental representation of a novel $SO(N)$ symmetry. Other options, based on different symmetries (possibly discrete) and multiplet content will be mentioned in Section 5.3. Besides the gauge and Yukawa couplings of the SM, the resulting model is defined by the scalar potential

$$V = m_H^2 H^\dagger H + \lambda_H (H^\dagger H)^2 + \frac{m_S^2}{2} S^2 + \frac{\lambda_S}{4} (S^2)^2 + \lambda_{HS} S^2 H^\dagger H, \tag{5.7}$$

which is the most general one based on symmetries and renormalizability. Here the SM and $O(N)$ indices have been suppressed for brevity. We will assume

$$m_H^2 < 0 \qquad m_S^2 > \frac{\lambda_{HS}}{\lambda_H} m_H^2 \tag{5.8}$$

to ensure that in vacuum the Higgs has a non-vanishing expectation value whereas $\langle S \rangle = 0$.

The basic idea, which is not new [4] but we believe we shall explore from a novel perspective, is this: a negative off-diagonal quartic, $\lambda_{HS} < 0$, provides H with a negative Debye mass at finite temperature, so that $\langle H \rangle \neq 0$ even at $T \gg m_W$. Of course this makes sense as long as the potential is bounded from below and as long as all the couplings remain

[4] Symmetry non-restoration was first realized within a 2-scalar model (essentially the small N version of the one we adopt below) by S. Weinberg [235]. The same picture was subsequently applied in different contexts by other authors. Ref. [236][237] implemented it in models with spontaneous CP violation; ref. [238] proposed that non-restoration might be used to avoid the monopole problem of GUTs, and the same application has been suggested in [239]. Spontaneous B violation at finite T was used in a model for baryogenesis via decays in [240], and a period of temporary color $SU(3)$ breaking has been employed in [241] in a model for EW baryogenesis. More recently, the idea of EW symmetry non-restoration has been put forward in [242] and applied in the same context as ours in [243].

perturbative in a range of energies above the weak scale.

At tree level, the conditions for a potential bounded from below are easily found to be

$$\lambda_H,\ \lambda_S > 0 \qquad \lambda_{HS} > -\sqrt{\lambda_H \lambda_S} \tag{5.9}$$

showing there exists a window $-\sqrt{\lambda_H\lambda_S} < \lambda_{HS} < 0$ where a negative λ_{HS} can in principle achieve our goal of symmetry non-restoration. In order to ensure the absence of instabilities within the EFT description below the scale Λ, the above conditions will have to be satisfied by the running couplings at all RG scales $\mu \lesssim \Lambda$.

The conditions ensuring our construction makes sense as a weakly coupled EFT can be established by a simple diagrammatic analysis to be

$$\begin{aligned} \epsilon_H &\equiv \frac{6\lambda_H}{16\pi^2} \ll 1, \\ \epsilon_S &\equiv \frac{\lambda_S N}{16\pi^2} \ll 1, \\ |\epsilon_{HS}| &\equiv \frac{2|\lambda_{HS}|\sqrt{N}}{16\pi^2} \ll 1, \end{aligned} \tag{5.10}$$

where we have allowed for the possibility that $N \gg 1$ be counted in. One direct way to derive these constraints is to consider s-wave $2 \to 2$ scattering in singlet states

$$|HH\rangle \equiv \frac{1}{\sqrt{4}}\left(|H_1H_1\rangle + \cdots + |H_4H_4\rangle\right) \qquad |SS\rangle \equiv \frac{1}{\sqrt{N}}\left(|S_1S_1\rangle + \cdots + |S_NS_N\rangle\right). \tag{5.11}$$

In the 2-dimensional Hilbert subspace with basis given by the $\ell = 0$ singlet states $\{|HH\rangle, |SS\rangle\}$ the S-matrix is

$$S \equiv e^{2i\delta} = \mathbb{1} + \frac{i}{8\pi}\mathcal{T} + \ldots \tag{5.12}$$

with

$$\mathcal{T} = \begin{pmatrix} 6\lambda_H & 2\lambda_{HS}\sqrt{N} \\ 2\lambda_{HS}\sqrt{N} & \lambda_S(N+2) \end{pmatrix} \tag{5.13}$$

and with the dots corresponding to terms beyond the tree approximation. By the above equation we deduce

$$\delta = \frac{1}{16\pi}\mathcal{T} + \ldots \tag{5.14}$$

and requiring, in the spirit of naive dimensional analysis (NDA), that all the eigenvalues of δ be $\lesssim \pi$, we derive constraints that are parametrically equivalent to eq. (5.10). The rule according to which $\sqrt{N}$ appears in eq. (5.13) is easily established: any SS singlet, either initial or final, counts $O(\sqrt{N})$, while HH counts $O(1)$. One is easily convinced that scattering in non-singlet channels gives weaker bounds, as the amplitudes are not enhanced by powers of N. Alternatively, the same constraints (5.10) can be derived by

considering the β-functions for $(\lambda_H, \lambda_S, \lambda_{HS})$, see (5.80) for the 1-loop approximation, and requiring that the relative change of any of them is less than $O(1)$ over one e-folding of RG evolution. Notice that the window of negative λ_{HS} allowed by stability, eq. (5.9), is described by $\epsilon_{HS}^2 \leq (2/3)\epsilon_H\epsilon_S$. Therefore, perturbativity of the diagonal quartics $\lambda_{H,S}$ plus stability guarantees perturbativity of the off-diagonal quartic λ_{HS}.

With eq. (5.10) satisfied, our model is a well defined weakly coupled EFT. Yet, remarkably, in the large-N limit there remains one class of calculable quantum effects that is not suppressed by the ϵ's. These have to do with the renormalization of the mass of H induced by the S-loop tadpole diagram in the left panel of Fig. 5.1. For instance, in vacuum this diagram gives a contribution to the RG evolution of m_H^2:

$$\mu\frac{d}{d\mu}m_H^2 = \frac{\lambda_{HS}N}{8\pi^2}m_S^2 \equiv \epsilon_{HS}\sqrt{N}m_S^2, \tag{5.15}$$

implying that for $\sqrt{N} \gtrsim 1/|\epsilon_{HS}|$ there remains a finite effect, for arbitrarily small $|\epsilon_{HS}|$. When compactifying some spacetime directions the same diagram contributes finite mass corrections of Casimir type. In particular, by considering the system at finite temperature and working with compactified euclidean time, this diagram provides a negative contribution to the thermal mass [5]

$$\delta m_H^2 = \frac{\lambda_{HS}N}{12}T^2 = \frac{2\pi^2}{3}\epsilon_{HS}\sqrt{N}T^2 < 0. \tag{5.16}$$

This result implies a very interesting property of this simple system at large-N: in the scaling limit $\epsilon_{H,S,HS} \to 0$, $N \to \infty$ with $\epsilon_{HS}\sqrt{N}$ fixed, the theory is free, in particular the S-matrix is trivial, yet at finite temperature a finite contribution to the H thermal mass survives. For instance, in the case $m_H^2 > 0$, where the symmetry is unbroken in the vacuum, an infinitesimally small $\epsilon_{HS} < 0$ at sufficiently large N can still generate a finite negative thermal mass and thus trigger symmetry breaking at finite temperature. At large-N we can thus have a free theory, which when heated up undergoes a phase transition. While this may seem paradoxical, one should consider that in order to heat up a system of $N \to \infty$ fields one must provide an energy density that diverges in the same way: the phase transition cannot occur at finite energy density.

The crucial implication of eq. (5.16), is that for sufficiently large N, eq. (5.6) can be ensured at all $T < \Lambda$ with ϵ_{HS} and ϵ_S small enough not to significantly perturb the SM dynamics and compatibly with stability, $\epsilon_{HS}^2 < (2/3)\epsilon_H\epsilon_S$. The main question concerns the minimal value of N for which all constraints are met. As we shall discuss in detail in the next subsections, for a combinations of factors, the needed value of N is quite large.

Before proceeding we would however like to make a few comments concerning naturalness.

[5] Here we consider $m_S \ll T$ and neglected refinements needed when S itself acquires a sizable thermal mass and which we shall discuss in detail later on.

Figure 5.1: LEFT: leading perturbative contribution to the Higgs thermal mass. The solid (dashed) lines correspond to S (H). RIGHT: a sample of diagrams re-summed by the large-N technique adopted in Section 5.1.5.

In our construction there is clearly no protection mechanism for the hierarchical separation between m_H, m_S and Λ. We do not want to try and further complicate the model to make this separation natural. Indeed, in the case of conventional approaches to the hierarchy, i.e. supersymmetry or compositeness, that would almost unavoidably force us to have to deal with the problem of new sources of flavor violation at the weak scale, which is what we wanted to avoid in the first place. One perhaps tenable perspective is that this separation of scales has an anthropic origin. The Higgs mass m_H could be pegged to its value by the atomic principle [244], that is by the necessity for m_W to be not too much above $\Lambda_{\rm QCD}$ in order to ensure nuclear and chemical complexity. The S mass m_S, on the other hand, could be pegged to the weak scale by the request of a sufficiently baryon-rich universe: $m_S > m_W$ would cause electroweak symmetry restoration at $m_W \lesssim T \lesssim m_S$ with corresponding sphaleron washout of the baryon asymmetry created at the scale Λ. Of course, we are aware of the weakness of such anthropic reasoning: once one invokes a multiverse of options it is well possible the baryon asymmetry can be generated by some other mechanism, for instance by leptogenesis. Still our simple argument does not seem fully implausible to us. [6]

On the quantitative side, notice that, given the physical cut-off Λ, we expect a finite correction to m_H^2

$$\delta m_H^2 \sim \frac{\lambda_{HS} N}{8\pi^2}\Lambda^2 \, = \epsilon_{HS}\sqrt{N}\Lambda^2. \tag{5.17}$$

Therefore the $\sqrt{N}$ enhancement of the S-tadpole, which is essential for the vacuum dynamics at finite T, implies a corresponding increase in the needed tuning. In practice we shall need $|\lambda_{HS}|N \sim 8$, so that the needed tuning is only a factor of a few worse than the one due to the top quark contribution.

5.1.3 Thermal vacuum dynamics in first approximation

Here we would like to give a first rough idea of the range of parameters where the model is compatible with our scenario for baryogenesis. In order to do so we approximate the

[6] In [245][246] anthropic arguments were similarly invoked to motivate the existence of a weak scale sector with all the ingredients necessary to realize EW baryogenesis.

thermal effective potential by the tree level potential with quartic couplings renormalized at a scale $\sim T$ and supplemented by the leading 1-loop thermal corrections to the scalar masses. In the next two sections we will perform a more accurate computation.

Now, according to the above approximation, the thermal potential is given by eq. (5.7) with the scalar masses replaced by

$$\begin{aligned} m_H^2 &\to m_H^2(T) = m_H^2 + \left[\frac{N}{12}\lambda_{HS} + \frac{1}{2}\lambda_H + \frac{3}{16}g^2 + \frac{1}{16}g'^2 + \frac{1}{4}y_t^2\right] T^2 \\ m_S^2 &\to m_S^2(T) = m_S^2 + \left[\frac{(2+N)}{12}\lambda_S + \frac{1}{3}\lambda_{HS}\right] T^2, \end{aligned} \tag{5.18}$$

where all couplings are understood to be evaluated at an RG scale $\sim T$. In the above equation we have included the thermal corrections induced by S as well as those by the leading SM couplings. Notice that the negative λ_{HS} contribution to $m_S^2(T)$ is not N-enhanced, so that in the scaling limit $N \to \infty$ with $\lambda_S N$ and $\lambda_{HS} N$ fixed it can be neglected. Moreover, it is easy to see that (5.8) is satisfied by the thermal masses in eq. (5.18), implying that S does not acquire a VEV. On the other hand, we have $\langle H \rangle = (0, h/\sqrt{2})^t$, with

$$\begin{aligned} \frac{h^2(T)}{T^2} &= -\frac{m_H^2(T)}{T^2\lambda_H} \simeq -\left[\frac{N}{12}\lambda_{HS} + \frac{1}{2}\lambda_H + \frac{3}{16}g^2 + \frac{1}{16}g'^2 + \frac{1}{4}y_t^2\right]\frac{1}{\lambda_H} \\ &\equiv \frac{\frac{N}{12}|\lambda_{HS}(T)| - A(T)}{\lambda_H(T)} \end{aligned} \tag{5.19}$$

where we have grouped all the SM contributions into the quantity $A(T) > 0$. We conservatively neglected the zero temperature mass; being negative, it helps making h/T bigger but also becomes quickly negligible as soon at $T \gtrsim 100$ GeV.

Now, given eq. (5.19), the requirement $h(T)/T \gtrsim 1.2$ in eq. (5.6) translates into a lower bound $|\lambda_{HS}(T)|N \gtrsim 12[A(T) + (1.2)^2\lambda_H(T)]$, which we shall impose for all $T < \Lambda$. This bound is strongest at $T \sim 100 - 1000$ GeV where it reads $|\lambda_{HS}|N \gtrsim 8$, since the SM couplings are larger there. Moreover, we must simultaneously impose the stability requirement $\lambda_{HS}^2(\mu) \leq \lambda_H(\mu)\lambda_S(\mu)$ at any scale $\mu < \Lambda$, to ensure our effective field theory makes sense. Combining these two constraints we obtain

$$\begin{aligned} N &\geq \frac{[N\lambda_{HS}(\mu)]^2}{\lambda_H(\mu)[N\lambda_S(\mu)]} = \frac{[N\lambda_{HS}(T)]^2}{\lambda_H(\mu)[N\lambda_S(\mu)]}\left(\frac{\lambda_{HS}(\mu)}{\lambda_{HS}(T)}\right)^2 \\ &\gtrsim \frac{[12]^2\left[A(T) + (1.2)^2\lambda_H(T)\right]^2}{\lambda_H(\mu)[16\pi^2\epsilon_S(\mu)]}\left(\frac{\lambda_{HS}(\mu)}{\lambda_{HS}(T)}\right)^2 \\ &\sim 800\left(\frac{0.01}{\epsilon_S(\mu)}\right)\frac{0.04}{\lambda_H(\mu)}\left(\frac{\lambda_{HS}(\mu)}{\lambda_{HS}(\mathrm{TeV})}\right)^2 \gtrsim 800\left(\frac{0.01}{\epsilon_S(\Lambda)}\right). \end{aligned} \tag{5.20}$$

The right hand side in the second line of Eq. (5.20) decreases with T as the SM couplings and $|\lambda_{HS}|$ respectively decrease and increase with T. Moreover that same expression

increases with μ, with the quantitatively most important effect being associated with the decrease in λ_H by a factor of about $3-4$ when running between the weak scale and $100-1000$ TeV. The final bound on N is therefore maximized at $T \sim$ TeV and for the largest $\mu \sim \Lambda$, which in our picture could be 100 to 1000 TeV. We conclude that unless $\epsilon_S(\Lambda)$ is rather strong, N must be in the hundreds. Notice also that the limiting values $N \sim 800$ and $|\lambda_{HS}|N \sim 8$ consistently correspond to a rather weak coupling $|\epsilon_{HS}| \sim 0.004$.

There are two ways in which the estimate (5.20) is inaccurate. The first is that in the region where $h/T \sim 1$ the thermal masses of the Higgs, $W^{\pm}$, Z, and top-quark are not negligible compared to the temperature, in which case the thermal loop correction to the Higgs potential starts being affected by Boltzmann suppressions causing a deviation from the simple expressions used in eq. (5.18). This effect tends to reduce $A(T)$ by a few 10% percent thus allowing slightly smaller values of N. The second reason why (5.20) is not fully accurate is that higher order corrections become significant already for $\epsilon_S \sim 0.01$. That is related to the well-known poor convergence of thermal perturbation theory. Whether or not these latter effects can help decrease the minimum value of N is an intricate question that necessitates the careful analysis presented in the following sections. In the next subsection we begin with a qualitative assessment of these issues.

5.1.4 3D or not 3D?

In order to set the stage for the refinement of the estimate of the previous section, we will here first discuss the issue of perturbation theory, which is the most important and subtle, and later address how a sizable h/T affects (5.20). The former issue was originally discussed many years ago in the context of gauge theories, see e.g. [247][248][249][250], while the second is a novel feature of the present framework.

The poorer convergence properties of perturbation theory in thermal field theory compared to QFT at zero temperature, stems from two joint facts. The first is the presence of light bosonic degrees of freedom (associated to the Matsubara zero modes) in the 3D effective theory below the scale $\sim \pi T$. The second is the IR relevance of their interactions, i.e. the presence of couplings of positive mass dimension that become strong at sufficiently low momenta. These facts can indeed lead to a complete breakdown of perturbation theory at finite temperature even for perfectly weakly coupled QFTs. One example of that is given by massless non-abelian gauge theories, which become strongly coupled at a scale roughly of the order of their effective 3D coupling $\propto g_3^2 = g_4^2 T$. This for instance implies the well know result that the free energy of hot QCD does not admit a series expansion in g_4^2 beyond the third order. Another example is given by systems in the neighborhood of a phase transition where some scalar is tuned to be light. The resulting 3D strong coupling can make it difficult to assess the order of the phase transition. Nevertheless, even without going to such limiting situations, it is always the case that perturbation

theory converges more slowly at finite temperature.

In order to better appreciate this issue let us consider the simple case of $\lambda\varphi^4$. The dimensionless 4D quartic coupling λ matches to a dimension 1 coupling $\lambda_{eff} = \lambda T$ in the 3D effective theory at scales $p \ll \pi T$. In this situation, the loop expansion parameter at external momentum p in the massless limit is proportional to λ_{eff}/p and grows strong at sufficiently small p. The presence of a finite mass saturates this IR growth. One can easily estimate the strength of the loop expansion parameter by comparing, for instance, the tree and 1-loop contributions to the 4-point function at small external momentum. [7] One finds that the loop expansion is roughly controlled by $\epsilon_{3D} = (3\lambda T)/(16\pi m_{\varphi,0}) = \epsilon_{4D}(\pi T/m_{\varphi,0})$, as dictated by dimensional analysis. In this equation $m_{\varphi,0}$ represents the mass of the zero mode including of course all thermal and quantum corrections. As we already mentioned, in the vicinity of a second order phase transition the effective mass $m_{\varphi,0}$ can be small enough to make the IR 3D dynamics strongly coupled even for an arbitrarily weak 4D quartic λ. However, even away from such situation and focussing on the high temperature regime, one finds that ϵ_{3D} is larger that its 4D counterpart. Indeed, at temperatures much larger than the 4D mass m_φ, the mass of the zero mode is dominated by thermal corrections, which at sufficiently weak coupling are well approximated by the 1-loop contribution

$$m^2_{\varphi,0} = \frac{\lambda}{4}T^2 + \dots \tag{5.21}$$

so that the 3D loop expansion parameters is

$$\epsilon_{3D} \equiv \frac{3\lambda T}{16\pi m_{\varphi,0}} \sim \frac{3\sqrt{\lambda}}{8\pi} = \frac{\sqrt{3}}{2}\sqrt{\epsilon_{4D}}, \tag{5.22}$$

where $\epsilon_{4D} = 3\lambda/(16\pi^2)$ is consistent with (5.10). This “square root" relation among the expansion parameter in respectively 3D and 4D is at the origin of the lower convergence rate of perturbation theory in thermal field theory. For instance, $\epsilon_{4D} = 0.01$, known by practice to be well within the perturbative region, corresponds to $\epsilon_{3D} \sim 0.1$, which can easily lead to poor convergence in the presence of upward numerical “accidents". Indeed our ϵ's are just rough indicators of the convergence of the perturbative series: perturbation theory can be safely applied when they are significantly $\ll 1$, but not necessarily so when they are just a few times smaller than 1. For instance the next loop order contribution within the effective 3D theory to eq. (5.21) [8] is of relative size $4 \times \epsilon_{3D}$, which is $\sim 40\%$ for a thoroughly weakly coupled 4D theory with $\epsilon_{4D} = 0.01$. The need for applying resummation techniques in thermal field theory, is related to these simple facts.

[7] We cannot proceed as in the previous section because in the Euclidean 3D theory there is no S-matrix. The estimate is however comparable to that implied by the S-matrix of the 3D Minkowskian theory that is obtained through Wick rotation.

[8] This contribution corresponds from the point of view of 4D diagrammatics to the resummation of the so called daisy diagrams.

We can easily adapt the above discussion to our model, focusing on the scalar sector, where the parameter ϵ_S can potentially be sizable. In principle one should also consider the light bosonic modes from the SM gauge sector. However, as it turns out, the overall contribution of these modes is subdominant and the associated higher order effects are thus not very significant. Indeed the leading SM contribution is by large the one from the top quark, which we do not expect to suffer from 3D effects given the fermionic Matsubara modes are all gapped. In view of that we shall neglect higher order effects from the gauge sector.

For what concerns the scalar sector there are two main differences with respect to the simple $\lambda\varphi^4$ case analyzed above. The first is that ours is a multi-field model. The second is that the zero mode of the Higgs field has negative squared mass at all T's, leading to symmetry breaking. Because of this second feature, the H spectrum separates into the radial mode h plus the (eaten) Goldstone bosons. Moreover, in the shifted vacuum, trilinear interactions appear. As in 3D the strength of the loop expansion parameter depends on the IR details (the thermal masses), one should in principle perform a detailed analysis to establish the regime of perturbativity of the theory. However, and not surprisingly, if the vector boson mass (eaten Goldstones) and Higgs are roughly of the same order then the estimates are quantitatively the same as in the absence of symmetry breaking when written in terms of the physical H mass. In particular, the trilinears that emerge in the broken phase do not change the overall picture qualitatively. [9] We will therefore simplify the present discussion considering the 3D effective theory for H and S in the absence of symmetry breaking.

The loop expansion parameters can be deduced by comparing leading and subleading contributions to some observable. In the absence of a 3D S-matrix, a surrogate of the analysis of section 5.1.2 is offered by the 2-point functions of linear combinations of the canonically normalized composite operators $O_S = S^2/\sqrt{N}$ and $O_H = H^\dagger H/\sqrt{4}$. Inspecting $\langle OO\rangle$ we find that the loop expansion parameters are:

$$\begin{aligned} \epsilon_H^{\rm 3D} &\equiv \frac{6\lambda_H}{16\pi}\frac{T}{m_{h,0}} \ll 1, \\ \epsilon_S^{\rm 3D} &\equiv \frac{\lambda_S N}{16\pi}\frac{T}{m_{s,0}} \ll 1, \\ |\epsilon_{HS}^{\rm 3D}| &\equiv \frac{2|\lambda_{HS}|\sqrt{N}}{16\pi}\frac{T}{\max[m_{h,0}, m_{s,0}]} \ll 1, \end{aligned} \tag{5.23}$$

where $m_{h,0}$, $m_{s,0}$ are the effective masses of the corresponding zero modes. [10] Comparing leading and subleading corrections to 4-point functions of S, H we find additional constraints. For example, requiring that the S and Higgs 1-loop contributions

[9] The stability constraint $\lambda_{HS}^2 < \lambda_H\lambda_S$ plays a role to ensure this result.

