VNIVERSITAT DE VALÈNCIA

Probing the top quark couplings within the ATLAS detector and EFT global fits

Tesi doctoral

Marcos Miralles López

IFIC (Universitat de València – CSIC)
Departament de Física Atòmica, Molecular i Nuclear
Programa de Doctorat en Física

Directora:
Dra. María Moreno Llácer

València, Juny 2023

La **Dra. María Moreno Llácer**, investigadora Ramón i Cajal del Institut de Física Corpuscular (UV–CSIC) i docent del departament de Física Atòmica, Molecular i Nuclear de la Universitat de València, com a directora de tesi,

Certifica:

Que la present memòria "Probing the top quark couplings within the ATLAS detector and EFT global fits" ha sigut realitzada sota la seua direcció o supervisió en l'Institut de Física Corpúscular, centre mixte de la Universitat de València i del CSIC, per **Marcos Miralles López** i constitueix la seua tesi per a optar al títol de Doctor per la Universitat de València una vegada cursats els estudis en el Doctorat en Física.

I per a que així conste, en compliment de la legislació vigent, presenta en el departament de Física Atòmica, Molecular i Nuclear de la Universitat de València la referida tesi doctoral, i firma el present certificat.

València, a 22 de maig de 2023

María Moreno Llácer

Declaration

This dissertation is the result of my own work, except where explicit reference is made to the work of others, and has not been submitted for another qualification to this or any other university.

Marcos Miralles López

A mi familia,

a la cual siempre he tenido cerca y de la cual he recibido el apoyo y el cariño durante toda mi vida para llegar hasta aquí, y lo que me queda.
Os quiero.

A mi perrita Luna,

que aunque no me entienda me ha acompañado con su monez y alegría incondicional.
∴

A mis amigos y compañeros,

Naseem, Paolo, Fabio, Mariam, Marta, Quique, Sergio, Lorenzo, Mariia, Jesús, Florencia, Josep, Pablo, Alfonso, Juan, Victoria, Marco, Montse y en especial a David y a Adrián por haber compartido conmigo mis penas y alegrías y su amistad que llevaré siempre conmigo.
Muchísimas gracias por todo.

Contents

"It's a warm summer evening
in ancient Greece . . . "

Preface

The work presented in this thesis targets the properties of the most massive elementary particle of the Standard Model (SM): the top quark. In particular, it scrutinises the electroweak interactions of the top quark with other SM particles using the proton–proton (pp) collision data delivered mainly by the Large Hadron Collider (LHC) and recorded by the ATLAS detector between 2015 and 2018.

The first part of this thesis focuses on the search of possible deviations of the SM expectations in several high–energy physics observables, in the context of an effective field theory (EFT) extension of the SM. This study is greatly motivated by the latest differential cross–section precision measurements provided by the ATLAS experiment, as well as including data from the CMS detector and other experiments such as Tevatron or LEP/SLC; which allow for a rigorous characterisation of the electroweak interactions of the top quark. The results of this global analysis are presented at next–to–leading order accuracy, including a careful study of the impact of correlations among the observables, and uncertainties in the EFT setup itself. The obtained limits on the set of chosen Wilson coefficients (C_i/Λ^2, where Λ is the new energy scale) present a good agreement between their central values and the SM expectation, with 95% probability bounds in the range from ± 0.35 to ± 8 TeV^{-2}. These results represent an improvement over previous studies and expose the need for novel observables that are able to provide complementary information to further constrain the EFT operator coefficients.

The second part of this thesis describes the first search for one of such observables: the leptonic charge asymmetry A_c^l of top–antitop quark pair production in association with a W boson ($t\bar{t}W$). Due to the unique properties of $t\bar{t}W$ production, the charge asymmetry is expected to be large and to showcase a significant sensitivity to a reduced set of four–fermion EFT operators, and to the chiral nature of possible new physics in this process. This search is based on $\sqrt{s} = 13$ TeV pp collision data collected by ATLAS during the 2015–2018 data–taking period, corresponding to an integrated luminosity of 139 fb^{-1}. The analysis targets a final–state topology with exactly three charged light leptons (electrons or muons). The selected events are distributed in multiple regions corresponding to positive and negative differences between the pseudorapidities of the charged leptons from top quark and antitop quark decays, and are used to extract the charge asymmetry using a profile–likelihood fit. At reconstruction level, the charge asymmetry is found to be $A_c^l(t\bar{t}W) = -0.12 \pm 0.14\,(\text{stat.}) \pm 0.05\,(\text{syst.})$. The result is unfolded to particle level in a fiducial region which is chosen to be close to the reconstruction level region to minimise acceptance effects. At particle level, it is found to be $A_c^l(t\bar{t}W)^{\text{PL}} = -0.11 \pm 0.17\,(\text{stat.}) \pm 0.05\,(\text{syst.})$.

In both cases, the extracted charge asymmetries are in agreement with the SM expectations of $A_c^l(t\bar{t}W)_{\mathrm{SM}} = -0.084\,^{+0.005}_{-0.003}\,(\mathrm{scale}) \pm 0.006\,(\mathrm{MC\ stat.})$ and $A_c^l(t\bar{t}W)^{\mathrm{PL}}_{\mathrm{SM}} = -0.063\,^{+0.007}_{-0.004}\,(\mathrm{scale}) \pm 0.004\,(\mathrm{MC\ stat.})$ respectively. The results are severely dominated by the statistical component of the uncertainty, motivating the study of the A_c^l once more data are available in the coming LHC data–taking periods.

One of the main challenges in the future of high–energy physics is the improvement of the accuracy at which Monte Carlo (MC) generators are able to simulate the experimental observables. In this respect, testing and validation work has been performed in this thesis to assess the MC modelling of the $t\bar{t}W$ sample by studying next–to–leading order multi–leg merged configurations and rare electroweak production mechanisms. In addition, the impact of the choice of the renormalization and factorization scales on the predicted cross–sections and selected observables is examined. These studies aim to improve the ATLAS MC simulation and reduce the systematic theoretical modelling uncertainties, which are already a limiting factor in several experimental LHC results.

The final part of this thesis is devoted to the study of the Yukawa coupling of the top quark, i.e. its coupling to the Higgs boson. Being the heaviest elementary particle, this coupling is the largest one in the SM and is expected to be the most sensitive to new physics effects. In this regard, it can be used to probe and extract limits on a well–known limitation of the SM: the insufficient $\mathcal{CP}$–violating terms to explain, for example, the matter–antimatter asymmetry of the early universe. In particular, this work concentrates on the validation and production of MC simulations that introduce a $\mathcal{CP}$–violating term in the coupling between the top quark and the Higgs boson using an EFT prescription. These samples are then used to build a hypothesis test to extract exclusion limits on the $\mathcal{CP}$ mixing angle α. This analysis also uses 139 fb^{-1} of $\sqrt{s} =$ 13 TeV pp collision data collected by ATLAS. Higgs bosons are identified via the diphoton decay channel, and their production in association with a top–antitop quark pair or a single top quark is studied. The $\mathcal{CP}$–odd hypothesis ($\alpha = 90°$) is excluded at 3.9σ and $|\alpha| > 43°$ is excluded at 95% confidence level.

1 The Standard Model

The Standard Model of elementary particles is the theory that presents the joint description of the electromagnetic (EM), weak and strong interactions. It is one of the major achievements in physical sciences in the last century. It is built upon the first discoveries in the early 20th century such as that of the proton in the atomic nucleus in 1919 [1] by E. Rutherford; the introduction of the concept of the relativistic wave equation for electrons by P. Dirac and E. Schrödinger [2, 3]*; the unification of the electromagnetic and weak forces through the presence of a weak–interacting boson (known as the *electroweak theory*) by S.L. Glashow, S. Weinberg and A. Salam [4–7]†; the formulation of a quantum field theory of strong interactions by H. Fritzsch and M. Gell-Mann [8] (introducing of quarks and gluons into the SM); the introduction of the Higgs mechanism by P. Higgs and F. Englert [9, 10]‡; among many other crucial steps towards its complete understanding. With these, the theoretical framework known as the Standard Model had been developed, and the elementary particles that constitute it had been defined.

The road to discover all these elementary particles, the building blocks of nature, started with the electron during the cathode ray tube experiment, back in 1897 by J.J. Thompson [11]. These experiments have seen great technological advances with the construction of machines and detectors in some of the most advanced scientific centers nowadays such as SLAC and Fermilab in the United States; DESY in Germany; or CERN in the France–Switzerland border. The work of thousands of professionals and researchers throughout the last decades has allowed the discovery of all the particles described by the SM. This culminated in the discovery of the Higgs boson in 2012, by both the ATLAS and CMS experiments [12, 13] at CERN, which brought major success to the SM theoretical framework.

The SM is a quantum field theory, which means that its fundamental objects, i.e. the elementary particles, are described by localised vibrations of their underlying quantum fields. It treats both matter and force fields with the same formalism as interactions are regarded to be mediated by the particles themselves. The dynamics and kinematics of such fields are described by the Lagrangian density $\mathscr{L}_0$. Moreover, the SM is a gauge theory, meaning that the Lagrangian is invariant under local transformations. The group of gauge transformations under which a Lagrangian is invariant is called a symmetry (gauge) group, and the SM is defined by the $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ symmetry group§. It is able to describe the strong, EM and weak forces by the exchange

*They both shared the Nobel Prize for this in 1933.

†They received the Nobel Prize in 1979.

‡Both were awarded the Nobel Prize in 2013 after the discovery of the Higgs boson.

§Definitions of these symmetry groups are given in the dedicated subsections below.

of spin–1 gauge fields: eight massless gluons for the strong interaction, a massless photon for the EM one, and three massive bosons ($W^{\pm}$ and Z) for the weak interaction. The matter content in the SM is given by the two groups of spin–1/2 fermions: leptons (charged leptons ℓ and neutrinos ν) and quarks (up– q_u and down–type q_d). These fermions are organised in three *flavour* generations (columns) and types (rows) as shown in Figure 1.1. Fermion fields can be either right– or left–handed. In the SM, left–handed fields transform as $SU(2)_L$ doublets and right–handed fields as $SU(2)_L$ singlets. Thus, one could represent the fermion content in each of the generations as[¶]

$$\begin{bmatrix} \nu_\ell & q_u \\ \ell^- & q_d \end{bmatrix} \equiv \begin{pmatrix} \nu_\ell \\ \ell^- \end{pmatrix}_L, \quad \begin{pmatrix} q_u \\ q_d \end{pmatrix}_L, \quad \ell^-_R, \quad q_{uR}, \quad q_{dR}, \tag{1.1}$$

plus their corresponding antiparticles[‖].

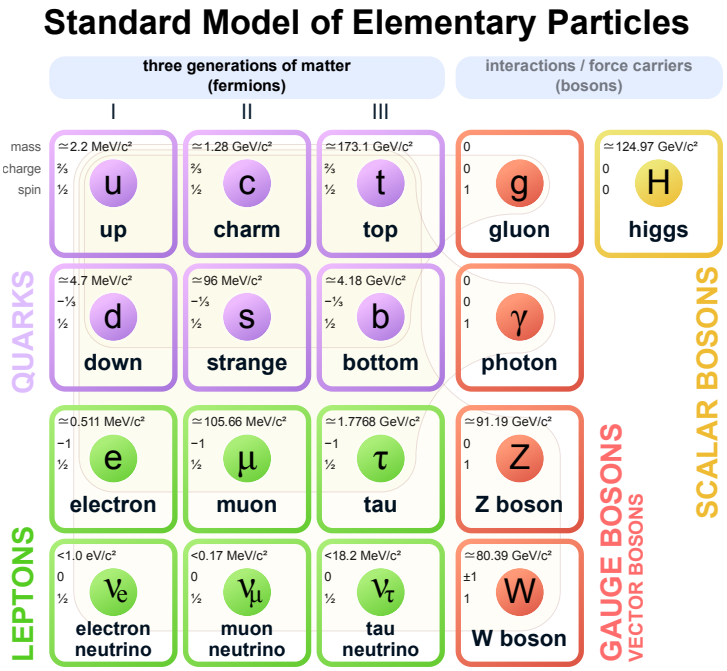


Figure 1.1: Fundamental particles described by the Standard Model including information about their properties. Source: Ref. [14].

[¶] Since neutrinos are massless in the SM, they have no right–handed field which is necessary to build a Dirac mass term in the Lagrangian.

[‖] Antiparticle fields are the result of applying the charge conjugation transformation, $\mathcal{C}$, over the associated particle fields. They possess the same mass, but opposite physical charge to their associated particle.

The dynamics of the SM Lagrangian depend on a set of free parameters whose numerical values have been obtained experimentally. The latest measurements of these parameters is summarised in Table 1.1.

The Spontaneous Symmetry Breaking (SSB) mechanism of the electroweak (EW) group to the electromagnetic group:

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \xrightarrow{SSB} SU(3)_C \otimes U(1)_{\mathrm{EM}}, \tag{1.2}$$

generates the fermion masses, along with those for the weak gauge bosons. The SSB mechanism also predicts a new physical scalar field, the Higgs boson, and quark mixing described by the Cabibbo–Kobayasi–Maskawa (CKM) matrix.

Parameters		Scheme	Value
Electron mass	m_e		0.511 MeV
Muon mass	m_μ		105.7 MeV
Tau mass	m_τ		1.78 GeV
Up quark mass	m_u	$\mu_{\overline{MS}} = 2$ GeV	$2.16^{+0.49}_{-0.26}$ MeV
Down quark mass	m_d	$\mu_{\overline{MS}} = 2$ GeV	$4.67^{+0.48}_{-0.17}$ MeV
Strange quark mass	m_s	$\mu_{\overline{MS}} = 2$ GeV	93^{+11}_{-5} MeV
Charm quark mass	m_c	$\mu_{\overline{MS}} = m_c$	1.27 ± 0.02 GeV
Bottom quark mass	m_b	$\mu_{\overline{MS}} = m_b$	$4.18^{+0.03}_{-0.02}$ GeV
Top quark mass	m_t	ATLAS combination [15]	172.69 ± 0.48 GeV
CKM mixing angle (1)	$\sin\theta_{12}$		0.22650 ± 0.00048
CKM mixing angle (2)	$\sin\theta_{23}$		$0.04053^{+0.00083}_{-0.00061}$
CKM mixing angle (3)	$\sin\theta_{13}$		$0.00361^{+0.00011}_{-0.00009}$
CKM $\mathcal{CP}$–violation phase	δ_{13}		$1.196^{+0.045}_{-0.043}$
Fine structure constant	α^{-1}		137.035 999 139(31)
Strong coupling constant	α_s	$\mu_{\overline{MS}} = M_Z$	0.1181(11)
Weak mixing angle	$\sin^2\theta_W$	$\mu_{\overline{MS}} = M_Z$	0.231 22(4)
Higgs boson mass	M_H		125.25 ± 0.17 GeV
Vacuum expectation value	v		≈ 246 GeV

Table 1.1: A set of 18 free parameters in the SM (values taken from PDG 2022 report [16]). Unlike the leptons, quarks are confined inside hadrons and are not observed as physical particles. This calls for any quark mass estimate to be referenced to the particular theoretical framework that is used to define it. The up, down and strange quark masses are given in a mass–independent minimal subtraction scheme such as the ($\overline{MS}$), at a scale $\mu \approx 2$ GeV. The charm and bottom quark masses are the *running* masses in the ($\overline{MS}$) scheme. The top quark mass is obtained from the best ATLAS combination of direct observations on top quark events. The CKM mixing angles are ordered from largest to smallest.

In the following sections, an introduction to how the SM Lagrangian is built, as well as a description of all fundamental particles and their properties, is given. Section 1.1 gives an overview of quantum electrodynamics (QED), and the mathematical formalism that enables matter and force field interactions known as the gauge principle. Section 1.2 describes the unification of the electromagnetic and weak forces, and the interaction between leptons and weak–interacting bosons. Section 1.3 describes the SSB of the EW sector via the Higgs mechanism. Analogously, Section 1.4 describes quantum chromodynamics (QCD), which is the theory that describes the strong force and introduces interactions between quarks and gluons into the SM. Finally, Section 1.5 gives an overview of the current limitations of the SM which can be studied using a wide range of new beyond the SM (BSM) physics models. Section 1.6 describes one of these BSM approaches that uses effective field theories to encapsulate possible new physics effects.

1.1 Quantum Electrodynamics

QED represents a combination of three main ideas developed in the last decades of modern physics: classical electrodynamics, quantum mechanics and special relativity. It is also the combination of Dirac's theory of the electron and the quantisation of the EM field into individual photons. With the insight of 1965 Nobel Prize winners R. Feynman, J. Schwinger and S. Tomonaga (among many others), QED was developed into a well–defined, renormalisable theory that has been tested with high precision.

1.1.1 Gauge principle

As aforementioned, the QED Lagrangian can be built first by considering the Lagrangian for a quantum field describing a free electron. Here, Dirac's Lagrangian for a free electron can be used:

$$\mathscr{L}_0 = \overline{\psi}(x)\left(i\gamma^\mu\partial_\mu - m\right)\psi(x), \tag{1.3}$$

where $\psi(x)$ is the Dirac electron, m is its mass, and γ^μ are the Dirac matrices. This Lagrangian is invariant under global $U(1)$ transformations,

$$\psi(x) \xrightarrow{U(1)} \psi'(x) \equiv e^{iQ\theta}\psi(x), \tag{1.4}$$

where θ is an arbitrary real constant. However if one allows this phase to depend on a local space–time coordinate, i.e. $\theta = \theta(x)$, then $\mathscr{L}_0$ is no longer invariant:

$$\partial_\mu\psi(x) \xrightarrow{U(1)} e^{i\theta}[\partial_\mu + iQ\partial_\mu\theta(x)]\psi(x). \tag{1.5}$$

Since in a gauge theory, the Lagrangian must be invariant under local $U(1)$ phase transformations (which is known as the gauge principle), the solution is to add a new spin–1 gauge field, $A_\mu(x)$, that transforms as

$$A_\mu(x) \xrightarrow{U(1)} A'_\mu(x) \equiv A_\mu(x) - \frac{1}{e}\partial_\mu\theta(x), \tag{1.6}$$

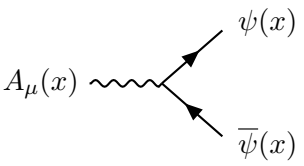
in order to counter the $\partial_\mu\theta(x)$ term in Equation 1.5. In addition, a covariant derivative is defined:

$$D_\mu\psi(x) \equiv \left[\partial_\mu + ieQA_\mu(x)\right]\psi(x). \tag{1.7}$$

With these two pieces, the resulting Lagrangian,

$$\mathscr{L} \equiv \overline{\psi}(x)\left(i\gamma^\mu D_\mu - m\right)\psi(x) = \mathscr{L}_0 - eQA_\mu(x)\overline{\psi}(x)\gamma^\mu\psi(x), \tag{1.8}$$

is invariant under $U(1)$ local transformations. Using the gauge principle has naturally given an interaction between the Dirac fermion and the $A_\mu(x)$ gauge field. This is no other than the well–known QED vertex:



Finally, to make $A_\mu(x)$ a true propagating field, one has to add the gauge–invariant kinetic term to the Lagrangian,

$$\mathscr{L}^{\text{QED}}_{\text{Kin.}} \equiv -\frac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x), \tag{1.9}$$

with $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ being the standard EM field strength which is also invariant under the $U(1)$ transformation in Equation 1.6*. Furthermore, the EM current J^{EM}_μ can be written as

$$eJ^{\text{EM}}_\mu = eQ\overline{\psi}\gamma_\mu\psi. \tag{1.10}$$

1.2 Electroweak interactions

To incorporate weak interactions into the QED Lagrangian, a more complex structure needs to be developed: several fermionic flavours, different properties

*As with the Dirac fermion, one could add a mass term to the $A_\mu(x)$ field: $\mathscr{L}_m = \frac{1}{2}m^2A^\mu A_\mu$. However, this term would break gauge invariance and therefore the $A_\mu(x)$ photon field remains massless. Experimentally, $m_\gamma < 1 \times 10^{-18}$ eV [16] which agrees with this prediction.

and transformations of right– and left–handed fields, the appearance of the massive gauge bosons $W^{\pm}$ and Z, and also the massless photon have to be accommodated.

The need for a weak interaction can be traced back to the observation of kaon K^+ decays into two final states with opposite parity, also known as the $\tau-\theta$ puzzle [17]. Additionally, a parity violating interaction was also needed to explain β decays of nuclei in polarised magnetic fields such as $n \to p\, e^-_L\, \overline{\nu}_{e,R}$ [18]. In this experiment, it was seen how left–handed fermions and right–handed antifermions $(\ell_L, \overline{\ell}_R)$ are preferred over right–handed fermions and left–handed antifermions $(\ell_R, \overline{\ell}_L)$, hence being the only participants in weak transitions. To achieve parity violation, these observations gave rise to the *vector minus axial–vector* $(V-A)$ theory of weak interactions.

1.2.1 Right– and left–handed fields

Dirac (fermion) fields exhibit two properties called chirality and helicity. The chirality operator γ^5 is Lorentz invariant and, in the high energy limit $(E \gg m)$, it shares the same physical meaning as the helicity operator. Therefore, in the chirality basis, the right– and left–handed projectors $P_{R,L} = \frac{1}{2}(\mathbb{1} \pm \gamma^5)$ select the right– and left–handed Dirac fields respectively. This means that right– and left–handed fermion fields can be written as

$$\ell_{R,L} = P_{R,L}\ \psi, \tag{1.11}$$

and their charge conjugates

$$\overline{\ell}_{R,L} = (P_{R,L}\ \psi)^\dagger \gamma^0 = \psi^\dagger P_{R,L} \gamma^0 = \psi^\dagger \gamma^0 P_{L,R} = \overline{\psi} P_{L,R}. \tag{1.12}$$

Furthermore, the $V-A$ weak interaction can be expressed in terms of these chirality states as

$$\overline{\ell}_L \gamma^\mu \ell_L = \overline{\psi} P_R \gamma^\mu P_L \psi = \overline{\psi} \gamma^\mu (P_L)^2 \psi = \overline{\psi} \gamma^\mu P_L \psi = \frac{1}{2} \overline{\psi} \gamma^\mu (\mathbb{1} - \gamma^5) \psi, \tag{1.13}$$

which shows that the weak interaction can only involve left–handed fermions, and right–handed antifermions. This is conducive to maximal parity ($\mathcal{P}$: left $\leftrightarrow$ right) and charge conjugation ($\mathcal{C}$: particle $\leftrightarrow$ antiparticle) violation. However, the combined transformation $\mathcal{CP}$ remains a good symmetry.