[10] In the case of interest $m_H^2(T) < 0$ and the origin is unstable. Since our perturbative estimates only make sense around an energy minimum, Eq. (5.23) should be interpreted as relations involving the zero mode masses around the vacuum, where the Higgs mass is $m_{h,0} \sim \sqrt{2}|m_H(T)|$.

to $\langle SSSS\rangle$ be smaller than the tree-level respectively give $\lambda_S NT/(8\pi m_{s,0}) \ll 1$ and $\lambda_{HS}^2 T/(8\pi m_{h,0}) \ll \lambda_S$, which when combined imply

$$\frac{\lambda_{HS}^2 N}{(8\pi)^2}\frac{T^2}{m_{h,0}m_{s,0}} \sim (\epsilon_{HS}^{3\mathrm{D}})^2 \frac{\max[m_{h,0}, m_{s,0}]^2}{m_{h,0}m_{s,0}} \ll 1. \tag{5.24}$$

This may be actually a stronger condition than the last one of (5.23) depending on the masses involved as well as the actual size of $\epsilon_{HS}^{3\mathrm{D}}$. We will see below that a precise determination of the expansion parameter associated to λ_{HS} is fortunately not necessary because its perturbativity is always guaranteed by stability and $\epsilon_{H,S}^{3\mathrm{D}} \ll 1$.

Using eq.(5.18) for the 3D effective masses we can finally estimate the 3D expansion parameters in analogy with the $\lambda\varphi^4$ we discussed above. Let us focus on ϵ_H^{3D} first. By eq. (5.6), $|m_{h,0}|/(\sqrt{2\lambda_H}T) \sim h(T)/T \gtrsim 1.2 - 1.5$, the first of eqs. (5.23) implies

$$\epsilon_H^{3D} \sim \frac{\sqrt{3\epsilon_H}}{4} \times \frac{T}{h(T)} \lesssim 0.02 - 0.03 \tag{5.25}$$

where we used the value of λ_H renormalized at the weak scale. At higher scales λ_H decreases so that the effective 3D coupling gets more suppressed. By the above equation we conclude that ϵ_H^{3D} is rather small and hence that we do not expect the need to chase higher order effects associated to this coupling. We expect similar conclusions concerning the 3D fate of the electroweak gauge couplings. Notice also that this is just an upper bound. There is in principle no limitation in making $h(T)/T$ bigger by taking N larger. Indeed for $h(T)/T$ big enough, the gap in the Higgs (as well as W, Z, t) gets larger than T, so that there is no significant 3D dynamics associated to this mode, as we shall discuss in more detail at the end of this section.

Regarding $\epsilon_S^{3\mathrm{D}}$ we similarly have [11]

$$\epsilon_S^{3\mathrm{D}} = \frac{\sqrt{3}}{2}\sqrt{\epsilon_S}. \tag{5.26}$$

In this case however we are not forced to go to very small values of ϵ_S. To the contrary, eq. (5.20) shows that the smallest N correspond to largish $\epsilon_S > 0.01$, which in turn implies $\epsilon_S^{3\mathrm{D}} \gtrsim 0.1$. Under these circumstances, while the perturbative series in 4D may still converge well, thermal 3D perturbation theory may require re-summation in order to make more firm statements. Luckily this can be done at leading order in the $1/N$ expansion, as we shall describe in the next section.

We should finally discuss perturbativity for ϵ_{HS}^{3D}. By using the stability constraint

[11] The relation coincides with eq.(5.22) as it is independent on N.

$\lambda_{HS}^2 \leq \lambda_H \lambda_S$ and eq. (5.23) we have

$$|\epsilon_{HS}^{3D}| \leq \sqrt{\frac{2}{3}}\sqrt{\epsilon_H^{3D}\epsilon_S^{3\mathrm{D}}}\sqrt{\frac{\min[m_{h,0}, m_{s,0}]}{\max[m_{h,0}, m_{s,0}]}} \leq \sqrt{\frac{2}{3}}\sqrt{\epsilon_H^{3D}\epsilon_S^{3\mathrm{D}}} \tag{5.27}$$

somewhere in between the other two expansion parameters. Similarly, (5.24) is automatically satisfied provided we combine stability and 3D perturbativity of $\lambda_{S,H}$.

Let us now discuss the impact of the sizable contribution to the thermal masses of h, W, Z, t that a large h/T could imply. In ordinary QFT, at small N and at weak coupling, thermal corrections to masses are parametrically smaller than the scale $\sim \pi T$ controlling the Matsubara spectrum. This statement applies to all corrections, including those induced via the effect of thermal loops on the VEVs of scalars. Given for instance a trilinear coupling g, the condition for a small thermal correction to masses is roughly $g^2T^2 \lesssim \pi^2T^2$, which parametrically coincides with the smallness of the loop expansion parameter $g^2/8\pi^2$. Notice in particular that at weak couplings the zero modes remain lighter that $\sim \pi T$, even if the relative effect is large, given the lightness of these modes at tree level. The consequence on the 3D dynamics are just as we discussed above. However, as our model clearly illustrates, the situation changes drastically when considering QFT at large-N, even at weak coupling. As we discussed, in the limit $N \to \infty$ with ϵ_S and ϵ_{HS} finite but perturbatively small, the tadpole correction of Fig. 5.1 to the thermal Higgs mass scales like $\epsilon_{HS}\sqrt{N} \to \infty$ and can consistently become much larger than πT. When that is the case and $\lambda_{HS} < 0$ the resulting VEV of h will also shift the masses of W, Z, t way above πT. More precisely that happens when $\epsilon_{HS}\sqrt{N} \gg 1$.

It is worth, to gain a better overview on the dynamics, to focus for a moment on the case $\epsilon_{HS}\sqrt{N} \gg 1$, even though this would require N much larger than the lowest allowed value. In this case, the heaviest SM states h, W, Z, t decouple from the thermal dynamics: their density and all their effects, including those on the effective potential, are Boltzmann suppressed. In particular they decouple from the 3D stochastic dynamics. Notice, in contrast, that in the N scaling we are considering ϵ_S is fixed: as long as perturbation theory applies, the mass $m_{s,0}$ of the zero mode of S, see eq. (5.18), remains $\ll \pi T$, and this mode dominates the thermal dynamics. When $\epsilon_{HS}\sqrt{N} \gg 1$, with $\epsilon_S, |\epsilon_{HS}| \ll 1$ the thermal spectrum has then the rough structure

$$m_h^2, m_W^2, m_Z^2, m_t^2 \gtrsim \pi^2|\epsilon_{HS}|\sqrt{N}T^2 \gg \pi^2T^2 \gg \pi^2\epsilon_S T^2 \sim m_s^2\,. \tag{5.28}$$

In order to systematically compute physical quantities one should then use an effective field theory approach in three steps: 1) integrate out the heavy SM states h, W, Z, t matching to an EFT for the remaining light (and nearly decoupled) SM states plus S; 2) run this EFT down to the compactification scale $\sim \pi T$; 3) match to a 3D EFT for S. Notice that the dynamics that determines the thermal masses of h, W, Z, t takes place at much lower energies, at or below T. In a sense the situation is similar to the case,

naturally realized in supersymmetry, where a flat direction field controlling the mass of heavy states is stabilized by some IR dynamics. The difference with respect to our case is that here the operator S^2 controlling the heavy masses is a genuine composite, in the sense that $\langle S \rangle = 0$. Indeed the large size of $\langle S^2 \rangle$ is not driven by the flatness of the potential but by the coherent pile-up of thermal fluctuations from N fields. We can however formulate the dynamics in a way that is not too different from the case of a flat direction. It just suffices to introduce an auxiliary field σ mediating the $(S^2)^2$ interaction, as it is standard in the study of large-N models like the Gross-Neveu model or the $O(N)$ model. As discussed in the next section, one can start from the equivalent Lagrangian eq. (5.29), which reduces to the original one by first integrating out σ. However, according to the EFT picture outlined above, it is more convenient to first integrate out h as well as W, Z, t, working around a background with non-vanishing σ. In such a way one derives an EFT redundantly written in terms of S and σ, where all degrees of freedom are light. This effective Lagrangian is exactly quadratic in S, and can be studied in a $1/N$ expansion by first integrating out all of S. We will partly illustrate that in the next section.

The extreme limit $\epsilon_{HS}\sqrt{N} \gg 1$ just described is formally interesting. However, the estimates in (5.20) reveal that typical values of $\epsilon_{HS}\sqrt{N}$ are usually far away from it. In these more realistic cases there is no big separation of scales, so we do not need to construct the effective theory as outlined in the previous paragraph. As we will see in the analysis of the next subsection, however, it is still useful to integrate-in the field σ (which controls the mass of h, W, Z, t) and keep all non-linear terms in this quantity; on the other hand, it suffices to work at one loop order in the SM couplings.

5.1.5 The effective potential

In this subsection we will find a more accurate approximation for the effective potential and to the conditions necessary to satisfy (5.6).

We begin by recalling that eq. (5.20) indicates that the region of parameter space where sphalerons are switched off at high T lies always at large N. Moreover the value of N is minimized for the largest ϵ_S, compatibly with perturbativity. These two joint facts prompt us to use the standard large-N methodology to re-sum all orders in ϵ_S. Indeed, neglecting for the moment the effects of λ_{HS}, the loop expansion in the S sector is easily seen to correspond to a series in $\lambda_S^p N^q \propto \epsilon_S^p N^{q-p}$ with $p \geq q$. Treating ϵ_S as a fixed and not necessarily small parameter, the large-N resummation consistently captures at leading order all the terms with $p = q$, with the next-to-leading terms ($p = q+1$) formally suppressed by $O(1/N)$. See the right of Fig. 5.1 for a sample of leading diagrams.

In the purely 4D case ($T = 0$), while the resummation of the leading $p = q$ series makes perfect sense, the dynamical regime $\epsilon_S \gtrsim O(1)$, where the effects would be most dramatic

and interesting, lies unfortunately in the UV, where the theory is out of control. [12] However, at finite temperature, as we already reviewed, the effective coupling becomes stronger in the 3D regime. In this situation, the resummation of the leading series in $1/N$ can capture more dramatic effects. For example, interpreting the mass $m_{s,0}$ of the zero mode of S as a free parameter, say by suitably tweaking the bare squared mass of S (in particular allowing it to be negative), the IR dynamics in 3D potentially displays various phases that are all tractable at large-N. In the strict $m_{s,0} = 0$ limit the system flows to a CFT (the so-called $O(N)$ model) in the far IR. The typical scale of the flow is $\sim \lambda_S NT/8\pi$, where the coupling λ becomes strong but the $1/N$ approximation remains reliable. For $0 < m_{s,0} \lesssim \lambda_S NT/8\pi$ we have a strongly coupled gapped phase, while for $m_{s,0} \gtrsim O(\lambda_S NT/8\pi)$ we get back to a weakly coupled regime, where however the effective coupling $\epsilon_S^{3D} \sim \lambda_S NT/(8\pi m_{s,0})$ is bigger than ϵ_S; in particular, in the generic high T phase we have seen that $\epsilon_S^{3D} \sim \sqrt{\epsilon_S}$. In our model (5.7) we are always in this latter case and therefore it is the less dramatic, but still quantitatively important, series in ϵ_S^{3D} that our large-N resummation shall capture. To be more quantitative, let us estimate the range of ϵ_S compatible with the perturbative definition of our model. For that purpose, observe that at leading order in $1/N$, and considering the realistic limit $\epsilon_H \ll \epsilon_S$, the beta function $\mu d\epsilon_S/d\mu = 2\epsilon_S^2$ develops a Landau pole at $\Lambda_L = \mu \exp[1/(2\epsilon_S(\mu))]$. Having a consistent theory below $\Lambda = 100$ TeV $< \Lambda_L$ then translates into the perturbativity requirement $\epsilon_S(m_t) \lesssim 0.07$. We confirmed numerically that the inclusion of all other couplings does not alter this upper bound significantly. In the following we will therefore stick to the domain $\epsilon_S(m_t) \lesssim 0.07$. Such a value corresponds to $\sqrt{\epsilon_S} \sim 0.25$ for which resummation of the leading 3D effects can make an almost $O(1)$ effect.

Having identified a reliable and systematic re-summation suitable for our model, let us now go back to our original task: finding the effective potential. An important simplification in this analysis comes from the fact that the vacuum of S is trivial at all temperatures whenever (5.8) holds. We have already shown this in the simplified analysis of Section (5.1.3), but it turns out it is in fact a general result, as argued in Appendix 5.A.

Now, the standard technique to resum the leading loop effects at large-N in the $O(N)$ model involves the introduction of an auxiliary field σ mediating the interactions of interest. In practice we shall add a trivial term $\frac{1}{4\lambda_S}\left(\sigma - \lambda_S S^2 - 2\lambda_{HS} H^\dagger H\right)^2$ to eq. (5.7), such that the part of the Lagrangian that depends only on the scalar fields becomes:

$$\begin{aligned} \mathcal{L} &= D_\mu H^\dagger D^\mu H - \left(m_H^2 + \frac{\lambda_{HS}}{\lambda_S}\sigma\right) H^\dagger H - \lambda_H \left(1 - \frac{\lambda_{HS}^2}{\lambda_H \lambda_S}\right) (H^\dagger H)^2 \\ &+ \frac{1}{2}\partial_\mu S \partial^\mu S - \frac{1}{2}(m_S^2 + \sigma) S^2 + \frac{1}{4\lambda_S}\sigma^2. \end{aligned} \tag{5.29}$$

Of course when integrating out σ this trivially gives back our Lagrangian. However, in

[12] The situation is famously reversed in Gross-Neveu model in 2D, where the coupling is asymptotically free and the strong regime lies controllably in the IR.

order to organize perturbation theory as an expansion in $1/N$, it is convenient to first integrate out S. The key property of eq. (5.29) is that S appears quadratically, so that it can be integrated out exactly to give:

$$\begin{aligned}\mathcal{L} \to \mathcal{L}_{\text{eff}} &= D_\mu H^\dagger D^\mu H - \left(m_H^2 + \frac{\lambda_{HS}}{\lambda_S}\sigma\right) H^\dagger H - \lambda_H \left(1 - \frac{\lambda_{HS}^2}{\lambda_H \lambda_S}\right) (H^\dagger H)^2 \\ &+ \frac{1}{4\lambda_S}\sigma^2 + N\Gamma[m_S^2 + \sigma, \partial],\end{aligned} \tag{5.30}$$

where $\Gamma[m_S^2 + \sigma, \partial]$ is a calculable function of σ and its derivatives (Note we also used the fact that $\langle S \rangle = 0$ in deriving (5.30). We will assume this is understood in the following.). The most salient property of the above result is that in the limit $N \to \infty$ with $\lambda_S N$ fixed, the action for σ in the second line is formally of order N. Neglecting for the moment the first line involving the interaction with H, this implies that $1/N$ is the loop expansion parameter for the self-interactions of σ. This is easily seen by going to a canonical basis $\sigma \to \sqrt{\lambda_S}\sigma \sim \sigma/\sqrt{N}$ and expanding Γ in powers of σ: the n-point couplings scale like $(1/\sqrt{N})^{n-2}$ as befits $1/N$ being the loop counting parameter. By the same rescaling, the $\sigma H^\dagger H$ coupling becomes $\lambda_{HS}/\sqrt{\lambda_S}$, which by the stability condition is $\leq \sqrt{\lambda_H}$. We thus conclude that, in the large-N limit with the SM couplings and ϵ_S and ϵ_{HS} fixed, σ interactions come in two classes: self interactions, with strength controlled by powers of $1/N$, and interactions with the Higgs, whose strength is comparable, at most, to that of the SM couplings.

Eq. (5.30) contains all the quantum fluctuations from S at leading order in $1/N$, with subleading effects associated to σ fluctuations not yet computed. When treating the couplings as the bare ones, the above Lagrangian should then not involve any UV divergence associate to the leading S loops. This is easily checked. Γ involves indeed a logarithmic UV divergent σ^2 piece that matches precisely the UV divergence in the tree coefficient $1/4\lambda_S$. Moreover, using the beta functions given in (5.80), one can easily check that the other two combinations of couplings λ_{HS}/λ_S and $\lambda_H - \lambda_{HS}^2/\lambda_S$ are RG invariant when considering the contributions that are purely induced by S-loops and leading in the $1/N$ expansion. It follows that the corresponding bare couplings are free of the associated UV divergences, as it should. This result means for instance that the two point function of σ from eq. (5.30) will be determined in the leading log approximation by λ_S renormalized at the largest scale among T, $\sqrt{\langle\sigma\rangle}$ and the external momentum ∂. In practice when considering thermal loops the relevant scale will be T. Notice however that Γ also contains finite corrections coming from $3D$ physics, whose relative size is controlled by $\sqrt{\epsilon_S}$ and which we will take into account.

With all the above comments in place we can now compute the effective potential. Motivated by the estimates in section 5.1.3 a reasonable approximation would be to work at leading order in the $1/N$ expansion (but all orders in ϵ_S) and at 1-loop in the SM couplings. Proceeding in the standard way we decompose the fields in their classical

background plus their fluctuations. Ignoring the eaten Goldstones, which we will treat separately, in the scalar sector we have

$$\sigma = \sigma_c + \delta\sigma \qquad\qquad H = \begin{pmatrix} 0 \\ \frac{h+\delta h}{\sqrt{2}} \end{pmatrix}. \tag{5.31}$$

When expanding the action at quadratic order in $\delta\sigma$ and δh we find a slight complication from a mixing term arising from the cubic interaction $\sigma H^\dagger H$. It is then convenient to integrate over the quantum fluctuations in two steps: we first perform a suitable shift of $\delta\sigma$ to eliminate the mixing and then compute the fluctuation determinant for the resulting diagonal quadratic action. The first step corresponds to

$$\delta\sigma \to \delta\sigma + \frac{\lambda_{HS}}{\lambda_S} G_{\sigma\sigma} h \delta h \tag{5.32}$$

with

$$G_{\sigma\sigma} = \left(\frac{1}{2\lambda_S} + \frac{N\delta^2\Gamma}{\delta\sigma\delta\sigma} \right)^{-1} \tag{5.33}$$

the $\delta\sigma$ propagator around the background. The δh self-energy gets corrected according to

$$\begin{aligned} \Sigma_{hh} \quad &\equiv \quad \partial^2 + m_H^2 + \frac{\lambda_{HS}}{\lambda_S}\sigma_c + 3\lambda_H \left(1 - \frac{\lambda_{HS}^2}{\lambda_H \lambda_S}\right) h^2 \\ &\to \quad \Sigma_{hh} + \frac{\lambda_{HS}^2}{\lambda_S^2} G_{\sigma\sigma} h^2. \end{aligned} \tag{5.34}$$

After this diagonalization we notice that the $\delta\sigma$ fluctuation determinant contributes a term that purely depends on σ_c and which is $1/N$ suppressed with respect to the leading action for σ. In accordance to our goals we neglect this term. The δh contribution, determined by the Tr ln of the self energy operator in eq. (5.34), corresponds to the SM Higgs contribution to the 1-loop effective potential, dressed by an infinite class of S loops, which resum the leading series in ϵ_S in the large-N limit. This computation is made involved by the non trivial momentum dependence of $G_{\sigma\sigma}$. In view of that, we have chosen to simplify our computation by making the approximation $G_{\sigma\sigma} \simeq 2\lambda_S(T)$. The error this entails corresponds to $O(\sqrt{\epsilon_S})$ relative to the h-loop contribution to the effective potential. In view of the smallness of the leading Higgs loop contribution compared to the combined effects of W, Z, t, we estimate this is a fair approximation. In principle, with some extra effort, this approximation could be dropped. We plan to reconsider this in the future, though it seems rather clear this is not going to change our estimates on the range of parameters by more than a few percent.

With all the above comments in place, in particular as concerns our approximations,

after some straightforward algebra, at last, the effective potential for h, σ_c reads

$$\begin{aligned} V_{\rm eff}(h,\sigma_c) &= \frac{1}{2}\left(m_H^2+\frac{\lambda_{HS}}{\lambda_S}\sigma_c\right)h^2+\frac{\lambda_H}{4}\left(1-\frac{\lambda_{HS}^2}{\lambda_H\lambda_S}\right)h^4 \\ &- \frac{1}{4\lambda_S}\sigma_c^2+NV_0(m_S^2+\sigma_c) \\ &+ V_0\left(m_H^2+\frac{\lambda_{HS}}{\lambda_S}\sigma_c+\lambda_H\left(3-\frac{\lambda_{HS}^2}{\lambda_H\lambda_S}\right)h^2\right) \\ &+ 3V_0\left(m_H^2+\frac{\lambda_{HS}}{\lambda_S}\sigma_c+\lambda_H\left(1-\frac{\lambda_{HS}^2}{\lambda_H\lambda_S}\right)h^2\right) \\ &+ 6V_1\left(\frac{g^2}{4}h^2\right)+3V_1\left(\frac{g^2+g'^2}{4}h^2\right)+12V_{1/2}\left(\frac{y_t^2}{2}h^2\right) \end{aligned} \tag{5.35}$$

where all the couplings should be taken renormalized at $\mu \sim T$ and the functions V_j are defined in Appendix 5.B. The contributions in the above equation are easily identified. Besides the "tree" level terms in the first two lines, the third line corresponds to the physical Higgs scalar fluctuation δh, the fourth to the eaten Goldstones in Landau ($\xi = 0$) gauge, the fifth to respectively W, Z and t. Notice that the difference in the argument of the function V_0 for Higgs and Goldstones is also determined by the mixing with σ we mentioned above.

Overall, (5.35) includes all 1-loop effects and re-sums *all powers* of ϵ_S at leading $1/N$ order, but ignores genuinely 2-loop contributions involving the SM couplings — since expected to be smaller than those solely controlled by ϵ_S — as well as corrections of order $\sqrt{\epsilon_S}$ on the already small Higgs contribution in the third line. All effects scaling as $\lambda_{HS}N$, which play a crucial role in symmetry non-restoration, come from $\sigma_c\lambda_{HS}/\lambda_S$ and have thus been taken into account in (5.35). An example of re-summed contributions is shown in the right panel of Fig. 5.1.

The true effective potential for h is found solving numerically the classical equations of motion for the constant configuration $\sigma_c(h^2)$ and replacing it back in (5.35), i.e.

$$V_{\rm eff}(h) = V_{\rm eff}(h, \sigma_c(h^2)). \tag{5.36}$$

Yet, before presenting our numerical results it is instructive to investigate the ideal limit in which only the second line of (5.35) is kept. This limit is useful to quantify for which values of ϵ_S our model is expected to be comfortably under perturbative control at finite T. The numerical solution σ_c in this simplified case, and for the high T regime $\sigma_c \gg m_S^2$, is shown in Fig. 5.2 as a function of $\epsilon_S(\mu = T)$. The exact result (solid line) approaches the expression $\sigma_c = \lambda_S N T^2/12$ (dotted line) found at leading order in an expansion in ϵ_S (i.e. the one we used for example in (5.18)) only for $\sigma_c/T^2 \ll 1$. Quantitatively, the leading order expression $\sigma_c = \lambda_S N T^2/12$ is larger by a factor $(1 - 2\sqrt{3\epsilon_S} + \mathcal{O}(\epsilon_S))$, that is $30 - 40\%$ already at $\epsilon_S \sim 0.01$ and gives us a measure of the perturbative domain at

high T. This result is especially important to us because physically σ_c represents the thermal mass squared of the singlet (see also (5.77)), whereas $\sigma_c\lambda_{HS}/\lambda_S$ is the dominant thermal correction to the Higgs mass squared, see (5.35). What Fig. 5.2 demonstrates is that when $\epsilon_S \gtrsim 0.01$ a numerical study of (5.35), including in particular the full function $NV_0(m_S^2+\sigma_c)$, is necessary to obtain a careful assessment of the effective potential at finite T.

We are finally ready to discuss our results. We derived $V_{\rm eff}(h)$ numerically according to (5.36) for a set of input values $\lambda_S(m_t), \lambda_{HS}(m_t), N$ and

$$m_s^2 = m_S^2 + \lambda_{HS}v^2. \tag{5.37}$$

We took $y_t(m_t) = 0.934-0.951$, $g(m_t) = 0.648$, $g'(m_t) = 0.359$, $g_s(m_t) = 1.17$, $\lambda_H(m_t) = 0.128$ and $m_t = 173$ GeV, $m_h = 125$ GeV. Consistently with our approximations in the effective potential, the RG evolution of the couplings is evaluated at 1-loop. The renormalization scale was fixed at the value $\mu = \sqrt{m_t^2+T^2}$ to minimize the logs in V_j at large T and avoid singularities at $T = 0$.

First we note that, as already anticipated when discussing (5.20), the function $h(T)/T$ increases at high T due to the decrease in $\lambda_H(\mu)$ and has a minimum at around the temperature at which the transition between high and low T occurs. This implies that the requirement that sphalerons are switched off at all $T < \Lambda = 100$ TeV is dominated by low temperatures, which motivates a posteriori the approximation made in (5.6). This behavior is shown in fig. 5.3. Our analysis also confirms the result (5.20) quantitatively: $h \gtrsim T$ requires large N and a sizable ϵ_S. Specifically, $N \gtrsim 100-200$ seems unavoidable if we require the description to stay perturbative up to $\Lambda = 100$ TeV.

In fig. 5.4 we show, for $\epsilon_S \in [0.01, 0.07]$ and $m_s = 0$, the value $N_{\rm min}$ corresponding to the smallest N that gives $h/T \geq 1.2-1.5$ compatibly with the stability condition $\lambda_{HS}^2 < \lambda_H\lambda_S$ all the way up to $T = 100$ TeV. As already evident from (5.20), $N_{\rm min}$ is obtained when maximizing ϵ_S. Moreover, for a given ϵ_S the minimum N is found at the maximally allowed $|\lambda_{HS}|$. The ratio $\lambda_{HS}^2/\lambda_H\lambda_S$ increases with the RG scale, so that that the stability constraint is dominated by large RG scales: demanding stability up to our UV cutoff translates into $|\lambda_{HS}(m_t)|/\sqrt{\lambda_S(m_t)\lambda_H(m_t)} \lesssim 0.4-0.6$. In this respect it is important to stress that the 1-loop approximation of the RG overestimates the drop in $\lambda_H(\mu \gg m_t)/\lambda_H(m_t)$ when $y_t(m_t) = 0.951$ is used, and hence over-constraints $|\lambda_{HS}|$, whereas an evolution closer to a 3-loop analysis is obtained for the lower value $y_t(m_t) = 0.934$. In this sense the choice $y_t(m_t) = 0.934$ appears to be more accurate. As the figure shows, the weaker drop in λ_H results in a weaker bound on N. Finally, we verified that for $N \gg N_{\rm min}$ and $\epsilon_S \to 0$ our numerical solution $h(T)$ approaches (5.19), as it should.

Additional information is provided by fig. 5.5. Taking the conservative value $y_t(m_t) =$

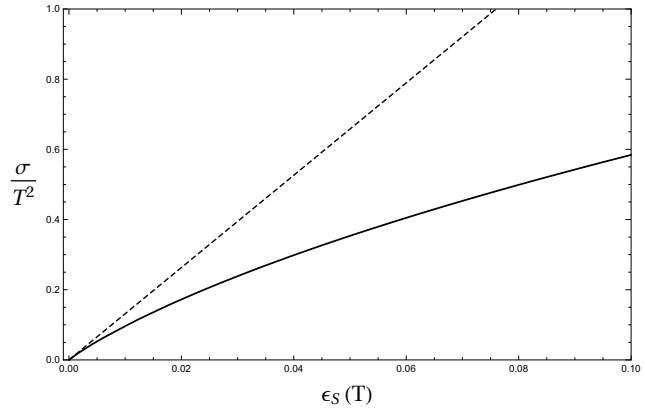


Figure 5.2: The solid curve shows the exact classical solution $\sigma(s^2 = 0)/T^2$ of (5.30) with $H \to 0$, as a function of $\epsilon_S(\mu = T)$. The dotted line is the approximate expression $\sigma/T^2 = \lambda_S N/12$ obtained at leading order in an expansion in $\epsilon_S \ll 1$.

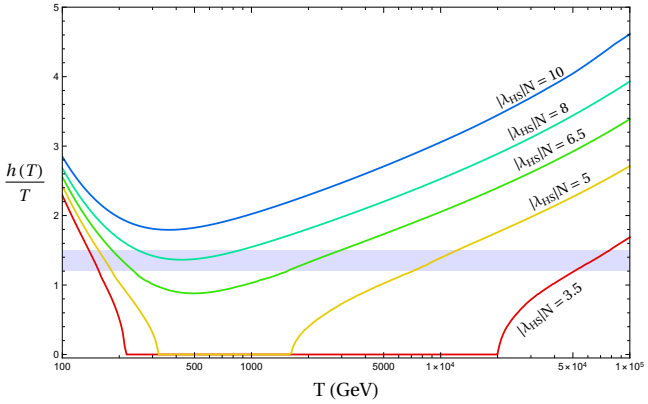


Figure 5.3: Here we show h/T as a function of the temperature for various values of $\lambda_{HS}N$ and $m_s = 0, \epsilon_S = 0.03, N = 1000$. The filled area identifies $1.2 \leq h/T \leq 1.5$. As discussed, these curves have a minimum at around $100 - 1000$ GeV, where the transition between the high-T and low-T behavior takes place. The growth at high temperatures is due to the running of λ_H.

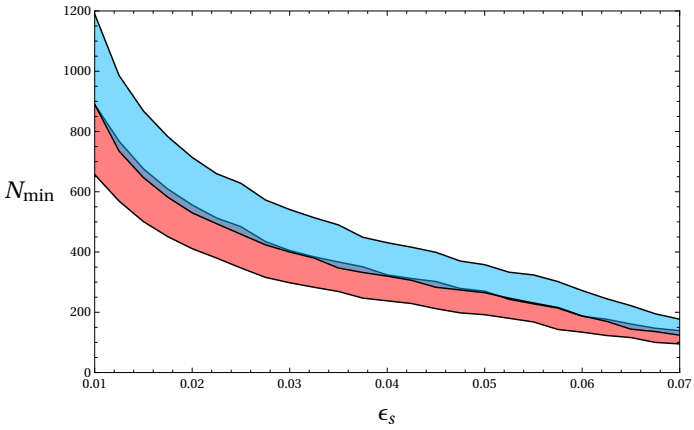


Figure 5.4: The colored area shows the minimum value of N for a given ϵ_S that gives $1.2 \leq h/T \leq 1.5$ for all $T \leq 100$ TeV compatibly with the stability requirement $\lambda_{HS}^2 \leq \lambda_H \lambda_S$. Here we took $m_s = 0$ and $y_t(m_t) = 0.934$ (red) 0.951 (blue).

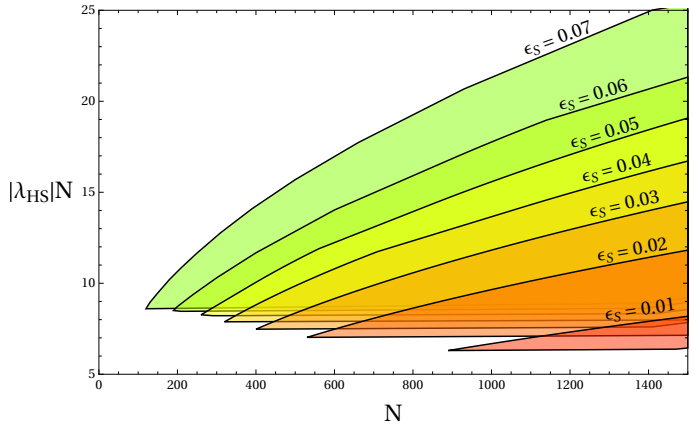


Figure 5.5: Region with $h/T \geq 1.2$ in the N, $\lambda_{HS} N < 0$ plane. Here $m_s = 0$, $y_t(m_t) = 0.951$.

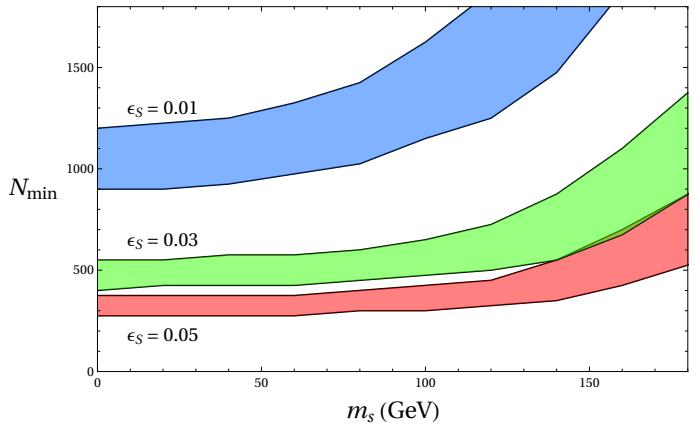


Figure 5.6: $N_{\min}$ as a function of m_s for a few representative values of ϵ_S (here $y_t(m_t) = 0.951$). In the colored region $1.2 \leq h/T \leq 1.5$ and the model is stable for all $T \leq 100$ TeV.

0.951, and again $m_s = 0$, the plot shows what region in the $N, \lambda_{HS} N < 0$ plane is selected by requiring $h/T \geq 1.2$ and stability up to the UV cutoff, for a few choices of ϵ_S. Note that $|\lambda_{HS}|N \gtrsim 6 - 9$, compatibly with what we found in Section 5.1.3, where the smaller values are obtained for smaller ϵ_S.

Another relevant question we can address with (5.35) concerns the allowed zero-temperature mass (5.37) of the scalar. Crucially, S cannot be much heavier than m_t because otherwise it would not be able to generate large enough thermal corrections to the Higgs mass to win over the positive top quark contribution. A more quantitative assessment of this expectation is given in Fig. 5.6, where we present the minimal value of N required to suppress sphaleron processes at all temperatures as a function of m_s, and for a few representative values of ϵ_S.

It is worth emphasizing that, while we have argued that $N \gtrsim 100 - 200$ is unavoidable within the minimal model (5.7) (see however possible extensions discussed in Section 5.3.2), N cannot be arbitrary large. The reason is that in scenarios of EW baryogenesis one expects [225]

$$\eta_b \sim c \sin\theta_{\rm CP} (20\alpha_w^5)/g_*, \tag{5.38}$$

where $20\alpha_w^5$ measures the strength of the non-perturbative sphaleron effects (as in section 5.1.1), c is a function of the wall dynamics, $\theta_{\rm CP}$ a CP-violating phase, and g_* the total number of relativistic degrees of freedom. In our scenario (5.38) is generated by the UV dynamics at $T \sim \Lambda$ and having $\sin\theta_{\rm CP} \sim 1$ is not a problem. However, with $g_* \sim N \gg 100$ the asymmetry would be suppressed by $\sim g_{*,\rm SM}/g_* \sim 100/N$ compared to standard scenarios of EW baryogenesis. For the typical values considered here, $N \sim 10^2 - 10^3$, the net effect is a mild $\mathcal{O}(1-10)$ suppression, but much larger values of N are clearly not viable. Specifically, from (5.38) we find the observed baryon asymmetry $\eta_b \sim 10^{-10}$ can only be reproduced provided $N/c \lesssim 10^4$.

5.2 Phenomenology

In the previous section we have seen that S has to have a mass around the weak scale. This opens the possibility of observing direct and indirect signatures of the new singlet, as discussed in the present section.

5.2.1 Collider constraints

We have already emphasized that S does not acquire a VEV because of our assumption (5.8): the $O(N)$ symmetry remains unbroken. S must thus be produced in pairs in accelerators, with rates proportional to $\lambda_{HS}^2 N$, and show up as missing energy. The most dramatic effect consists in the generation of an invisible Higgs branching ratio, and is

Figure 5.7: The two 1-loop diagrams contributing to the Higgs trilinear.

kinematically allowed only for $m_s^2 < m_h^2/4$:

$$\mathrm{BR}(h \to \mathrm{inv}) = \frac{\Gamma(h \to SS)}{\Gamma_{\mathrm{tot}}} = \frac{\lambda_{HS}^2 v^2 N}{8\pi m_h \Gamma_{\mathrm{tot}}} \left(1 - 4\frac{m_s^2}{m_h^2}\right)^{1/2}. \tag{5.39}$$

Requiring $\mathrm{BR}(h \to \mathrm{inv}) < 10\%$ and conservatively assuming $|\lambda_{HS}|N = 10$ (see Fig. 5.5) we get the severe bound $N \gtrsim 10^6$. To avoid it in the following we will stick to the regime $m_s > m_h/2$.

Direct production via vector boson fusion or associated production has been considered in [251]. According to that study, for $m_h/2 < m_s < \mathcal{O}(m_t)$ there exist realistic chances of discovering S via direct production at a future 100 TeV pp collider as long as $\lambda_{HS}\sqrt{N} \gtrsim 0.2 - 0.4$, which for $|\lambda_{HS}|N = 10$ reads $N \lesssim \mathcal{O}(10^3)$.

Besides direct processes, virtual effects can be important. Because the singlet does not mix with H, all couplings of the Higgs boson to the SM vector bosons and fermions are standard at tree level. Yet, 1-loop diagrams with virtual S's modify them offering an interesting indirect probe of the model. The most important effect in the $N \to \infty$ limit concerns the Higgs wave function and corresponds to a universal reduction by a factor $(1 - \delta Z_H/2)$ of on-shell Higgs couplings with respect to their SM value. One finds

$$\begin{aligned} \delta Z_H &= \frac{\lambda_{HS}^2 N v^2}{8\pi^2 m_s^2} \int_0^1 dx \frac{x(1-x)}{1 - \frac{m_h^2}{m_s^2} x(1-x)} \\ &= \frac{\lambda_{HS}^2 N v^2}{48\pi^2 m_s^2} f(m_h^2/m_s^2), \end{aligned} \tag{5.40}$$

where $f(m_h^2/m_s^2)$ is a decreasing function of m_s, of order $f = 3.5(1.5)$ for $m_s = 70(100)$ GeV and quickly approaching $f = 1$ for larger masses. For $m_s = \mathcal{O}(m_h)$, and using $|\lambda_{HS}|N = 10$ as a benchmark value, $\delta Z_H \sim (\lambda_{HS}^2 N v^2)/(48\pi^2 m_s^2) \sim 1/N$. Future linear colliders might be able to test $N \sim 100 - 1000$ [252] and thus a sizable portion of the relevant parameter space.

Other corrections are less important. Consider for instance the corrections to the Higgs trilinear arising at 1-loop from the diagrams in Figure 5.7. As S is light, the result can only roughly be described as a momentum independent correction to the trilinear. We

are here nonetheless only interested in a rough estimate of the size of the effects. The triangle diagram in the left of Fig. 5.7 gives a correction $\sim \lambda_{HS}^3 N v^2/(24\pi^2 \lambda_H m_s^2)$ relative to the SM trilinear. By taking $|\lambda_{HS}|N = 10$ this becomes $\sim (100/N)\delta Z_H \sim 100/N^2$. This is only comparable to the wave function effect in the limiting case $N \sim 100$ and moreover, in view of the expected experimental sensitivity on the trilinear, not relevant. The diagram in the right of Fig. 5.7 gives a momentum dependent correction of the same size $\sim 1/N$ as that induced by the wave function correction. This is also below the quoted future sensitivities, modulo the unlikely possibility that the momentum dependence of the correction could be used to boost up the signal.

5.2.2 Dark matter constraints and simple fixes

Within our framework S is exactly stable. Unfortunately it cannot be identified with the dark matter because such a possibility appears to be in conflict with current direct and indirect dark matter searches.

Let us see why this is the case. For our mechanism to work S must have been in thermal equilibrium at temperatures above the weak scale. This happens for $|\lambda_{HS}| \gtrsim 10^{-8}$, or equivalently $N \lesssim 10^8 (|\lambda_{HS}|N) \sim 10^9$. Its present-day energy density ρ_S is thus determined by standard freeze-out dynamics. The main number-changing processes setting ρ_S are annihilations into SM particles controlled by the interaction $\lambda_{HS} H^\dagger H S^2$. In the regime of interest, $m_s > m_h/2$, the Higgs propagator is always off resonance, and the largest annihilation rate is into Higgs pairs (for $m_s > m_h$) and vector boson pairs (for $m_s > m_W$). The thermally averaged cross section is $\langle \sigma(S_i S_j \to \mathrm{SM}) v\rangle \equiv \delta_{ij} \langle \sigma_{\mathrm{ann}} v\rangle$ with $\sigma_{\mathrm{ann}} v$ maximized above threshold by $\langle \sigma_{\mathrm{ann,max}} v\rangle \sim \lambda_{HS}^2/(8\pi m_s^2)$. The density n_i of each of the $i = 1, \cdots, N$ scalar components is determined solving the Boltzmann equation $dn_i/dt + 3Hn_i \approx -\sum_j \langle \sigma(S_i S_j \to \mathrm{SM}) v\rangle (n_i n_j - n_i^{\mathrm{eq}} n_j^{\mathrm{eq}})$ with initial condition $n_i \sim n_i^{\mathrm{eq}} \sim T^3/\pi^2$ at high $T > m_s$. Because by symmetry this equation is the same for each n_i one finds that the total number density $\sum_i n_i = n_{\mathrm{tot}} = N n_i$ satisfies

$$\frac{dn_{\mathrm{tot}}}{dt} + 3H(T) n_{\mathrm{tot}} \approx -\frac{\langle \sigma_{\mathrm{ann}} v\rangle}{N} [(n_{\mathrm{tot}})^2 - (n_{\mathrm{tot}}^{\mathrm{eq}})^2], \tag{5.41}$$

which essentially coincides with the evolution of the density for a single scalar dark matter candidate with an effective annihilation cross section $\langle \sigma_{\mathrm{ann}} v\rangle/N$. The solution of (5.41) is a function $n_{\mathrm{tot}}(T)$ that decreases in time until the freeze-out temperature $T_{\mathrm{fo}}/m_s \sim 1/30 - 1/20$ that occurs when $n_{\mathrm{tot}}(T_{\mathrm{fo}}) \sim N H(T_{\mathrm{fo}})/\langle \sigma_{\mathrm{ann}} v\rangle$. From that time on S was out of equilibrium and $n_{\mathrm{tot}}(T)$ continued to decrease mainly due to the expansion of the universe. For all $T < T_{\mathrm{fo}}$ the energy density per unit entropy remains approximately constant and given by $\rho_S/s = m_s n_{\mathrm{tot}}/s = \sqrt{180/\pi} N (m_s/T_{\mathrm{fo}})/[\langle \sigma_{\mathrm{ann}} v\rangle M_{\mathrm{Pl}} \sqrt{g_*(T_{\mathrm{fo}})}]$, where $H(T) \sim \sqrt{g_*} T^2/M_{\mathrm{Pl}}$ and $g_* \sim 100$ (at decoupling the large number of S degrees of freedom contribute negligibly to the entropy). We can now compare the latter with the observed dark matter density, which itself is of order five times larger than that of

baryons, i.e. $\rho_{\rm DM}/s \sim 5 m_p \eta_b$ with $m_p \sim 1$ GeV the proton mass and $\eta_b \sim 10^{-10}$ the baryon density per unit entropy. Combining everything together we arrive at

$$\begin{aligned} \frac{\rho_S}{\rho_{\rm DM}} &\sim \frac{N}{\langle\sigma_{\rm ann} v\rangle M_{\rm Pl}} \frac{1}{m_p \eta_b} \\ &\gtrsim \frac{N}{\langle\sigma_{\rm ann,max} v\rangle M_{\rm Pl}} \frac{1}{m_p \eta_b} \sim \left(\frac{N}{50}\right)^3 \left(\frac{m_s}{100\ \text{GeV}}\right)^2 \left(\frac{10}{\lambda_{HS} N}\right)^2 . \end{aligned} \tag{5.42}$$

In many respects S behaves analogously to the singlet scalar dark matter ($N = 1$) studied extensively in the literature (see e.g. [253–255] and references therein). For example, similarly to what we saw for the thermal abundance, the annihilation rates relevant to indirect dark matter searches are the same as those of a single scalar up to a suppression $1/N$. As a consequence, once the cross section for annihilation into SM particles at freeze-out is fixed to the value required to explain the dark matter of the universe, indirect signatures are effectively the same as for the standard $N = 1$ case. The absence of observational evidence of such signatures already represents an important constraint on the model [253–255].

However, when considering direct detection signals the situation is much worse compared to the well-studied singlet scalar dark matter model because in our scenario the number of expected events is increased by a factor $N \gg 1$.

Direct detection experiments set limits on the cross section for spin-independent scattering of S off a nucleon n, which we write as $\sigma_{\rm DD}(S_i n \to S_i n)$ per each component i, and the signal is proportional to the total present-day number density: $N n_i \sigma_{\rm DD} = n_{\rm tot} \sigma_{\rm DD} = (\rho_S/m_s)\sigma_{\rm DD}$. An explicit calculation gives [253–255]

$$\sigma_{\rm DD} = \frac{\lambda_{HS}^2}{\pi m_s^2} \frac{\mu_{nS}^2 m_n^2}{m_h^4} f_N^2, \tag{5.43}$$

where μ_{nS} is the reduced mass of the S-nucleon system, m_n the nucleon mass, and $f_N \sim 0.3$ a nuclear form factor. Assuming S constitutes the totality of the total dark matter, the current experimental bound for a singlet of mass not too far from m_h roughly reads $\sigma_{\rm DD} < 10^{-46}$ cm^2, see [256]. This translates into

$$N \gtrsim 2 \times 10^3, \tag{5.44}$$

which is much larger than required for S to be the dark matter, see (5.42). In our minimal scenario a large N implies a small annihilation rate and consequently a larger abundance, and ultimately a signal stronger by a factor $\rho_S/\rho_{\rm DM}$ in direct detection experiments. More precisely, the direct detection cross section and the annihilation rate are related by $\sigma_{\rm DD} \sim (\mu_{nS}^2 m_n^2/m_h^4)\langle\sigma_{\rm ann,max} v\rangle$. The actual bound in our model is

$(\rho_S/\rho_{\rm DM})\sigma_{\rm DD} < 10^{-46}$ cm^2 and reads:

$$10^{-46}\ \text{cm}^2 > \frac{\rho_S}{\rho_{\rm DM}}\sigma_{\rm DD} \gtrsim N\frac{\mu_{nS}^2 m_n^2}{m_h^4}\frac{1}{m_p \eta_b M_{\rm Pl}} \sim N \times 10^{-44}\ \text{cm}^2, \tag{5.45}$$

which is impossible to satisfy for any $N \geq 1$ by at least a couple orders of magnitude.