Two additional results may be extracted: mass terms must mix right– and left–handed fields as

$$\overline{\psi}\psi = \overline{\ell}_R \ell_L + \overline{\ell}_L \ell_R, \tag{1.14}$$

and the EM current from Equation 1.10 does not mix both components, i.e.

$$J^{\mathrm{EM}}_\mu = Q \overline{\psi} \gamma_\mu \psi = Q(\overline{\ell}_L \gamma_\mu \ell_L + \overline{\ell}_R \gamma_\mu \ell_R). \tag{1.15}$$

1.2.2 Choosing the symmetry group

In analogy to the EM current, charged currents (CCs) can be obtained to describe transitions between pair of fermions whose charge differ by one unit. Such is the case of the muon decay $\mu^- \to \nu_\mu\ e^-\ \overline{\nu}_e$, as well as of the neutron decay mentioned in Section 1.2.

From Equation 1.13, the generic CCs for these processes can be written as

$$\begin{aligned} j_\mu \equiv j_\mu^+ &= \overline{\ell}_L \gamma_\mu \nu_L, \\ j_\mu^\dagger \equiv j_\mu^- &= \overline{\nu}_L \gamma_\mu \ell_L. \end{aligned} \tag{1.16}$$

By introducing a weak isospin ($T = 1/2$) doublet structure for the left–handed fields,

$$\psi_L = \begin{pmatrix} \nu \\ \ell \end{pmatrix}_L, \begin{pmatrix} q_u \\ q_d \end{pmatrix}_L \qquad \begin{matrix} T_3 = +1/2 \\ T_3 = -1/2 \end{matrix} \quad , \tag{1.17}$$

and the usual Pauli matrices τ_i, the CC can be rewritten as

$$J_\mu^\pm = \overline{\psi}_L \gamma_\mu \tau_\pm \psi_L, \tag{1.18}$$

where $\tau_\pm = \frac{1}{2}(\tau_1 \pm i\tau_2)$. Since neutrinos are massless, the right–handed part of the charged lepton $\psi_R \equiv \ell_R$ is accommodated in a weak isospin singlet ($T = 0$). By choosing this representation, a third current ($J_\mu^3 = \overline{\psi}_L \gamma_\mu \tau_3 \psi_L$) is expected to exist that does not change the charge and is thus called the neutral current (NC). The combination of the CC and the NC give the weak isospin triplet of weak currents,

$$J_\mu^i = \overline{\psi}_L \gamma_\mu \frac{\tau_i}{2} \psi_L, \qquad i = 1, 2, 3, \tag{1.19}$$

with algebra

$$[\tau_i, \tau_j] = i\epsilon_{ijk}\tau^k. \tag{1.20}$$

So far, the representations of left–handed fields as $SU(2)_L$ doublets and right–handed fields as $SU(2)_L$ singlets have been obtained. The J_μ^i currents are invariant under the non-abelian global $SU(2)_L$ transformation

$$\psi_L \xrightarrow{SU(2)} \psi'_L = U_L \psi_L, \tag{1.21}$$

with

$$U_L \equiv \exp\left\{ i\frac{\tau_i}{2}\alpha^i \right\}, \qquad U_L U_L^\dagger = U_L^\dagger U_L = 1, \qquad \det U_L = 1. \tag{1.22}$$

In order to combine the weak and EM interactions, an $SU(2)_L$–invariant $U(1)$ current must be constructed since the EM current J_μ^{EM}, from Equation 1.15, is not invariant under $SU(2)_L$ transformations. Such current takes the form

$$J_\mu^Y = \overline{\ell}_R \gamma_\mu Y_R \ell_R + \overline{\psi}_L \gamma_\mu Y_L \psi_L, \tag{1.23}$$

where the hypercharges Y_R and Y_L are the conserved charged operators of the $U(1)_Y$ symmetry. Note that they differ for right– and left–handed fermions. With this, $J_\mu^{\rm EM}$ can be written as

$$J_\mu^{\rm EM} = Q(\overline{\ell}_L\gamma_\mu\ell_L + \overline{\ell}_R\gamma_\mu\ell_R) = \underbrace{\overline{\nu}_L\gamma_\mu\frac{1}{2}\nu_L - \overline{\ell}_L\gamma_\mu\frac{1}{2}\ell_L}_{J_\mu^3} + \underbrace{\frac{1}{2}\overline{\ell}_R\gamma_\mu Y_R\ell_R + \frac{1}{2}\overline{\psi}_L\gamma_\mu Y_L\psi_L}_{\frac{1}{2}J_\mu^Y}, \tag{1.24}$$

from which the relations $Y_R = 2Q$ and $Y_L = 2Q\mp 1$ can be extracted. By taking into account the third component of the weak isospin T_3 for left–handed (doublet: $T_3 = \pm\frac{1}{2}$) and right–handed (singlet: $T_3 = 0$) fields, these relations can be combined into the Gell-Mann–Nishijima formula*:

$$Q = T_3 + \frac{1}{2}Y. \tag{1.25}$$

Following this formula, the quantum numbers for all leptons and quarks are summarised in Table 1.2.

	T	T_3	Q	Y
ν_L	1/2	1/2	0	−1
ℓ_L	1/2	−1/2	−1	−1
ν_R	0	0	0	0
ℓ_R	0	0	−1	−2

	T	T_3	Q	Y
$q_{u,L}$	1/2	1/2	2/3	1/3
$q_{d,L}$	1/2	−1/2	−1/3	1/3
$q_{u,R}$	0	0	2/3	4/3
$q_{d,R}$	0	0	−1/3	−2/3

Table 1.2: Weak quantum numbers for leptons (left) and quarks (right). As expected the right–handed neutrino ν_R does not carry $SU(2)_L$ or $U(1)_Y$ charges, and is therefore decoupled from EW interactions.

Therefore, the correct symmetry group for EW interactions has already been chosen to be

$$SU(2)_L \otimes U(1)_Y. \tag{1.26}$$

1.2.3 The $SU(2)_L \otimes U(1)_Y$ EW Lagrangian

Similarly to Section 1.1.1, to construct an invariant Lagrangian under $SU(2)_L \otimes U(1)_Y$, the gauge principle has to be applied. In first place, four new spin–1 gauge fields have to be introduced: an $SU(2)_L$ triplet $\widetilde{W}_\mu(x) \equiv \frac{\tau_i}{2}W_\mu^i(x)$ and a $U(1)_Y$ singlet $B_\mu(x)$ associated to the weak isospin and hypercharge respectively. In order to keep the Lagrangian invariant, these gauge fields must

*$T_i = \frac{\tau_i}{2}$ and Y are the associated charge operators of the J^i and J^Y currents respectively. They satisfy the $SU(2)_L \otimes U(1)_Y$ algebra $[T_i, T_j] = i\epsilon_{ijk}T^k$ and $[T_i, Y] = 0$.

transform as

$$B_\mu(x) \to B'_\mu(x) \equiv B_\mu(x) - \frac{1}{g'}\partial_\mu\beta(x), \tag{1.27}$$

$$\widetilde{W}_\mu(x) \to \widetilde{W}'_\mu(x) \equiv U_L(x)\widetilde{W}_\mu(x)U_L^\dagger(x) + \frac{i}{g}\partial_\mu U_L(x)U_L^\dagger(x), \tag{1.28}$$

where $U_L(x) \equiv \exp\left\{i\frac{\tau_i}{2}\alpha^i(x)\right\}$ and since $B_\mu(x)$ transforms under a $U(1)$ symmetry, $\beta(x)$ is equivalent to $\theta(x)$ for QED in Section 1.1.1. In second place, the covariant derivative takes the form

$$D_\mu \equiv \partial_\mu + ig\frac{\tau_i}{2}W_\mu^i(x) + ig'\frac{Y}{2}B_\mu(x). \tag{1.29}$$

Using the respective weak isospin and hypercharge currents from Equations 1.19 and 1.23, the resulting (massless) EW Lagrangian is

$$\mathscr{L}_{\text{EW}} = \mathscr{L}_0 - gJ_\mu^i W_i^\mu(x) - \frac{g'}{2}J_\mu^Y B^\mu(x). \tag{1.30}$$

The final pieces are the gauge–invariant kinematic terms for the gauge fields,

$$\mathscr{L}_{\text{Kin.}}^{\text{EW}} \equiv -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W_{\mu\nu}^i W_i^{\mu\nu}, \tag{1.31}$$

where $B_{\mu\nu} \equiv \partial_\mu B_\nu - \partial_\nu B_\mu$ and $W_{\mu\nu}^i \equiv \partial_\mu W_\nu^i - \partial_\nu W_\mu^i - g\epsilon_{ijk}W_\mu^j W_\nu^k$ are their corresponding gauge field strengths. Some final remarks: this kinetic lagrangian gives rise to cubic and quartic self–interactions among the gauge fields. The strength of such interaction is commonly given the same $SU(2)_L$ coupling g. This is also the case for all left–handed fermion interactions via charged or neutral currents. Unlike the QED Lagrangian, gauge symmetry forbids a mass term since it would communicate right– and left–handed fields (see Equation 1.14), that transform differently, and it would lead to the breaking of the symmetry. Hence, the $SU(2)_L \otimes U(1)_Y$ EW Lagrangian contains massless fields, both for fermions and bosons.

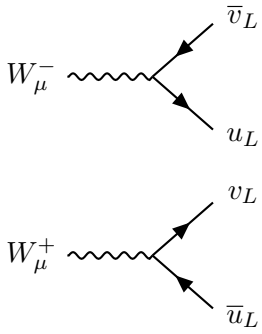
1.2.4 Recovering the "real" bosons

Thus far, a description of EW interactions between all fermions in the SM, their quantum numbers and representations have been obtained. Nonetheless, the usual four weak gauge bosons, i.e. $W^\pm$, Z and the photon, have to be related to the W_μ^i and B_μ gauge fields from the previous section.

Comparing the currents in Equation 1.18 to the EW Lagrangian in Equation 1.30, the charged vector bosons describing CCs can be defined as

$$W_\mu^\pm \equiv \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp iW_\mu^2\right), \tag{1.32}$$

in such a way that



(1.33)

where $u_L \equiv (\nu, q_u)_L^{\mathrm{T}}$ and $v_L \equiv (l, q_d)_L^{\mathrm{T}\dagger}$ since both leptons and quarks behave in the same way. This shows that the universality of the lepton and quark EW interactions is a direct consequence of the chosen gauge symmetry.

On the other hand, the neutral W_μ^3 and B_μ fields are still unmatched. To obtain the neutral vector bosons, Z and γ, the following transformation matrix can be used:

$$\begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} \equiv \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix}, \tag{1.34}$$

where θ_W is the weak mixing angle whose value is shown in Table 1.1.

Substituting these relations into the EW Lagrangian (Equation 1.30) for the neutral current, and using Equations 1.19 and 1.23 gives

$$\begin{aligned} \mathscr{L}_{\mathrm{NC}} = -g J_\mu^3 W_3^\mu - \frac{g'}{2} J_\mu^Y B^\mu = &- \left(g \sin\theta_W J_\mu^3 + g' \cos\theta_W \frac{1}{2} J_\mu^Y \right) A^\mu \\ &- \left(g \cos\theta_W J_\mu^3 - g' \sin\theta_W \frac{1}{2} J_\mu^Y \right) Z^\mu. \end{aligned} \tag{1.35}$$

Recalling that $J_\mu^{\mathrm{EM}} = J_\mu^3 + \frac{1}{2} J_\mu^Y$, the QED interaction Lagrangian, $\mathscr{L}_{\mathrm{QED}} = -e J_\mu^{\mathrm{EM}} A^\mu$, can be recovered from the A^μ piece. This implies that

$$g \sin\theta_W = g' \cos\theta_W = e, \tag{1.36}$$

which links the three couplings together. The other piece corresponds to weak neutral currents mediated by the Z^μ field. Rearranging the relation between the currents and substituting into this second piece yields

$$\begin{aligned} \mathscr{L}_{\mathrm{NC}}^Z &= -\frac{g}{\cos\theta_W} \left(J_\mu^3 - \sin^2\theta_W J_\mu^{\mathrm{EM}} \right) Z^\mu \\ &= -\frac{g}{2\cos\theta_W} Z^\mu \sum_f \overline{\psi}_f \gamma_\mu \left(v_f - a_f \gamma^5 \right) \psi_f, \end{aligned} \tag{1.37}$$

[†]Here, the label T indicates that the vector representation has been transposed.

where f represents any fermion, $v_f = T_{3,f} - 2Q_f \sin^2\theta_W$ and $a_f = T_{3,f}$. Table 1.3 presents a summary of the neutral coupling strengths for all fermions. Hence, the Lagrangian that describes the EW interactions is a combination of the above, i.e.

$$\mathscr{L}_{\text{EW}} \equiv \mathscr{L}_0 + \mathscr{L}_{\text{CC}} + \mathscr{L}_{\text{NC}}^{Z} + \mathscr{L}_{\text{QED}} + \mathscr{L}_{\text{Kin.}}^{\text{EW+QED}} + \text{h.c.} \tag{1.38}$$

	ν	ℓ	q_u	q_d
$2v_f$	1	$-1 + 4\sin^2\theta_W$	$1 - \frac{8}{3}\sin^2\theta_W$	$-1 + \frac{4}{3}\sin^2\theta_W$
$2v_a$	1	-1	1	-1

Table 1.3: Neutral current couplings to fermions.

1.3 Spontaneous Symmetry Breaking

The EW Lagrangian from Equation 1.38 contains massless fields for both the gauge bosons and the fermions. However, the physical $W^{\pm}$ and Z gauge bosons are indeed heavy massive objects which makes this Lagrangian stray away from reality. To include a mass term, the symmetry would have to be broken. The basic premise is that non–symmetric physical states can be obtained from a symmetric Lagrangian that breaks the symmetry and gives mass to the fermions and the bosons. In particular, when considering the lowest energy state (ground state) of a system described by a symmetric Lagrangian $\mathscr{L}$, in case this state is degenerate (there are multiple eigenstates describing the ground state), arbitrarily selecting one of those eigenstates would give a ground state that stops sharing the symmetries of $\mathscr{L}$. This way of obtaining an asymmetric ground state is known as SSB, where the asymmetry is not due to a non–invariant term in $\mathscr{L}$, but rather to an arbitrary selection of one of the degenerate ground states.

1.3.1 The Goldstone theorem

Ferromagnetism presents a similar SSB example. In a ferromagnetic material, the system is rotationally invariant and the ground state spins are aligned in some arbitrary direction, creating a non–vanishing magnetisation. No matter where the magnetisation orientation points to, the system retains its properties. Furthermore, this degenerate ground state is asymmetric and excited states obtained from it also share this asymmetry. In quantum field theory, the ground state is the vacuum. Moreover, the vacuum has to be degenerate and create a non–vanishing asymmetric quantity to break the symmetry.

This quantity is commonly taken to be the vacuum expectation value (v) of a complex scalar field $\phi(x) = \frac{1}{\sqrt{2}}[\phi_1(x) + i\phi_2(x)]$ with Lagrangian

$$\mathscr{L} = \partial_\mu \phi^\dagger \partial^\mu \phi - V(\phi), \qquad V(\phi) = \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2, \tag{1.39}$$

where $\mathscr{L}$ is invariant under global phase transformations of $\phi(x)$. Because the potential must be bounded from below (in order to have a stable ground state), i.e. $\lambda > 0$, two options are left depending on the sign of μ^2:

- $\mu^2 > 0$: unique vacuum state for $\phi = 0$. A massive scalar boson with mass μ and quartic coupling λ,
- $\mu^2 < 0$: continuous vacuum state along the $\Phi_1\Phi_2$ plane with $\phi_1^2 + \phi_2^2 = v^2 = -\frac{\mu^2}{\lambda}$.

It is then common to choose one vacuum state ($\phi_1 = v$, $\phi_2 = 0$), which spontaneously breaks the symmetry, and to perform small excitations over this vacuum parameterised as $\phi(x) \equiv \frac{1}{\sqrt{2}}[v + \eta(x) + i\xi(x)]$. Thus, the Lagrangian takes the form

$$\mathscr{L} = \frac{1}{2}(\partial_\mu \xi)^2 + \frac{1}{2}(\partial_\mu \eta)^2 + \mu^2 \eta^2 + \dots \quad , \tag{1.40}$$

where $\eta(x)$ describes a massive particle of mass $m_\eta^2 = -2\mu^2$, and $\xi(x)$ remains massless*. This result from SSB is well known as the Goldstone Theorem [19–21]. This states that, if a Lagrangian is invariant under a continuous symmetry which is spontaneously broken, there are new massless spin–0 particles (Nambu–Goldstone bosons) equal to the number of broken generators. Figure 1.2 shows the shape of the scalar potential and the process of SSB by selecting an arbitrary ground state.

1.3.2 The Higgs mechanism

Goldstone's Theorem becomes a crucial element if a $SU(2)_L$ doublet of complex fields (with $Y = 1$) is introduced as

$$\phi(x) \equiv \begin{pmatrix} \phi^{(+)}(x) \\ \phi^{(0)}(x) \end{pmatrix} \tag{1.41}$$

into the Lagrangian:

$$\mathscr{L}_{\text{SSB}} = (D_\mu \phi)^\dagger D^\mu \phi - \mu^2 \phi^\dagger \phi - \lambda(\phi^\dagger \phi)^2, \tag{1.42}$$

where D_μ is the covariant derivative from the $SU(2)_L \otimes U(1)_Y$ Lagrangian in Equation 1.29. $\phi(x)$ is also known as the Higgs doublet. By choosing it in this

*This is an unsurprising result since $\xi(x)$ describes transitions between the degenerate vacuum states with the same energy along a flat potential.

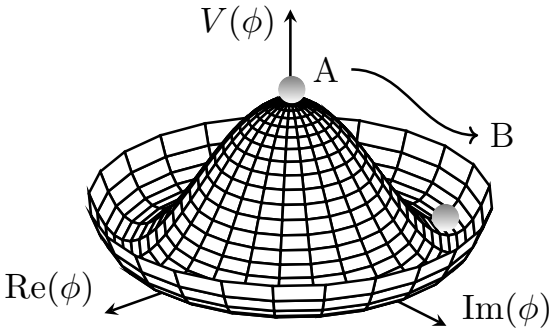


Figure 1.2: Shape of the scalar potential for $\mu^2 < 0$. Selecting an arbitrary ground state ($A \to B$), spontaneously breaks the symmetry.

form, only the neutral field $\phi^{(0)}(x)$ can acquire a vacuum expectation value since the electric charge remains as a conserved quantity. Upon choosing a ground state ($\lambda > 0, \mu^2 < 0$), the $SU(2)_L \otimes U(1)_Y$ group is spontaneously broken into the $U(1)_{\mathrm{QED}}$ group which remains a symmetry of the chosen vacuum. Excitations over this vacuum state are parameterised as

$$\phi(x) = \exp\left\{i\frac{\tau_i}{2}\theta^i(x)\right\}\frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v+H(x)\end{pmatrix} \xrightarrow[\theta^i=0]{\text{Unitarity Gauge}} \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v+H(x)\end{pmatrix}, \tag{1.43}$$

where four new fields, $\theta^i(x)$ and $H(x)$, appear. The three fields $\theta^i(x)$, corresponding to the predicted three massless Goldstone bosons, can be rotated away through a $SU(2)_L$ transformation (unitarity gauge: $\theta^i(x) = 0$) leaving only the Higgs scalar field $H(x)$.

Substitution of Equation 1.43 into Equation 1.42 leads to mass terms for the $W^{\pm}$ and Z weak bosons, as well as for the Higgs boson:

$$\begin{aligned}\mathscr{L}_{\mathrm{SSB}} \xrightarrow{\theta^i=0} \mathscr{L}_H = \frac{1}{2}\partial_\mu H \partial^\mu H + (v+H)^2\left\{\frac{g^2}{4}W^+_\mu W^{-\mu} + \frac{g^2}{8\cos^2\theta_W}Z_\mu Z^\mu\right\}\\ + \mu^2 H^2 + \mathcal{O}(H^3).\end{aligned} \tag{1.44}$$

Therefore, the masses of the heavy bosons in the SM are given by

$$M_W = \frac{1}{2}vg, \qquad \frac{M_W}{M_Z} = \cos\theta_W, \qquad M_H = \sqrt{-2\mu^2}. \tag{1.45}$$

Through this particular case of SSB, known as the Higgs mechanism, the weak bosons of the theory have acquired the desired masses while the photon remains massless since $U(1)_{\mathrm{QED}}$ is still a symmetry of the system. In addition, a new real scalar particle $H(x)$ appears which also exhibits interactions with these

bosons and among itself. As a final remark, it is interesting to count degrees of freedom (d.o.f.) before and after SSB. The d.o.f. introduced by the three of the four scalar fields in $\phi(x)$ (Eq. 1.41) are needed to give masses to the weak gauge bosons. These are no other than the three SSB massless Goldstone bosons which get "absorbed" by each of the weak bosons as an additional longitudinal polarisation and become massive. The Higgs boson can then be seen as the aftermath of having too many scalar fields before SSB. This is summarised in Table 1.4.

$\mathscr{L}$	Fields	d.o.f
massless $\mathscr{L}_{\rm EW}$ (Eq. 1.38), $\phi(x)$	$\phi(x)$ doublet, complex scalar	4
	$W^{\pm}_{\mu}, Z_{\mu}$ massless spin–1 vector	$3 \times 2 = 6$
$\mathscr{L}_{\rm EW} + \mathscr{L}_{H}$, $H(x)$	$H(x)$ real scalar	1
	$W^{\pm}_{\mu}, Z_{\mu}$ massive spin–1 vector	$3 \times 3 = 9$

Table 1.4: Summary of the number of d.o.f before (top) and after (bottom) SSB in the EW sector. The total number of d.o.f. remains constant.