We conclude that the minimal model (5.7) is not compatible with direct dark matter searches. In the reminder of this section we will propose three classes of extensions. We will entertain the possibility of adding new light particles that constitute the dark matter (Section 5.2.2), or breaking the stabilizing $O(N)$ symmetry, thereby making S decay on cosmological scales and completely removing the dark matter from the model (Section 5.2.2), or finally introducing a new confining force (Section 5.2.3). The main lesson to be drawn from the following discussion is that it is possible to build simple extensions of the toy model (5.7) that are compatible with all observations. However, at this earlier stage we have no reason to prefer one of the following realizations over any other, so a detailed analysis of concrete scenarios is left for future work. It is nevertheless worth mentioning that the additional particles introduced in realistic models may affect the numerical analysis in Section 5.1.5, so that a proper model-dependent analysis of the finite-T effective potential may be required. Furthermore, the richer structure of the realistic models potentially makes the prospects of detection a bit more exciting compared to the original model of eq. (5.7).

Extension 1: adding a dark matter candidate

As our first extension we introduce new fermions that will play the role of the dark matter. For definiteness we promote S to the traceless symmetric 2-index representation of a global $SO(n)$. With this slight modification the expression S^2 in the potential (5.7) should be interpreted as $\text{tr}[S^2]$; in addition another quartic $\lambda_S'\text{tr}[S^4]$ as well as a new soft trilinear $\mu_D\text{tr}[S^3]$ are in principle allowed. With $\mu_D^2 n \ll 16\pi^2 m_t^2$ (this condition may be enforced by an approximate Z_2 symmetry) and $\lambda_S' \ll \lambda_S n$ the results of the previous sections, in particular the suppression of sphalerons at high T, are left unchanged provided we identify $N = n(n+1)/2 - 1$.

Now, if no new ingredients are added, our estimate (5.42) shows that $\rho_S/\rho_{\rm DM} \gtrsim 1$ for $N \gtrsim 50$. Our plan here is to render S unstable and let its decay products be the dark matter. This we achieve adding a pair of Weyl fermions χ with mass $m_\chi < m_s/2$ in the fundamental of $SO(n)$, and introducing the coupling $y\chi_i\chi_i S_{ij}$. Note that χ_i is exactly stable due to the $SO(n)$ symmetry and an accidental Z_2. In this new set up $S \to \chi\chi$ would quickly deplete the scalar population at a temperature $T_{\rm decay} \lesssim m_s$, provided at this scale the decay rate is comparable or larger than $H(T = T_{\rm decay})$. We will assume $T_{\rm decay} \gtrsim T_{\rm BBN} \sim 1$ MeV to avoid affecting BBN. This corresponds to requiring $y^2 \gtrsim 4\pi\sqrt{g_*}T_{\rm BBN}^2/(m_s M_{\rm Pl})$.

With a sizable y, however, χ gets thermalized at $T \gtrsim m_s$ via reactions such as $\Gamma(\chi\chi \leftrightarrow SS), \Gamma(\chi\chi \leftrightarrow SSS)$, the latter being mediated by an off-shell S; having no lighter states to decay into, nor efficient annihilation channels, the present-day relic χ density would then be way too large to be identified with the dark matter. To ensure χ be a compelling dark matter candidate we impose the conservative conditions $\Gamma(SSS \to \chi\chi) \sim y^2\lambda_S^2 n^3 T/(4\pi)^5 \lesssim H(m_s)$, $\Gamma(SS \to \chi\chi) \sim y^4 n^2 T/(4\pi)^3 \lesssim H(m_s)$, which guarantee χ does not thermalize, and assume its primordial number density is negligible. In practice, for a coupling in the range

$$4\pi\sqrt{g_*}\frac{T_{\rm BBN}^2}{m_s M_{\rm Pl}} \lesssim y^2 \lesssim \min\left[4\pi\sqrt{g_*}\frac{m_s}{M_{\rm Pl}}\frac{\sqrt{N}}{\epsilon_S^2}, 4\pi\sqrt{\sqrt{g_*}\frac{m_s}{M_{\rm Pl}}\frac{4\pi}{N}}\right] \tag{5.46}$$

χ was never thermalized at any $T < \Lambda$ and the present-day density is entirely controlled by $S \to \chi\chi$. The decay will generate a density of order $\rho_\chi/s \sim (\rho_S/s)(m_\chi/m_s)$, where the suppression of m_χ/m_s arises from the fact that the energy density of χ approximately scales as radiation from $T \sim m_s$ down to $T \sim m_\chi$, and only after that threshold χ started to behave as cold dark matter. Choosing the masses such that $(\rho_S/\rho_{\rm DM})(m_\chi/m_s) \sim 1$ the fermion can thus be identified with the dark matter. The phase space bound $m_\chi \gtrsim$ keV [257] combined with (5.42) tell us that the initial ρ_S cannot be arbitrarily large: the viable regime is in practice limited to $N \lesssim 10^3 - 10^4$. The conflict with direct dark matter searches that the minimal model suffers from are here evaded: there is virtually no hope of detecting χ because all interactions between the dark matter and the SM are mediated by the Planck-suppressed coupling y, see (5.46).

Extension 2: breaking $O(N)$ softly

Our second class of extensions has an unstable S, there is no dark matter candidate and the direct/indirect detection constraints are trivially satisfied.

Let us consider for instance the most general set of soft terms we can add to (5.7):

$$\delta V_{\rm soft} = a_i\frac{\mu^3}{g_S}S_i + b_{ij}\frac{\mu^2}{2}S_iS_j + c_{ijk}g_S\mu S_iS_jS_k + d_ig_S\mu S_i|H|^2, \tag{5.47}$$

where μ is a mass scale, g_S a coupling and $a_i, b_{ij}, c_{ijk}, d_i$ arbitrary real matrices. We assume that $a_i, b_{ij}, c_{ijk}, d_i$ are of order unity and $\mu \ll |m_H|, m_S$, so that our analysis of the physics at finite temperature in the $O(N)$ symmetric model is not affected by the soft terms in (5.47). Our main task here is to identify what values of μ are large enough to trigger S decay on cosmological scales and simultaneously small enough to be consistent with collider constraints.

Note that a_i, d_i are invariant under two (in general inequivalent) $O(N-1) \subset O(N)$, and when combined leave an $O(N-2)$ invariant. This means that in their presence only

two components S_i will be able to decay while additional soft terms are necessary to destabilize the remaining $N-2$ ones. Similarly, b_{ij} leaves intact a Z_2^N subgroup and is not enough by itself for our purpose either. Finally, c_{ijk} breaks $O(N)$ completely.

The field fluctuations δh and $\delta S_i = S_i - \langle S_i \rangle$ obtain a mass matrix that at leading order in $\mu/\Delta m$, where $\Delta m^2 \equiv m_h^2 - m_s^2$, has the following structure

$$\sim \begin{pmatrix} m_h^2 + \mathcal{O}(\mu^2) & d_i \mu g_S v \\ d_i \mu g_S v & \delta_{ij} m_s^2 + b'_{ij} \mu^2 \end{pmatrix}, \tag{5.48}$$

where $b'_{ij}\mu^2 = b_{ij}\mu^2 + 6 c_{ijk} g_S \mu \langle S_k \rangle$ with $\langle S_i \rangle \sim d_i \mu g v^2/m_s^2 + \mathcal{O}(\mu^3)$. The diagonalization may be done in two steps: we first act with a rotation of angle $\sim d_i \mu g_S v/\Delta m^2$ to reduce (5.48) to a block diagonal form and then diagonalize the $N \times N$ block through a $O(N)$ orthogonal matrix R. Carrying out this procedure we find the mixing angles between δS_i and the Higgs are given by

$$\sin\theta_i \sim R_{ij} d_j \frac{\mu g_S v}{\Delta m^2}. \tag{5.49}$$

Assuming that the order one coefficients a_i, b_{ij}, c_{ijk} and d_i are independent, we expect b'_{ij} is also independent from d_i. Furthermore, we choose the coupling g_S such that $(g_S v)^2 \lesssim \Delta m^2$. [13] In this case *all* physical components δS_i have a non-vanishing mixing with the Higgs that allow for their decay into pairs of SM fermions or virtual gauge bosons. The decay width is of order $\Gamma_i \sim \sin^2\theta_i \Gamma_{\rm tot}$, where we used the fact that $m_{s_i} \sim m_h$ is a reasonable approximation and denoted by $\Gamma_{\rm tot}$ the Higgs width. In order not to spoil the successes of standard BBN cosmology we have to require all singlet components are removed from the plasma well before temperatures of order $T_{\rm BBN} \sim 1$ MeV. Since $\Gamma_{\rm tot} \sim$ MeV this gives $\theta_i^2 \gg \sqrt{g_*}{\rm MeV}/M_{\rm Pl} \sim 10^{-21}$.

There is potentially another concern to take into account, however. If the scalars decay when their energy density dominate the universe, then a large entropy can be injected, in turn leading to a dilution of the baryon asymmetry. To avoid this we will conservatively require all singlet components decay before the temperature T_* at which $\rho_S(T) = \rho_S(T_{\rm fo})(T/T_{\rm fo})^3$ starts to dominate over radiation. Recalling that in a standard cosmology matter-radiation equality takes place at around $T_{\rm eq} \sim 1$ eV, using (5.42) and the benchmark values $\lambda_{HS} N = 10$, $m_s = 100$ GeV, we estimate

$$T_* \sim T_{\rm eq} \left(\frac{N}{50}\right)^3 \sim 1 \text{ MeV} \left(\frac{N}{5000}\right)^3. \tag{5.50}$$

For the values of N considered in this study $T_* < T_{\rm BBN}$, so this request is weaker than

[13] If $(g_S v)^2 \gtrsim \Delta m^2$ the b' contribution can become negligible compared to the off-diagonal terms in (5.48). In this case the mass matrix would be nearly degenerate and the rotation R would orient d_i in the direction of a single component of the δS_i fields, allowing only for its mixing with an angle $\sim \sqrt{N} \mu g_S v/(\Delta m^2)$, where we used the fact that a vector with random entries of order unity has $|d| \sim \sqrt{N}$. Since we want all the components of S to decay, we choose to stay away from this region.

demanding BBN remains standard.

At the same time, bounds from collider experiments tell us that the mixing cannot be too large. In fact, too large mixing angles would imply the physical Higgs has couplings to the SM particles suppressed with respect to those predicted by the SM by a factor $\cos\theta$, with

$$\cos^2\theta \sim 1 - \sum_i \sin^2\theta_i \sim 1 - N\left(\frac{\mu g_S v}{\Delta m^2}\right)^2. \tag{5.51}$$

To be consistent with experimental data, we require $N(\mu g_S v/\Delta m^2)^2 \ll 10\%$.

The previous two conditions are satisfied in the region

$$10^{-21} \ll \left(\frac{\mu g_S v}{\Delta m^2}\right)^2 \ll \frac{10^{-1}}{N}. \tag{5.52}$$

We see that with values $N \sim 100 - 1000$ there is a large portion of parameter space that allows us to remain compatible with both dark matter and collider bounds.

5.2.3 A more ambitious fix: confining S and composite dark matter

In this section we will explore a scenario where S is promoted to the anti-symmetric representation of a dark $SO(n)$ gauge group. Contrary to the previous model, the gauge symmetry here forbids a scalar trilinear and the renormalizable Lagrangian remains invariant under an accidental Z_2 symmetry:

$$S \to -S. \tag{5.53}$$

We will assume that the symmetry (5.53) is preserved by our UV completion at the scale $\sim \Lambda$. We do this in order to obtain a qualitatively different scenario compared to the one discussed in Section 5.2.2. [14]

Our earlier discussion on symmetry non-restoration goes through qualitatively unchanged provided we identify $N = n(n-1)/2 \sim n^2/2$. Yet, the calculation of the relic abundance of S should take into account a new ingredient: the dark gauge bosons.

We estimate the population of S at freeze-out observing that the dominant annihilation

[14] In the absence of a Z_2 symmetry S is expected to decay into exotic gauge bosons and SM particles. For example, taking S in the 2-index symmetric of $SO(n)$ allows a trilinear coupling S^3, whose main effect would be to trigger a (1-loop) width into dark gluons at the renormalizable level. The Z_2-symmetric scenario discussed in this section reveals a richer phenomenology, which we find interesting to investigate.

channel is into dark gauge bosons:

$$\frac{\rho_S}{s} \sim \frac{8\pi m_s^2}{g_D^4(m_s)} \frac{m_s/T_{\rm fo}}{M_{\rm Pl}\sqrt{g_*(T_{\rm fo})}}, \tag{5.54}$$

where the relevant number of relativistic degrees now should include the dark gluons as well, i.e. $g_*(T_{\rm fo}) \sim n^2/2 \sim N$. With $m_s \sim 100$ GeV, a relic density comparable to the dark matter is obtained for $g_D^4(m_s)\sqrt{N} \sim 10^{-3} - 10^{-2}$. The present-day abundance of S may be however significantly affected due to novel non-perturbative effects at $T \ll m_s$. In fact, at temperatures of order the strong coupling scale, which we may estimate using the 1-loop beta function as $m_D \sim m_s \exp[-24\pi^2/(21 g_D^2(m_s)(n-2))] \ll m_s$, the dark gluons and S confine into $SO(n)$-singlet states with large self-couplings. It is therefore important to understand what implications this new phase can have on the relic density of the new particles.

First of all, let us consider the bound states — with masses $\sim m_D$ — that are formed by pure gluons, which we will refer to as dark glueballs. The CP-even dark glueballs are unstable and can decay into light SM particles via the effective interaction

$$\frac{g_D^2(n-2)}{96\pi^2} \frac{\lambda_{HS}}{m_s^2} H^\dagger H F_D^{\mu\nu} F_{D\mu\nu}, \tag{5.55}$$

where $F_D^{\mu\nu}$ and g_D are the field strength and coupling of the dark gauge group. Denoting by Φ the interpolating field for the CP-even dark glueballs, the decay may be estimated observing that at the confinement scale naive dimensional analysis — combined with a large-n counting — suggests $g_D^2(m_D)n \sim 16\pi^2$ and $F_D^{\mu\nu}F_{D\mu\nu} \sim m_D^3 \Phi n/4\pi$. In practice the operator (5.55) interpolates a mixing of order $\theta \sim (\lambda_{HS}\sqrt{N} v m_D^3)/(4\pi m_s^2 m_h^2) \ll 1$ between the CP-even dark glueballs and the Higgs boson. The main decay channels the dark glueballs acquire are into SM fermions f, with a width of order $\Gamma(\text{dark glue} \to f\bar{f}) \sim \theta^2 m_D (m_f/v)^2/(4\pi)$.

We conservatively require the glueballs disappear as soon as they become non-relativistic. This request is motivated by the following consideration. At early times the universe expansion is controlled by dark radiation due to the large number $g_* \sim N$ of new degrees of freedom. By continuity we therefore expect the density of non-relativistic glueballs would dominate over SM radiation for a long time before they decay. If this was the case we would experience a significant dilution of η_b due to either a large entropy production from their out-of-equilibrium decay or a prolonged period of dark glueball-dominated expansion. We evade these undesirable complications conservatively requiring the glueballs disappear as soon as they formed. This gives the order of magnitude bound $m_D \gtrsim (N/1000)^{1/5}(m_f/\text{GeV})^{-2/5}$ GeV, which tells us that the dark dynamics should confine at a scale not too far from the GeV. The proximity between the required m_D and the weak scale, or more precisely the S mass, implies the new gauge symmetry should be rather strong already at $T \sim m_s$. Using our 1-loop estimate of m_D this requirement

approximately reads $g_D^2(m_s)n \sim 3$.

The lightest CP-odd dark glueball is presumably heavier than the CP-even one, and can therefore annihilate into the CP-even states. The associated cross section times initial relative velocity (roughly of order $v_{\rm in} \sim \sqrt{T/m_D} \sim 1$ at the temperatures relevant to this discussion) is $\sim \pi/[N^2 m_D^2]$. The power of N arises from the fact that trilinear (quartic) glueball couplings are of order $1/n$ ($1/n^2$), so the amplitude squared scales as $(1/n^2)^2 \sim 1/N^2$. For the typical parameters we are interested in, $m_D \sim 1$ GeV and $N \sim 10^3$ (see (5.44)), these processes should be efficient enough to deplete the CP-odd glueball abundance below the observed dark matter. The resulting relics would also eventually decay into SM particles via CP-violating loops involving the Higgs and SM fermions and gauge bosons. The time scale could be quite large, causing potential trouble at later stages of cosmological evolution (BBN, CMB, etc.). Without resorting to a detailed analysis, we notice however that all problems can be avoided by assuming the topological vacuum angle $\bar{\theta}_D$ of the dark gauge group is non-vanishing. In that case the CP-odd glueballs mix with angle $\bar{\theta}_D$ with the CP-even ones and decay on a time scale $1/[\bar{\theta}_D^2 \Gamma(\text{dark glue} \to f\bar{f})] \lesssim 10^{10}[10^{-8}/\bar{\theta}_D]^2$years. [15]

The introduction of the dark gauge bosons opens a new decay channel for the Higgs; because we have seen that the dark glueballs typically have long lifetimes, this new channel corresponds to an invisible width for the Higgs and we have to make sure it is not too large. In the limit $m_s \gg m_h$, and neglecting radiative dark gluon corrections for simplicity, we obtain:

$$\Gamma(h \to \text{dark gluons}) = \left(\frac{g_D^2(n-2)}{16\pi^2}\right)^2 \frac{m_h^3}{288\pi} \left(\frac{\lambda_{HS} v}{m_s^2}\right)^2 n(n-1). \tag{5.56}$$

As an order of magnitude estimate, we will adopt the very same expression even in the realistic regime $m_s \sim m_h$. Taking as representative input values $m_s = 100$ GeV, $|\lambda_{HS}|N = 10$, and $g_D^2(m_s)(n-2) = 3$ (as required to obtain a confinement scale around the GeV), we find that (5.56) is $< 10\%$ of the total Higgs width as soon as $n \gtrsim 20$, i.e. $N \gtrsim 200$.

Reannihilation after confinement and Sglueball dark matter

Next, we want to understand what happens to the population of exotic scalars. S confines into glueball-like states of mass $\sim m_s$ which belong to one of two classes of hadrons. In the first class we find mesonic states, like $\text{Tr}[SS]$, that carry no conserved Z_2 charge. These are Coulombian in nature, analogously to the quarkonium in QCD, have mass $2m_s - \mathcal{O}(\alpha_D^2 m_D)$ and a characteristic size of order the Bohr radius $r_{SS} \sim 1/[\alpha_D m_s]$. Here

[15] It is easy to convince oneself this additional source of CP violation does not affect baryogenesis: it becomes potentially relevant only after sphalerons have already shut off and no B-violating process is active.

$\alpha_D = g_D^2 n/(4\pi) \sim 0.2 - 0.3$ is the 't Hooft coupling renormalized at a scale of order r_{SS} — somewhere in between $1/m_D$ and $1/m_s$. The absence of any conserved charge implies they quickly decay, with rates $\sim \alpha_D^5 m_s$ similarly to charmonium, essentially removing part of the original S's from the plasma. The other class of S-hadrons is built out of an odd number of S's, and carry a non-trivial charge (5.53). The stable configuration, which we will call *Sglueball*, is composed of one S and gluons. It has mass $\sim m_s$ and binding energy $\sim m_D$. Its stability, however, does not immediately guarantee that the *Sglueball*s represent a significant fraction of the dark matter, as they can efficiently re-annihilate at late times through the strong interaction process

$$Sglueball + Sglueball \to \ [SS] + \text{glueballs}. \tag{5.57}$$

followed by the cosmologically fast decay of the mesons into glueballs.

To establish how many *Sglueball*s are left today we need to quantify how efficient the latter process is. A crude estimate can be obtained by a simple dimensional argument. At small n the only available parameter is m_D, which controls both the size of the *Sglueballs* and the mass splittings. In this limit the cross section is thus expected to be geometrical $\sim \pi/m_D^2$. At large n the trilinear coupling between two *Sglueballs* and unstable mesons $[SS]^*$ scales like $1/n$, while the spectrum is basically fixed. We thus expect that, as long as the resonant processes *Sglueballs*+*Sglueballs*$\to [SS]^*$ are kinematically accessible, the cross section simply features a reduction by $1/n^2$: $\sigma_{Sg} \sim \pi/(m_D^2 n^2)$. If the resonant reactions are not allowed, on the other hand, the generic large-n expectation is that $\sigma_{Sg} \sim \pi/(m_D^2 n^4)$, corresponding to a two body ($[SS]$ + glueball) final state.