1.3.3 Fermion masses

A mass term of the type $\mathscr{L}_m = -m\overline{\psi}\psi$ is not allowed since it would break the gauge symmetry. However, making use of the Higgs doublet, the solution is to introduce new gauge–invariant fermion–scalar couplings as

$$\begin{aligned}\mathscr{L}_{\rm Y} &= -G_d \begin{pmatrix}\overline{q}_u & q_d\end{pmatrix}_L \phi \, (q_d)_R - G_u \begin{pmatrix}\overline{q}_u & q_d\end{pmatrix}_L \phi^c \, (q_u)_R - G_\ell \begin{pmatrix}\overline{\nu} & \ell\end{pmatrix}_L \phi \, \ell_R + \text{h.c.} \\ &= -\left(1 + \frac{H}{v}\right) \left\{ m_d \overline{q}_d q_d + m_u \overline{q}_u q_u + m_\ell \overline{\ell}\ell \right\},\end{aligned} \tag{1.46}$$

where $\phi^c \equiv i\tau_2 \phi^\dagger$ ($Y = -1$) and G_f denotes the Yukawa coupling for each fermion to the Higgs boson. These parameters are arbitrary but are also fixed by the masses, i.e.

$$m_f = \frac{G_f v}{\sqrt{2}}. \tag{1.47}$$

Figure 1.3 shows the couplings of fermions and gauge bosons[†] to the Higgs boson measured by the ATLAS experiment [22]. The Higgs boson couples strongly to particles with high mass; this is explicitly shown for fermions in Equation 1.47. Because of its mass, the top quark has the largest Yukawa coupling to the Higgs ($G_t \sim 1$) which makes it a crucial particle to study.

[†]For the gauge bosons, V, the coupling to the Higgs is proportional to m_V^2

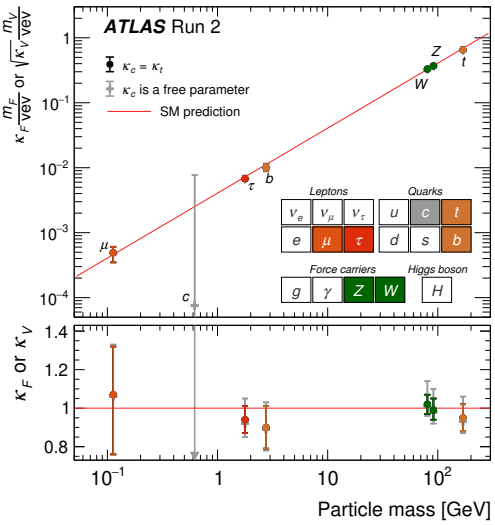


Figure 1.3: Measured values of the couplings for fermions and gauge bosons as a function of their masses for a vacuum expectation value of $v = 246$ GeV by the ATLAS experiment. The SM prediction is shown as the red line. The vertical bar on each point denotes the 68% confidence interval. The lower panel shows the values of the coupling strength modifiers κ_V (for bosons) and κ_F (for fermions). They are defined as the measured cross–section or partial decay value normalised to the SM expectation. Source: Ref. [22] (Fig. 5).

1.3.4 Quark mixing: CKM matrix

In the SM, fermions are grouped as families which share very similar properties other than their masses. With this in mind, the most general Yukawa Lagrangian can be written as

$$\mathscr{L}_{\mathrm{Y}} = -G_d^{ij}\,\overline{Q}_{Li}\,\phi\,(q_d)_{Rj} - G_u^{ij}\,\overline{Q}_{Li}\,\phi^c\,(q_u)_{Rj} - G_\ell^{ij}\,\overline{L}_{Li}\,\phi\,\ell_{Rj} + \text{h.c.} \quad , \tag{1.48}$$

where G_f^{ij} are 3×3 complex matrices and i, j are the generation indices. After SSB, the mass matrices for all fermions can be expressed as

$$\mathbf{M}_f^{ij} = \frac{G_f^{ij} v}{\sqrt{2}}. \tag{1.49}$$

These $\mathbf{M}$ matrices can be diagonalised by unitary matrices, $\mathbf{V}_{L,R}^f$, like $\mathbf{N}_{\text{diag.}}^f = \mathbf{V}_L^{f\dagger}\mathbf{M}_f\mathbf{V}_R^f$. Note that this would also hold for any number N_G of fermion generations, in that event $\mathbf{N}^f = \text{diag}(m_1, \ldots, m_{N_G})$. Upon defining the physical mass eigenstates,

$$u_{L,R} \equiv \mathbf{V}_{L,R}^u (q_u)_L, \quad d_{L,R} \equiv \mathbf{V}_L^d (q_d)_{L,R}, \quad \ell'_{L,R} \equiv \mathbf{V}_{L,R}^\ell \ell_{L,R}, \tag{1.50}$$

$\mathscr{L}_{\rm Y}$ takes the form

$$\mathscr{L}_{\rm Y} = -\left(1+\frac{H}{v}\right)\left\{\overline{d}\,\mathbf{N}^d\,d + \overline{u}\,\mathbf{N}^u\,u + \overline{\ell}'\,\mathbf{N}^\ell\,\ell'\right\}. \tag{1.51}$$

In terms of these mass eigenstates, neutral currents do not change and do not display mixing between fermion flavours. In other words, there are no flavour–changing neutral currents (FCNCs) in the SM Lagrangian (GIM mechanism [23][‡]). On the other hand, charged currents do exhibit flavour mixing. In order to write these CCs in terms of the mass eigenstates, the $N_G \times N_G$ unitarity mixing matrix $\mathbf{V}$, called the CKM matrix [24, 25], is introduced. It appears only in the quark sector, i.e.

$$\mathscr{L}_{\rm CC} = -\frac{g}{2\sqrt{2}}W^-_\mu\left\{\sum_{ij}\overline{u}_{iL}\gamma^\mu\mathbf{V}^{ij}d_{jL} + \sum_{\ell}\overline{\nu}_{\ell L}\gamma^\mu\ell_L\right\} + \text{h.c.}\,, \quad \mathbf{V} = \mathbf{V}^u_L\mathbf{V}^{d\dagger}_L. \tag{1.52}$$

The CKM matrix connects up–type quarks to down–type quarks. In the SM, the absence of a right–handed ν_R field prevents the inclusion of a lepton mixing matrix due to the massless nature of neutrinos.

In general, a $N_G \times N_G$ unitarity matrix has N_G^2 real parameters. These can be reduced to $\frac{1}{2}N_G(N_G-1)$ angles and $\frac{1}{2}(N_G-1)(N_G-2)$ phases [25]. For $N_G = 2$, $\mathbf{V}$ is determined by a single parameter: the Cabibbo angle θ_C, which defines the Cabibbo rotation matrix [24]. For $N_G = 3$, the CKM matrix is described by three angles $(\theta_{12}, \theta_{23}, \theta_{13})$ and one complex phase (δ_{13}). The standard choice to parametrise the CKM matrix is [16, 26]

$$\begin{aligned}\mathbf{V} &= \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb}\end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13}\end{pmatrix},\end{aligned} \tag{1.53}$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$ with i, j labelling the quark generations. δ_{13} is the only complex phase in the SM, consequently the only source of $\mathcal{CP}$ violation which advocated the existence of a third generation of quarks, $Q = (t, b)$ [24]. With only two generations, $\mathcal{CP}$ violating effects observed in the kaon system [27] could not be explained. These third generation quarks were later discovered at Fermilab in 1978 (bottom quark) [28, 29] and 1995 (top quark) [30, 31].

[‡]The GIM mechanism, proposed by S. Glashow, J. Iliopoulos and L. Maiani, postulated the existence of a fourth quark to suppress FCNCs in the $K \to \mu^+\mu^-$ decay, and therefore the introduction of the second generation of quark doublets $Q = (c, s)$.

1.3.5 The SM Higgs boson

In the SM, the Higgs particle is a massive scalar boson with zero spin, positive parity, no electric charge, and no colour charge. It interacts with massive particles and is highly unstable, quickly decaying into other particles. After an extensive 40–year search, the Higgs particle was finally discovered in 2012 by the ATLAS and CMS experiments at the LHC at CERN laboratory [12, 13].

At hadron colliders, the SM Higgs boson is primarily produced through the loop–induced gluon–fusion process (ggF). This process involves the interaction of two gluons resulting in the production of a Higgs boson ($gg \to H$). Among the various contributions to this process, the top quark loop provides the leading contribution [32]. Other production processes, such as vector–boson fusion ($qq \to qqH$) and production in association with vector bosons ($q\bar{q} \to H + W/Z$) or top quarks ($gg, q\bar{q} \to t\bar{t}H$ and tHq), have much smaller production rates, suppressed by more than one order of magnitude compared to ggF production [33–38]. These processes produce distinct final–state event topologies.

Higgs bosons may decay into a pair of vector bosons or a pair of fermions. The dominant decay modes (in descending order) are: a pair of bottom quarks ($b\bar{b}$) 58%, W bosons (WW^*) 22%, τ–leptons ($\tau^+\tau^-$) 6%, Z bosons (ZZ^*) 3%, charm quarks ($c\bar{c}$) 3%, photons ($\gamma\gamma$) 0.2% and muons ($\mu^+\mu^-$) 0.02%. These decay modes can be measured experimentally via the dominant ggF production mode, as well as the other sub–leading production modes. Combinations of specific production and decay modes are conducive to different analysis sensitivities to the targetted decay, as a better background rejection can be achieved experimentally.

For example, the main decay mode of the Higgs boson into bottom quarks ($H \to b\bar{b}$) can be observed in the strongly boosted $H + W/Z$ process [39]. Strongly boosted refers to the situation where the Higgs boson has a large transverse momentum, making the SM cross–section of this production mode similar to that of ggF in this regime [40, 41]. Similarly, the Higgs boson decay into τ–leptons can be observed in the vector–boson fusion process [42], where the event topology of this production mode allows of a better rejection of the $Z \to \tau\tau$ background [43].

1.4 Quantum Chromodynamics

During the 1950's, a large number of particles called hadrons were rapidly being discovered, adding to the complexity of the physics picture at the time. Due to the amount of new hadrons and the known similarities among them in terms of their quantum numbers, the existence of more fundamental particles that would constitute these hadrons was postulated: the quarks. Hadrons can be classified into mesons and baryons, assuming that mesons are $M \equiv q\bar{q}$ states

and baryons $B \equiv qqq$ states. In 1961, both M. Gell-Mann and Y. Ne'eman were able to classify these using the eightfold way [44,45]. In short, this classification made use of the strangeness $(S)^*$ and electric charge properties of hadrons to arrange them into: an octet and a singlet for the mesons, and an octet and a decuplet for baryons. As a matter of fact, the baryon decuplet predicted the existence of the Ω^- hadron, with $S = -3$, which had not been discovered yet. Its discovery in 1964 [46] gave M. Gell-Mann the 1969 Physics Nobel Prize for his work. Finally, it was understood that the structure of these hadron groups could be explained by the existence three quark flavours (u, d, s) that constitute these hadrons.

The first hint of an additional quantum number for quarks is commonly taken to be the Δ^{++} baryon. In order to satisfy Fermi–Dirac statistics, the introduction of colour as a new quantum number associated with the symmetry group $SU(3)_C$ was realised for quarks [47, 48]. This way, each quark would have $N_C = 3$ different colours q^α, $\alpha = 1, 2, 3$ (red, green, blue). Due to the non–observation of colourful free states, quarks must be confined within colour–singlet bound states. This is known as colour confinement and is one of the main results of QCD. The other being asymptotic freedom which describes the reduction of the interaction force among quarks and gluons as the energy scale increases (shorter distances) [49, 50].

The construction of the QCD Lagrangian is very similar to that of the EW Lagrangian in Section 1.2.3. Considering the quark field q_f^α where f denotes the quark flavour and α its colour, the free Lagrangian is

$$\mathscr{L}_0 = \sum_f \overline{q}_f (i\gamma^\mu \partial_\mu - m_f) q_f. \tag{1.54}$$

These fields are invariant under the global $SU(3)_C$ transformations

$$q_f^\alpha \xrightarrow{SU(3)} (q_f^\alpha)' = U_\beta^\alpha q_f^\beta, \qquad UU^\dagger = U^\dagger U = 1, \qquad \det U = 1, \tag{1.55}$$

where

$$U = \exp\left\{ i \frac{\lambda^a}{2} \theta_a \right\}, \qquad a = 1, 2, \ldots, 8. \tag{1.56}$$

Here, $\frac{\lambda^a}{2}$ are the eight generators of $SU(3)_C$ and the λ^a matrices satisfy the commutation relation

$$\left[\frac{\lambda^a}{2}, \frac{\lambda^b}{2} \right] = i f_{abc} \frac{\lambda^c}{2}, \tag{1.57}$$

where f_{abc} are the $SU(3)_C$ structure constants, which are real and totally anti-symmetric. Analogously to the construction of the QED Lagrangian, invariance

*The term strangeness predates the discovery of the s quark, although it is now defined as $S = -(n_s - n_{\overline{s}})$, where n_s and $n_{\overline{s}}$ indicate the number of strange quarks and antiquarks in a hadron, respectively.

under local $SU(3)_C$ transformations is required, which leads to the definition of a covariant derivative,

$$D^\mu \equiv \partial^\mu + ig_S \frac{\lambda^a}{2} G_a^\mu(x) \equiv \partial^\mu + ig_S G^\mu(x), \tag{1.58}$$

where $G_a^\mu(x)$ fields are the eight gauge bosons (gluons) that transform as

$$G_a^\mu(x) \xrightarrow{SU(3)} (G_a^\mu)'(x) = G_a^\mu(x) - \frac{1}{g_S}\partial^\mu \theta_a(x) - f^{abc}\theta_b(x) G_c^\mu(x). \tag{1.59}$$

This last term was not present in the QED Lagrangian since it is an Abelian theory; it eventually leads to gluon self–interactions. Adding the covariant derivative into the Lagrangian, along with a kinetic term for the gluons, gives the QCD Lagrangian

$$\mathscr{L}_{\text{QCD}} \equiv \sum_f \overline{q}_f (i\gamma^\mu D_\mu - m_f) q_f - \frac{1}{4} G_a^{\mu\nu} G^a_{\mu\nu}, \tag{1.60}$$

where $G_a^{\mu\nu} = \partial^\mu G_a^\nu - \partial^\nu G_a^\mu - g_S f^{abc} G_b^\mu G_c^\nu$. Some final remarks, $\mathscr{L}_{\text{QCD}}$ exhibits interactions between quarks and gluons as well as cubic and quartic gluon self–interactions with the same strong coupling strength $\alpha_s(Q) = g_S^2(Q)/4\pi$ which depends on (runs with) the energy scale Q. In addition, quarks can emit gluons. The first experimental results consist of the observation of a 3–jet topology interpreted as three gluons or a $q\overline{q}g$ final state (gluon bremsstrahlung) at the PLUTO, TASSO, MARK–J experiments at DESY and SLAC [51, 52]. Finally, a mass term for gluons is forbidden by gauge symmetry, thus remaining as massless spin–1 gauge bosons.

1.5 Limitations of the Standard Model

Despite being the most successful theory in describing fundamental particles and how they interact with an impressive list of high–precision experimental results that exhibit perfect agreement with the predictions, the SM is inherently an incomplete theory: there are some physical phenomena and experimental results that it is not able to explain. These require new BSM physics that introduce subtle modifications to the SM to accommodate those effects and explain the observations:

- **Gravity:** The SM does not include any interactions with gravity. Including a new massless boson with $s = 2$, called the *graviton*, which would mediate interactions between SM particles and gravity, leads to a non–renormalisable theory at high energy scales. Furthermore, the interplay between the SM and general relativity presents itself as another difficult theoretical challenge.

- **Dark matter and dark energy:** Astronomical and cosmological observations to date predict that the SM particle content, what is referred to as ordinary matter, constitutes about a 5% of the total universe matter. To complete the picture, a new kind of weakly–interacting matter, called *dark matter* (DM), should constitute about a 25% of the universe, and *dark energy* should account for the remaining 70%. The SM does not provide a suitable candidate for dark matter which mainly interacts gravitationally and does not emit or absorb light. Dark energy, a constant energy density for the vacuum, that could model the accelerated expansion of the universe, remains as an even bigger mystery.

- **Neutrino masses:** The SM predicts massless neutrinos. However, evidence of neutrino oscillations [53] contradicts this prediction. As they propagate, neutrinos oscillate between their three different mass eigenstates which are linear combinations of their known three flavour eigenstates, i.e. those in Figure 1.1. The Seesaw mechanisms are able to accommodate neutrino masses. These mechanisms depend on whether neutrinos are Majorana or Dirac particles* [54–60]. Massive neutrinos would imply the existence right–handed neutrinos which have not been observed. In addition, current limits on their masses are extraordinarily small which challenges the fact that the same mechanism would give mass to all SM particles.

- **Matter–antimatter asymmetry:** The conservation laws in the SM would predict an equal amount of matter and antimatter. Yet, no anti–matter is naturally present in the universe. This imbalance can be explained by baryon number† non–conservation, and by both $\mathcal{C}$ and $\mathcal{CP}$ violation [61]. The only source of $\mathcal{CP}$ violation in the SM comes from the δ_{13} complex phase in the CKM matrix which is not sufficient to explain this asymmetry. As a result, additional sources of $\mathcal{CP}$ violation are searched for in BSM extensions.

Currently, there is not a single SM extension that is able to incorporate and explain all observations at all the energy scales that are being considered here. While this theory might be very far away from being realised, one of the widely–used approaches is to consider EFTs to parametrise unknown physics at unaccessible energies by current experiments. It is also paramount to develop new powerful collider machines capable of measuring these high–energy

*If the neutrino and antineutrino are considered as different particles, they are called Dirac particles, while if they are the same, they are called Majorana particles.

†The baryon number is a quantum number of a system. It is defined as $B = \frac{1}{3}(n_q - n_{\bar{q}})$, with n_q and $n_{\bar{q}}$ being the number of quarks and antiquarks, respectively.

phenomena to discriminate between the different BSM models and constrain the free parameters in them.

1.6 Effective Field Theories

All physical theories are effective theories (when properly formulated). In fact, the SM could be seen as an EFT of a higher energy scale theory such as that of the Planck scale ($\Lambda_{\text{Planck}} \sim 10^{19}$ GeV), where quantum mechanics meets gravity. This vision of the description of all physical systems brings forth several observations that motivate the use of EFTs to make calculations on them, even if the exact theoretical framework is not known.

Over the last centuries, several discoveries have been made over an impressive wide range of scales. From experiments that probe the fundamental particles of nature at extraordinarily small distances (10^4 GeV) to astronomical scales as large as the size of our observable universe (10^{-24} GeV). The fact that precise models with accurate predictions have been developed at all these different length, energy and time scales, is a clear indication that Nature decouples. That is, physical phenomena that occur at higher (lower) energy scales than the energy E at which calculations are being made can have their scales set to infinity (zero) due to this decoupling. In this limit, their effects can be neglected or reintroduced later, if needed, in perturbation theory. Observations at this energy E always have finite precision. Given that precision, only certain degrees of freedom, symmetries and dynamics have to be considered.

On the contrary, this argument can be taken in detriment of the SM: it advocates that, although it provides extremely precise predictions at the working energy scales, important effects at higher or lower scales are being neglected and possibly not being understood, making it an incomplete theory. However, the limitations of the SM are well known and some have already been tackled, at their working energy scales, as EFTs of the SM. Such is the case of Chiral Perturbation Theory [62], the Heavy Quark Effective Theory [63] or the Standard Model Effective Field Theory (SMEFT) [64–68].

Using EFTs, allows us to simplify complex multi–scale calculations into a series of simpler single–scale calculations. They are also necessary if the dynamics of the exact theory are unknown or incomputable. An EFT is defined by an effective Lagrangian, which in turn is completely specified by the following three ingredients:

- Degrees of freedom: to analyze a particular physical system, it is important to identify and isolate those elements that are most relevant. The choice of these variables is crucial since they have to capture the physics which is most important for the problem at hand.

- Symmetries: the selected symmetries determine the dynamics of the system. Moreover, all terms compatible with the chosen symmetries should have to be considered in the Lagrangian which could add to the complexity of the EFT.

- Expansion parameters: due to the possible infinite number of terms in the Lagrangian, EFTs can be realised by featuring one or more expansion parameters. Typically, these parameters are given as the ratio of the working energy scale E over the high energy scale Λ at which further effects are neglected (E/Λ). For this to work, a power counting scheme is adopted, where for each term in the Lagrangian a definite order in the expansion parameter is assigned. This ensures that calculations can be done at the desired level of accuracy and that, at a specified order in the expansion, the Lagrangian has a finite number of terms. This allows for an order–by–order renormalisation.

A brilliant example of how an EFT can jump–start the understanding of a particular theory is Fermi's Theory of weak interactions [69]. In the low–energy limit, the muon decay $\mu^- \to e^- \bar{\nu}_e \nu_\mu$ can be approximated by a 4–fermion interaction. In this limit, the W boson field can be left out of the theory since there is not enough energy to produce a physical W boson, i.e. $E \ll M_W$. This effective contact interaction was able to perfectly describe the muon decay without the need – at the time – of unknown, and experimentally inaccessible W propagator fields. E. Fermi was able to resolve and explain the β decay by introducing this contact interaction, proportional to the effective Fermi coupling strength constant G_F, which has now been measured to great precision: $G_F = 1.166\ 378\ 7(6) \times 10^{-5}\ \text{GeV}^{-2}$ [16]. Of course, once the full (higher–energy) EW theory was developed, the process was better understood through the mediation of a W boson. This proves the point that, in absence of a better understanding of the complete theory, the proper use of an EFT can not only provide accurate predictions, but also motivate searches and advances in the current theory.

1.6.1 SMEFT

Nowadays, one of the most popular EFT frameworks is the SMEFT, an EFT constructed out of SM fields. It is used to analyse deviations from the SM and searches of BSM physics. It is defined as a *bottom–up* EFT, where the basis of terms in the Lagrangian are built without making any connection to a ultraviolet complete theory. However, this comes at the expense of having a large number of possible terms that are compatible with the chosen symmetries, which are those from the SM.

In the SMEFT, all new physics is assumed to be above the high energy scale Λ, where $\Lambda = 10^3$ GeV typically. New physics is also assumed to follow the

SM symmetries and therefore be Lorentz and gauge invariant. Then, all the possible compatible terms are built into the effective Lagrangian at each order in the mass dimension:

$$\mathscr{L}_{\text{SMEFT}} = \mathscr{L}_{\text{SM}} + \sum_{\forall i, \mathcal{D}\geq 5} \frac{C_i O_i^{(\mathcal{D})}}{\Lambda^{\mathcal{D}-4}}, \tag{1.61}$$

where $O_i^{(\mathcal{D})}$ are the dimension–$\mathcal{D}$ operators constructed with the SM fields, and C_i are the coupling strength constants where the information on any heavy degrees of freedom is hidden away (also known as Wilson coefficients). The new physics scale Λ appears in the expansion parameter $(1/\Lambda^{\mathcal{D}-4})$ to suppress the higher dimension operators and to make the Wilson coefficients dimensionless. Additionally, it ensures that all terms in the expansion are of dimension–4 like the SM Lagrangian.

Following this prescription, only one dimension–5 operator can be built following the SM symmetries. This operator is related to the neutrino Majorana mass and violates lepton number. It also generates a mass term for left–handed neutrinos like $m_\nu \simeq (c/\Lambda)v^2$. Given the current neutrino mass bounds, the required energy scale is very large: taking $c \sim 1$ and $m_\nu \sim 1$ eV implies that $\Lambda \simeq (v^2/m_\nu) \simeq 10^{13}$ GeV. Dimension–7 operators also violate lepton number conservation and are comparatively suppressed by other orders in the expansion.

There are a total of 59 independent dimension–6 operators, assuming baryon number conservation [70]. To set experimental bounds on them, each Wilson coefficient ideally needs a dedicated observable. This is usually done in the frame of a global fit in the coefficient parameter space. Deviations of the coefficients from zero in the fit are indications of new physics in the corresponding effective interaction vertex. To assess the effect of each operator on a given observable, X, the SMEFT predictions can be written as an expansion in orders of $1/\Lambda^2$:

$$X = X_{\text{SM}} + \frac{1}{\Lambda^2}\sum_i C_i X_i^{(1)} + \frac{1}{\Lambda^4}\sum_{ij} C_i C_j X_{ij}^{(2)} + \mathcal{O}(\Lambda^{-4}). \tag{1.62}$$

This expansion contains the SM part, as well as the interferences of the effective dimension–6 operators with the SM (referred to as *linear* terms), which are proportional to C_i/Λ^2, and the *quadratic* terms proportional to C_iC_j/Λ^4. The effects of dimension–8 operators and the so–called *double insertions** terms of dimension–6 operators that contribute at Λ^{-4} order are not included.

*Double insertions represent interference between SM diagrams with diagrams with two effective couplings. These contribute as Λ^{-4} terms in the total amplitude.

1.7 The top quark as a probe for precision and BSM physics

The top quark is the most massive fundamental particle predicted by the SM. It is a spin–1/2, up–type quark (charge $+2/3e$) in the third generation of elementary particles. Its discovery was announced in 1995 by the CDF and D0 collaborations at the Tevatron collider [30, 31]. The combination of its high mass, short lifetime, and the need of high–energy particle accelerators for its production, contributed to why the top quark was the last quark to be discovered. This event not only further validated the SM theoretical framework, but also opened up an entirely new avenue for top quark–related measurements and analyses that are now a fundamental piece of any modern–day scientific programme.

At the LHC, production of top quarks happens through two main mechanisms: the production of a top and antitop quark pair ($t\bar{t}$ production) via strong interactions, and the production of a single–top quark associated to other particles via EW interactions. Moreover, the top quark decays almost exclusively into a W boson and a bottom quark (tWb vertex), which grants it a unique signature in the final state. Over the years, the LHC has been able to produce a large number of top quarks, and many analysis have been performed to obtain the most precise measurements of its properties. These include, but are not limited to, the top quark mass, the top quark width and branching ratios, its polarisation, and spin correlations and production asymmetries (in the case of $t\bar{t}$ production). These precision measurements have served as an excellent way to consolidate the SM predictions [16]. Further details on the production mechanisms, as well as on the top quark's properties and decay properties, are given in Chapter 3.

The work in this thesis focuses on the properties of the top quark, as well as its couplings with other SM particles, and also in the context of BSM physics. Due to its high mass, it is uniquely sensitive to new physics at higher energy scales. By studying its (effective) couplings with other SM particles, insight into the structure of the SM and its possible extensions can be gained. Its couplings to the EW sector are of particular interest and can be examined with a variety of experimental observables coming not only from the LHC, but also from other legacy measurements from the Tevatron and LEP/SLC machines.

With all, the LHC is the world's largest top quark "factory", allowing the study its coupling properties with unprecedented precisions. A complete understanding of the machine's operation, collision phenomenology, as well as of the identification and reconstruction of the collision products is essential to accurately study top quark physics. In turn, this provides a remarkable opportunity to push both the experimental and theoretical frontiers of particle physics.

2 CERN, LHC and the ATLAS experiment

With the objective of pushing the SM and other BSM theories to their limits, it is crucial to develop and construct powerful machines that can target the relevant physics and do so with the highest precision possible. To this extent, the LHC, located at the CERN laboratory, is the largest and most powerful collider experiment built so far. The energy potential that has been achieved in its collisions allows for the production and further study of the most massive particles in the SM, and could unveil evidences about BSM heavier particles. The top quark and other massive particles, such as bosons, are produced at the LHC at a formidable rate making it the best scenario to study them with unparalleled accuracy.

In collider experiments, two beams are accelerated at ultra–relativistic energies to collide at specific points. At the LHC ring, there are four of such points where state–of–the–art detectors are located to identify the remnant particles after the collision. The ATLAS detector [71, 72] is the largest one of these four detectors, and also one of the two multi–purpose experiments located in the LHC ring. The LHC has been providing pp collision data for over a decade at different running energies and this has led to very precise measurements in particle physics, including the discovery of the Higgs boson in 2012.

In this chapter, a description of the experimental setup is given. Section 2.1 covers the LHC in detail and the other facilities at CERN that work along with it to accelerate particles to (almost) light speeds. Details on pp collision phenomenology are also given in Section 2.2. Section 2.3 presents the ATLAS experiment, including a description of its subsystems and the techniques used to identify and measure the properties of the particles produced in collisions. Furthermore, Section 2.4 details the ATLAS trigger system (an impressive piece of *online* event rejection technology), and Section 2.5 describes the object reconstruction techniques.

2.1 The Large Hadron Collider

The LHC [73, 74] is located at the CERN laboratory, across the border between France and Switzerland, and is the world's largest and highest–energy particle collider, allowing us to probe the frontiers of knowledge. The accelerator is placed in a 27–kilometer circular ring, at about 100 meters underground, which provides a natural shielding against background radiation from different sources. Construction of the underground cavities was realised between 1984 and 1989 for the Large Electron–Positron Collider (LEP), followed by the construction and installation of the LHC machine and the detectors themselves.

The collider tube consists of two adjacent beam pipes where mainly protons*

*Heavy ion collisions also recorded, mostly ionised lead nuclei, but also xenon.

are accelerated in opposite directions. These beam pipes are ultra–high vacuum environments that avoid collisions with gas particles in the beam. Particles are accelerated using electric fields in radio-frequency cavities. In addition, the beams are bent and kept focused inside the ring with powerful superconducting dipole magnets and quadrupole magnets, respectively. These superconducting magnets require operating temperatures of about 1.9 K, making the LHC the largest cryogenic facility in the world, and the CERN laboratory one of the focal points of ultra–modern technology development.

2.1.1 The journey of a proton

Prior to being injected into the main accelerator, protons are accelerated through a set of machines in the CERN accelerator complex that progressively increase their energy. These machines boost the beam energy and inject it into one another until they reach the LHC, where each of the two beams carry half of the desired centre–of–mass (COM) energy. This thesis uses the dataset recorded during the 2015–2018 period (Run 2) which accelerated the proton beams up to 6.5 TeV. An energy of 6.8 TeV was recently achieved, in July 5^{th} 2022, which marked the start of the ongoing third period of data–taking (Run 3).

Since 2020, the LINear ACcelerator 4 (Linac 4) is the first step in the acceleration process[†], boosting ionised hydrogen atoms to 160 MeV. These are then injected into the first circular booster of the chain: the Proton Synchrotron Booster. In the injection process, the ions are stripped of their electrons, leaving only protons. Protons are bunched together with a spacing between bunches of 25 ns. These are accelerated to 2 GeV and injected into the Proton Synchrotron, which pushes the beam energy up to 26 GeV. Protons are then injected into the Super Proton Synchrotron, which boosts them to 450 GeV and prepares them for final injection into the LHC. Once in the LHC ring, protons take approximately 20 minutes to ramp–up to their maximum collision energy and are brought to collide at the four main interaction points. Figure 2.1 shows the whole CERN accelerator complex, including other machines that provide collisions to other experiments outside the LHC ring.

The LHC program has spanned over a decade now and is planned to record data for the next 20 years approximately. Until now, pp collision data has been collected at different COM energies. The first data–taking period, from 2010 to 2013 (Run 1), collided protons at $\sqrt{s} = 7$ and 8 TeV, while Run 2 did so at $\sqrt{s} = 13$ TeV. The third data–taking period, which collides protons at $\sqrt{s} = 13.6$ TeV, started in the summer of 2022 and is expected to last at least four years. After each data–taking period, the LHC remains shutdown for a long period – called Long Shutdown – in order to do repairs and improve-

[†]Linac 2 has provided protons to CERN's accelerator complex since 1978, and was retired and replaced by Linac 4 in 2020. Therefore, it was still in operation during Run 2.

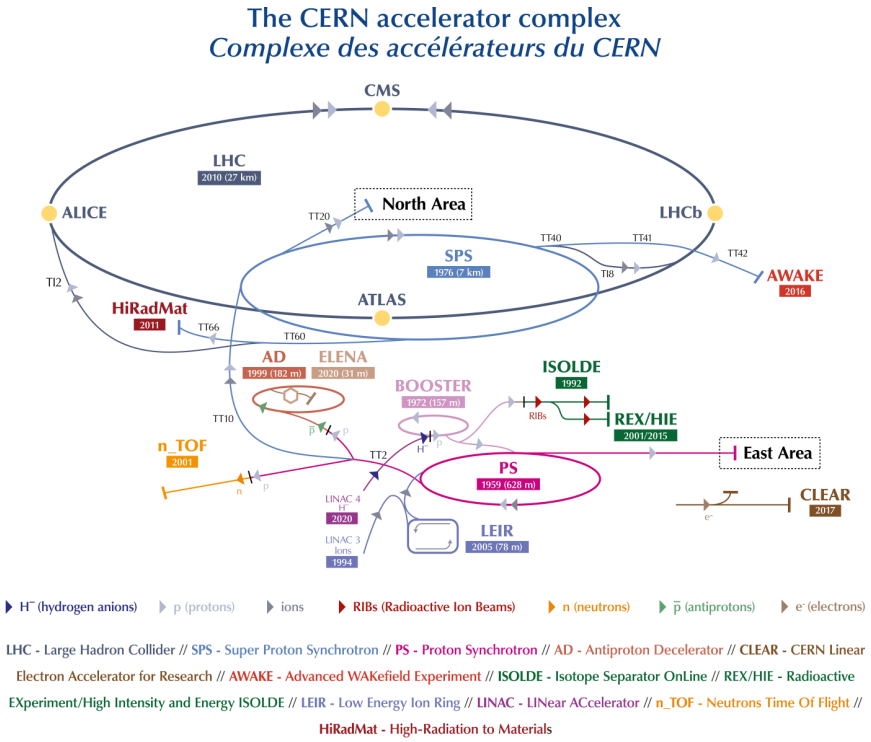


Figure 2.1: The CERN accelerating complex. Source: Ref. [75].

ments in both the accelerator and the detectors. The next long shutdown after Run 3 (estimated to be around 2026) is of great importance since the LHC will undergo a significant upgrade to increase the amount of recorded data. This next upgrade is called the High–Luminosity LHC (HL–LHC) [76]. Luminosity is a measure of the amount of collisions per unit of area and time and is one of the most important parameters in collider physics. It is further detailed in Section 2.2.4. Around 157 fb^{-1} of data were delivered by the LHC [77] during Run 2 and another 3000 fb^{-1} are expected during the HL–LHC.

As aforementioned, other heavy ions are also collided at the LHC over shorter running periods. Pb–Pb collisions have been recorded at $\sqrt{s} = 2.76$, 5.02 TeV and p–Pb at $\sqrt{s} = 5.02$ TeV. On the other hand, Xe–Xe collisions have been recorded at $\sqrt{s} = 5.44$ TeV.

2.1.2 The LHC detectors

The LHC ring is divided in eight arc sections connected by eight straight *insertion* sections. These insertions host not only the main four detectors where the

beams collide, but also the beam dump, insertion and collimation (cleaning) systems, as well as the radio–frequency cavities. At the main four interaction points reside the particle detectors. One of them, the ATLAS detector, is detailed in Section 2.3 as the results in this thesis use the data collected by it during Run 2. The other main detectors that contribute to the rich physics program that the LHC offers are detailed below.

- **CMS:** With ATLAS, the CMS (Compact Muon Solenoid) [78] experiment is the other general–purpose detector at the LHC. For this reason, the physics studied with it is similar to that of ATLAS, ranging from high–precision SM measurements to searches of new physics. From an experimental point of view, having two different experiments targeting the same physics is a source of cross–validation of the measurements and discoveries at the LHC.

- **LHCb:** Arguably the most peculiar detector, the LHCb (Large Hadron Collider beauty) [79] experiment covers only a small portion of the whole solid angle. The detector is placed in the forward region, close to the beam line, allowing it to specialise in b–physics owing to the fact that b–hadrons a primarily produced in this region. Accurate measurements of $\mathcal{CP}$ violation parameters are being performed here, contributing to the understanding of the matter–antimatter asymmetry.

- **ALICE**: The ALICE (A Large Ion Collider Experiment) [80] experiment focuses mainly on heavy–ion physics. This allows for the study of strongly interacting matter at extreme energy densities, what is known as the quark–gluon plasma. This state of matter was expected to be abundant at the early universe; its study under laboratory conditions allows the deciphering of questions about strong interactions and how matter is organised.

Apart from the main four experiments, there are other smaller experiments at the LHC ring. These are: TOTEM (TOTal Elastic and diffractive cross section Measurement) [81] which shares the CMS interaction point and provides complementary measurements of pp cross–sections, elastic scattering, and diffraction processes; MoEDAL (Monopole and Exotics Detector at the LHC) [82] which searches for magnetic monopoles and highly ionising stable massive particles; LHCf (Large Hadron Collider forward) [83] which shares the ATLAS interaction point and provides complementary measurements in the forward region of collisions to help explain the origin of ultra-high-energy cosmic rays; FASER (ForwArd Search ExpeRiment) [84] which again shares the ATLAS cavern and is designed to search for light and extremely weakly interacting particles; and finally, the SND (Scattering and Neutrino Detector) [85] detector which focuses in neutrino physics measurements from LHC collisions.

2.1.3 The LHC computing Grid

All the events recorded by all the experiments are stored at the Worldwide LHC Computing Grid (WLCG) [86] for them to be analysed. The need for such a large Grid–based computer network is fundamental to handle the massive amounts of data produced by experiments (this is estimated to be 15 petabytes per year). This infrastructure connects over 150 computing centres in more than 40 different countries around the world. The Grid is composed of several *tiers*. Tier–0 is located at CERN, and contains the first copies of the collected data. It also performs the first pass reconstruction and distribution to Tier–1 centres. These centres share the total amount of data, perform a large–scale reprocessing and distribute them to the next tier. One of them is located at PIC, in Barcelona, and forms part of the WLCG Spanish cloud [87]. Tier–2 sites (about 160) are located at universities or scientific institutes and usually have enough storage space to cover the specific undergoing analyses. Within the spanish territory, Tier–2 facilities are located at IFIC (Valencia), IFAE (Barcelona) and UAM (Madrid). Finally, Tier–3 refers to local computing clusters that individual scientist can access for their analyses.

2.2 Phenomenology of proton–proton collisions

The LHC program is mainly focused on the scrutiny of pp collisions. It is therefore paramount to give an insight on the properties of the proton and the phenomenology of these high–energy collisions. In 1969, an experiment at SLAC reported the first *inelastic* scattering of electrons and protons (Hydrogen nuclei) [88]. In this context, the inelastic collision observations led to the conclusion of protons having an internal structure, thus supporting the proposed quark model (and the SM) at that time.

Now, it is well established that protons are bound states of quarks and gluons (also known as partons). They are formed by *valence* quarks (uud) that communicate via the strong interaction (gluon exchange). Moreover, these partons inside the proton continuously emit virtual gluons that split into quark–antiquark pairs (called *sea* quarks). At the LHC, collisions are regarded to happen via deep inelastic scattering processes. Here, *deep* emphasises the high–energy scales that are being used. These allow the probing of small distances inside the proton, resolving more of its radiative structure, and effectively resulting in scattering collisions between the partons inside both proton beams (known as the *hard* process).

2.2.1 The parton distribution functions

Quarks, antiquarks and gluons share a fraction of the total proton momentum which can be probed as a function of the energy scale. For instance, at low

energies ($Q^2 \sim 1$ GeV), the three valence quarks become more dominant in the proton, while at higher energies the momentum fraction associated with gluons and antiquarks increases. The dynamics of the proton and the parton momentum fraction probability distributions, at a given Q^2, are encoded in what is known as parton density functions (PDFs).

The PDFs cannot be analytically determined due to the complexity of non–perturbative QCD. They are extracted from global fits to data from several experiments and extrapolated to new energy scales by solving the DGLAP perturbative QCD (pQCD) equations [89]. Three of the active groups that provide these PDFs for LHC physics are the CTEQ [90], the MSTW [91] and the NNPDF [92] collaborations. Figure 2.2 shows an example of the PDFs of quarks, antiquarks and gluons for two different energy scales $\mu^2 \equiv Q^2$ coming from a global fit from the NNPDF Collaboration. At higher energies, the gluon fraction increases which can impact the kinematics of some processes at the LHC.

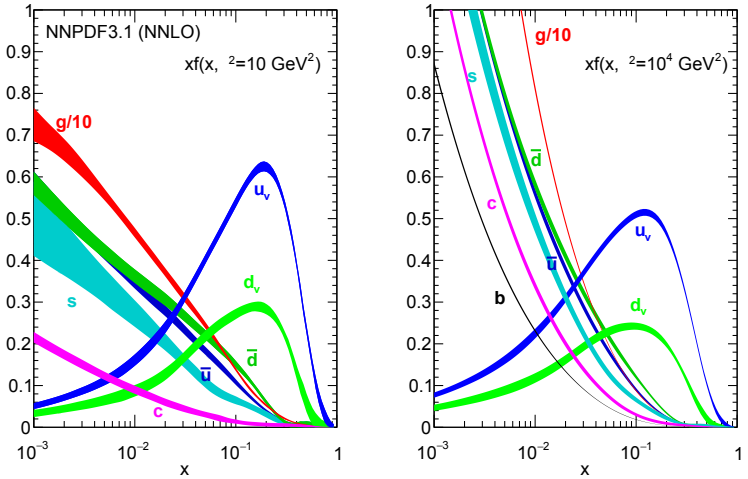


Figure 2.2: Parton distribution functions of a proton at $\mu^2 = 10$ GeV^2 (left) and $\mu^2 = 10^4$ GeV^2 (right) for the NNPDF3.1 set. Uncertainty bands cover the 68% confidence level. Source: Ref. [93] (Fig. 2).

2.2.2 High–energy *pp* collision formalism

There are two types of scattering processes in *pp* collisions: the *soft* and the aforementioned *hard* processes (low– and high–energy processes, respectively). While QCD is the underlying theory for both, the approach to resolve them is quite different. As commented in Section 1.4, two of the main results coming from QCD are asymptotic freedom and colour confinement. Asymptotic

freedom is conducive to the treatment of quarks and gluons as free particles at high–energy (short distance) scales, allowing for the use of pQCD calculations. These typically describe partonic collision cross–sections which are process–dependent. On the other hand, low–energy (long distance) interactions between the partons exhibit low momentum transfer and are dominated by non–perturbative QCD effects, which are not completely understood*. They are encapsulated in the PDFs of the protons, which are universal: they are determined from a process (or group or processes) and can be used for others.

Owing to the decoupling nature of physics at different scales, it is possible to separate the short and long distance physics present in pp collisions. The hard interaction cross–section can be described using the QCD *factorisation* theorem: the inclusive cross–section for a $pp \to X$ process is then given by the convolution of the PDFs $f_{a,b}$ and the partonic cross–section $\hat{\sigma}_{ab\to X}$ as

$$\sigma_{pp\to X} = \sum_{a,b} \int dx_a dx_b f_a(x_a, \mu_F^2) f_b(x_b, \mu_F^2) \hat{\sigma}_{ab\to X}(x_a, x_b, \mu_R^2, \mu_F^2), \qquad (2.1)$$

where a and b are the incoming partons and $x_{a,b}$ are the momentum fractions of the partons inside the corresponding protons. A depiction of the factorisation theorem in hadronic pp collisions is shown in Figure 2.3. The factorisation scale μ_F defines the boundary between low– and high–energy physics and hence the energy scale at which pQCD calculations are valid from. The partonic cross–section can be expanded as a power series of α_s in pQCD as

$$\begin{aligned}\hat{\sigma}_{ab\to X} = \big[\hat{\sigma}_{\mathrm{LO}}(x_a, x_b, \mu_F^2) + \alpha_s(\mu_R^2)\hat{\sigma}_{\mathrm{NLO}}(x_a, x_b, \mu_F^2) \\ + \alpha_s^2(\mu_R^2)\hat{\sigma}_{\mathrm{NNLO}}(x_a, x_b, \mu_F^2) + \dots\big]_{ab\to X},\end{aligned} \qquad (2.2)$$

where α_s depends on the renormalisation scale μ_R†. Upon considering the infinite terms in the expansion series, the physical cross–section is independent from the μ_R and μ_F scales. However, this is not feasible and the truncation of the series leads to a dependence of the total cross–section with both scales. This dependence decreases as higher order corrections are included in the calculations. Both scales are usually chosen to be equal and of the order of the momentum scale of the hard process. Moreover, the impact of higher–order terms beyond the truncation of the series can be estimated by varying these scales (typically by a factor two and a half of their nominal values) and assigning this variation as an uncertainty on the calculation.

Apart for the hard process, there are other many soft sub–processes that accompany the hard interaction and that are essential to obtain an accurate modelling of the collisions. These soft interactions are collected in what is known as

*These soft interactions are what is called the underlying event (UE).

†In renormalisation schemes, an energy scale μ_R is introduced to help vanish large logarithms emerging from loop corrections to the total cross–section.