We can understand these claims more quantitatively by considering the partial wave decomposition of the annihilation cross section

$$\sigma_{fi} = \frac{\pi}{k_i^2} \sum_{\ell=0}^{\infty} (2\ell+1) |S_{fi}^{(\ell)}|^2 \tag{5.58}$$

where by i we indicate the initial two *Sglueball* state, by k_i their center of mass momentum and by f any state they could annihilate into. At large n we expect $S_{fi}^{(\ell)}$ to be well described by Breit-Wigner amplitudes mediated by narrow $[SS]^*$ meson states

$$|S_{fi}^{(\ell)}|^2 \sim \sum_{r_\ell} \frac{\Gamma_i^{(\ell,r_\ell)} \Gamma_f^{(\ell,r_\ell)}}{(E_i - M^{(\ell,r_\ell)})^2 + (\Gamma_{\rm tot}^{(\ell,r_\ell)}/2)^2}, \tag{5.59}$$

with r_ℓ any additional label identifying the intermediate states. The spectrum and the widths are controlled by m_D and by n. In particular, the typical expectation for the *Sglueball* and $[SS]^*$ masses is respectively $M_{Sg} = m_s + c_{Sg} m_D$ and $M^{(\ell,r_\ell)} = 2m_s + c^{(\ell,r_\ell)} m_D$, with the c's $O(1)$ coefficients. Given the initial energy $E_i = 2\sqrt{M_{Sg}^2 + k_i^2}$

the Breit-Wigner is thus controlled by

$$E_i - M^{(\ell,r_\ell)} \simeq (2c_{Sg} - c^{(\ell,r_\ell)})m_D + \frac{k_i^2}{M_{Sg}} \equiv \bar{c}^{(\ell,r_\ell)} m_D + E_k \tag{5.60}$$

with $E_k \sim T$ the kinetic energy of the initial *Sglueball* pair. On the other hand, $\Gamma_i^{(\ell,r_\ell)} \equiv \Gamma([SS]^* \to Sglueball{+}Sglueball)$ and $\Gamma_f^{(\ell,r_\ell)} \equiv \Gamma([SS]^* \to [SS]{+}glueballs)$, as well as the total width, are all of order $\Gamma \sim m_D/n^2$. Now, as explained below, reannihilation is dominantly taking place around $T \sim m_D/(20 \div 30)$ so that the thermal breadth of the energy denominator (5.60) is larger than the resonance width as soon as $n^2 \gtrsim 30$. Thus, in our case, resonant annihilation happens for resonances that are within the thermal breadth of the initial state energy. Assuming resonances are spaced by $O(m_D)$ the *probability* for that to happen is $p_r \sim T/m_D$. Such resonant partial waves give a contribution to the cross section

$$\sim \frac{\pi}{k_i^2}(2\ell+1)\frac{\Gamma}{T} \sim \frac{\pi}{k_i^2}(2\ell+1)\frac{m_D}{T}\frac{1}{n^2} \tag{5.61}$$

while non-resonant partial waves give

$$\sim \frac{\pi}{k_i^2}(2\ell+1)\frac{\Gamma^2}{m_D^2} \sim \frac{\pi}{k_i^2}(2\ell+1)\frac{1}{n^4}. \tag{5.62}$$

To complete our estimate we should finally take into account that, given the initial momentum k_i and given the typical range $1/m_D$ of the interaction, we expect only partial waves up to $\ell_{max} \sim k_i \times 1/m_D$ to significantly contribute. A fraction $p_r \sim T/m_D$ of such waves will be resonant, so that, as long as $p_r \ell_{max} \gg 1$, we statistically expect a number $\sim \ell_{max} T/m_D$ of resonant channels. Combined with eq.(5.61), this leads to an estimate

$$\sigma_{fi}\Big|_{p_r\ell_{max}\gg 1} \sim \frac{\pi}{k_i^2}\ell_{max}^2\frac{1}{n^2} \sim \frac{\pi}{m_D^2}\frac{1}{n^2}\,. \tag{5.63}$$

On the other hand, for $p_r\ell_{max} \lesssim 1$ there is a good chance no channel is resonant, so that, according to eq.(5.62) a more likely estimate for this case is

$$\sigma_{fi}\Big|_{p_r\ell_{max}\lesssim 1} \sim \frac{\pi}{k_i^2}\ell_{max}^2\frac{1}{n^4} \sim \frac{\pi}{m_D^2}\frac{1}{n^4}, \tag{5.64}$$

where the two powers of angular momentum arise from summing over all partial waves as indicated in (5.58). The parameters of our scenario make non-resonant annihilation much more plausible. Indeed $\ell_{max} \sim \sqrt{Tm_s}/m_D$ so that

$$p_r\ell_{max} \sim \left(\frac{T}{m_D}\right)^{3/2}\left(\frac{m_s}{m_D}\right)^{1/2} \sim 0.1 \tag{5.65}$$

Nonetheless, without a complete non perturbative control of our model, it seems impossible to reach a definite conclusion. We shall thus discuss both possibilities.

Having obtained a rough estimate of the cross section for (5.57) (see (5.61) and (5.62)) we can proceed to calculate the present-day abundance of *Sglueballs*. The latter is the result of a complicated system of coupled Boltzmann equations involving the densities of *Sglueballs*, mesons, and glueballs. Yet, we expect the dominant phase of S depletion takes place at temperatures when the inverse reactions in (5.57) are negligible. A number of considerations suggest that the critical temperature T_D at which this condition starts to hold lies in the range [16] $m_D/(20 \div 30) < T_D < m_D$. In the following we will therefore adopt the intermediate value $T_D = m_D/10$ as a reference. For $T < T_D$ the *Sglueballs* yield Y approximately follows the equation $dY/dt = -s\langle\sigma_{Sg} v_{\rm in}\rangle Y^2$. We found that $\langle\sigma_{Sg} v_{\rm in}\rangle \sim \sqrt{T/m_s}\pi/(N^a m_D^2)$, where $a = 2$ generically and $a = 1$ for the special case of resonant annihilation. The current energy density encapsulated in the *Sglueballs* is finally estimated as

$$\begin{aligned}\frac{\rho}{s} \sim m_s Y_0 &= \frac{m_s Y_D}{1 + Y_D \int_{t_D}^{t_0} dt\; s\langle\sigma_{Sg} v_{\rm in}\rangle} \sim \frac{m_s}{\int_{t_D}^{t_0} dt\; s\langle\sigma_{Sg} v_{\rm in}\rangle} \sim \frac{N^a m_D m_s}{\pi M_{\rm Pl}\sqrt{g_*(T_D)}} \frac{m_D}{T_D}\sqrt{\frac{m_s}{T_D}} \\ &\sim 10^{3(a-2)}\, \frac{\rho_{\rm DM}}{s} \left(\frac{N}{10^3}\right)^a \left(\frac{m_D}{1\text{ GeV}}\right)^2 \left(\frac{m_s}{100\text{ GeV}}\right) \left(\frac{0.1\text{ GeV}}{T_D}\right)^{3/2}, \end{aligned} \tag{5.66}$$

where we have taken the *Sglueball* mass to be m_s and neglected their small binding energy $\mathcal{O}(m_D) \ll m_s$. The primordial *Sglueballs* yield at T_D, denoted by Y_D, is typically a fraction of order unity of the freeze-out S population. In the second line of (5.66) we used the fact that $\int_{t_D}^{t_0} dt\; s\langle\sigma_{Sg} v_{\rm in}\rangle > 1/Y_D$ for values of N satisfying $N \ll 10^8$. The main message to be qualitatively inferred from (5.66) is that the *Sglueballs* may be the dark matter. This conclusion may however be invalidated in the (unlikely) presence of a resonance (i.e. if $a = 1$).

An extension with mesonic dark matter

Here we show that there exists a minimal variation of the present model in which the issue of resonant vs non-resonant annihilation does not arise and the new scalars can be robustly argued to be the dark matter.

We introduce another scalar S' of mass $m_{s'} \sim m_s$ in the adjoint of $SO(n)$ preserving a $Z_2 \times Z_2'$ symmetry, and argue that the $[SS']$ meson becomes a potential dark matter candidate. These are quarkonium-like states, with binding energy $\sim \alpha_D^2 \mu_{SS'}$ and size $r_{SS'} \sim 1/[\alpha_D \mu_{SS'}]$, where $\mu_{SS'}$ is the reduced mass of the SS' system. Because of the sizable binding energy, $\alpha_D^2 \mu_{SS'} > m_D$, the decay into a "*Sglueball*+S'*glueball*" pair is

[16] The upper bound may be estimated as follows. Because the $[SS]$ mesons decay instantaneously (in Hubble units) as soon as $T < m_D$, one may naively expect that the inverse reactions in (5.57) would be very unlikely already at $T < m_D$. However, during meson de-excitation and decay many glueballs are injected in the bath, and these latter may occasionally find an unstable meson and dissociate it back into *Sglueballs*. This tells us that T_D must be safely below the glueballs mass scale. A lower bound on T_D may instead be identified observing that if we wait a bit longer, at $T < m_D/(20 \div 30)$, the glueballs have frozen out and the rate for the inverse (5.57) reaction is smaller than the expansion rate.

expected to be kinematically forbidden. As a consequence, the $[SS']$-meson is exactly stable. The population of our dark matter candidate is mainly governed by annihilation into pairs of unstable $SS, S'S'$ mesons (plus possibly dark glueballs):

$$[SS'] + [SS'] \to [SS] + [S'S'] \tag{5.67}$$

For $m_{s'} \neq m_s$ this channel is always energetically favored because the masses $M_{SS,S'S',SS'}$ of the mesons satisfy $\Delta M \equiv (M_{SS'} + M_{SS'}) - (M_{SS} + M_{S'S'}) \sim \frac{\alpha_D^2}{2} \frac{(m_s - m_{s'})^2}{m_s + m_{s'}} > 0$. The reaction rate importantly depends on the final phase space, which is simply determined by the relative velocity of the mesons in the center of mass frame. Indicating by $\mu_{SS'}$ the reduced mass of the final state, in the non-relativistic limit we have

$$v_{\rm fin}^2 = v_{\rm in}^2 \frac{M_{SS'}}{2\mu_{SS'}} + 2\frac{\Delta M}{\mu_{SS'}} \simeq v_{\rm in}^2 \frac{(m_s + m_{s'})^2}{4 m_s m_{s'}} + \alpha_D^2 \frac{(m_s - m_{s'})^2}{m_s m_{s'}} [1 + \mathcal{O}(\alpha_D^2)]. \tag{5.68}$$

Of course we have $v_{\rm in} = v_{\rm fin}$ when $m_{s'} = m_s$, otherwise $v_{\rm fin} \sim \alpha_D$. The rate for (5.67) is dominated by the s-wave channel (the typical angular momentum is indeed $\ell \sim \mu_{SS'} v_{\rm in} r_{SS'} \sim v_{\rm in}/\alpha_D \ll 1$). The cross section is therefore set by the Bohr radius of the initial states, $\sigma_D \propto (\pi r_{SS'}^2)(v_{\rm fin}/v_{\rm in})$, up to a ratio of final and initial relative velocities arising from the incoming flux ($\propto v_{\rm in}$) and the phase space ($\propto v_{\rm fin}$). In addition, since these are genuinely $2 \to 2$ reactions there is a $(1/n^2)^2$ factor from large-n counting. Overall we obtain that the quantity relevant to our calculation, i.e. an average of the cross section times relative velocity of the initial states, is given by

$$\langle \sigma_D v_{\rm in} \rangle \sim \pi r_{SS'}^2 \times \langle v_{\rm fin} \rangle \times \frac{1}{N^2} \sim \frac{\pi}{N^2 \alpha_D \mu_{SS'}^2} \times \begin{cases} 1 & m_{s'} \neq m_s \\ \frac{\langle v_{\rm in} \rangle}{\alpha_D} & m_{s'} = m_s. \end{cases} \tag{5.69}$$

This is a factor $m_D^2/(\alpha_D \mu_{SS'}^2) < 1$ ($m_{s'} \neq m_s$) or $\langle v_{\rm in} \rangle m_D^2/(\alpha_D^2 \mu_{SS'}^2) < 1$ ($m_{s'} = m_s$) times the result (5.62) obtained above for the nonresonant *Sglueball* annihilation. A calculation similar to (5.66), which may be repeated here with $T_D \sim m_D$, then suggests that the stable $[SS']$ meson may be the totality of the dark matter for the typical choice $\mu_{SS'} \sim 50$ GeV, $m_D \sim 1$ GeV, $N \sim 10^3$ (consistent with (5.44)), and finally $\alpha_D \sim 0.2 - 0.3$. (Note that the term Y_D in the denominator of (5.66) is not negligible if the abundance at freeze-out was already comparable to that of the present-day dark matter.) The above estimates are only qualitative, however. We plan to come back to this interesting topic in the future.

We have thus shown that the introduction of a dark gauge symmetry allows us to evade the dark matter constraints consistently with all current bounds. We should however warn the reader that, as opposed to the extensions discussed in Section 5.2.2 and 5.2.2, these models introduce novel elements that can affect our analysis of the thermal effective potential. In particular, the sizable $g_D^2(m_s)n \sim 3$ required to ensure a fast enough

decay rate for the dark glueballs also generates a non-negligible positive contribution to the thermal S mass and thus tends to suppress the negative corrections to $m^2_{h,0}(T)$. Simultaneously, a non-negligible gauge coupling helps pushing the Landau pole of λ_S to higher scales, facilitating an extension of our models beyond $\Lambda > 100$ TeV. While our earlier results on the effective Higgs potential at finite T will remain qualitatively correct, a careful assessment of $N_{\rm min}$ should take these novel elements into account.

5.3 Comments on alternative models

We would like to argue now that the same qualitative conclusions we found for the $O(N)$ model extend to any large-N realization of the low energy sector. Furthermore, it is possible to build small N sectors by introducing new EW-charged scalars.

5.3.1 Large-N models

We have found that a large-N dynamics offers a tractable framework in which a parametrically large Higgs VEV $h(T)/T$ can be obtained while maintaining perturbativity. It is a simple exercise to show that essentially all we have seen for the specific $O(N)$ model of (5.7) generalizes to a wider class of nearly-conformal large-N dynamics coupled to $H^\dagger H$. The interaction does not necessarily involve fundamental scalars, and could in principle be built out of a fermionic pair or a gauge field strength. In these latter cases it may thus be possible to relax the stability constraint, that has played a key role in our analysis, and even avoid the introduction of a new hierarchy problem if the new coupling to the Higgs is non-renormalizable, i.e. it has dimension $d \gtrsim 4$. Unfortunately, in order to overcome the resulting suppression $(T/\Lambda)^{d-4}$ in the thermal Higgs mass — and thus maintain the $B+L$ asymmetry down to low T — we would be forced to take extremely large values of N, which is in tension with the considerations made below (5.38). Overall, the option of a non-renormalizable coupling to $H^\dagger H$ seemed less convincing to us.

Even sticking to scenarios in which a renormalizable coupling to $H^\dagger H$ is built out of fundamental scalars, strictly speaking there is no necessity of having an approximate $O(N)$ symmetry: all is needed for our program to be realized is many degrees of freedom coupled to the Higgs mass operator. Consider for instance a model with N scalars S_i and potential

$$V_4 = \lambda'_{HS} H^\dagger H \sum_i S_i^2 + \frac{\lambda'_S}{4} \sum_i S_i^4 + \lambda' \sum_{ij} S_i^2 S_j^2, \tag{5.70}$$

where we may assume $\lambda' \sim \lambda'^2_{HS}/16\pi^2 \ll \lambda'_S$ for simplicity. In this alternative model, while we still have $\delta m_H^2 \sim \lambda'_{HS} N T^2/12$ at high T, the N dependence of the constraints from stability and perturbativity are modified. Yet the analogue of eq. (5.20) giving the lower bound on N is remarkably unaffected with respect to the original model of eq. (5.7).

Indeed the loop expansion parameter for the S quartic is now given by $\epsilon'_S \equiv \lambda'_S/16\pi^2$ which does not carry any N dependence. On the other hand considering the potential along the direction $S_i^2 \sim S_j^2$ one obtains the stability constraint $(\lambda'_{HS}N)^2 < \lambda_H(\lambda'_S N)$. This second result can be conveniently written as

$$N \gtrsim \frac{(\lambda'_{HS}N)^2}{\lambda_H \lambda'_S} \tag{5.71}$$

which leads to the same constraint of eq. (5.20), once we notice that the loop expansion parameter is now $\epsilon'_S = \lambda'_S/16\pi^2$ rather than $\epsilon_S = \lambda_S N/16\pi^2$. The parameter ϵ'_S also practically replaces ϵ_S when computing the effective potential, the S_i thermal masses in particular. So we again find no qualitative new feature arising from (5.70). This simple alternative shows that, in scenarios where our mechanism is realized through couplings to the Higgs bilinear $H^\dagger H$, the lower bound on N is structurally robust.

It would be interesting to investigate other applications for the large number of degrees of freedom coupled to $H^\dagger H$ that these models feature. A possible direction to explore is a potential connection with the scenario of [245][258], where $N \gg 1$ scalars are introduced to address the cosmological constant problem and the hierarchy problem respectively. See also [259] for a connection to the dark matter problem.

5.3.2 Models with smaller N

It is possible to realize simple perturbative scenarios that support significantly smaller values of N. The key point is that we do not necessarily need to couple the new light degrees of freedom, say S, to the SM Higgs at all. Our mechanism would work equally well if we introduced a new EW-charged scalar that interacts with S via a negative quartic, but with the SM mainly via the EW force. As a concrete example we may introduce an inert H' at the weak scale, having the same SM charges as H but carrying an unbroken Z_2 parity. It is the latter scalar that acquires a large VEV at finite T and suppresses $B + L$ washout. Still, its zero-temperature mass squared may be chosen to be positive so that H' does not contribute to EW symmetry breaking today, and all precision EW observables remain essentially unaffected.

The discussion of Sections 5.1.3, 5.1.4, 5.1.5 applies to this scenario as well provided we make a few important adjustments. First, the role of the Higgs boson is now played by H', so its quartic $\lambda_{H'}$ is an unknown parameter. Second, the Z_2 symmetry forbids a tree-level interaction between H' and the SM fermions, in particular the top quark. This eliminates the largest of the positive corrections in (5.19) and makes A smaller. Furthermore, it implies the new quartic $\lambda_{H'}$ does not receive large renormalization effects and thus stays almost constant (or maybe increases a bit) up to the UV cutoff. The stability condition stays approximately unchanged through the RG scale, rather than becoming significantly stronger at high scales as it was in our model. Finally, the function $h'(T)/T$ now decreases

with T, so the constraint $h'/T > 1.2$ we derived using (5.5) becomes in fact conservative. All these effects tend to decrease the minimum N required to avoid washout of the primordial asymmetry. For example, taking $\lambda_{H'}(m_t) = \lambda_H(m_t)/2$ and $\epsilon_S = 0.01$ the rough estimate (5.20) now gives $N \gtrsim 60$. Repeating the numerical analysis of Section 5.1.5 for $\lambda_{H'}(m_t) = \lambda_H(m_t)/2$, $\epsilon_S(m_t) = 0.05$, and still imposing the conservative constraint $h'(T)/T > 1.2$, we find that $N_{\rm min} \sim 20$. This is a significant difference compared to the $O(N)$ model considered in this chapter. Less exotic scenarios may now be envisioned. For example, $N = 24 \gtrsim N_{\rm min}$ can be obtained by simply promoting S to the adjoint representation of an $SU(5)$ gauge group, along the lines followed in Section 5.2.2.

The constraints discussed in Section 5.2 should be carefully re-assessed in models with additional EW-charged scalars. As concerns to our previous example, we note that the inert doublet H' is typically a small fraction of the dark matter, and that current bounds from dark matter experiments are easily accommodated. Also, collider constraints are rather weak provided we choose the H' couplings to the Higgs such that the mass splitting between its components is $\Delta \ll m_W$, but nevertheless sufficiently large to allow a fast decay of the charged scalar. This ensures the main signatures at colliders involve missing energy, that are poorly constrained (see e.g. [260]). Furthermore, with $\Delta \ll m_W$ the new contribution to the EW parameters is safely below current bounds.

5.4 Electro-weak baryogenesis at $T_c \gg 100$ GeV

In this section we sketch two examples of physics for the EW phase transition occurring at the critical temperature $T_c \propto \Lambda$. Our main goal is to illustrate the feasibility of our program. In particular, focussing on our basic $O(N)$ model we will show its compatibility with a realistic scenario for baryogenesis at the scale Λ.

5.4.1 Weakly-coupled sectors

We start by considering scenarios in which the dynamics at $\sim \Lambda$ is weakly-coupled. One simple option is to add a new scalar ϕ with potential couplings

$$\Delta V = \frac{m_\phi^2}{2}\phi^2 + \lambda_{H\phi}\phi^2 H^2 + \lambda_{\phi S}\phi^2 S^2 + \frac{\lambda_\phi}{4}\phi^4 \tag{5.72}$$

with $m_\phi^2 = \Lambda^2 > 0$, $\lambda_{H\phi}, \lambda_\phi > 0$, $\lambda_{\phi S} < 0$. As in the previous section, we can consistently assume $m_S^2(T)$ is always positive, so that $\langle S \rangle = 0$ at all temperatures. The main role of S is to induce *negative* thermal masses for the other two scalars: $m_{H,\phi}^2(T) \sim m_{H,\phi}^2 - |\lambda_{HS,\phi S}| N T^2/12$.

To get the basic picture we can study the thermal vacuum dynamics in the same approximation used in Sec. 5.1.3, where the potential is determined by the quartic

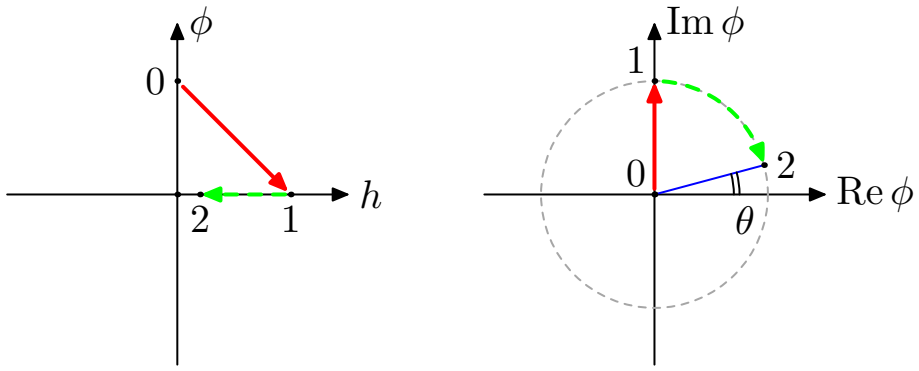


Figure 5.8: Schematic representation of the two kind of phase transitions studied in this section. The red arrow represents a first order phase transition at $T \sim T_c \propto \Lambda$, while the green dashed arrow shows a continuous change of the VEVs with the temperature $T < \Lambda$.

couplings renormalized at $\mu = T$ and by the 1-loop thermal masses. At high T, where $m^2_{H,\phi}(0)$ can be neglected, and assuming $\lambda_{H\phi} > \sqrt{\lambda_H \lambda_\phi}$ one finds that there exist two distinct vacua separated by a barrier: $\mathbf{0} \equiv (\phi \neq 0, h = 0)$ and $\mathbf{1} \equiv (\phi = 0, h \neq 0)$. We choose the parameters such that the deepest minimum at high T is $\mathbf{0}$, which simply requires

$$\frac{m^4_\phi(T \gg T_c)}{\lambda_\phi} > \frac{m^4_H(T \gg T_c)}{\lambda_H}. \tag{5.73}$$

Defining $|\lambda_{HS}/\lambda_{\phi S}|\sqrt{\lambda_\phi/\lambda_H} = \zeta > 0$ and assuming the S loop tadpoles dominate the thermal masses, the above request is equivalent to $\zeta < 1$. Because $m^2_\phi > 0$ by assumption, as T decreases $m^2_\phi(T)/T^2$ gets smaller in absolute value and eventually becomes positive when $T \leq T_\phi \sim 12 m^2_\phi/(|\lambda_{\phi S}|N)$. At the critical temperature $T_c^2 = T_\phi^2/(1-\zeta)[1 + \mathcal{O}(m^2_H/m^2_\phi)]$ the solution $\mathbf{1}$ starts to become deeper and a first order phase transition $\mathbf{0} \to \mathbf{1}$ can take place. For $T \ll T_c$ the physics is described by the model of Section 5.1.2, and $h(T)$ continuously decreases with T down to its zero-temperature value. A schematic picture of the phase diagram of this model is depicted in the left panel of Figure 5.8.

The transition $\mathbf{0} \to \mathbf{1}$ can be strongly first order, however, only if it is rather fast and occurs in the temperature range $T_\phi < T < T_c$. Indeed, after $T < T_\phi$ the fields can continuously roll down to $h \neq 0$, thus avoiding the jump off the barrier. It is the strength of the couplings that controls the temperature T_n at which bubbles of the true vacuum start to nucleate, and thus the onset of the phase transition, as well as its overall duration. While an accurate investigation is necessary to establish the region of parameter space in which EW baryogenesis actually occurs, we observe that one can in principle tune $1 - \zeta \sim T^2_\phi/T^2_c \ll 1$ in such a way that the barrier persists for a long time. [17] We are

[17] Actually, with $1 - \zeta \ll 1$ the two vacua start almost degenerate at high T; hence a small drop in T is

thus confident that a non-vanishing region of parameter space where $T_\phi < T_n < T_c$ can be engineered and our program realized.