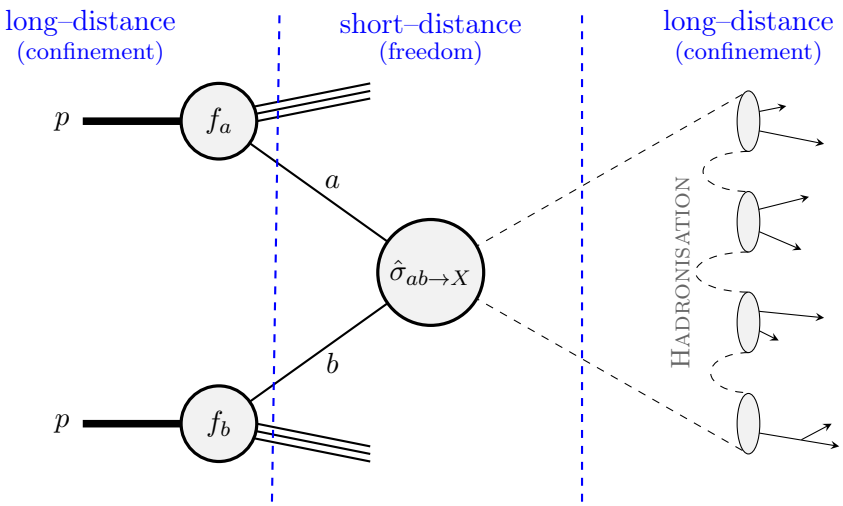


Figure 2.3: Schematic representation of a hard process in pp collisions.

the underlying event. They include secondary parton collisions from other beam remnants after the hard scattering and initial and final state radiation (ISR and FSR). These soft interactions, mainly described by non–perturbative QCD, are modelled by parton shower (PS) Monte Carlo generators such as PYTHIA [94] and HERWIG [95]. These models come with a set of *tune* parameters that are determined from fits to LHC data and are sometimes modified depending on the target physics. Parton showers also take care of the final–state *hadronisation* of quarks and gluons and their corresponding decays into other particles. This is depicted by the grey "blobs", the hadrons, on the rightmost part of Figure 2.3.

2.2.3 Monte Carlo generators

Monte Carlo event generators have the demanding task of simulating physical events that occur in the collisions. They are an indispensable tool in any particle physics analysis to estimate signal and background event rates in both SM and BSM scenarios. In order to model the complete hard interaction, MC generators are usually split into matrix element (ME) calculations followed by the description of the regions dominated by the soft and collinear parton emissions, which is undertaken by the above–mentioned parton shower programs. ME calculations can provide accurate descriptions of the parton scattering process in pQCD, including higher–order corrections in α_s and interference terms with the external radiated quarks and gluons. However, they cannot resolve parton emissions in the soft regime.

For this reason, the description of the hard process should combine both types of calculations. Interfacing ME calculations with PS MC generators is

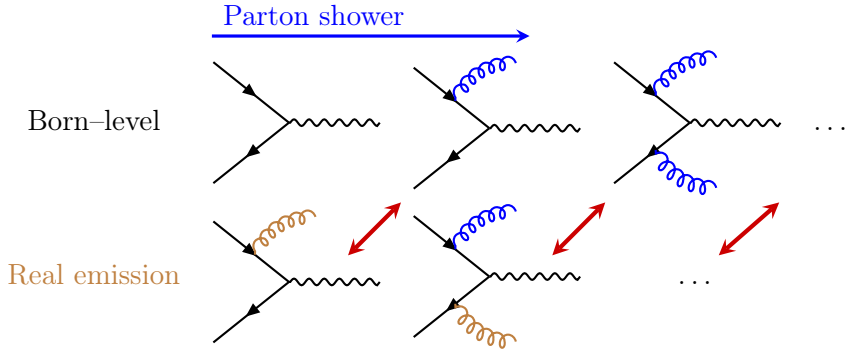


Figure 2.4: Contributions of real (brown) and soft (blue) gluon emissions to the Born–level amplitude. The diagrams connected by the red arrows would lead to double–counting and have to be appropriately removed by the matching/merging techniques.

done following a universal format imprinted in the "Les Houches Accord" [96]. Special care has to be taken when combining both techniques as both calculations may overlap in certain kinematic regimes leading to double–counting. Figure 2.4 shows an example, at LO, of how interfacing both calculations can lead to this overlap. *Matching* techniques take advantage of the use of the best description of the calculation in a given regime in order to avoid it. Similarly, *merging* techniques take care of the double–counting when combining Born–level calculations with a different number of radiated parton multiplicities. An important parameter in matching and merging algorithms is the matching/merging scale μ_Q. Additional partons are tested to determine whether their energy is above μ_Q. In this case, the simulation is done using a ME with an additional parton emission. On the contrary, the additional parton is left to be simulated by the PS.

The most commonly used MC generators are listed below:

- MadGraph5_aMC@NLO [97,98] (shortened to aMC@NLO from now on) began as a LO ME generator that feeded the simulated partonic events to a PS program such as Pythia or Herwig. This is known as the ME+PS method[‡]. In lue of the improvements on NLO matching in the MC@NLO [99] program, together with the aMC@NLO [98] development, it is now capable of performing NLO+PS[§] calculations for an extensive amount of SM processes at the LHC. It is also one of the most

[‡]The ME+PS method includes the LO diagram contributions as well as the emission of a certain number of additional partons in the MEs. It makes use of an analogous matching scale to μ_Q to separate the phase–space.

[§]The NLO+PS method promotes the accuracy of the ME+PS one by including the full NLO in QCD calculation, counting the real emission of an additional parton at Born–level accuracy and all virtual loop corrections to the LO process.

commonly used MC generators for BSM physics simulations, and in particular for EFT extensions of the SM, with a wide range of different UFO models available for it.

- POWHEG [100, 101] was developed as an alternative NLO matching technique, implemented in the POWHEG–BOX [102] framework. Therefore, it is also capable of NLO+PS calculations when interfaced with a PS like PYTHIA or HERWIG. The matching in POWHEG is controlled by the h_{damp} parameter, which is analogous to the μ_Q scale defined above. One peculiarity about POWHEG is that it provides events with (almost) only positive weights[¶], in contrast to other NLO MC generators that provide a mix of both positive and negative weights.

- SHERPA [103] is regarded as a general purpose MC (GPMC) generator since it incorporates both the ME generation [104] and its own PS model based on the Catani-Seymour dipole factorisation scheme [105]. It is capable of NLO+PS calculations where the virtual QCD corrections for the MEs at NLO accuracy are provided by the OPENLOOPS 2 libraries [106–109].

- PYTHIA [94] is often used to simulate the PS on events coming from the ME generators. Nonetheless, it is considered as a GPMC because it can also handle simple $1 \to 2$ and $2 \to 1$ processes at LO accuracy (including vector boson and Higgs boson production and decay, among others). These calculations not only include the soft emissions which are built into all PS generators, but also an additional hard parton emission thanks to the ME corrections (MECs) developed in Refs. [110–112]. Meaning that the same accuracy is achieved by MECs and ME+PS if only one additional parton is considered, with the advantage of MECs not having any matching scale parameter.

- HERWIG [95] is an alternative PS to PYTHIA since it shares the same functionality described above. The current version series, HERWIG 7 [113–115], is also capable of NLO+PS calculations using its own variants of the MC@NLO and POWHEG matching for several processes.

In events with large final–state multiplicities at the LHC, in particular with several partonic emissions (also known as *multi–leg* setups), it is crucial to develop high–precision simulations to model the data. While a complete NNLO+PS calculation is desirable, it is still far away – from the theoretical point of view – for many processes. An intermediate step is to improve the NLO

[¶]This is defined in its acronym: POsitive Weight Hardest Emission Generator.

generators to provide these multi–leg setups at NLO accuracy. This is accomplished by multi–leg merging algorithms such as FxFx [116] (aMC@NLO), MePs@NLO [117] (Sherpa), UNLOPS [118] or MiNLO [119].

2.2.4 Luminosity

On the subject of high–energy collisions, the *instantaneous* luminosity $\mathcal{L}$ is a vital parameter to measure the rate of the interactions (events) at the LHC. It is defined as

$$\mathcal{L} = \frac{1}{\sigma}\frac{dN}{dt}, \tag{2.3}$$

where dN/dt is the number of events per unit of time and σ is the cross–section of a particular process. The instantaneous luminosity can be expressed as a function of the beam parameters, assuming that the proton bunches are Gaussian–distributed, as

$$\mathcal{L} = f\frac{N_1 N_2 N_b}{4\pi\sigma_x\sigma_y}, \tag{2.4}$$

where f is the revolution frequency of the bunches, N_1 and N_2 are the number of particles per bunch, N_b represents the number of bunches and $\sigma_{x,y}$ are the Gaussian widths of the beam in the horizontal and vertical planes, respectively.

Recording the total *integrated* luminosity over a period of time (i.e. $\mathcal{L}_{\text{int.}} = \int \mathcal{L}dt$) is relevant since it provides a direct connection with the total number of observed events $N_{\text{obs.}}$ for given process and its cross–section σ:

$$N_{\text{obs.}} = \mathcal{L}_{\text{int.}}\sigma. \tag{2.5}$$

2.2.5 Pile–up

Pile–up is a direct consequence of the beam properties and the phenomenology of pp collisions. It is defined as the mean number of collisions per bunch crossing and can severely impact the data–acquisition efficiency. There are two distinct pile–up mechanisms: *in–time* pile–up, which refers to the multiple number of hard collisions that may occur per bunch crossing; and *out–of–time* pile–up, which refers to the bunch spacing being smaller than the detector response time, hence recording collisions coming from different bunches simultaneously.

Pile–up inevitably scales with the instantaneous luminosity, a challenging undesired effect that looks to be mitigated from the experimental and theoretical points of view [120, 121]. Figure 2.5 shows the total integrated luminosity[‖]

[‖]The total Run 2 integrated luminosity recorded by ATLAS is 147 fb^{-1}. However, the amount of that data that is flagged as "good for physics" is slightly lower due to inherent detector inefficiencies. After the ATLAS data quality assessment, the recorded "good" data has an integrated luminosity of 139 fb^{-1}.

and pile–up for the Run 2 data–taking period recorded by the ATLAS experiment.

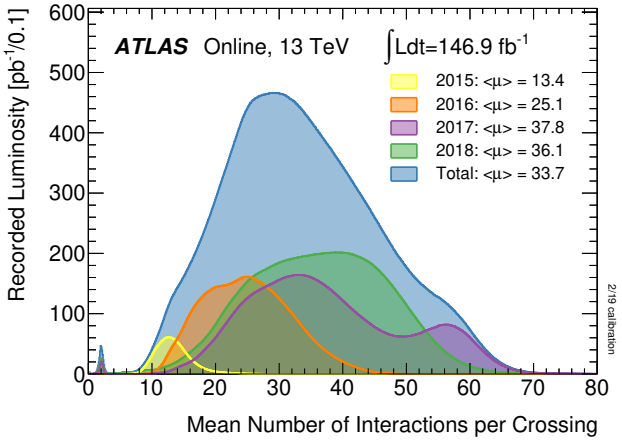


Figure 2.5: Luminosity-weighted distributions of the number of interactions per bunch crossing over the 2015–2018 data–taking period recorded by the ATLAS experiment. Total integrated luminosities are shown along with the average number of interactions per year. Source: Ref. [122].

2.3 The ATLAS detector

Soon after construction started for the LHC and the physics case was built, the discussion to determine what type of experiments would be able to exploit, and fullfil, the challenges ahead kicked–off. The culmination of years of detector research and development by thousands of physicists, engineers, technicians and other supporting staff was the 1992 Evian conference where proto–collaborations, that would later form the well–known current LHC ring collaborations, presented their "Expressions of Interest", detailing their experiment concept designs and ambitions for the coming decades of thrilling *pp* collisions [123].

At the Evian meeting, two of these proto–collaborations, ASCOT [124] and EAGLE [125], presented "similar" proposals incorporating a toroidal magnet configuration for the muon spectrometer (MS) system. It was agreed at that time that the combination of both collaborations under a single name, combining resources and expertise from both groups, was the way forward. In October 1992, ASCOT and EAGLE joint forces producing a successful "Letter of Intent" (see Figure 2.6) for the LHC under the name of the ATLAS Collaboration.

Nowadays, the ATLAS Collaboration [71,72] involves more than 5500 members working together from all around the world and almost 3000 scientific au-

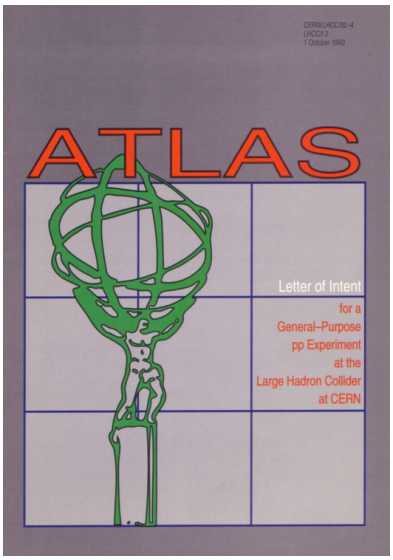


Figure 2.6: The ATLAS Letter of Intent cover page [126]. Submitted on October 1st 1992 and marking the start of the ATLAS Collaboration.

thors. It is not only one of the largest scientific collaborations in the globe, but also the biggest high–energy collider experiment ever constructed spanning over 46 m long, 25 m high, with a 25 m diameter, and a total mass of over 7000 tonnes. The ATLAS detector is built to detect all the particles with excellent resolution and in the complete solid angle coming from pp collisions from the LHC accelerator ring. It is organised in several layers with cylindrical symmetry, each providing complementary information about the particles travelling through them. The innermost layer is the inner detector (ID), or tracking system, that reconstructs the traces and trajectories of the charged particles traversing it. The next layer is formed by the calorimeters, of which there are two types, and that measure the energy of the particles. Depending on the type of particle, either electrons or photons, or hadrons, the read–out happens at either the electromagnetic calorimeter (ECAL) or the hadronic calorimeter (HCAL), respectively. Finally, the muon spectrometer aims to reconstruct the muons that are not stopped in the other detector layers. Figure 2.7 shows a cut–away view of the current ATLAS detector, with the caveat that during Run 2, the New Small Wheel (NSW) MS was not yet installed. In the following section, an overview of the ATLAS subsystems is given.

2.3.1 ATLAS detector geometry

The detector is built to be symmetrical in the forward and backward direction from the interaction point (IP). Moreover, the IP marks the origin of the co-

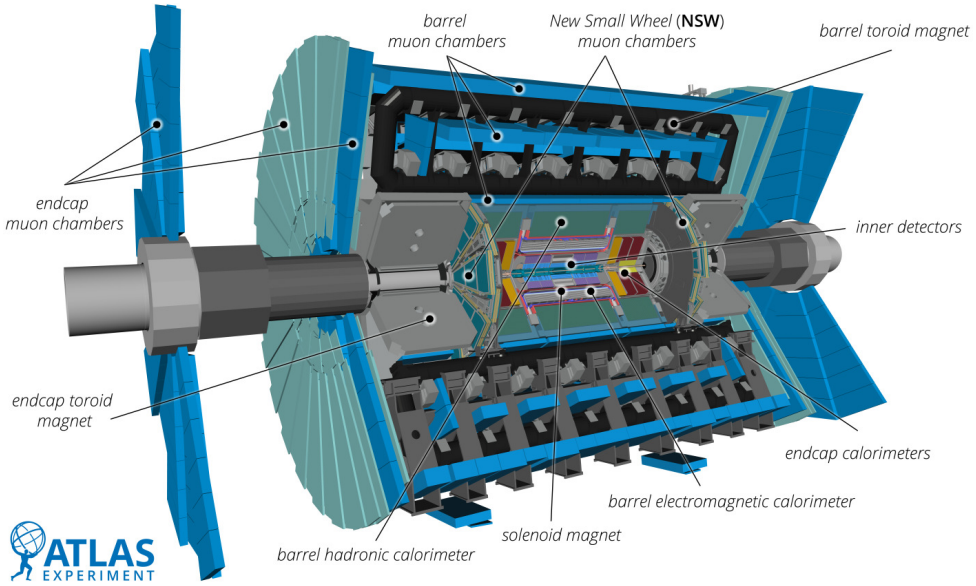


Figure 2.7: The ATLAS detector schematic view. Source: Ref. [127].

ordinate system (x, y, z) with the z–axis running along the beam pipe and the x–y plane transverse to the beam direction. This plane is used to define some of the most important kinematic variables such as the transverse momentum p_{T} or the transverse missing momentum $E_{\mathrm{T}}^{\mathrm{miss}}$. Working with polar coordinates is also beneficial. In particular, the azimuthal angle ϕ is defined around the beam axis (in the x–y plane) and the polar angle θ measures the distance to the beam direction. The latter is typically not used, instead it is transformed into the pseudorapidity η defined as

$$\eta = -\ln \tan(\theta/2). \tag{2.6}$$

Following this transformation, $\eta = 0$ ($\theta = 90°$) indicates a direction perpendicular to the beam pipe and $\eta \to \infty$ ($\theta = 0°$) would be along the beam pipe. The pseudorapidity is used to measure the coverage of the detector subsystems (and particle direction) which is limited to about $|\eta| < 5$ ($\theta = 1°$). Additionally, the pseudorapidity–based angular distance between two particles is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

2.3.2 The Inner Detector

The ID [128] is the most central ATLAS subsystem to the IP and is composed of three main subdetectors: the pixel detector, the Semiconductor Tracker (SCT), and the Transition Radiation Tracker (TRT). Its goal is to precisely measure the

origin and trajectories of the charged particles that traverse it. To do so, it is embedded in an homogenous 2 T magnetic field that bends the particle trajectories, which allows for the determination of their charge sign and momentum. It is made up of high–granularity* semiconductor silicon pixel (pixel detector) and strip detectors and continuous tracking straw tube gaseous detectors. The combination of these technologies results in excellent pattern recognition and precise momentum measurements. The ID has two types of detector geometries: the *barrel* region that is made up of concentric cylinders around the beam pipe; and the two *end–cap* regions which consist of perpendicular disks to the beam direction. The combination of both provide a coverage up to $|\eta| < 2.5$. Not to mention that the detector material, electronics, and cooling components have to withstand the high levels of radiation, with minimal deterioration to their efficiencies, as well as being kept as small as possible to avoid secondary particle emissions, and energy losses in the interactions with the incoming particles. Figure 2.8 shows a schematic view of the ID barrel region.

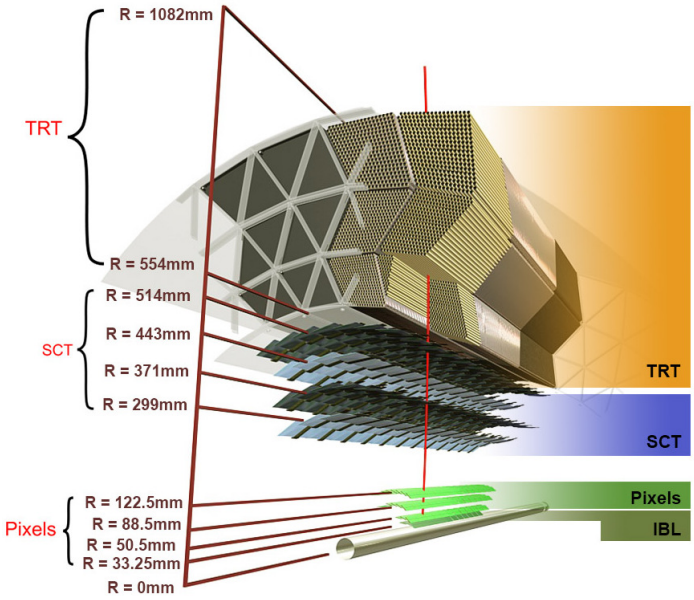


Figure 2.8: The ATLAS Inner Detector barrel region schematic view. From innermost to outermost: the pixel layers, the four cylindrical layers of the SCT and the straw layers of the TRT are shown. Source: Ref. [129] (Fig. 1).

The pixel detector lies at the innermost part of the ID. It contains 1774 identical silicon pixel detectors arranged in three barrel layers and three end–cap disks at each side of the barrel. With these, it covers $|\eta| < 2.5$ and is

*The signals are measured with a precision of almost 10 µm to determine the particle kinematics.

designed to record three hits per reconstructed track. For Run 2, a new layer of pixel detectors, the Insertable B-Layer (IBL) [130], was installed at the core of the pixel detector. It contains 280 additional pixel modules, extending the coverage to $|\eta| < 2.7$, and providing an additional hit per track.

The SCT is the next subdetector of the ID. It follows a similar geometry as the pixel detector, but uses long and narrow double–sided microstrip silicon detectors instead, hence covering a larger area. It contains 4088 modules in four barrel layers and nine end–cap disks at each end with a coverage of $|\eta| < 2.5$. A total of eight hits can be recorded per track.

The TRT is the last subdetector of the ID. It is formed by about 300 000 straw tubes arranged in the barrel and end–cap sections. It provides about thirty hits per reconstructed track and covers $|\eta| < 1$ in the barrel section and $|\eta| < 2$ in the end–caps. Detection is accomplished when traversing particles ionise the Xenon gas inside the straws which generates the signal. Moreover, the gaps between the straws are filled with transition radiation detectors which emit radiation that depends on the type of particle. This gives the TRT the ability to distinguish the electrons (and positrons) from other heavier charged particles [131].

2.3.3 The Calorimeters

The next part of the ATLAS detector are the calorimeters [132, 133]. Their main role is to stop the particles and to measure their energy with a total coverage of $|\eta| < 4.9$ between the barrel and end–cap sections. Calorimeters are made up of a dense absorber (passive) material, which is large enough to absorb most of the energy of the showering particles; and an active material, which is intertwined with it, detects the particles from the shower, and produces the read–outs. Figure 2.9 shows the layout of the calorimetry system.

The ECAL is a high–granularity calorimeter positioned just after the ID. It covers $|\eta| < 3.2$ and is focused on measuring particles that interact electromagnetically such as electrons and photons. The barrel section has an accordion geometry, ensuring the complete azimuthal angle coverage, with lead absorbers and liquid Argon (LAr) for the active material. On the other hand, the end–caps are organised in two coaxial wheels. A cryostat around the ECAL is needed to keep the low operating temperatures.

The HCAL surrounds the ECAL and targets the energy measurements of strongly interacting particles (mainly hadrons) with a lower granularity than the ECAL. Several components are part of the HCAL. The Tile Calorimeter covers the surrounding barrel section with a coverage of $|\eta| < 1.7$, using steel and and plastic scintillators for the absorber and active material, respectively. Then, the Hadronic End–cap Calorimeters (HEC) cover the end–caps ($|\eta| < 3.2$) which use LAr for the active material and copper as the absorber. They are

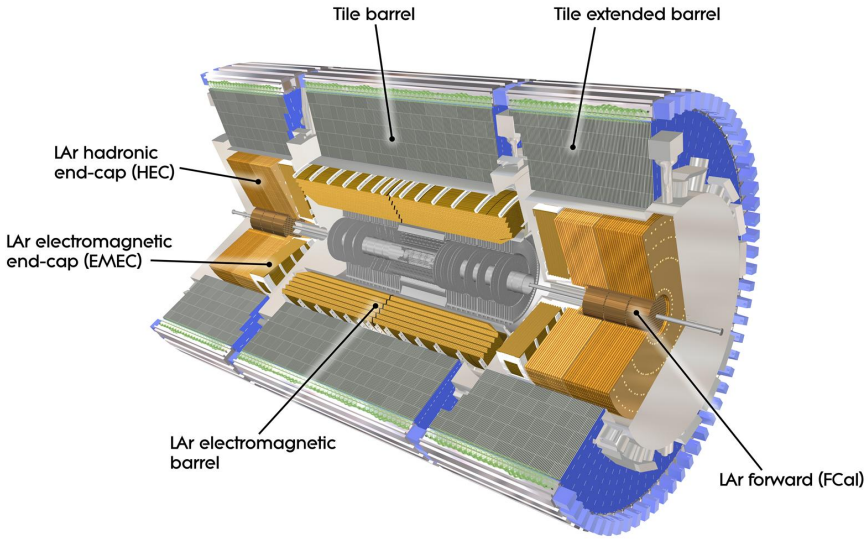


Figure 2.9: The ATLAS calorimetry system schematic view. Source: Ref. [134].

inside the same ECAL cryostat to maintain the LAr operating temperatures. Finally, the forward Calorimeters (FCal), embedded inside the end–cap LAr calorimeters, extend the coverage to $|\eta| < 4.9$ and provide complementary measurements to the $E_{\mathrm{T}}^{\mathrm{miss}}$ (see Section 2.5.6). They are composed by various modules made of copper and tungsten as the absorber material and LAr as the active material.

2.3.4 The Muon Spectrometer

The outermost part of the ATLAS detector is reserved for the muon spectrometers [135]. These aim to measure the momentum and trajectory of the muons that leave little signal on the other subsystems. Muons are bent by a magnetic field, exerted by the barrel and end–cap superconducting air–core toroids, of 3.9 and 4.1 T respectively. Figure 2.10 shows the layout of the MS system and the toroidal magnets. It is formed by three cylindrical barrel sections and six end–cap sections with a coverage of $|\eta| < 2.7$.

Four different types of technologies are used to trigger and reconstruct the system when a muon is detected. The Monitored Drift Tubes (MDTs) provide high–precision measurements in the central barrel sections and end–caps. The Cathode Strip Chambers (CSCs) complement the MDTs in the forward region and posses a higher granularity to withstand the larger amounts of particle flux and background rates. Moreover, the Resistive Plate Chambers (RPCs) and

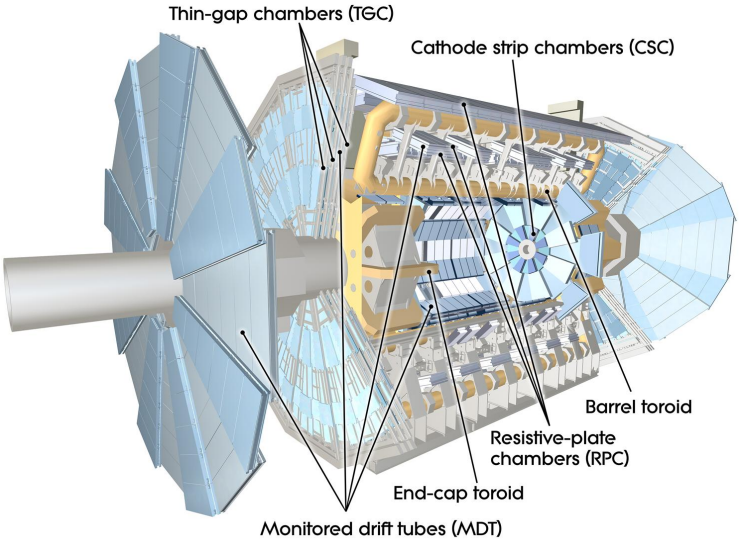


Figure 2.10: The ATLAS muon spectrometer schematic view. Source: Ref. [136].

Thin Gap Chambers (TGCs) detectors provide triggers in both the barrel and end–cap sections in addition to measurements in the orthogonal direction to the muon bending.

2.4 ATLAS trigger system

Given the short proton bunch spacing, the multiple possible collisions per bunch, and the thousands of particles and showers of particles per collision, the amount of data to be recorded is estimated to be about 60 million MB/s. The ATLAS data acquisition and trigger systems are used to ensure optimal data–taking and to provide an event selection tool for those events with distinguishing characteristics for physics analyses. Thus, reducing the amount of stored data and making its quantity more manageable. ATLAS used a two–staged online event trigger system for Run 2: the Level 1 (L1) hardware–based trigger, and the High Level Trigger (HLT) which requires software-based event reconstruction [137].

The L1 trigger, integrated into the detector, reads information from the calorimeters and the MS, and takes a decision on whether to keep or reject and event every 2.5 µs. It reduces the read–out rate from the 40 MHz down to about 100 kHz. These events are then transferred to the HLT, which operates on 40 000 dedicated CPU cores that perform a detailed analysis of each event with information from specific detector regions (in about 200 µs). The HLT can further reduce the read–out event rate to about 1000 Hz. These events are

then stored at the different tiered WLCG sites for offline event reconstruction.

2.5 ATLAS object reconstruction

After the selection of the events by the HLT, an offline object reconstruction software is executed. This reconstruction aims to build the particles being used by all physics analysis such as leptons, photons jets, as well as the event $E_{\mathrm{T}}^{\mathrm{miss}}$, by combining information from all detector subsystems. This section gives an overview of how these particles and their kinematic properties are reconstructed by the ATLAS experiment.

2.5.1 Track and vertex reconstruction

The first step is the reconstruction of the tracks recorded by the ID which allows for accurate particle momentum and charge identification. This is done using an *inside–out* tracking strategy which follows the particle's propagation through the ID [138]. This is complemented by the *outside–in* strategy where track information is built starting from the TRT, and is matched to other tracks in the pixel and SCT strip detectors [139].

The inside–out approach is divided in a few steps. In first place, track *seeds* are found coming from sets of three space–points* in the pixel and strip silicon detectors which are then evaluated with the combinatorial Kalman Filter [140]. This filter combines the original seeds with subsequent space-points from the remaining silicon detectors that are compatible with the original seed. In second place, resolution of overlapping track candidates and rejection of incorrect combinations is performed by an ambiguity solver software. This software scores each track based on a set of quality criteria and keeps the best tracks [141]. Finally, the tracks are extended onto the TRT. In case of a successful track extension, the whole track is re–fitted using the additional hits on the TRT.

Following the inside–out tracking, the outside–in approach takes place where unused TRT tracks are matched to the leftover hits by the inside–out approach. These TRT initial track seeds are typically chosen in regions where successive ECAL showers develop. Unmatched TRT tracks to the SCT are also kept as possible sources of photon conversion reconstruction assignments.

Once all tracks have been obtained, the location of the original hard scatter vertex, known as the primary vertex (PV), is found using a two–step approach [142]. The first step is vertex finding, where tracks are associated to vertex candidates. This is followed by an iterative vertex fitting procedure that

*Given the high density of particles in the inner layers of the ID, the hits in the pixel or strip silicon detectors are grouped in clusters that may include energy deposits from various particles. From these clusters, the three–dimensional particle position and its uncertainty is enclosed in a space–point.

determines the best vertex candidate position. The output of the vertex finding algorithm is a list of three–dimensional vertices positions along with their covariance matrices. The PV of the event is selected as the candidate vertex with the largest $\sum p_{\mathrm{T}}^2$[†] of all of its constituent tracks, while other candidates are considered as pile–up vertices.

2.5.2 Electron and photon reconstruction

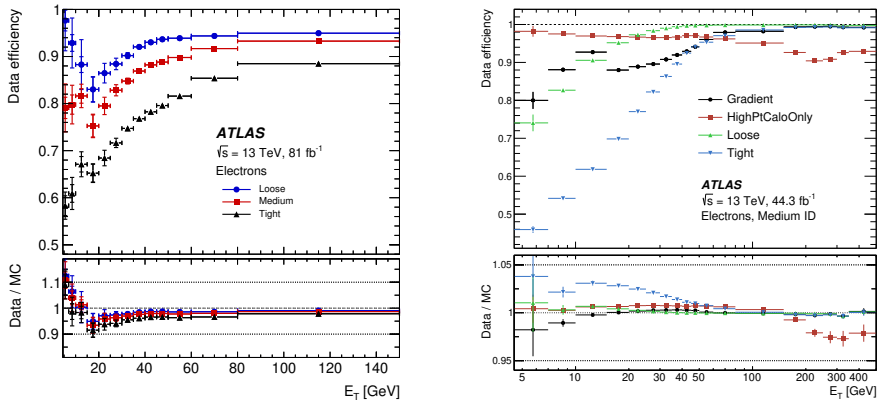
Electrons are generally reconstructed starting from the energy deposits in the ECAL and matching these with the tracks at the ID. On the other hand, photons leave no trace in the ID and are thus characterised only by their ECAL showers [143]. In addition, emission of bremsstrahlung photons coming from electrons/positrons or conversions of photons to an electron–position pair contribute to these signatures, making the reconstruction more challenging.

ATLAS electron and photon reconstruction is done using dynamical clusters of adjustable size called *superclusters* which can characterise the energy from bremsstrahlung photons and converted photons. Superclusters are formed following a set of criteria such as the cluster minimum transverse energy E_{T} of 1 GeV, their $\Delta\eta \times \Delta\phi$ grid extension, and, in case of electrons, their matching to ID tracks [144]. After the supercluster reconstruction, electrons and photons are calibrated and their positions and energies are corrected. The calibration is based on energy scale and resolution corrections between data and simulation from $Z \to e^+e^-$ events; and is validated using $J/\psi \to e^+e^-$ and radiative Z boson decays [145].

Electrons may be identified as *prompt* when originating from vector boson decays, or *non–prompt* when originating from weak decays of secondary long–lived particles such as b–hadrons or from photon conversions. To improve the purity of the prompt electrons and photons, and reject non–prompt ones, several identification requirements are defined. Electron identification requirements combine information from the ID tracks, the ECAL shower shape, and the previous matching of superclusters to ID tracks into a likelihood discriminant. There are three working points (WPs) for this discriminant labelled as *Tight*, *Medium* or *Loose* based on prompt electron signal efficiency and background rate rejection. The Tight working point carries the hardest selection and a reduced efficiency with respect to the others, but provides a superior discrimination against the other non–prompt electrons. In addition, isolation requirements are imposed on electrons to improve this discrimination. These generally quantify and set thresholds on the maximum activity around the electron candidate, using transverse energy (calorimeter-based) and momentum (track–based) sums in ΔR cones around the electron candidate, but excluding it from the calculation [143].

[†]The candidate vertex must have at least two associated tracks with $p_{\mathrm{T}} > 500$ MeV.

After the reconstruction, identification and isolation requirements, electron efficiencies are measured using the tag–and–probe method [146], for both data and simulation, in $Z \to e^+e^-$ and $J/\psi \to e^+e^-$ events. Figure 2.11 shows the electron identification efficiencies for the three WPs (left) and the isolation efficiencies for four different methods at a Medium electron identification point (right). The bottom panel shows the ratio of data to the $Z \to e^+e^-$ and $J/\psi \to e^+e^-$ simulation. These ratios are used as the scale factors (SFs) that are later applied as correction factors to simulated events, in order to match efficiencies in data. For electrons, photons and the other objects described in the following sections, the associated uncertainties on the calibrations and the SFs are propagated through the analyses and accounted for as additional sources of detector–related systematic uncertainty (see, e.g. Section 5.6.1). In particular for electrons, their SFs are typically obtained separately for the reconstruction, identification, and isolation efficiencies.



(a) (b)

Figure 2.11: (a) Electron identification efficiency for the three WPs and (b) electron isolation efficiency using a Medium electron identification, as a function of E_{T}, in $J/\psi \to e^+e^-$ ($E_{\mathrm{T}} < 15$ GeV) and $Z \to e^+e^-$ ($E_{\mathrm{T}} > 15$ GeV) events. The inner uncertainties are statistical, and the total uncertainties are the statistical and systematic uncertainties are added in quadrature. For both plots, the bottom panel shows the data–to–simulation ratios. Source: Ref. [143] (Figs. 17a and 23a).

2.5.3 Muon reconstruction

Muons are reconstructed from information coming from the ID and MS and occasionally from the calorimeters [147]. Reconstruction at the ID is analogous to that for the electrons, while at the MS the muon reconstruction starts with a short straight-line local track pattern search in each muon chamber, in order to form *segments*. Then, muon tracks candidates are formed by means of com-

bining segments from different MS layers (starting from the middle layer and adding the inner and outer layer segments). MS muon tracks may be combined with ID tracks, if available. For this reason, several muon types are defined depending on which detector subsystems are used for the reconstruction:

- *Combined* muons are provided by matching MS tracks to ID tracks and then performing a global fit on MS and ID hits. Muons outside of the ID acceptance ($|\eta| > 2.5$) form a subset of the combined muons called silicon–associated forward muons.
- *Inside–out* muons are provided by a complementary inside-out algorithm that extrapolates ID hits to loosely-aligned MS hits.
- *Muon–extrapolated* are MS reconstructed muons without any match in the ID. Their parameters are extrapolated towards the IP to assess their compatibility, hence extending the coverage to $|\eta| < 2.7$.
- *Segment–tagged* muons are provided by extrapolating ID tracks to at least one MS segment. Muon parameters are taken from the ID tracks. These are mainly low p_{T} muons or muons that traverse low–acceptance MS regions.
- *Calorimeter–tagged* muons are identified by matching ID tracks to energy deposits in the calorimeter compatible with a minimum–ionising particle. Again, muon parameters are taken from the ID tracks. These are mainly muons that traverse partially instrumented MS regions, i.e. $|\eta| < 0.1$.

Analogously to electrons, muons also have several identification and isolation WPs. Muon identification is based on a set of quantities like the p_{T} or charge–to–momentum ratios that characterise the track quality [147]. Three main WPs, *Tight*, *Medium* or *Loose*, are defined in the same manner as for electrons. An additional WP, *High–*p_{T}, is optimised for muons with $p_{\mathrm{T}} > 100$ GeV.

After the reconstruction, muons are calibrated to achieve a permille and percent precision on the muon momentum scale and resolution respectively. These corrections are obtained by comparing $Z \to \mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ events in data and MC. Likewise, the tag–and–probe method is used to estimate the efficiencies of the reconstruction, identification, and isolation requirements, in order to obtain the corresponding SFs. Figure 2.12 shows the muon identification efficiency for the Medium identification point. The bottom panel shows the ratio of data to the $Z \to \mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ simulations.

2.5.4 Jet reconstruction

Jets are collections of collimated particles that originate from the hadronisation of the emitted partons in the hard scatter. These particles interact with

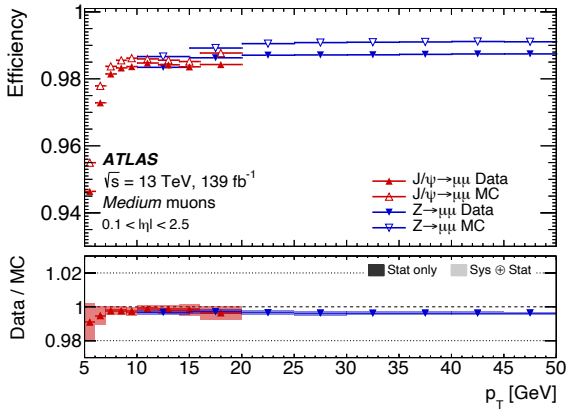


Figure 2.12: Muon reconstruction and identification efficiency for the Medium WP, in $Z \to \mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ events, as a function of the muon p_{T}. The panel at the bottom shows the ratio of the measured to predicted efficiencies, with statistical and systematic uncertainties. Source: Ref. [147] (Fig. 13).

the detector and produce clustered energy deposits in the HCAL. Tracking information from the ID is combined with calorimetry measurements to build the hadronic jet objects. This exploits the higher granularity of the ID tracks and improves the jet energy and angular resolutions, especially in the low–energy regime. This is done by the *particle flow* (PF) algorithm which represents an improvement compared jet clustering algorithms that only used calorimetry data in the central region of the detector [148]. Then, jet finding algorithms run over these final–state objects and select which clusters are more likely to form a jet, based on their reconstructed parameters, and can therefore be traced back to the original parton.

Several jet finding algorithms are available for jet reconstruction, but the most widely–used one in ATLAS is the sequential anti–k_t algorithm [149]. This algorithm is based on distance measures to form to create the jet cones: such as the distance between two objects d_{ij} and the distance of an object to the beam d_{iB} which can be written as

$$d_{ij} = \min(k_{t,i}^{-2}, k_{t,j}^{-2}) \frac{\Delta_{ij}^2}{R^2}, \qquad d_{iB} = k_{t,i}^{-2}, \tag{2.7}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, R is the radius of the cone that defines the jet, and $k_{t,i}$, y_i and ϕ_i are the transverse momentum, the rapidity and azimuthal angle of the object i, respectively. The anti–k_t algorithm follows an iterative method to reconstruct the jets. First, a database is created with all possible d_{ij} and d_{iB}, computed for all available objects, and the smallest pair is

compared: if $d_{ij} < d_{iB}$, both objects (i and j) are merged into one jet and are removed from the database; if on the contrary $d_{ij} > d_{iB}$, then object i becomes a jet and is removed from the database. This is repeated until all objects have been assigned. This procedure automatically combines soft particles close to energetic ones in ΔR which make these jets infrared and collinear (IRC) safe[‡].

Jets can be classified in two types, depending on the jet radius R. *Small–R* jets, with $R = 0.4$ in the anti–k_t algorithm, include jets originating from quarks and gluons and are the most commonly used jet objects in ATLAS. *Large–R* jets, with $R = 1.0$ in the anti–k_t algorithm, target decays of heavy hadronic particles that are sufficiently boosted so that their products are close together.

2.5.4.1 Jet calibration

Reconstructed jets have to be calibrated to account for several effects such as pile–up, energy leaks, or undetectable sources of energy from the interaction of the detector and the incoming particles. A calibration of the jet energy scale (JES), which adjusts the reconstructed jet energy, momentum and direction, is performed; bringing it closer to the *truth* jet[§] [150]. The calibration procedure consists on several stages which can be broadly summarised in three major steps. First, pile–up contributions (see Section 2.2.5) are subtracted by removing the excess energy from additional pp collisions. Then, the jet energy and direction are corrected to the truth jet coming from the MC record. Moreover, the *global sequential calibration* technique further improves jet momentum resolution and the associated uncertainties [150]. Finally, the resulting calibration factors are applied to the MC simulation and compared to data to further refine them (also known as *in situ* calibration).

Complementary jet energy resolution (JER) measurements are done in processes where the jet momentum can be precisely obtained, like back–to–back di–jet events. Precise knowledge of the JER is decisive in high jet multiplicity final–states measurements, in both SM and BSM physics, as well as affecting the $E_{\mathrm{T}}^{\mathrm{miss}}$ which can play a crucial role in some searches. In order to match the JER of simulations to data, a smearing procedure on the simulated events is applied where necessary[¶] [150].

An additional effort is done to suppress the contribution coming from pile–up jets. Tracking information matched to each jet can be used to identify and reject those not originating from the hard scatter event. In ATLAS, the jet vertex tagger (JVT) discriminant is used to classify these pile–up jets [151].

[‡]IRC safety refers to a final jet collection that cannot be modified by arbitrary soft and collinear gluon emissions.

[§]Truth jets are reconstructed using stable final–state particles and exclude muons, neutrinos, and particles from pile-up interactions. They are geometrically matched to reconstructed jets within $\Delta R < 0.3$.

[¶]In regions where jet p_{T} resolution is lower in data than in MC, smearing is not applied.

The efficiency of this discriminant in selecting jets coming from the hard scatter is evaluated using a tag–and–probe method in $Z \to \mu^+\mu^- +$jets events and SFs are derived from the data–to–simulation comparisons.

2.5.5 Flavour tagging

Identifying jets being initiated by bottom quarks (known as b–jets) from the hard scattering interaction is a crucial step in the ATLAS reconstruction process. These b–jets can be used to discriminate and reconstruct promptly decaying parent particles that are otherwise not detected, such as top quarks. They are characterised by a long–lived b–hadron that gives place to a displaced vertex inside the jet volume. ATLAS uses several algorithms for b–jet identification, also known as b–tagging algorithms, which exploit the b–hadron and the b–quark fragmentation properties [152]. This identification is done in two main steps. In first place, low–level algorithms take advantage of charged particle tracks and of the displaced vertex to identify b–hadron decays inside jets. Secondly, this information is given to high–level algorithms that make use of multivariate classifiers to increase the b–jet identification efficiency.

In ATLAS, the two most common b–tagging algorithms being used are: the MV2 algorithm [153], based on boosted decision tree (BDT) discriminant; and the DL1 algorithm [153], based on deep neural network. Both algorithms provide a multidimensional output for not only b–jets, but also charm jets (c–jets) and light–flavoured jets probabilities. The performance of these algorithms is evaluated by defining a set of single–cut WPs, corresponding to a fixed selection requirement on the b–tagging algorithm discriminant distribution. Based on these WP selections, the MV2 and DL1 output distributions are divided in five *pseudo–continuous* bins for 85%, 77%, 70% and 60% efficiencies to select a b–jet.

The performance of these algorithms, for each WP, is also assessed using data–to–simulation comparisons in $t\bar{t}$ production enriched events, and SFs are obtained from the ratios. In order to be used in ATLAS analyses, several transformations and corrections are applied to them [152].

2.5.6 Missing transverse momentum

The $E_{\mathrm{T}}^{\mathrm{miss}}$ is a key quantity in particle physics since it allows for the energy measurement of invisible[‖] particles in the transverse plane. The $E_{\mathrm{T}}^{\mathrm{miss}}$ relies on total momentum conservation and can be computed from the negative sum of

[‖] Invisible particles refers to particles that escape detection such as neutrinos, long–lived particles or DM candidates.

all reconstructed and calibrated objects by ATLAS as

$$E_{\mathrm{T}}^{\mathrm{miss}} = \underbrace{-\sum_{\text{electrons}} p_{\mathrm{T}}^{e} - \sum_{\text{muons}} p_{\mathrm{T}}^{\mu} - \sum_{\text{photons}} p_{\mathrm{T}}^{\gamma} - \sum_{\text{taus}} p_{\mathrm{T}}^{\tau} - \sum_{\text{jets}} p_{\mathrm{T}}^{j}}_{\text{hard term}} \underbrace{- \sum_{\substack{\text{unused} \\ \text{tracks}}} p_{\mathrm{T}}^{\text{tracks}}}_{\text{soft term}}, \tag{2.8}$$

where the soft term includes particles with tracks that are associated with the hard scatter PV but with no match to any reconstructed object.