CP-violation at the bubble walls, necessary to generate a baryon asymmetry during the transition, may be introduced in several ways. For example one option is through vector-like fermions with SM quantum numbers, whose mass matrix (with eigenvalues $\sim \Lambda$) and Yukawa couplings to the Higgs jointly violate CP. The key point is that these states are heavy enough to avoid constraints from the non-observation of rare flavor and CP violation at low energies.

5.4.2 Strongly-coupled sectors

It is interesting to investigate the alternative possibility that the heavy physics at Λ be strongly-coupled. A seemingly natural assumption is then that both H, S emerge as composite states. However, a more careful look reveals this might not be so simple to realize. Indeed, the required large number of degrees of freedom in S must be associated to a large number $N_{\rm fund}$ of fundamental constituents at the scale Λ. In such a framework a generic expectation is that all non-perturbative effects, including bubble nucleation, are exponentially suppressed by $e^{-\mathcal{O}(N_{\rm fund})}$. This would be a disaster for us, because it would mean that at $T = \mathcal{O}(\Lambda)$ bubble formation is too slow to drive a 1st order phase transition. In this situation the phase transition would either never complete (see for instance [261]), or complete at low temperatures where the CP violating sources are decoupled, or complete at the disappearance of the barrier as a smooth crossover. A more realistic possibility would be to have H, S emerge as composites from two distinct dynamics: a "small N" Higgs dynamics governing baryogenesis and a large-N singlet dynamics designed to preserve the asymmetry. [18]

We here consider a simplified picture in which the Higgs is a composite Nambu-Goldstone boson of a new strong dynamics at a scale $\Lambda = g_H f$, whereas S is taken to be elementary. We will argue that under reasonable assumptions, combining this picture with our scalar S with a negative coupling to the Higgs results in a strongly first order EW phase transition.

The picture is as follows. At $T > T_c$ the EW symmetry is unbroken. At $T = T_c$ the strong Higgs dynamics undergoes a first order phase transition associated to a symmetry breaking pattern $\mathcal{G} \to \mathcal{H}$. For definiteness we will have in mind the minimal $O(5)/O(4)$ scenario, [6] but our results straightforwardly generalize to more complicated cosets.

enough for **1** to dominate, i.e. $T_c \gg T_\phi \sim m_\phi$.

[18] One possible way to couple the two sectors, and ultimately generate the desired coupling $\lambda_{HS} H^\dagger H S^2$ within a framework that does not suffer from the *big* hierarchy problem, is to postulate the UV (small N) Higgs dynamics and the (large-N) S dynamics possess scalar operators $O_{H,S}$ having scaling dimension close to two and being odd under some Z_2 symmetry. In that case the nearly marginal operator $\mathcal{L}_{\rm int} = \bar{\lambda}_{HS} O_H O_S$ would be allowed and, after confinement and symmetry braking at a scale Λ, would interpolate $\lambda_{HS} H^\dagger H S^2$.

The Higgs is a Nambu-Goldstone mode of the coset $\mathcal{G}/\mathcal{H}$, that we parametrize via $s_h \equiv \sin(h/f)$. For $T \leq T_c$ the composite Higgs is an exact flat direction of the strong dynamics. However, couplings to the SM and to S break $\mathcal{G}$ explicitly, thus lifting such a degeneracy. Under the assumption that the largest T-dependent effects at temperatures $T \leq T_c$ are controlled by the $\mathcal{G}$-breaking coupling κ between S and the Higgs — as in our toy model (5.7) — the full effective potential will typically acquire the form

$$V(h, s=0) = \frac{m_h^2 f^2}{8\xi}\left[-2\xi s_h^2 + s_h^4 + \mathcal{O}(s_h^6)\right] - \frac{\kappa}{4} N T^2 f^2 s_h^2 \left[1 + \mathcal{O}(1/N)\right], \tag{5.74}$$

where we used the fact that the couplings are engineered such that $\xi \equiv \langle s_h^2(T=0)\rangle \ll 1$, as usual in composite Higgs models. (In (5.74) we assumed the coupling between the Nambu-Goldstone Higgs and S is parametrized by $\sim \kappa s_h^2 f^2 S^2$, but one may consider more general options.)

The vacuum solution derived from (5.74) reads

$$\langle s_h^2(T)\rangle = \begin{cases} 1 & T_* < T < T_c \\ \xi + (1-\xi)\frac{T^2}{T_*^2} & T < T_*, \end{cases} \tag{5.75}$$

where $T_*^2 = (1-\xi) m_h^2/(\kappa N \xi)$. We see that S destabilizes the EW-symmetric vacuum $h^2 = 0$ (solution **0**) as soon as the phase transition occurs, so that the EW symmetry abruptly goes from unbroken to broken right at the critical temperature, where $h^2 \sim f^2$ (solution **1**), before approaching its zero-temperature value $h^2 \ll f^2$ (solution **2**). Specifically, the W-mass $m_W^2(T) = g^2 f^2 \langle s_h^2(T)\rangle/4$ *jumps* from $m_W = 0$ at $T > T_c$ to some non-zero $m_W(T)$ at all $T < T_c$. A schematic picture of the phase diagram of the present scenario, $\mathbf{0} \to \mathbf{1} \to \mathbf{2}$, is shown in the right panel of Figure 5.8.

The sphaleron shut-off condition (5.6) in the present scenario becomes $f^2\langle s_h^2(T)\rangle/T^2 \gtrsim 1$. Using (5.75) we find that this is satisfied for all $T_* < T < T_c$ provided the transition in the exotic dynamics is strongly first order, in the sense that $T_c \lesssim f$. To ensure $B+L$ preservation at $T < T_*$ we should instead require

$$\frac{f^2}{T^2}\langle s_h^2(T)\rangle = \frac{f^2}{T^2}\xi + \frac{\kappa N}{2\lambda_H} \gtrsim 1. \tag{5.76}$$

Recalling that (5.74) was derived in the simplified limit $\kappa N \gg y_t^2$ in which the SM thermal loops are neglected, the latter condition is simply a re-writing of the request that (5.19) be larger than unity.

To realize EW baryogenesis at $T_c \lesssim f$ we further postulate the existence of sizable CP-violating couplings for the Higgs at those temperatures. These necessarily arise from the new interactions that are introduced in order to reproduce the Yukawa couplings of the SM. For example, in modern incarnations of the Composite Higgs, these emerge via

interactions like $yq\mathcal{O}_\Psi$, where q is a SM fermion and $\mathcal{O}_\Psi$ a family of composite fermionic operators of the strong Higgs dynamics (see, e.g. [262][6]). After symmetry breaking $\mathcal{O}_\Psi$ interpolate heavy vector-like resonances of mass $\Lambda = g_H f$ and the coupling generates a q/Ψ mixing, which ultimately results in the SM Yukawas. The phases in the coupling between q, Ψ and the bubble wall, $\sim yf/\Lambda = y/g_H$, are unsuppressed at $T \sim T_c$, so EW baryogenesis can take place efficiently. On the other hand, all corrections to low energy rare processes are controlled by powers of $E/\Lambda \ll 1$ (where E is the characteristic energy) and can be within current bounds given the large Λ considered here (see [43][47] for details).

5.5 Conclusions

The absence of any indication of CP-violation beyond the SM puts significant pressure on standard realizations of EW baryogenesis. Here we have demonstrated that it is possible to build unconventional scenarios where all the relevant dynamics — CP-violation and a strongly first order EW phase transition — takes place at a new threshold $\Lambda \gg m_W$. The physics threshold Λ may well be associated to the fundamental scale of flavor violation and safely lie in the range $\Lambda \sim 100 - 1000$ TeV. The emerging picture suggests a connection between flavor and CP violation, the EW phase transition, and baryogenesis, and might feature interesting correlations in the corresponding indirect signatures (rare flavor- and CP-violating processes, primordial gravitational waves) all being characterized by $\Lambda \gg m_W$.

This new class of scenarios for EW baryogenesis rests on the existence of a new sector at the weak scale whose defining role is to prevent washout of the primordial $B + L$ asymmetry. Our setup essentially removes the CP-problem of the ordinary scenarios for EW baryogenesis, but retains their predictivity: there must exist new physics at the weak scale. The key novelty is that such new physics, i.e. our sector at the Fermi scale, can be SM-neutral and CP-conserving, and therefore more easily compatible with observations. All CP-odd phases beyond the SM are instead associated to physics at Λ and decouple from low energy experiments, in agreement with observation. This setup paves the way to a multitude of novel realizations of EW baryogenesis. In Section 5.4 we have sketched two concrete scenarios, in which the low-energy sector is combined with a complete picture at Λ that includes CP-violation and a strongly first order EW phase transition, but many other realizations may be considered.

We have studied in detail a specific low-energy model consisting of a single scalar field in the fundamental representation of a new global $O(N)$, see (5.7), meant to represent an existence proof of our low-energy sector. The parameter N is constrained by the requirement of perturbativity and stability of the EFT. Both constraints are significantly exacerbated by the largeness of the top quark Yukawa and the significant RG running of the Higgs quartic, and at the end of the day they force N to be above 100. We have

seen, however, that such values may be dramatically reduced when non-minimal Higgs sectors at the weak scale are considered, see Section 5.3.

Our $O(N)$ model reveals an interesting $N \to \infty$ scaling in which qualitatively important effects are present at finite T despite the fact that the zero-temperature dynamics is arbitrarily weakly-coupled. We studied in detail the effective potential at finite T resulting from (5.7) and identified the regions of parameter space where our program is realized. Our large-N expansion has been employed in order to obtain a reliable, and systematically improvable, approximation of $V_{\rm eff}$.

The minimal model (5.7) also has an exactly stable dark matter candidate that is in conflict with current direct detection experiments. We have illustrated various extensions where this bound is evaded without affecting our main conclusions. A particularly appealing option appears to arise when gauging a subgroup of the global symmetry of the minimal model. We have discussed a specific incarnation with gauged $SO(n)$ and with matter consisting of just a real scalar S in the adjoint representation. We have identified a range of parameters that is interesting for baryogenesis, dark matter, and late cosmology (BBN and CMB): n is somewhat large $n \gtrsim 50$ and confinement happens at a scale of order 1 GeV, parametrically well below the mass of S, $m_s \sim 100$ GeV, see Section 5.2.3. Using generic properties of confining large-N gauge theories we have argued that a bound state made of one S and gluons, the *Sglueball*, could be a plausible dark matter candidate. Besides the standard perturbative calculable freeze-out of the scalar S, which happens prior to confinement, the large-N suppression of rehannihilation after confinement plays crucial a role. The resulting novel scenario for dark matter is certainly worth a more detailed and careful study.

5.A Proof of $\langle S\rangle = 0$

In this appendix we demonstrate that the true vacuum of (5.7), subject to (5.8), satisfies $s^2 = \langle S\rangle^2 = 0$. We will work at leading non-trivial order in $1/N$. Because s is not a flat direction, subleading $1/N$ effects cannot impact the conclusion qualitatively.

To prove our claim, let us first simplify our discussion by considering an ideal model without the Higgs, i.e. set $H = 0$ in (5.30). In this case the effective potential, that in general depends on the space-time independent value s^2, is simply obtained integrating out σ from $\mathcal{L}_{\rm eff} = -\frac{1}{2}(m_S^2+\sigma)s^2 + \frac{1}{4\lambda_S}\sigma^2 + N\Gamma[m_S^2+\sigma] + \mathcal{O}(1/N)$. As already emphasized in the text, the large-N approximation corresponds to an expansion in loops of the auxiliary field. The leading diagrams contributing to the effective potential for s are therefore found solving the classical equation of motion of $\partial\mathcal{L}_{\rm eff}/\partial\sigma = 0$. The effective potential for s^2 at leading order in $1/N$ is finally given by plugging the space-time independent classical solution $\sigma_c(s^2)$ back in $V_{\rm eff}(s) \equiv -\mathcal{L}_{\rm eff}(s, \sigma_c(s^2))$. By taking the total derivative

of $V_{\rm eff}$ with respect to s we see that

$$\frac{dV_{\rm eff}}{ds} = \left(\frac{\partial}{\partial s} + \frac{\partial \sigma_c}{\partial s}\frac{\partial}{\partial \sigma_c}\right) V_{\rm eff} = \frac{\partial}{\partial s} V_{\rm eff} = (m_S^2 + \sigma_c)s. \tag{5.77}$$

Superficially, eq. (5.77) tells us that $dV_{\rm eff}/ds = 0$ has two possible solutions: $s = 0$ or $m_S^2 + \sigma_c = 0$. However, it turns out that $m_S^2 + \sigma_c = 0$ is unphysical. Indeed, in this latter case (and after having properly renormalized m_S^2) we have $-\partial \mathcal{L}_{\rm eff}/\partial \sigma_c = s^2/2 + m_S^2/(2\lambda_S) - N\Gamma'[0] = s^2/2 + m_S^2/(2\lambda_S) + NT^2/24 = 0$, which cannot be satisfied under the hypothesis (5.8). We thus conclude that $s = 0$ is the only stable solution at this order. Also, differentiating twice the effective potential one finds that $d^2V_{\rm eff}/ds^2(s = 0) \equiv m_S^2(T) = m_S^2 + \sigma_c(s = 0) \sim m_S^2 + \lambda_S NT^2/12$ represents the full thermal mass. The direction s is not flat and next to leading perturbative $\mathcal{O}(1/N)$ corrections will not spoil our conclusion.

An analogous result extends to the case with the Higgs field (and the full SM) included. The basic reason is that the SM loops are truly perturbative, and in particular not enhanced by powers of N, and can thus be neglected at leading order. As it was argued regarding the $1/N$ corrections, their inclusion cannot alter our arguments qualitatively. This observation allows us to simplify our analysis by considering the Lagrangian (see (5.30)):

$$\begin{aligned} \mathcal{L}_{\rm eff} &= -\frac{1}{2}\left(m_H^2 + \frac{\lambda_{HS}}{\lambda_S}\sigma\right)h^2 - \frac{\lambda_H}{4}\left(1 - \frac{\lambda_{HS}^2}{\lambda_H \lambda_S}\right)h^4 \\ &- \frac{1}{2}(m_S^2 + \sigma)s^2 + \frac{1}{4\lambda_S}\sigma^2 + N\Gamma[m_S^2 + \sigma] + \mathcal{O}(1/N, \text{SM loops}). \end{aligned} \tag{5.78}$$

It then immediately follows that (5.77) remains approximately correct. Furthermore, the equation of motion at $\sigma_c = -m_S^2$ is now given by $\lambda_S s^2 + m_S^2 + \lambda_S NT^2/12 + \lambda_{HS}h^2 = 0$: replacing the solution h^2, found differentiating (5.78), one can easily verify that this is incompatible with eq (5.8) at any T. This again ensures that $s = 0$ is the only consistent solution.

5.B Thermal potentials and beta functions

Within a 1-loop approximation, the integration of a particle of spin $j = 0, 1/2, 1$ and (field-dependent) mass-squared M^2 results in the following contribution to the finite-T effective potential:

$$\begin{aligned} V_j(M^2) &= (-)^{2j}\frac{1}{64\pi^2}(M^2)^2[\ln(M^2/\mu^2) - c_j] \\ &+ (-)^{2j}T\int \frac{d\vec{p}}{(2\pi)^3}\ \ln\left[1 - (-)^{2j}\exp\left(-\frac{1}{T}\sqrt{\vec{p}^2 + M^2}\right)\right]. \end{aligned} \tag{5.79}$$

We expressed (5.79) in the $\overline{\text{MS}}$ scheme, with μ the renormalization point; $c_{0,1/2} = 3/2$ for scalars and fermions whereas $c_1 = 5/6$ for vectors.

An explicit 1-loop computation shows that (still in the $\overline{\text{MS}}$ scheme):

$$\begin{aligned} 8\pi^2\mu\frac{d\lambda_H}{d\mu} &= (8+N_H)\lambda_H^2 + N\lambda_{HS}^2 + \mathcal{O}(g^2\lambda_H, y_t^2\lambda_H, g^4, y_t^4) \\ 8\pi^2\mu\frac{d\lambda_S}{d\mu} &= (8+N)\lambda_S^2 + N_H\lambda_{HS}^2 \\ 8\pi^2\mu\frac{d\lambda_{HS}}{d\mu} &= \lambda_{HS}\left[(2+N_H)\lambda_H + (2+N)\lambda_S + 4\lambda_{HS} + \mathcal{O}(g^2, y_t^2)\right], \end{aligned} \tag{5.80}$$

where $N_H = 4$. These equations were used in Section 5.1.5 as a non-trivial check of the consistency of our effective potential (5.35). In fact we verified that, generalizing V_{eff} to include $s \neq 0$, (5.80) follow from the RG invariance of the effective potential.

Bibliography

[1] S. Chen, A. Glioti, G. Panico and A. Wulzer, *Parametrized classifiers for optimal EFT sensitivity*, *JHEP* **05** (2021) 247 [2007.10356].

[2] S. Chen, A. Glioti, R. Rattazzi, L. Ricci and A. Wulzer, *Learning from radiation at a very high energy lepton collider*, *JHEP* **05** (2022) 180 [2202.10509].

[3] A. Glioti, R. Rattazzi and L. Vecchi, *Electroweak Baryogenesis above the Electroweak Scale*, *JHEP* **04** (2019) 027 [1811.11740].

[4] G. Panico and A. Wulzer, *The Composite Nambu-Goldstone Higgs*, vol. 913. Springer, 2016, 10.1007/978-3-319-22617-0, [1506.01961].

[5] R. Contino, *The Higgs as a Composite Nambu-Goldstone Boson*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Physics of the Large and the Small*, pp. 235–306, 2011, 1005.4269, DOI.

[6] K. Agashe, R. Contino and A. Pomarol, *The Minimal composite Higgs model*, *Nucl. Phys. B* **719** (2005) 165 [hep-ph/0412089].

[7] S. R. Coleman, J. Wess and B. Zumino, *Structure of phenomenological Lagrangians. 1.*, *Phys. Rev.* **177** (1969) 2239.

[8] C. G. Callan, Jr., S. R. Coleman, J. Wess and B. Zumino, *Structure of phenomenological Lagrangians. 2.*, *Phys. Rev.* **177** (1969) 2247.

[9] R. Barbieri, B. Bellazzini, V. S. Rychkov and A. Varagnolo, *The Higgs boson from an extended symmetry*, *Phys. Rev. D* **76** (2007) 115008 [0706.0432].

[10] D. B. Kaplan, *Flavor at SSC energies: A New mechanism for dynamically generated fermion masses*, *Nucl. Phys. B* **365** (1991) 259.

[11] G. Giudice, C. Grojean, A. Pomarol and R. Rattazzi, *The Strongly-Interacting Light Higgs*, *JHEP* **06** (2007) 045 [hep-ph/0703164].

[12] H. Georgi, *Generalized dimensional analysis*, *Phys. Lett. B* **298** (1993) 187 [hep-ph/9207278].

[13] R. Barbieri, D. Buttazzo, F. Sala, D. M. Straub and A. Tesi, *A 125 GeV composite Higgs boson versus flavour and electroweak precision tests*, *JHEP* **05** (2013) 069 [1211.5085].

[14] C. Abel et al., *Measurement of the Permanent Electric Dipole Moment of the Neutron*, *Phys. Rev. Lett.* **124** (2020) 081803 [2001.11966].

[15] D. A. Demir, M. Pospelov and A. Ritz, *Hadronic EDMs, the Weinberg operator, and light gluinos*, *Phys. Rev. D* **67** (2003) 015007 [hep-ph/0208257].

[16] J. de Vries, G. Falcioni, F. Herzog and B. Ruijl, *Two- and three-loop anomalous dimensions of Weinberg's dimension-six CP-odd gluonic operator*, *Phys. Rev. D* **102** (2020) 016010 [1907.04923].

[17] G. D'Ambrosio, G. F. Giudice, G. Isidori and A. Strumia, *Minimal flavor violation: An Effective field theory approach*, *Nucl. Phys. B* **645** (2002) 155 [hep-ph/0207036].

[18] M. Redi and A. Weiler, *Flavor and CP Invariant Composite Higgs Models*, *JHEP* **11** (2011) 108 [1106.6357].

[19] M. Redi, *Leptons in Composite MFV*, *JHEP* **09** (2013) 060 [1306.1525].

[20] O. Domenech, A. Pomarol and J. Serra, *Probing the SM with Dijets at the LHC*, *Phys. Rev. D* **85** (2012) 074030 [1201.6510].

[21] S. Alioli, M. Farina, D. Pappadopulo and J. T. Ruderman, *Precision Probes of QCD at High Energies*, *JHEP* **07** (2017) 097 [1706.03068].

[22] CMS collaboration, A. M. Sirunyan et al., *Search for new physics in dijet angular distributions using proton–proton collisions at $\sqrt{s} = 13$ TeV and constraints on dark matter and other models*, *Eur. Phys. J. C* **78** (2018) 789 [1803.08030].

[23] ATLAS collaboration, M. Aaboud et al., *Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev. D* **99** (2019) 072009 [1901.03584].

[24] W. Altmannshofer and P. Stangl, *New physics in rare B decays after Moriond 2021*, *Eur. Phys. J. C* **81** (2021) 952 [2103.13370].

[25] K. Agashe, R. Contino, L. Da Rold and A. Pomarol, *A Custodial symmetry for $Zb\bar{b}$*, *Phys. Lett. B* **641** (2006) 62 [hep-ph/0605341].

[26] Particle Data Group collaboration, R. L. Workman, *Review of Particle Physics*, *PTEP* **2022** (2022) 083C01.

[27] M. Redi, *Composite MFV and Beyond*, *Eur. Phys. J. C* **72** (2012) 2030 [1203.4220].

[28] A. Paul and D. M. Straub, *Constraints on new physics from radiative B decays*, *JHEP* **04** (2017) 027 [1608.02556].

[29] L. Vecchi, *The Natural Composite Higgs*, 1304.4579.

[30] R. Barbieri, D. Buttazzo, F. Sala and D. M. Straub, *Less Minimal Flavour Violation*, *JHEP* **10** (2012) 040 [1206.1327].

[31] J. D. Wells and Z. Zhang, *Effective theories of universal theories*, *JHEP* **01** (2016) 123 [1510.08462].

[32] R. Barbieri, A. Pomarol, R. Rattazzi and A. Strumia, *Electroweak symmetry breaking after LEP-1 and LEP-2*, *Nucl. Phys. B* **703** (2004) 127 [hep-ph/0405040].

[33] J. de Blas, M. Ciuchini, E. Franco, S. Mishima, M. Pierini, L. Reina et al., *The Global Electroweak and Higgs Fits in the LHC era*, *PoS* **EPS-HEP2017** (2017) 467 [1710.05402].

[34] CMS collaboration, A. Tumasyan et al., *Search for new physics in the lepton plus missing transverse momentum final state in proton-proton collisions at* $\sqrt{s}$ *= 13 TeV*, 2202.06075.

[35] J. Ellis, M. Madigan, K. Mimasu, V. Sanz and T. You, *Top, Higgs, Diboson and Electroweak Fit to the Standard Model Effective Field Theory*, *JHEP* **04** (2021) 279 [2012.02779].

[36] J. de Blas et al., *Higgs Boson Studies at Future Particle Colliders*, *JHEP* **01** (2020) 139 [1905.03764].

[37] R. K. Ellis et al., *Physics Briefing Book: Input for the European Strategy for Particle Physics Update 2020*, 1910.11775.

[38] G. Panico, L. Ricci and A. Wulzer, *High-energy EFT probes with fully differential Drell-Yan measurements*, *JHEP* **07** (2021) 086 [2103.10532].

[39] R. Contino, M. Ghezzi, C. Grojean, M. Muhlleitner and M. Spira, *Effective Lagrangian for a light Higgs-like scalar*, *JHEP* **07** (2013) 035 [1303.3876].

[40] G. Durieux, A. Irles, V. Miralles, A. Peñuelas, R. Pöschl, M. Perelló et al., *The electro-weak couplings of the top and bottom quarks — Global fit and future prospects*, *JHEP* **12** (2019) 98 [1907.10619].

[41] J. Mrazek, A. Pomarol, R. Rattazzi, M. Redi, J. Serra and A. Wulzer, *The Other Natural Two Higgs Doublet Model*, *Nucl. Phys. B* **853** (2011) 1 [1105.5403].

[42] LHCb collaboration, R. Aaij et al., *Observation of CP Violation in Charm Decays*, *Phys. Rev. Lett.* **122** (2019) 211803 [1903.08726].

[43] B. Keren-Zur, P. Lodone, M. Nardecchia, D. Pappadopulo, R. Rattazzi and L. Vecchi, *On Partial Compositeness and the CP asymmetry in charm decays*, *Nucl. Phys. B* **867** (2013) 394 [1205.5803].

[44] A. Cerri et al., *Report from Working Group 4: Opportunities in Flavour Physics at the HL-LHC and HE-LHC*, *CERN Yellow Rep. Monogr.* **7** (2019) 867 [1812.07638].

[45] Belle-II collaboration, W. Altmannshofer et al., *The Belle II Physics Book*, *PTEP* **2019** (2019) 123C01 [1808.10567].

[46] R. Barbieri, *Flavour and Higgs compositeness: present and "near" future*, 1910.00371.

[47] M. Frigerio, M. Nardecchia, J. . Serra and L. Vecchi, *The Bearable Compositeness of Leptons*, *JHEP* **10** (2018) 017 [1807.04279].

[48] Mu3e collaboration, G. Hesketh, S. Hughes, A.-K. Perrevoort and N. Rompotis, *The Mu3e Experiment*, in *2022 Snowmass Summer Study*, 4, 2022, 2204.00001.

[49] Mu2e collaboration, R. H. Bernstein, *The Mu2e Experiment*, *Front. in Phys.* **7** (2019) 1 [1901.11099].

[50] W. Buchmuller and D. Wyler, *Effective Lagrangian Analysis of New Interactions and Flavor Conservation*, *Nucl. Phys. B* **268** (1986) 621.

[51] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085 [1008.4884].

[52] D. Atwood and A. Soni, *Analysis for magnetic moment and electric dipole moment form-factors of the top quark via $e^+e^- \to t\bar{t}$*, *Phys. Rev. D* **45** (1992) 2405.

[53] M. Diehl and O. Nachtmann, *Optimal observables for the measurement of three gauge boson couplings in $e^+e^- \to W^+W^-$*, *Z. Phys. C* **62** (1994) 397.

[54] I. Dunietz, H. R. Quinn, A. Snyder, W. Toki and H. J. Lipkin, *How to extract CP violating asymmetries from angular correlations*, *Phys. Rev. D* **43** (1991) 2193.

[55] A. S. Dighe, I. Dunietz and R. Fleischer, *Extracting CKM phases and $B_s - \bar{B}_s$ mixing parameters from angular distributions of nonleptonic B decays*, *Eur. Phys. J. C* **6** (1999) 647 [hep-ph/9804253].

[56] S. Banerjee, R. S. Gupta, J. Y. Reiness, S. Seth and M. Spannowsky, *Towards the ultimate differential SMEFT analysis*, 1912.07628.

[57] I. Anderson et al., *Constraining Anomalous HVV Interactions at Proton and Lepton Colliders*, *Phys. Rev. D* **89** (2014) 035007 [1309.4819].

[58] K. Kondo, *Dynamical Likelihood Method for Reconstruction of Events With Missing Momentum. 1: Method and Toy Models*, *J. Phys. Soc. Jap.* **57** (1988) 4126.

[59] P. Artoisenet, V. Lemaitre, F. Maltoni and O. Mattelaer, *Automation of the matrix element reweighting method*, *JHEP* **12** (2010) 068 [1007.3300].

[60] F. Fiedler, A. Grohsjean, P. Haefner and P. Schieferdecker, *The Matrix Element Method and its Application in Measurements of the Top Quark Mass*, *Nucl. Instrum. Meth. A* **624** (2010) 203 [1003.1316].

[61] T. Martini and P. Uwer, *Extending the Matrix Element Method beyond the Born approximation: Calculating event weights at next-to-leading order accuracy*, *JHEP* **09** (2015) 083 [1506.08798].

[62] T. Martini and P. Uwer, *The Matrix Element Method at next-to-leading order QCD for hadronic collisions: Single top-quark production at the LHC as an example application*, *JHEP* **05** (2018) 141 [1712.04527].

[63] J. Brehmer, K. Cranmer, G. Louppe and J. Pavez, *Constraining Effective Field Theories with Machine Learning*, *Phys. Rev. Lett.* **121** (2018) 111801 [1805.00013].

[64] K. Cranmer, J. Pavez and G. Louppe, *Approximating Likelihood Ratios with Calibrated Discriminative Classifiers*, 1506.02169.

[65] P. Baldi, K. Cranmer, T. Faucett, P. Sadowski and D. Whiteson, *Parameterized neural networks for high-energy physics*, *Eur. Phys. J. C* **76** (2016) 235 [1601.07913].

[66] M. Stoye, J. Brehmer, G. Louppe, J. Pavez and K. Cranmer, *Likelihood-free inference with an improved cross-entropy estimator*, 1808.00973.

[67] J. Brehmer, G. Louppe, J. Pavez and K. Cranmer, *Mining gold from implicit models to improve likelihood-free inference*, *Proc. Nat. Acad. Sci.* (2020) 201915980 [1805.12244].

[68] J. Brehmer, K. Cranmer, G. Louppe and J. Pavez, *A Guide to Constraining Effective Field Theories with Machine Learning*, *Phys. Rev. D* **98** (2018) 052004 [1805.00020].

[69] J. Brehmer, F. Kling, I. Espejo and K. Cranmer, *MadMiner: Machine learning-based inference for particle physics*, *Comput. Softw. Big Sci.* **4** (2020) 3 [1907.10621].

[70] J. Neyman and E. S. Pearson, *On the Problem of the Most Efficient Tests of Statistical Hypotheses*, *Phil. Trans. Roy. Soc. Lond. A* **231** (1933) 289.

[71] Particle Data Group collaboration, M. Tanabashi et al., *Review of Particle Physics*, *Phys. Rev. D* **98** (2018) 030001.

[72] A. Falkowski, M. Gonzalez-Alonso, A. Greljo and D. Marzocca, *Global constraints on anomalous triple gauge couplings in effective field theory approach*, *Phys. Rev. Lett.* **116** (2016) 011801 [1508.00581].

[73] D. R. Green, P. Meade and M.-A. Pleier, *Multiboson interactions at the LHC*, *Rev. Mod. Phys.* **89** (2017) 035008 [1610.07572].

[74] A. Butter, O. J. P. Éboli, J. Gonzalez-Fraile, M. Gonzalez-Garcia, T. Plehn and M. Rauch, *The Gauge-Higgs Legacy of the LHC Run I*, *JHEP* **07** (2016) 152 [1604.03105].

[75] R. Franceschini, G. Panico, A. Pomarol, F. Riva and A. Wulzer, *Electroweak Precision Tests in High-Energy Diboson Processes*, *JHEP* **02** (2018) 111 [1712.01310].

[76] G. Panico, F. Riva and A. Wulzer, *Diboson Interference Resurrection*, *Phys. Lett.* **B776** (2018) 473 [1708.07823].

[77] A. Azatov, J. Elias-Miro, Y. Reyimuaji and E. Venturini, *Novel measurements of anomalous triple gauge couplings for the LHC*, *JHEP* **10** (2017) 027 [1707.08060].

[78] A. Azatov, D. Barducci and E. Venturini, *Precision diboson measurements at hadron colliders*, *JHEP* **04** (2019) 075 [1901.04821].

[79] J. Baglio, S. Dawson and S. Homiller, *QCD corrections in Standard Model EFT fits to WZ and WW production*, *Phys. Rev. D* **100** (2019) 113010 [1909.11576].

[80] M. J. Duncan, G. L. Kane and W. W. Repko, *A New Standard Model Test for Future Colliders*, *Phys. Rev. Lett.* **55** (1985) 773.

[81] K. Hagiwara, R. Peccei, D. Zeppenfeld and K. Hikasa, *Probing the Weak Boson Sector in $e^+e^- \to W^+W^-$*, *Nucl. Phys. B* **282** (1987) 253.

[82] G. Cowan, *Statistical data analysis*. Oxford University Press, USA, 1998.

[83] M. Anthony and P. L. Bartlett, *Neural Network Learning: Theoretical Foundations*. Cambridge University Press., 1999.

[84] R. T. D'Agnolo and A. Wulzer, *Learning New Physics from a Machine*, *Phys. Rev. D* **99** (2019) 015014 [1806.02350].

[85] D. Liu, A. Pomarol, R. Rattazzi and F. Riva, *Patterns of Strong Coupling for LHC Searches*, *JHEP* **11** (2016) 141 [1603.03064].

[86] G. Cuomo, L. Vecchi and A. Wulzer, *Goldstone Equivalence and High Energy Electroweak Physics*, *SciPost Phys.* **8** (2020) 078 [1911.12366].

[87] A. Kusina et al., *nCTEQ15 - Global analysis of nuclear parton distributions with uncertainties*, *PoS* **DIS2015** (2015) 041 [1509.01801].

[88] D. Clark, E. Godat and F. Olness, *ManeParse : A Mathematica reader for Parton Distribution Functions*, *Comput. Phys. Commun.* **216** (2017) 126 [1605.08012].

[89] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079 [1405.0301].

[90] C. Degrande, G. Durieux, F. Maltoni, K. Mimasu, A. Vasquez, E. Vryonidou et al., *no*, .

[91] T. Sjostrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **05** (2006) 026 [hep-ph/0603175].

[92] T. Sjostrand, S. Mrenna and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2007) 852 [0710.3820].

[93] E. Conte, B. Fuks and G. Serret, *MadAnalysis 5, A User-Friendly Framework for Collider Phenomenology*, *Comput. Phys. Commun.* **184** (2013) 222 [1206.1599].

[94] U. Baur, T. Han and J. Ohnemus, *Amplitude zeros in W+- Z production*, *Phys. Rev. Lett.* **72** (1994) 3941 [hep-ph/9403248].

[95] L. J. Dixon and Y. Shadmi, *Testing gluon selfinteractions in three jet events at hadron colliders*, *Nucl. Phys. B* **423** (1994) 3 [hep-ph/9312363].

[96] no, *Poisson distribution. Encyclopedia of Mathematics.*, .

[97] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan et al., *Pytorch: An imperative style, high-performance deep learning library*, in *Advances in Neural Information Processing Systems 32* (H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox and R. Garnett, eds.), pp. 8024–8035. Curran Associates, Inc., 2019.

[98] 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].

[99] J. P. Delahaye, M. Diemoz, K. Long, B. Mansoulié, N. Pastrone, L. Rivkin et al., *Muon Colliders*, 1901.06150.

[100] "The international muon collider collaboration." https://muoncollider.web.cern.ch.

[101] D. Buttazzo, D. Redigolo, F. Sala and A. Tesi, *Fusing Vectors into Scalars at High Energy Lepton Colliders*, *JHEP* **11** (2018) 144 [1807.04743].

[102] R. Franceschini, *High(est) energy lepton colliders*, *Int. J. Mod. Phys. A* **36** (2021) 2142016.

[103] R. Capdevilla, F. Meloni, R. Simoniello and J. Zurita, *Hunting wino and higgsino dark matter at the muon collider with disappearing tracks*, *JHEP* **06** (2021) 133 [2102.11292].

[104] H. Al Ali et al., *The Muon Smasher's Guide*, 2103.14043.

[105] D. Buttazzo, R. Franceschini and A. Wulzer, *Two Paths Towards Precision at a Very High Energy Lepton Collider*, *JHEP* **05** (2021) 219 [2012.11555].

[106] T. Han, D. Liu, I. Low and X. Wang, *Electroweak couplings of the Higgs boson at a multi-TeV muon collider*, *Phys. Rev. D* **103** (2021) 013002 [2008.12204].

[107] N. Bartosik et al., *Preliminary Report on the Study of Beam-Induced Background Effects at a Muon Collider*, 1905.03725.

[108] R. Capdevilla, D. Curtin, Y. Kahn and G. Krnjaic, *Discovering the physics of $(g-2)_\mu$ at future muon colliders*, *Phys. Rev. D* **103** (2021) 075028 [2006.16277].

[109] M. Ruhdorfer, E. Salvioni and A. Weiler, *A Global View of the Off-Shell Higgs Portal*, *SciPost Phys.* **8** (2020) 027 [1910.04170].

[110] M. Chiesa, F. Maltoni, L. Mantani, B. Mele, F. Piccinini and X. Zhao, *Measuring the quartic Higgs self-coupling at a multi-TeV muon collider*, *JHEP* **09** (2020) 098 [2003.13628].

[111] A. Costantini, F. De Lillo, F. Maltoni, L. Mantani, O. Mattelaer, R. Ruiz et al., *Vector boson fusion at multi-TeV muon colliders*, *JHEP* **09** (2020) 080 [2005.10289].

[112] R. Capdevilla, D. Curtin, Y. Kahn and G. Krnjaic, *A No-Lose Theorem for Discovering the New Physics of $(g-2)_\mu$ at Muon Colliders*, 2101.10334.

[113] N. Bartosik et al., *Detector and Physics Performance at a Muon Collider*, *JINST* **15** (2020) P05001 [2001.04431].

[114] W. Yin and M. Yamaguchi, *Muon $g-2$ at multi-TeV muon collider*, 2012.03928.

[115] J. Kalinowski, T. Robens, D. Sokolowska and A. F. Zarnecki, *IDM Benchmarks for the LHC and Future Colliders*, *Symmetry* **13** (2021) 991 [2012.14818].

[116] G.-y. Huang, F. S. Queiroz and W. Rodejohann, *Gauged $L_\mu - L_\tau$ at a muon collider*, *Phys. Rev. D* **103** (2021) 095005 [2101.04956].

[117] W. Liu and K.-P. Xie, *Probing electroweak phase transition with multi-TeV muon colliders and gravitational waves*, *JHEP* **04** (2021) 015 [2101.10469].

[118] T. Han, S. Li, S. Su, W. Su and Y. Wu, *Heavy Higgs bosons in 2HDM at a muon collider*, *Phys. Rev. D* **104** (2021) 055029 [2102.08386].

[119] S. Bottaro, A. Strumia and N. Vignaroli, *Minimal Dark Matter bound states at future colliders*, *JHEP* **06** (2021) 143 [2103.12766].

[120] T. Li, M. A. Schmidt, C.-Y. Yao and M. Yuan, *Charged lepton flavor violation in light of the muon magnetic moment anomaly and colliders*, *Eur. Phys. J. C* **81** (2021) 811 [2104.04494].

[121] P. Asadi, R. Capdevilla, C. Cesarotti and S. Homiller, *Searching for leptoquarks at future muon colliders*, *JHEP* **10** (2021) 182 [2104.05720].

[122] M. Sahin and A. Caliskan, *Excited muon production at muon colliders via contact interaction*, 2105.01964.

[123] J. Chen, C.-T. Lu and Y. Wu, *Measuring Higgs boson self-couplings with 2 → 3 VBS processes*, *JHEP* **10** (2021) 099 [2105.11500].

[124] G. Haghighat and M. Mohammadi Najafabadi, *Search for lepton-flavor-violating ALPs at a future muon collider and utilization of polarization-induced effects*, 2106.00505.

[125] S. Bottaro, D. Buttazzo, M. Costa, R. Franceschini, P. Panci, D. Redigolo et al., *Closing the window on WIMP Dark Matter*, 2107.09688.

[126] P. Bandyopadhyay, A. Karan and C. Sen, *Discerning Signatures of Seesaw Models and Complementarity of Leptonic Colliders*, 2011.04191.

[127] C. Sen, P. Bandyopadhyay, S. Dutta and A. KT, *Displaced Higgs production in Type-III Seesaw at the LHC/FCC, MATHUSLA and Muon collider*, 2107.12442.

[128] T. Han, W. Kilian, N. Kreher, Y. Ma, J. Reuter, T. Striegl et al., *Precision Test of the Muon-Higgs Coupling at a High-energy Muon Collider*, 2108.05362.

[129] P. Bandyopadhyay, A. Karan and R. Mandal, *Distinguishing signatures of scalar leptoquarks at hadron and muon colliders*, 2108.06506.

[130] R. Dermisek, K. Hermanek and N. McGinnis, *Di-Higgs and tri-Higgs boson signals of muon g-2 at a muon collider*, *Phys. Rev. D* **104** (2021) L091301 [2108.10950].

[131] S. Qian, C. Li, Q. Li, F. Meng, J. Xiao, T. Yang et al., *Searching for heavy leptoquarks at a muon collider*, 2109.01265.

[132] R. Ruiz, A. Costantini, F. Maltoni and O. Mattelaer, *The Effective Vector Boson Approximation in High-Energy Muon Collisions*, 2111.02442.

[133] M. Chiesa, B. Mele and F. Piccinini, *Multi Higgs production via photon fusion at future multi-TeV muon colliders*, 2109.10109.

[134] W. Liu, K.-P. Xie and Z. Yi, *Testing leptogenesis at the LHC and future muon colliders: a Z' scenario*, 2109.15087.

[135] D. Buttazzo and P. Paradisi, *Probing the muon $g-2$ anomaly with the Higgs boson at a muon collider*, *Phys. Rev. D* **104** (2021) 075021.

[136] L. Di Luzio, R. Gröber and G. Panico, *Probing new electroweak states via precision measurements at the LHC and future colliders*, *JHEP* **01** (2019) 011 [1810.10993].

[137] T. Han, Z. Liu, L.-T. Wang and X. Wang, *WIMPs at High Energy Muon Colliders*, *Phys. Rev. D* **103** (2021) 075004 [2009.11287].

[138] J. Chen, T. Li, C.-T. Lu, Y. Wu and C.-Y. Yao, *The measurement of Higgs self-couplings through $2 \to 3$ VBS in future muon colliders*, 2112.12507.

[139] R. Capdevilla, D. Curtin, Y. Kahn and G. Krnjaic, *Systematically Testing Singlet Models for $(g-2)_\mu$*, 2112.08377.

[140] R. Franceschini and X. Zhao, *to appear*.

[141] M. Ciafaloni, P. Ciafaloni and D. Comelli, *Bloch-Nordsieck violating electroweak corrections to inclusive TeV scale hard processes*, *Phys. Rev. Lett.* **84** (2000) 4810 [hep-ph/0001142].

[142] M. Ciafaloni, P. Ciafaloni and D. Comelli, *Electroweak Bloch-Nordsieck violation at the TeV scale: 'Strong' weak interactions?*, *Nucl. Phys. B* **589** (2000) 359 [hep-ph/0004071].

[143] V. S. Fadin, L. N. Lipatov, A. D. Martin and M. Melles, *Resummation of double logarithms in electroweak high-energy processes*, *Phys. Rev. D* **61** (2000) 094002 [hep-ph/9910338].

[144] M. Melles, *Subleading Sudakov logarithms in electroweak high-energy processes to all orders*, *Phys. Rev. D* **63** (2001) 034003 [hep-ph/0004056].

[145] J.-y. Chiu, F. Golf, R. Kelley and A. V. Manohar, *Electroweak Sudakov corrections using effective field theory*, *Phys. Rev. Lett.* **100** (2008) 021802 [0709.2377].

[146] J.-y. Chiu, F. Golf, R. Kelley and A. V. Manohar, *Electroweak Corrections in High Energy Processes using Effective Field Theory*, *Phys. Rev. D* **77** (2008) 053004 [0712.0396].

[147] J.-y. Chiu, A. Fuhrer, R. Kelley and A. V. Manohar, *Factorization Structure of Gauge Theory Amplitudes and Application to Hard Scattering Processes at the LHC*, *Phys. Rev. D* **80** (2009) 094013 [0909.0012].

[148] J.-y. Chiu, A. Fuhrer, R. Kelley and A. V. Manohar, *Soft and Collinear Functions for the Standard Model*, *Phys. Rev. D* **81** (2010) 014023 [0909.0947].

[149] A. Manohar, B. Shotwell, C. Bauer and S. Turczyk, *Non-cancellation of electroweak logarithms in high-energy scattering*, *Phys. Lett. B* **740** (2015) 179 [1409.1918].

[150] A. V. Manohar and W. J. Waalewijn, *Electroweak Logarithms in Inclusive Cross Sections*, *JHEP* **08** (2018) 137 [1802.08687].

[151] C. W. Bauer, N. Ferland and B. R. Webber, *Combining initial-state resummation with fixed-order calculations of electroweak corrections*, *JHEP* **04** (2018) 125 [1712.07147].

[152] M. Ciafaloni, P. Ciafaloni and D. Comelli, *Towards collinear evolution equations in electroweak theory*, *Phys. Rev. Lett.* **88** (2002) 102001 [hep-ph/0111109].

[153] P. Ciafaloni and D. Comelli, *Electroweak evolution equations*, *JHEP* **11** (2005) 022 [hep-ph/0505047].

[154] B. Fornal, A. V. Manohar and W. J. Waalewijn, *Electroweak Gauge Boson Parton Distribution Functions*, *JHEP* **05** (2018) 106 [1803.06347].

[155] C. W. Bauer, N. Ferland and B. R. Webber, *Standard Model Parton Distributions at Very High Energies*, *JHEP* **08** (2017) 036 [1703.08562].

[156] C. W. Bauer, D. Provasoli and B. R. Webber, *Standard Model Fragmentation Functions at Very High Energies*, *JHEP* **11** (2018) 030 [1806.10157].

[157] T. Han, Y. Ma and K. Xie, *High energy leptonic collisions and electroweak parton distribution functions*, *Phys. Rev. D* **103** (2021) L031301 [2007.14300].

[158] J. Chen, T. Han and B. Tweedie, *Electroweak Splitting Functions and High Energy Showering*, *JHEP* **11** (2017) 093 [1611.00788].

[159] S. Dawson, *The Effective W Approximation*, *Nucl. Phys. B* **249** (1985) 42.

[160] Z. Kunszt and D. E. Soper, *On the Validity of the Effective W Approximation*, *Nucl. Phys. B* **296** (1988) 253.

[161] P. Borel, R. Franceschini, R. Rattazzi and A. Wulzer, *Probing the Scattering of Equivalent Electroweak Bosons*, *JHEP* **06** (2012) 122 [1202.1904].