Systematic uncertainties on the $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement are usually estimated by propagating the energy scale and the resolution uncertainties of both the objects in the hard term, as well as those coming from the soft term. Those from the soft term include the modelling of the UE and their impact on the p_{T} scale and resolution of unclustered energy [154]. They are estimated from data–to–simulation $Z \to \ell^{+}\ell^{-}$ events using observables that measure the contribution of the soft term to the total $E_{\mathrm{T}}^{\mathrm{miss}}$.

2.5.7 Objects overlap removal

Ambiguities between independently reconstructed electrons, muons, and jets can arise. A sequential procedure, known as the overlap removal (OLR), is performed to resolve these ambiguities and avoid the double counting of particle candidates. It is applied before the scale factors are determined for all objects, so any mis-modelling of the overlap removal would be accounted for by the scale factors and their associated uncertainties. The following criteria is applied: muon candidates are preferred over electron candidates if they share the same track; the closest jet candidate within a distance $\Delta R = 0.2$ of an electron candidate is discarded; if the electron–jet distance is between 0.2 and 0.4, the electron candidate is removed; finally, for muon–jet distances $\Delta R < 0.4$, the muon candidate is discarded if the jet has more than two associated tracks, otherwise the jet candidate is removed.

3 The Top Quark

As mentioned in Section 1.7, the high mass and properties of the top quark makes its study a crucial task at the LHC. Top quark physics not only encapsulates the precise study of its intrinsic properties, but also the way it interacts and couples to other massive particles. Section 3.1 covers an overview of its prediction and discovery, while Section 3.2 details the several production modes of top quarks in high–energy physics experiments. Section 3.3 covers the top quark decay modes and lists the different decay channels. Additional observables that probe the top quark and decay properties with high precision such as the top quark mass, the W boson helicity fractions, the top quark polarisation or asymmetries are given in Sections 3.4, 3.5, 3.6, and 3.7, respectively. Section 3.8 gives details on the SM couplings of the top quark to other SM particles, while Section 3.10 expands on it by describing effective couplings involving EW interactions. This serves as the founding blocks of the first work carried out in this thesis, where the latest measurements involving top quark EW couplings are scrutinised in the context of a global EFT fit.

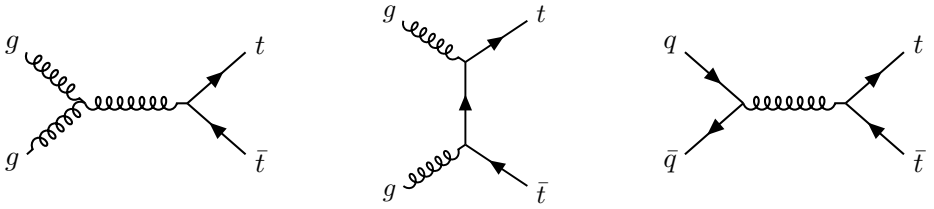
3.1 Towards finding the top quark

The top quark's existence was notably motivated by the discovery $\mathcal{CP}$ violation in the kaon system [27] and the need of a third family of quarks to allow for a $\mathcal{CP}$–violating phase in the CKM matrix. With the discovery of the bottom quark in 1964 at Fermilab [28, 29] and the EW force carriers (W and Z bosons) in 1983 at CERN [155], the existence of a sixth quark was unquestioned. The race for the discovery of the last piece of the fermion sector of the SM developed in the 80s with no return for none of most powerful accelerators at DESY or CERN at the time. This meant that the top quark had to have an extremely high mass, unreachable by those colliders. It was not until the early 90s, when the Tevatron 1.8 TeV proton–antiproton collider at Fermilab came into operation which had enough energy to observe the top quark [30, 31].

3.2 Top quark production modes

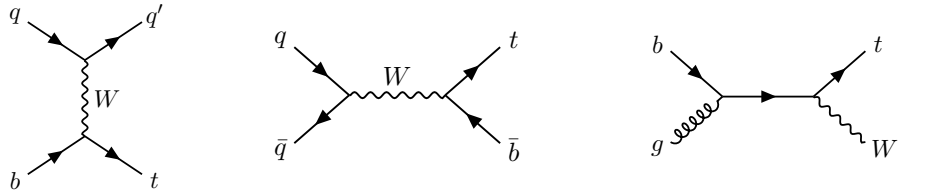
As previously discussed, production of top quarks happens through two main mechanisms at hadron colliders: $t\bar{t}$ production via strong interactions (see Figure 3.1) and single–top quark production via EW interactions (see Figure 3.2). The former can happen via gluon–gluon (gg) fusion and quark–antiquark annihilation ($q\bar{q}$). As a result of the different collision energies and initial hadrons, the main mechanism for $t\bar{t}$ production differs between Tevatron and the LHC. At Tevatron, $q\bar{q}$ annihilation contributes to 85% of the $t\bar{t}$ production cross–

section*, while at the LHC the gg fusion channel is the dominant one contributing to about 90% of the total cross–section at 13 TeV. The most precise calculations of the cross–section for $t\bar{t}$ production are done at next–to–next–to–leading order in QCD with next–to–next–to–leading logarithm soft gluon resummations (NNLO+NNLL). These are $\sigma_{t\bar{t}} = 7.2^{+0.1}_{-0.2}$ (scale) $^{+0.2}_{-0.1}$ (PDF) pb for Tevatron, and $\sigma_{t\bar{t}} = 832^{+20}_{-29}$ (scale) $\pm$ 35 (PDF) pb for the LHC at 13 TeV [156, 157]. Both ATLAS and CMS have performed measurements of the $t\bar{t}$ cross–section at the several running centre of mass energies of the LHC during the last decade [158], finding good agreement with the theoretical predictions as shown in the top panel of Figure 3.3.



(a) gg fusion (b) $q\bar{q}$ annihilation

Figure 3.1: Feynman diagrams at LO for $t\bar{t}$ production in hadron colliders. The first two diagrams correspond to gg fusion and the last one to $q\bar{q}$ annihilation.



(a) t–channel (b) s–channel (c) tW associated production

Figure 3.2: Feynman diagrams at LO for single–top quark production in hadron colliders for the three main mechanisms: t–channel (left), s–channel (middle) and tW associated production (right).

Besides $t\bar{t}$ pair production, the production of a single–top quark associated to other particles is also possible via three mechanisms: t–channel, s–channel, and tW associated production. The contributions to the total single–top quark production for the LHC at $\sqrt{s} = 13$ TeV are about 73%, 3% and 24% respectively. The most precise calculations of the cross–sections are done

*This production mode requires around 350 GeV of collision energy at the minimum. In 1992, only the Tevatron with a centre of mass energy of 1.8 TeV had enough energy to produce the $t\bar{t}$ pair. During Run I (1992–1996), Tevatron collided 900 GeV proton beams with 900 GeV antiproton beams.

at NNLO [159] for t–channel, and at next–to–leading order (NLO) including NNLL resummations for s–channel and tW [160, 161]. The bottom panel of Figure 3.3 shows the latest ATLAS and CMS measurements [158] of the single–top quark production cross–section for each channel, which are in agreement with the theoretical predictions.

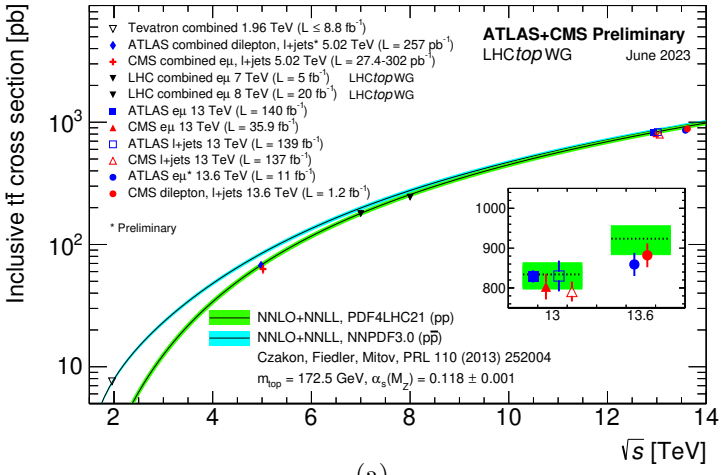


Figure 3.3: (a) Comparison of the measured $t\bar{t}$ production cross–section as a function of the COM energy with the theoretical predictions at NNLO+NNLL accuracy. (b) Comparison of the measured single–top quark production cross–section for each channel, at different COM energies, with the theoretical predictions based on NLO, NLO+NNLL and NNLO (t–channel only). In both, theory bands cover uncertainties due to renormalisation and factorisation scales, as well as those from parton density functions and the strong coupling. All calculations are done with $m_t = 172.5$ GeV. Source: Ref. [158].

3.3 Top quark decay

The top quark decay is completely specified by the contents of the CKM matrix. Given that $|V_{tb}| >> |V_{ts}|, |V_{td}|$, the top quark decays via the EW interaction almost exclusively to a W boson and a bottom quark. Precise measurements of the ratio of the branching ratio $\Gamma(t \to Wb)/\Gamma(t \to Wq,\ q = b, s, d) = 0.957 \pm 0.034$ [16], which is proportional to $|V_{tb}|^2$, can be used to constrain this CKM matrix element. The full width of the top quark has been extracted to be $\Gamma_t = 1.42^{+0.19}_{-0.15}$ GeV [16]. It translates to a mean lifetime of about 5×10^{-25} s, which is smaller than the formation time of QCD bound state hadrons, $\mathcal{O}(10^{-24}$ s), and implies that top quarks decay before hadronisation. This allows for the study of a "bare" quark; which transfers its quantum numbers, such as the spin information, to its decay products. Besides, this unique feature of the top quark decay helps in its study at particle detectors as its properties can be studied through its decay products.

The reconstruction of the top quark – or that of its decay products – depends on the decay mode of the W boson. Figure 3.4 presents all possible $t\bar{t}$ decay modes from which the corresponding branching fractions can easily be extracted.

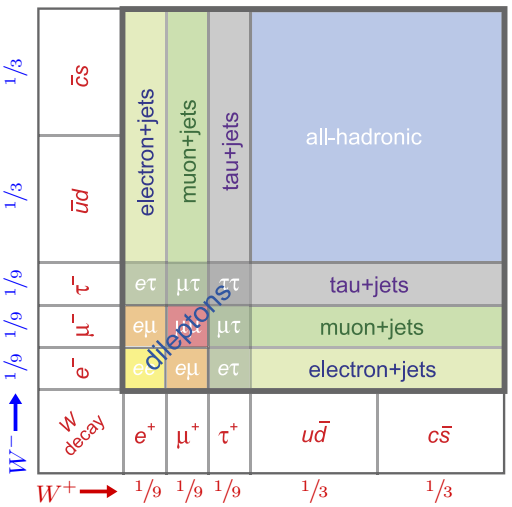


Figure 3.4: Top quark pair decay channels. They are displayed in a 9×9 grid where the size of each of the coloured areas represents the branching fraction of each decay channel over the total decay. By looking at this grid in one dimension, branching fractions for single–top quark production can also be extracted. Source: Ref. [162] (Fig. 3)

The usual channels are separated in *all hadronic* ($36/81 \sim 44\%$) when both W bosons decay to a quark–antiquark pair, *semileptonic* or *lepton+jets* ($36/81 \sim 44\%$) when one W boson decays to a charged lepton and its corresponding

neutrino and the other decays to quarks, and *dileptonic* ($9/81 \sim 11\%$) when both W bosons decay to charged leptons and their corresponding neutrinos. These expected values are extremely close to the experimental branching ratios found in Ref. [16]. Discrepancies arise from small hadronic corrections and the fact that the τ lepton cannot be identified by current detectors[†]; rather its decay products are detected which ultimately complicates the branching ratio determination. Considering this fact, many analyses only use *light* charged leptons (electrons and muons) and exclude τ–hadronic leptons from their fiducial volumes.

3.4 Top quark mass

The top quark mass is a free parameter of the SM, and can be measured either *directly* from the reconstruction of its decay products or *indirectly* by comparing the measured $t\bar{t}$ production cross–sections to those predicted by the theory. Currently, the most precise measurements use the former method, together with kinematic templates of data and MC simulations, to obtain the best–fit top quark mass [163]. Controversially, this extracted mass cannot be used as an input parameter for precise calculations since its relation with a well defined (renormalisable) field theory parameter is not clear. It is usually referred to as the Monte Carlo mass m_t^{MC} which can be later related to specific mass schemes, such as the $\overline{\mathrm{MS}}$ or pole scheme, which can then be used in theoretical calculations [164].

Precise measurements of the top quark mass, alongside the W and Higgs boson mass, are crucial to assess the self–consistency and vacuum stability of the SM, and also new physics effects [16, 165–167]. The current understanding and compatibility between these measurements can be seen in Figure 3.5.

3.5 W boson helicity

The measurement of the W boson helicity fractions in the top quark decay is another stress test of the SM. The W boson can be emitted in three different helicity states[‡]: either negative, zero (longitudinal) or positive. The helicity fractions are defined as the partial decay rate of each state over the total rate, $F_{L,0,R} = \Gamma_{L,0,R}/\Gamma(t \to Wb)$, where F_L, F_0 and F_R are the left–handed, longitudinal and right–handed helicity fractions, respectively. As mentioned in Section 1.2.4, the W boson can only couple to left–handed fermions which means

[†]Owing to its high mass, the τ lepton promptly decays and does not leave a trace in the detector's electromagnetic calorimeters. It is reconstructed from its (hadronic) decay products such as pions.

[‡]Helicity is the projection of the spin of a particle (s) onto its direction of momentum (p): $h = s \cdot \hat{p}$. A particle has positive helicity (right–handed) when its spin direction aligned in the same direction of its motion, and negative helicity (left–handed) if opposite.

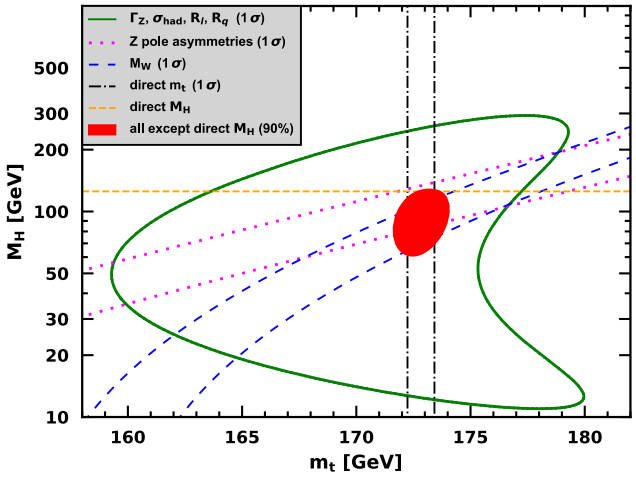


Figure 3.5: Fit result and 1σ uncertainties in m_H as a function of m_t for various inputs, and the 90% confidence level (CL) region (red) allowed by all data. $\alpha_s(m_Z) = 0.1185$ is assumed except for the fits including the Z lineshape. The width of the horizontal dashed band is not visible on the scale of the plot. Source: Ref [16].

that F_R is naturally suppressed in the SM since it would involve a right–handed bottom quark. A summary of the helicity fractions and their theoretical predictions can be seen in Figure 3.6. Experimentally, the helicity fractions have been found to agree with the SM prediction [168–170]. These measurements are very sensitive to the Wtb vertex and therefore become exceptional probes to constrain non–SM couplings between the W boson, the bottom and the top quarks [171, 172].

3.6 Top quark polarisation and spin correlation

In virtue of its fermionic nature and short life–time, the top quark's spin information is transferred to its decay products[§], and therefore its polarisation and $t\bar{t}$ spin correlations can be measured. At the LHC, the dominating gg fusion mechanism in $t\bar{t}$ production is symmetric under parity which is conducive to unpolarised (anti)top quarks at LO[¶]. However, the spins of the top and antitop quarks are expected to be strongly correlated [174].

The top quark spin cannot be measured directly, but the angular distributions of the decay products of the top quark are correlated with the chosen spin

[§]The life–time is also shorter than the spin decorrelation time scale, $\mathcal{O}(10^{-21}$ s) [16].

[¶]Small longitudinal polarisation arise from EW corrections and transverse polarisation from absorptive terms at one–loop [173].

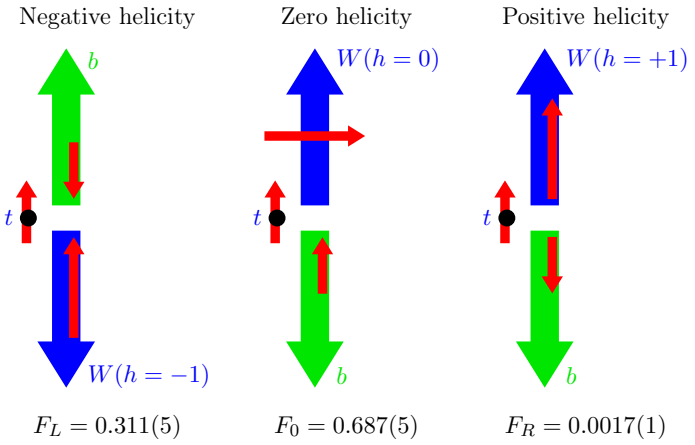


Figure 3.6: Possible states for the W boson polarisation in top quark decays. The red arrows indicate the spin direction, while the green and blue arrows represent the momentum direction of the bottom quark and W boson, correspondingly, in the top quark rest frame (black circle). The numbers at the bottom show the predicted helicity fractions at NNLO in QCD for $m_t = 172.8 \pm 1.3$ GeV [170].

axis $\hat{z}$ as

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta_a} = \frac{1}{2}(1 + P_z \kappa_a \cos\theta_a), \tag{3.1}$$

where Γ is the decay width, θ_a is the angle between the decayed particle and the chosen spin axis in the (anti)top quark rest frame, κ_a is the analysing power of the decayed particle and P_z is the (anti)top quark polarisation in the spin axis direction. The charged lepton has maximal spin analysing power for the top quark, $\kappa_{\ell^+} \approx 1$ [175], and is reversed for the antitop quark, i.e. $\kappa_{\ell^+} = -\kappa_{\ell^-}$. The absolute azimuthal opening angle between the two charged leptons $\Delta\phi_{\ell^+\ell^-}$ in the laboratory frame – following from a dileptonic decay of the $t\bar{t}$ system – is one of the more straightforward variables to measure these spin properties on. Both ATLAS and CMS have performed measurements of the top quark polarisation and $t\bar{t}$ spin correlations in the dileptonic channel at $\sqrt{s} = 13$ TeV finding a good agreement with the SM expectations [176, 177]. As a final remark, these observables register a high sensitivity to BSM models that modify the existing couplings or add new vertices. Such effect can be parametrised in an EFT to constrain the relevant effective couplings, which is also explored in Ref. [176].

In single–top quark production, especially in the dominant t–channel at Born–level, top quarks are produced with their spin completely aligned along the direction of the down–type quarks [178, 179]. While for single–antitop quark production the polarisation is in the direction opposite to down–type quark. The most recent results of the top quark polarisation in single–top

t–channel production [180] find a good agreement with the NNLO$_{\rm QCD}$ SM prediction [181]. The $\mathcal{CP}$ properties of the tWb vertex can also be examined in single–top quark production [171]: measurements of sizeable polarisation vector components in the transverse plane to the spectator quark direction would be a hint of $\mathcal{CP}$ violation. However, the results from Ref. [180] are consistent with the SM $\mathcal{CP}$–conserving hypothesis.

3.7 Top quark asymmetries

The $t\bar{t}$ production via $q\bar{q}$ annihilation mechanism can lead to forward–backward asymmetries, $A_{\rm FB}^{t\bar{t}}$, upon counting the number of top quarks that are emitted within the hemisphere centered on the incoming quark beam (forward), relative to those produced within the incoming antiquark beam (backward). In the SM, the top quark pair production is symmetric under the top and antitop exchange at leading order (LO) in QCD, thus, such asymmetry is zero. Nonetheless, interferences due to diagrams at NLO (for $q\bar{q}$ and $qg/\bar{q}g$ initial states) contribute to $A_{\rm FB}^{t\bar{t}}$ causing a preferred emission of the top quark in the direction of the incoming quark and the antitop quark in the direction of the incoming antiquark*.

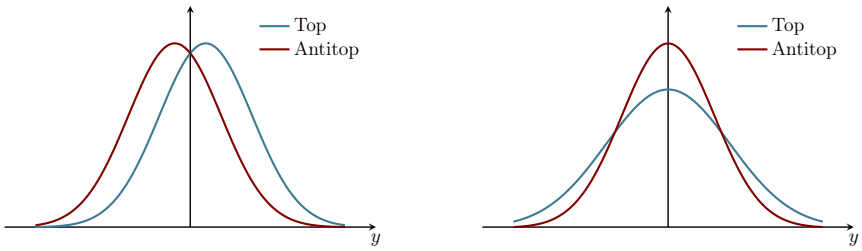
Since $t\bar{t}$ production at Tevatron is dominated by $q\bar{q}$ annihilation (through the nature of the antisymmetric $p\bar{p}$ collisions), this asymmetry ($A_{\rm FB}^{t\bar{t}}$) is expected to be around 10% and can be measured using the direction of the incoming beams [182]. However, this is not the case at the LHC: given that the symmetric gg fusion mechanism dominates (which is not conducive to the interferences described above), any asymmetry caused by the sub–leading $q\bar{q}$ annihilation contribution is washed out from the $t\bar{t}$ production cross–section. Furthermore, the LHC pp collisions are symmetric, meaning that it is not possible to determine the direction of the incoming quark or antiquark, which complicates the definition of $A_{\rm FB}^{t\bar{t}}$.