[162] T. Kinoshita, *Mass singularities of Feynman amplitudes*, *J. Math. Phys.* **3** (1962) 650.

[163] T. D. Lee and M. Nauenberg, *Degenerate systems and mass singularities*, *Phys. Rev.* **133** (1964) B1549.

[164] P. Ciafaloni and D. Comelli, *The Importance of weak bosons emission at LHC*, *JHEP* **09** (2006) 055 [hep-ph/0604070].

[165] A. Denner and S. Pozzorini, *One loop leading logarithms in electroweak radiative corrections. 1. Results*, *Eur. Phys. J. C* **18** (2001) 461 [hep-ph/0010201].

[166] R. Franceschini et al., *The CLIC Potential for New Physics*, 1812.02093.

[167] ALEGRO collaboration, E. Adli et al., *Towards an Advanced Linear International Collider*, 1901.10370.

[168] A. Denner and S. Pozzorini, *One loop leading logarithms in electroweak radiative corrections. 2. Factorization of collinear singularities*, *Eur. Phys. J. C* **21** (2001) 63 [hep-ph/0104127].

[169] M. Ciafaloni, P. Ciafaloni and D. Comelli, *Anomalous Sudakov Form Factors*, *JHEP* **03** (2010) 072 [0909.1657].

[170] A. Denner, B. Jantzen and S. Pozzorini, *Two-loop electroweak next-to-leading logarithmic corrections to massless fermionic processes*, *Nucl. Phys. B* **761** (2007) 1 [hep-ph/0608326].

[171] J. H. Kuhn, S. Moch, A. A. Penin and V. A. Smirnov, *Next-to-next-to-leading logarithms in four fermion electroweak processes at high-energy*, *Nucl. Phys. B* **616** (2001) 286 [hep-ph/0106298].

[172] L. Ricci, R. Torre and A. Wulzer, *On the W&Y interpretation of high-energy Drell-Yan measurements*, *JHEP* **02** (2021) 144 [2008.12978].

[173] CLICdp collaboration, H. Abramowicz et al., *Top-Quark Physics at the CLIC Electron-Positron Linear Collider*, *JHEP* **11** (2019) 003 [1807.02441].

[174] G. Durieux and O. Matsedonskyi, *The top-quark window on compositeness at future lepton colliders*, *JHEP* **01** (2019) 072 [1807.10273].

[175] M. Farina, G. Panico, D. Pappadopulo, J. T. Ruderman, R. Torre and A. Wulzer, *Energy helps accuracy: electroweak precision tests at hadron colliders*, *Phys. Lett. B* **772** (2017) 210 [1609.08157].

[176] C. Degrande, G. Durieux, F. Maltoni, K. Mimasu, E. Vryonidou and C. Zhang, *Automated one-loop computations in the standard model effective field theory*, *Phys. Rev. D* **103** (2021) 096024 [2008.11743].

[177] D. B. Kaplan and H. Georgi, $SU(2) \times U(1)$ *Breaking by Vacuum Misalignment*, *Phys. Lett. B* **136** (1984) 183.

[178] G. Banelli, E. Salvioni, J. Serra, T. Theil and A. Weiler, *The Present and Future of Four Top Operators*, *JHEP* **02** (2021) 043 [2010.05915].

[179] G. Durieux, M. Perelló, M. Vos and C. Zhang, *Global and optimal probes for the top-quark effective field theory at future lepton colliders*, *JHEP* **10** (2018) 168 [1807.02121].

[180] T. Appelquist, B. A. Dobrescu and A. R. Hopper, *Nonexotic Neutral Gauge Bosons*, *Phys. Rev. D* **68** (2003) 035012 [hep-ph/0212073].

[181] L. D. Landau and E. M. Lifshitz, *Quantum Electrodynamics*, vol. 4 of *Course of Theoretical Physics*. Butterworth-Heinemann, 1982.

[182] V. V. Sudakov, *Vertex parts at very high energies in quantum electrodynamics*, *Zh. Eksp. Teor. Fiz.* **30** (1956) 87.

[183] A. Wulzer, *An Equivalent Gauge and the Equivalence Theorem*, *Nucl. Phys. B* **885** (2014) 97 [1309.6055].

[184] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, *Phys. Lett.* **8** (1964) 214.

[185] A. Ali, L. Maiani and A. D. Polosa, *Multiquark Hadrons*. Cambridge University Press, 6, 2019, 10.1017/9781316761465.

[186] L. Maiani and A. Pilloni, *GGI Lectures on Exotic Hadrons*, 7, 2022, 2207.05141.

[187] JPAC collaboration, M. Albaladejo et al., *Novel approaches in Hadron Spectroscopy*, 2112.13436.

[188] Belle collaboration, S. K. Choi et al., *Observation of a narrow charmonium-like state in exclusive $B^{\pm} \to K^{\pm}\pi^{+}\pi^{-}J/\psi$ decays*, *Phys. Rev. Lett.* **91** (2003) 262001 [hep-ex/0309032].

[189] Belle collaboration, S. K. Choi et al., *Observation of a resonance-like structure in the $\pi^{\pm}\psi'$ mass distribution in exclusive $B \to K\pi^{\pm}\psi'$ decays*, *Phys. Rev. Lett.* **100** (2008) 142001 [0708.1790].

[190] LHCb collaboration, R. Aaij et al., *Observation of an exotic narrow doubly charmed tetraquark*, 2109.01038.

[191] LHCb collaboration, R. Aaij et al., *Study of the doubly charmed tetraquark T_{cc}^{+}*, 2109.01056.

[192] BaBar collaboration, B. Aubert et al., *Study of the $B \to J/\psi K^{-}\pi^{+}\pi^{-}$ decay and measurement of the $B \to X(3872)K^{-}$ branching fraction*, *Phys. Rev. D* **71** (2005) 071103 [hep-ex/0406022].

[193] CDF collaboration, D. Acosta et al., *Observation of the narrow state $X(3872) \to J/\psi\pi^{+}\pi^{-}$ in $\bar{p}p$ collisions at $\sqrt{s} = 1.96$ TeV*, *Phys. Rev. Lett.* **93** (2004) 072001 [hep-ex/0312021].

[194] D0 collaboration, V. M. Abazov et al., *Observation and properties of the $X(3872)$ decaying to $J/\psi\pi^{+}\pi^{-}$ in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, *Phys. Rev. Lett.* **93** (2004) 162002 [hep-ex/0405004].

[195] CMS collaboration, S. Chatrchyan et al., *Measurement of the X(3872) Production Cross Section Via Decays to $J/\psi\pi^+\pi^-$ in pp collisions at $\sqrt{s}$ = 7 TeV*, *JHEP* **04** (2013) 154 [1302.3968].

[196] LHCb collaboration, R. Aaij et al., *Observation of $X(3872)$ production in pp collisions at $\sqrt{s} = 7$ TeV*, *Eur. Phys. J. C* **72** (2012) 1972 [1112.5310].

[197] E. Braaten and H.-W. Hammer, *Interpretation of Neutral Charm Mesons near Threshold as Unparticles*, *Phys. Rev. Lett.* **128** (2022) 032002 [2107.02831].

[198] A. Esposito, L. Maiani, A. Pilloni, A. D. Polosa and V. Riquer, *From the line shape of the X(3872) to its structure*, *Phys. Rev. D* **105** (2022) L031503 [2108.11413].

[199] E. S. Swanson, *Short range structure in the X(3872)*, *Phys. Lett. B* **588** (2004) 189 [hep-ph/0311229].

[200] N. A. Tornqvist, *Isospin breaking of the narrow charmonium state of Belle at 3872-MeV as a deuson*, *Phys. Lett. B* **590** (2004) 209 [hep-ph/0402237].

[201] F.-K. Guo, C. Hanhart, U.-G. Meißner, Q. Wang, Q. Zhao and B.-S. Zou, *Hadronic molecules*, *Rev. Mod. Phys.* **90** (2018) 015004 [1705.00141].

[202] L. D. Landau and E. M. Lifshits, *Quantum Mechanics: Non-Relativistic Theory*, vol. v.3 of *Course of Theoretical Physics*. Butterworth-Heinemann, Oxford, 1991.

[203] L. Maiani, F. Piccinini, A. D. Polosa and V. Riquer, *Diquark-antidiquarks with hidden or open charm and the nature of X(3872)*, *Phys. Rev. D* **71** (2005) 014028 [hep-ph/0412098].

[204] L. Maiani, F. Piccinini, A. D. Polosa and V. Riquer, *The Z(4430) and a New Paradigm for Spin Interactions in Tetraquarks*, *Phys. Rev. D* **89** (2014) 114010 [1405.1551].

[205] Y. A. Smorodinsky, *Normalization of the deuteron wave function*, *Dokl.Akad.Nauk SSSR.* **60** (1948) 217.

[206] C. Bignamini, B. Grinstein, F. Piccinini, A. D. Polosa and C. Sabelli, *Is the X(3872) Production Cross Section at Tevatron Compatible with a Hadron Molecule Interpretation?*, *Phys. Rev. Lett.* **103** (2009) 162001 [0906.0882].

[207] A. Esposito, A. L. Guerrieri, L. Maiani, F. Piccinini, A. Pilloni, A. D. Polosa et al., *Observation of light nuclei at ALICE and the X(3872) conundrum*, *Phys. Rev. D* **92** (2015) 034028 [1508.00295].

[208] P. Artoisenet and E. Braaten, *Production of the X(3872) at the Tevatron and the LHC*, *Phys. Rev. D* **81** (2010) 114018 [0911.2016].

[209] A. Esposito, E. G. Ferreiro, A. Pilloni, A. D. Polosa and C. A. Salgado, *The nature of X(3872) from high-multiplicity pp collisions*, *Eur. Phys. J. C* **81** (2021) 669 [2006.15044].

[210] LHCb collaboration, R. Aaij et al., *Observation of Multiplicity Dependent Prompt $\chi_{c1}(3872)$ and $\psi(2S)$ Production in pp Collisions*, *Phys. Rev. Lett.* **126** (2021) 092001 [2009.06619].

[211] S. Weinberg, *Evidence That the Deuteron Is Not an Elementary Particle*, *Phys. Rev.* **137** (1965) B672.

[212] G.-Y. Chen, W.-S. Huo and Q. Zhao, *Identifying the structure of near-threshold states from the line shape*, *Chin. Phys. C* **39** (2015) 093101 [1309.2859].

[213] S. Fleming, M. Kusunoki, T. Mehen and U. van Kolck, *Pion interactions in the $X(3872)$*, *Phys. Rev. D* **76** (2007) 034006 [hep-ph/0703168].

[214] M. Jansen, H. W. Hammer and Y. Jia, *Light quark mass dependence of the X(3872) in an effective field theory*, *Phys. Rev. D* **89** (2014) 014033 [1310.6937].

[215] E. Braaten, *Galilean-invariant effective field theory for the X(3872)*, *Phys. Rev. D* **91** (2015) 114007 [1503.04791].

[216] E. Braaten, L.-P. He and J. Jiang, *Galilean-invariant effective field theory for the $X(3872)$ at next-to-leading order*, *Phys. Rev. D* **103** (2021) 036014 [2010.05801].

[217] M. Suzuki, *The X(3872) boson: Molecule or charmonium*, *Phys. Rev. D* **72** (2005) 114013 [hep-ph/0508258].

[218] S. Weinberg, *Lectures on Quantum Mechanics*. Cambridge University Press, 2013.

[219] R. Jackiw, *Delta-function potentials in two-and three-dimensional quantum mechanics*, *M.A.B. Bég memorial volume* (1991) 25.

[220] R. Jackiw, *Diverse Topics in Theoretical and Mathematical Physics*, Advanced Series in Mathematical Physics. World Scientific, 1995.

[221] E. Braaten and M. Kusunoki, *Low-energy universality and the new charmonium resonance at 3870-MeV*, *Phys. Rev. D* **69** (2004) 074005 [hep-ph/0311147].

[222] L. Maiani, A. D. Polosa and V. Riquer, *The new resonances Zcs(3985) and Zcs(4003) (almost) fill two tetraquark nonets of broken SU(3)f*, *Sci. Bull.* **66** (2021) 1616 [2103.08331].

[223] S. Dimopoulos and L. Susskind, *On the Baryon Number of the Universe*, *Phys. Rev. D* **18** (1978) 4500.

[224] V. A. Kuzmin, V. A. Rubakov and M. . E. Shaposhnikov, *On the Anomalous Electroweak Baryon Number Nonconservation in the Early Universe*, *Phys. Lett. B* **155** (1985) 36.

[225] A. G. Cohen, D. B. Kaplan and A. . E. Nelson, *WEAK SCALE BARYOGENESIS*, *Phys. Lett. B* **245** (1990) 561.

[226] C. Grojean, G. Servant and . J. D. Wells, *First-order electroweak phase transition in the standard model with a low cutoff*, *Phys. Rev. D* **71** (2005) 036001 [hep-ph/0407019].

[227] ACME collaboration, V. Andreev et al., *Improved limit on the electric dipole moment of the electron*, *Nature* **562** (2018) 355.

[228] V. Cirigliano, Y. Li, . S. Profumo and M. J. Ramsey-Musolf, *MSSM Baryogenesis and Electric Dipole Moments: An Update on the Phenomenology*, *JHEP* **01** (2010) 002 [0910.4589].

[229] M. D'Onofrio, K. Rummukainen and . A. Tranberg, *Sphaleron Rate in the Minimal Standard Model*, *Phys. Rev. Lett.* **113** (2014) 141602 [1404.3565].

[230] F. R. Klinkhamer and N. S. Manton, *A Saddle Point Solution in the Weinberg-Salam Theory*, *Phys. Rev. D* **30** (1984) 2212.

[231] B. Kleihaus, J. Kunz and Y. Brihaye, *The Electroweak sphaleron at physical mixing angle*, *Phys. Lett. B* **273** (1991) 100.

[232] N. S. Manton, *Topology in the Weinberg-Salam Theory*, *Phys. Rev. D* **28** (1983) 2019.

[233] G. D. Moore, C.-r. Hu and B. Muller, *Chern-Simons number diffusion with hard thermal loops*, *Phys. Rev. D* **58** (1998) 045001 [hep-ph/9710436].

[234] P. B. Arnold and L. D. McLerran, *Sphalerons, Small Fluctuations and Baryon Number Violation in Electroweak Theory*, *Phys. Rev. D* **36** (1987) 581.

[235] S. Weinberg, *Gauge and Global Symmetries at High Temperature*, *Phys. Rev. D* **9** (1974) 3357.

[236] R. N. Mohapatra and G. Senjanovic, *Broken Symmetries at High Temperature*, *Phys. Rev. D* **20** (1979) 3390.

[237] R. N. Mohapatra and G. Senjanovic, *Soft CP Violation at High Temperature*, *Phys. Rev. Lett.* **42** (1979) 1651.

[238] P. Salomonson, B. S. Skagerstam and A. Stern, *On the Primordial Monopole Problem in Grand Unified Theories*, *Phys. Lett. B* **151** (1985) 243.

[239] G. R. Dvali, A. Melfo and G. Senjanovic, *Is There a monopole problem?*, *Phys. Rev. Lett.* **75** (1995) 4559 [hep-ph/9507230].

[240] S. Dodelson and L. M. Widrow, *BARYON SYMMETRIC BARYOGENESIS*, *Phys. Rev. Lett.* **64** (1990) 340.

[241] M. J. Ramsey-Musolf, P. Winslow and G. White, *Color Breaking Baryogenesis*, *Phys. Rev. D* **97** (2018) 123509 [1708.07511].

[242] P. Meade and H. Ramani, *Unrestored Electroweak Symmetry*, *Phys. Rev. Lett.* **122** (2019) 041802 [1807.07578].

[243] I. Baldes and G. Servant, *High scale electroweak phase transition: baryogenesis ⁀ extbackslash& symmetry non-restoration*, *JHEP* **10** (2018) 053 [1807.08770].

[244] V. Agrawal, S. M. Barr, J. F. . Donoghue and D. Seckel, *Viable range of the mass scale of the standard model*, *Phys. Rev. D* **57** (1998) 5480 [hep-ph/9707380].

[245] N. Arkani-Hamed, S. Dimopoulos and S. Kachru, *Predictive landscapes and new physics at a TeV*, hep-th/0501082.

[246] L. Senatore, *Hierarchy from baryogenesis*, *Phys. Rev. D* **73** (2006) 043513 [hep-ph/0507257].

[247] P. H. Ginsparg, *First Order and Second Order Phase Transitions in Gauge Theories at Finite Temperature*, *Nucl. Phys. B* **170** (1980) 388.

[248] A. D. Linde, *Infrared Problem in Thermodynamics of the Yang-Mills Gas*, *Phys. Lett. B* **96** (1980) 289.

[249] D. J. Gross, R. D. Pisarski and . L. G. Yaffe, *QCD and Instantons at Finite Temperature*, *Rev. Mod. Phys.* **53** (1981) 43.

[250] T. Appelquist and R. D. Pisarski, *High-Temperature Yang-Mills Theories and Three-Dimensional Quantum Chromodynamics*, *Phys. Rev. D* **23** (1981) 2305.

[251] D. Curtin, P. Meade and C.-T. Yu, *Testing Electroweak Baryogenesis with Future Colliders*, *JHEP* **11** (2014) 127 [1409.0005].

[252] S. Dawson et al., *Working Group Report: Higgs Boson*, in *Community Summer Study 2013: Snowmass on the Mississippi*, 10, 2013, 1310.8361.

[253] J. M. Cline, K. Kainulainen, P. . Scott and C. Weniger, *Update on scalar singlet dark matter*, *Phys. Rev. D* **88** (2013) 055025 [1306.4710].

[254] A. Beniwal, F. Rajec, C. Savage, P. Scott, C. Weniger, M. White et al., *Combined analysis of effective Higgs portal dark matter models*, *Phys. Rev. D* **93** (2016) 115016 [1512.06458].

[255] GAMBIT collaboration, P. Athron et al., *Status of the scalar singlet dark matter model*, *Eur. Phys. J. C* **77** (2017) 568 [1705.07931].

[256] XENON collaboration, E. Aprile et al., *First Dark Matter Search Results from the XENON1T Experiment*, *Phys. Rev. Lett.* **119** (2017) 181301 [1705.06655].

[257] S. Tremaine and J. E. Gunn, *Dynamical Role of Light Neutral Leptons in Cosmology*, *Phys. Rev. Lett.* **42** (1979) 407.

[258] N. Arkani-Hamed, T. Cohen, . R. T. D'Agnolo, A. Hook, H. D. Kim and D. Pinner, *Solving the Hierarchy Problem at Reheating with a Large Number of Degrees of Freedom*, *Phys. Rev. Lett.* **117** (2016) 251801 [1607.06821].

[259] T. Cohen, R. T. D'Agnolo and . M. Low, *Freezing in the hierarchy problem*, *Phys. Rev. D* **99** (2019) 031702 [1808.02031].

[260] A. Belyaev, G. Cacciapaglia, I. P. Ĩvanov, F. Rojas-Abatte and M. Thomas, *Anatomy of the Inert Two Higgs Doublet Model in the light of the LHC and non-LHC Dark Matter Searches*, *Phys. Rev. D* **97** (2018) 035011 [1612.00511].

[261] P. Creminelli, A. Nicolis and . R. Rattazzi, *Holography and the electroweak phase transition*, *JHEP* **03** (2002) 051 [hep-th/0107141].

[262] R. Contino, Y. Nomura and A. Pomarol, *Higgs as a holographic pseudoGoldstone boson*, *Nucl. Phys. B* **671** (2003) 148 [hep-ph/0306259].

Curriculum Vitae

Alfredo Glioti

Contact Information

Chemin des Paleyres 16
1006, Lausanne, VD
Switzerland

Phone: +39 377 12 55 730
E-mail: alfredo.glioti@epfl.ch

Academic experience

École Polytechnique Fédérale de Lausanne, Lausanne, VD, Switzerland

PhD in Physics — 1/10/2017 — current
- Topic: Quantum Field Theory and Physics Beyond the Standard Model
- Supervisor: Prof. Riccardo Rattazzi

Teaching assistant
- Quantum Field Theory I – Prof. Riccardo Rattazzi — 10/2021
- Quantum Field Theory I – Prof. Riccardo Rattazzi — 10/2020
- Quantum Field Theory II – Prof. Riccardo Rattazzi — 2/2020
- Quantum Field Theory I – Prof. Riccardo Rattazzi — 10/2019
- Quantum Field Theory II – Prof. Riccardo Rattazzi — 2/2019
- Quantum Field Theory I – Prof. Riccardo Rattazzi — 10/2018
- General Relativity and Cosmology II – Prof. Mikhail Shaposhnikov — 2/2018

Publications

[1] C. Aime et al., *Muon Collider Physics Summary*, 2203.07256.

[2] Muon Collider collaboration, J. de Blas et al., *The physics case of a 3 TeV muon collider stage*, 2203.07261.

[3] S. Chen, A. Glioti, R. Rattazzi, L. Ricci and A. Wulzer, *Learning from radiation at a very high energy lepton collider*, *JHEP* **05** (2022) 180 [2202.10509].

[4] S. Chen, A. Glioti, G. Panico and A. Wulzer, *Parametrized classifiers for optimal EFT sensitivity*, *JHEP* **05** (2021) 247 [2007.10356].

[5] A. Glioti, R. Rattazzi and L. Vecchi, *Electroweak Baryogenesis above the Electroweak Scale*, *JHEP* **04** (2019) 027 [1811.11740].

Education

University of Perugia, Perugia, Italy

Master's degree (*Laurea Magistrale*) in Physics — 1/10/2015 — 28/9/2017
- Thesis: Force-Free Electrodynamics approach to Black Hole Magnetospheres
- Supervisor: Prof. Gianluca Grignani

Bachelor's degree (*Laurea Triennale*) in Physics — 1/10/2012 — 29/9/2015
- Thesis: Potential, Fields and the Aharonov-Bohm Effect
- Supervisor: Prof. Simone Pacetti

Talks

ML4Jets 2021, Heidelberg, Germany — 7/2021
Higgs and Effective Field Theory 2021, Hefei, China — 4/2021
Higgs Hunting 2019, Orsay/Paris, France — 7/2019

Courses

Supersymmetry - Prof. Luca Vecchi — 9/2020
Deep Learning - Prof. François Fleuret — 3/2019
Before and Behind the Standard Model - Prof. Andrea Wulzer — 3/2019
Conformal Field Theory and Gravity - Prof. João Penedones — 9/2018
Advanced Quantum Field Theory - Prof. Alessandro Vichi — 9/2017
Gauge Theories and the Standard Model - Prof. Riccardo Rattazzi — 9/2017

SCHOOLS	GGI Lectures on the Theory of Fundamental Interactions, Doctoral School	1/2019
	Mass: From the Higgs to Cosmology, Doctoral School	7/2018
LANGUAGES	Italian: Native language.	
	English: Fluent (speaking, reading, writing).	
	French: Intermediate (reading); basic (speaking, writing).	
COMPUTER SKILLS	Programming languages: Mathematica, Python.	
	Knowledge of the Deep Learning package PyTorch.	