Figure 3.7 presents a sketch of how top and antitop quark rapidity† distributions compare between Tevatron and LHC. These shapes are a direct consequence of what is described above. For Tevatron, since top quarks are preferably emitted in the direction of the incoming quark, a larger number of them is expected in the positive y region (forward) while the opposite hap-

*The main contribution to the asymmetry appears at NLO in QCD due to interference terms of order α_s^3 to the total cross–section, that are odd under the interchange $t \leftrightarrow \bar{t}$, while keeping the initial partons fixed. These are: interferences between Born–level $q\bar{q} \to t\bar{t}$ and the one–loop (box and crossed) corrections, which are antisymmetric under the $t \leftrightarrow \bar{t}$ interchange; and also interferences between initial and final state radiation diagrams. Moreover, antisymmetric interferences terms of order α_s^3 for qg ($\bar{q}g$) initial states also contribute to the asymmetry at a smaller rate.

†The rapidity is defined as $y = \frac{1}{2}\ln\frac{E+p_z}{E-p_z}$ where E is the energy and p_z the momentum in the direction of the beam axis.

pens for antitop quarks. For the LHC, both distributions are expected to be symmetric but have different widths because the quarks inside the proton have different PDFs than the antiquarks. To be more specific, valence quarks inside the proton carry a larger momentum fraction than virtual sea antiquarks resulting in the emission of more energetic top quarks, closer to the beam direction (larger y). The difference between these distributions defines a central–forward asymmetry between top and antitop quarks, also called *charge asymmetry*.



(a) Tevatron $p\bar{p}$ collisions (b) LHC pp collisions

Figure 3.7: Sketch of $t\bar{t}$ production at Tevatron and LHC as a function of the rapidity y of the top and antitop quarks.

3.7.1 Top quark charge asymmetry

As hinted in above, top and antitop quarks are produced centrally (symmetric rapidity) in pp collisions at the LHC which leads to vanishing asymmetries in comparison with Tevatron [183–186]. Despite these small asymmetries, the large $t\bar{t}$ production statistics available at the LHC can be exploited to obtain event selections that are optimised to increase the fraction of $q\bar{q}$ events. At the LHC, a central–forward asymmetry is defined as

$$A^t_{c,y} = \frac{N(\Delta|y_t| > 0) - N(\Delta|y_t| < 0)}{N(\Delta|y_t| > 0) + N(\Delta|y_t| < 0)}, \tag{3.2}$$

where $\Delta|y_t| = |y_t| - |y_{\bar{t}}|$ is the difference in absolute rapidity between the top quark and the antitop quark. ATLAS and CMS have provided inclusive and differential measurements at $\sqrt{s} = 13$ TeV [187, 188]. In the case of ATLAS, it is measured using both the dileptonic and lepton+jets channels and is found to be $A^t_{c,y} = 0.0068 \pm 0.0015$ (stat. $\oplus$ syst.) which is 4.7 standard deviations away from zero. Both ATLAS and CMS values are in agreement with the SM prediction (NNLO_{QCD} + NLO_{EW}[‡]) [189].

As previously announced, the charge asymmetry is enhanced in fiducial volumes where the $q\bar{q}$ initiated fraction of events is increased. This promotes

[‡]NLO_{EW} refers to the $\mathcal{O}(\alpha^2) + \mathcal{O}(\alpha_s^2\alpha)$ terms, where the latter dominate. For the LHC, the ratio of the asymmetries with or without the EW corrections is $\frac{\text{QCD+EW}}{\text{QCD}} \sim 1.14$.

the study of the charge asymmetry in a boosted $t\bar{t}$ pair regime where the $q\bar{q} \to t\bar{t}$ process is favoured. Several kinematic variables such as the invariant mass of the $t\bar{t}$ pair $m_{t\bar{t}}$, the transverse momentum of the top quark p_{T}^{t} or the $t\bar{t}$ pair $p_{\mathrm{T}}^{t\bar{t}}$, or the longitudinal boost of the $t\bar{t}$ system along the beam axis $\beta_{t\bar{t}}$ grant access to this regime. As a matter of fact, both ATLAS and CMS have performed measurements of the charge asymmetry as a function of these variables finding not only an increase in the central value of the asymmetry, but also good agreement with the SM. Results have been done at $\sqrt{s} = 7$ and 8 TeV [190], and also at $\sqrt{s} = 13$ TeV [187, 188]. In addition, several BSM models grant an enhancement of the charge asymmetry in the boosted regime [191], making this observable a valuable probe of these models and yet another way to study top quark interactions (in $t\bar{t}$ production) in detail.

Leptonic charge asymmetries can also be measured as a complement to the discussed $t\bar{t}$ asymmetries. Using the charged leptons to compute the asymmetry is perhaps a simpler task since lepton kinematics are precisely measured at detectors and it avoids the full reconstruction of the (anti)top quark's kinematics. In $t\bar{t}$ production, the SM predicts the lepton–based asymmetry to be smaller than the reported $t\bar{t}$ one. This is because the decayed leptons do not typically follow the direction of their parent (anti)top quarks, which dilutes the asymmetry defined in Eq. 3.2, and makes the measurement more challenging.

The leptonic charge asymmetry is defined as

$$A_{c,\eta}^{\ell} = \frac{N(\Delta|\eta_{\ell}| > 0) - N(\Delta|\eta_{\ell}| < 0)}{N(\Delta|\eta_{\ell}| > 0) + N(\Delta|\eta_{\ell}| < 0)}, \tag{3.3}$$

where $\Delta|\eta_{\ell}| = |\eta_{\ell^+}| - |\eta_{\ell^-}|$ is the difference in absolute pseudorapidity[§] (defined in Section 2.3.1) between the antilepton decaying from the top quark, and the lepton decaying from the antitop quark. Consequently, this asymmetry has to be measured necessarily in the $t\bar{t}$ dileptonic channel. Both ATLAS and CMS have reported measurements of this asymmetry at $\sqrt{s} = 7$ TeV [192, 193], and 13 TeV for ATLAS [187], finding good agreement with the SM ($\mathrm{NLO_{QCD}}$ + $\mathrm{NLO_{EW}}$) [185]. In particular, the latter ATLAS inclusive measurement reports $A_{c,\eta}^{\ell} = 0.0054 \pm 0.0012$ (stat) $\pm$ 0.0023 (syst.) which is compatible with the SM prediction, $A_{c,\eta}^{\ell}(\mathrm{SM}) = 0.0040 \pm 0.0002$ (scale) [185], where the uncertainty is estimated using the usual renormalisation and factorisation scale variations with $\mu_R = \mu_F = m_t = 173.34$ GeV[¶]. Differential measurements of $A_{c,\eta}^{\ell}$ in the boosted $t\bar{t}$ regime also give good agreement with the SM [187].

[§]For kinematical reasons, the asymmetries measured with the pseudorapidity instead of the rapidity of the particle are about 10–30% larger but otherwise retain the same features. Additionally, the measurement of the lepton pseudorapidity at particle detectors is direct, hence why it gets used over the rapidity to define the asymmetry.

[¶]Calculations of $A_{c,\eta}^{\ell}$ at $\sqrt{s} = 13$ TeV, while not reported in Ref. [185], have been specifically computed by the same authors in the phase-space of Ref. [187].

3.8 Top quark couplings

The previous sections have focused on $t\bar{t}$ production since it is one of the processes at the LHC with the largest cross–section and thus an excellent way to study the SM with high precision, and in an extended kinematic regime. However, production of a $t\bar{t}$ pair is also possible in association with other particles, albeit with smaller cross–sections and reduced statistics. At the LHC, the large amounts of data give access to the detailed examination of the couplings of the top quark with other particles such as the photon, the Z or the Higgs boson; with the coupling to the latter one being, without doubt, one of the cornerstones of modern physics. Furthermore, the role of the top quark in high-energy extensions of the SM is crucial since, at higher energies, couplings between lighter particles can usually be ignored. In the same way, the Higgs boson and its coupling to the top quark is also fundamental in both SM and BSM studies. This subsection discusses $t\bar{t}$ and single–top quark production in association with bosons (referred to as $t\bar{t}+X$ and tXq* processes, with $X=\gamma, Z, W, H$) and the relevant top–quark electroweak couplings.

Over the total LHC operation, processes involving a $t\bar{t}$ pair or a single–top in association with additional photons, Z, W or Higgs bosons have been observed. Figure 3.8 displays the measured cross–section of all the observed process at different COM energies. All SM predictions (grey bands) exhibit a good agreement with the observed values.

Due to the different couplings of the bosons to the other particles in the SM, in $t\bar{t}+X$ processes, the boson can be emitted as initial or final–state radiation. Table 3.1 displays example Feynman diagrams at LO for each of the four $t\bar{t}+X$ processes and whether the boson can be emitted as ISR and/or FSR. A single–top quark can also be produced in association with bosons. One of these processes, tW production, has already been presented in Section 3.2, while others are tZq, $t\gamma q$ and tHq production (via t–channel). This opens up the possibility to study top quark couplings to bosons, especially in the context on an EFT, where many of these processes are sensitive to particular sets of operators.

While $t\bar{t}+X$ production is mainly mediated by QCD interactions, the EW nature of associate single–top quark production leads to small QCD corrections, making these processes a somewhat "cleaner" way to study top quark couplings in the EW sector [194]. Figure 3.9 presents example LO Feynman diagrams for associated single–top quark production (via t–channel) and their three possible production mechanisms. In particular, the boson scattering mechanism gives access to other triple boson couplings, e.g. WWZ, WWH or $WW\gamma$. Moreover, the strong top quark polarisation in t–channel production, makes

*These single–top quark processes, which have been presented as tXq, stand for both single–top and single–antitop quark production (via t–channel).

these processes complementary probes to other top quark measurements. Finally, tXq processes are sensible to the sign of the t–X coupling in question, while $t\bar{t} + X$ are sensitive to the square of the coupling.

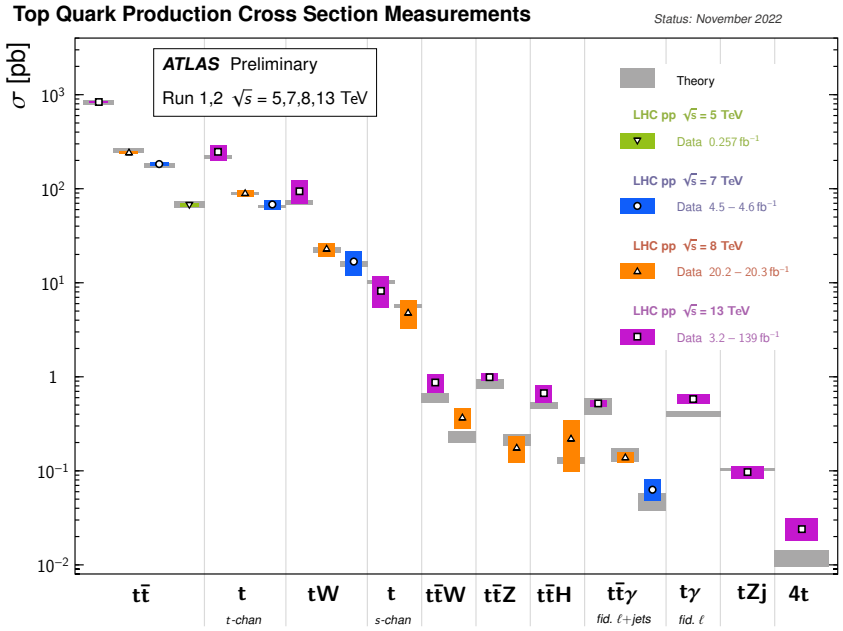
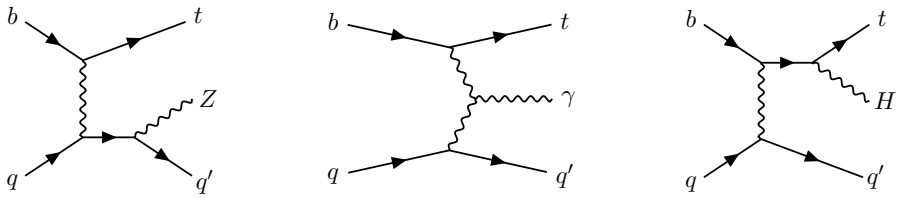


Figure 3.8: Summary of the top quark associated processes cross–sections measured by the ATLAS detector. Source: Ref. [158].



(a) tZq (b) $t\gamma q$ (c) tHq

Figure 3.9: Example Feynman diagrams at LO for associate single–top quark production via t–channel. The three different production mechanisms shown (from left to right: final–state quark radiation, boson scattering, or top quark radiation) are accessible by each of the tZq, $t\gamma q$ and tHq processes. Equivalent Feynman diagrams can be obtained if an antitop quark is produced instead of a top quark.

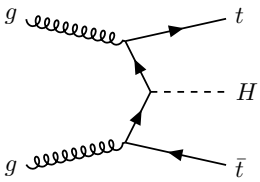
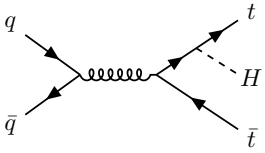
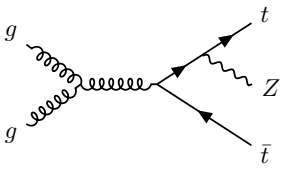
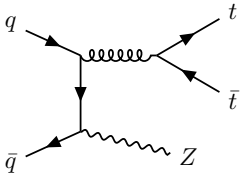
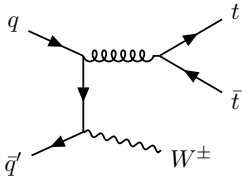
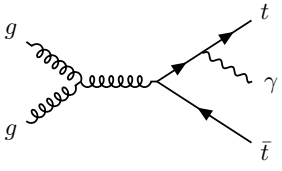
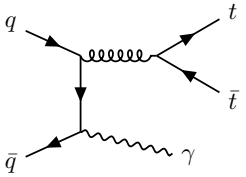
Process	gg initiated	$q\bar{q}$ initiated	ISR boson	FSR boson
$t\bar{t}H$			✗†	✓
$t\bar{t}Z$			✓	✓
$t\bar{t}W$	✗		✓	✗
$t\bar{t}\gamma$			✓	✓

Table 3.1: Example Feynman diagrams at LO for $t\bar{t}+X$ ($X = \gamma, Z, W, H$) production at the LHC. For $t\bar{t}W$ production, $\bar{q}'$ indicates a quark of different flavour from that of the other initial–state quark.

†The Higgs boson can be radiated as ISR from incoming light quarks (u, d, c, s, b) but the Yukawa couplings are very weak; hence the cross–section is low.

These associate production processes probe the different couplings of the top–quark to the SM bosons. Taking the general definitions of the SM couplings detailed throughout Chapter 1, these can be specified by substituting the fermion fields by top quark fields:

$$\text{top}-Z \text{ coupling (from Eq. 1.37)}: \quad \frac{g}{4\cos\theta_W} Z_\mu \bar{t}_L \gamma^\mu \left(1 - \tfrac{8}{3}\sin^2\theta_W - \gamma^5\right) t_L, \tag{3.4}$$

$$\text{top}-W \text{ coupling (from Eq. 1.33)}: \quad \frac{g}{2\sqrt{2}} W^-_\mu V_{tq} \bar{q}_L \gamma^\mu t_L, \tag{3.5}$$

$$\text{top}-\gamma \text{ coupling (from Eq. 1.10)}: \quad \tfrac{2}{3} e A_\mu \bar{t}\gamma^\mu t \tag{3.6}$$

$$\text{top}-H \text{ coupling (from Eq. 1.46)}: \quad \frac{G_t}{\sqrt{2}} H \bar{t} t, \tag{3.7}$$

$$\text{top}-g \text{ coupling (from Eq. 1.60)}: \quad g_S G^\mu_a \bar{t}^\alpha \gamma_\mu \left(\frac{\lambda^a}{2}\right)_{\alpha\beta} t^\beta. \tag{3.8}$$

3.8.1 top$-Z$ coupling

The top$-Z$ can be accessed directly via the $t\bar{t}Z$ and tZq production processes and is sensitive to the weak isospin.

- $t\bar{t}Z$ production: it is of particular interest in BSM searches since deviations of the top$-Z$ coupling could imply the existence of new effects in the EW symmetry breaking mechanism. Differential distributions also serve as discriminators between the different MC generator models and tune parameters used in them. $t\bar{t}Z$ production is also an irreducible background in (multi–lepton) SM measurements like $t\bar{t}H$, and also in BSM searches [195, 196]. The ATLAS and CMS Collaborations have performed inclusive and differential cross–section measurements of the $t\bar{t}Z$ process [197, 198], reaching about a 9% precision, and finding good agreement with the SM calculation ($\text{NLO}_{\text{QCD+EW}}$ + NNLL) [199, 200].

- tZq production: the ATLAS and CMS Collaborations have performed inclusive and differential cross–section measurements [201, 202], respectively, of the tZq process, reaching about a 11–15% precision, and finding good agreement with the SM calculation (NLO_{QCD}) [201].

3.8.2 top$-W$ coupling

The top$-W$ coupling can be accessed directly via single–top quark production processes and studying the top quark decay ($t \to Wb$). Additionally, the tWb vertex can be probed using the other associated single–top production processes.

- tW production: it is intriguing due to its interference with $t\bar{t}$ production[‡], sensitivity in BSM models, and contributions to many analyses as an irreducible background. The ATLAS and CMS Collaborations have performed inclusive and differential cross–section measurements of the tW process [204, 205], reaching about a 10% precision, and finding good agreement with the SM calculation (NNLO) [159].

3.8.3 top$-\gamma$ coupling

The top$-\gamma$ coupling can be accessed directly via the $t\bar{t}\gamma$ and $t\gamma q$ production processes and is sensitive to the top quark electric charge and magnetic dipole moments [206, 207]. What is more, the photon can be radiated from either ISR or FSR, or can produced via triple gauge coupling $WW\gamma$ in single–top quark production, making the studies of this coupling a challenging task.

- $t\bar{t}\gamma$ production: deviations in the kinematics of the radiated photon, such as its transverse momentum, could be conducive to new physics effects involving anomalous dipole moments of the top quark [208, 209]. EFT interpretations can provide a better understanding of $t\bar{t}$ pair production mechanism, focusing on the $t\bar{t}$ spin correlations and the charge asymmetry [210]. The ATLAS and CMS Collaborations have performed inclusive and differential cross–section measurements of the $t\bar{t}\gamma$ process [211–214], reaching about a 4–8% precision, and finding good agreement with the SM calculation ($\mathrm{NLO_{QCD}}$) [215, 216].

- $t\gamma q$ production: ATLAS has provided the first observation of this process which is compatible within 2σ of the SM prediction [217]. The measured cross–section is about 40% higher than the SM prediction ($\mathrm{NLO_{QCD}}$) [217], which falls in line with the $t\gamma q$ evidence CMS result [218].

3.8.4 top$-H$ coupling

The top$-H$ Yukawa coupling can be accessed directly via the $t\bar{t}H$ and tHq production processes. In addition, the study of this interaction can probe other Higgs properties such as the $\mathcal{CP}$ structure of the coupling. While the SM

[‡]Both tW and $t\bar{t}$ production can lead to the same $WbWb$ final–state. When considering NLO (and beyond) corrections to the tW process, special care has to be taken the double resonant contribution from $t\bar{t}$ production. This can be accomplished in two well–known ways: the diagram removal (DR) and diagram subtraction (DS) schemes [203]. In short, the DR approach removes all doubly resonant contributions from the tW cross–section calculation which, in turn, violates gauge invariance. On the other hand, the DS scheme introduces subtraction terms that work towards cancelling these $t\bar{t}$ resonances while preserving gauge invariance.

predicts a $\mathcal{CP}$–conserving ($\mathcal{CP}$–even) interaction, the measurement of a $\mathcal{CP}$–violating ($\mathcal{CP}$–odd) contribution to the coupling would be a sign of BSM physics and could give some insights into the matter–antimatter asymmetry in the early universe. Both the $t\bar{t}H$ and tHq processes are sensitive to the presence of this $\mathcal{CP}$–odd component and can be used to constrain it (see Chapter 6) [219–222].

- $t\bar{t}H$ production: this Higgs production mode can be used to target many of the Higgs decay channels. The ATLAS and CMS Collaborations have performed cross–section combinations of some of these analyses [223, 224], reaching about a 20–25% precision, and finding good agreement with the SM calculation ($\text{NLO}_{\text{QCD+EW}}$ + NNLL) [199]. Cross–section measurements in several bins of the Higgs boson p_{T} (also referred to as Simplified Template Cross–Sections) have also been performed [22].

- tHq production: this process complements $t\bar{t}H$ as it gives access to the determination of the sign of the Yukawa coupling [225, 226]. Moreover, the destructive interference between the diagrams with top$-H$ and $W-H$ couplings leads to a minimal cross–section in the SM [227]. Deviations in the relative coupling strengths could lead to a notable increase in the cross–section which can be parametrised and studied in an EFT framework [228]. Due to this small cross–section, only upper limits of tHq production have been extracted in some $t\bar{t}H$ analyses as both processes lead to similar final states and are difficult to disentangle [219, 229].

3.8.5 top$-g$ coupling

The top$-g$ coupling can be accessed directly via $t\bar{t}$ and $t\bar{t}+X$ production, although in the latter the sensitivity is typically smaller. $t\bar{t}$ production has been already discussed in Section 3.2. This has also been studied extensively at Tevatron [230].

3.9 $t\bar{t}W$ production

$t\bar{t}W$ production is one of the most unique signatures that are available to study at the LHC. State–of–the–art cross–section calculations are especially complex, as large corrections arise from higher powers of both the strong and weak couplings [199]. Thus, measurements of the $t\bar{t}W$ process represent a sensitive test of the predictions oft the QCD and the EW sector of the SM, as well as their interplay. Both the inclusive and differential cross–section measurements are very relevant, as they can provide indirect hints of BSM physics in scenarios where at least one of the SM couplings is modified [231]. Furthermore, it is an irreducible background of some rare SM processes such as $t\bar{t}H$ and four–top quark ($t\bar{t}t\bar{t}$) production [223, 224, 232, 233]; and it can also play a role in BSM

searches such as supersymmetric squark or gluino production or vector-like quarks [195, 196].

$t\bar{t}W$ production can be seen as a standard $q\bar{q} \to t\bar{t}$ LO production diagram, with an additional emission of a W boson in the initial state. However, the presence of this W boson gives it some unique properties compared to the other $t\bar{t}+X$ processes. For instance, a gg initial state is prohibited at LO (not available until NNLO) and, due to the $V-A$ structure of weak interactions, only left–handed quarks can populate the initial state (see Table 3.1). These properties limit the number of EFT couplings that impact the $t\bar{t}W$ cross–section, and alter the kinematics of the $t\bar{t}$ pair; making it an exceptional probe for both SM measurements and BSM searches.

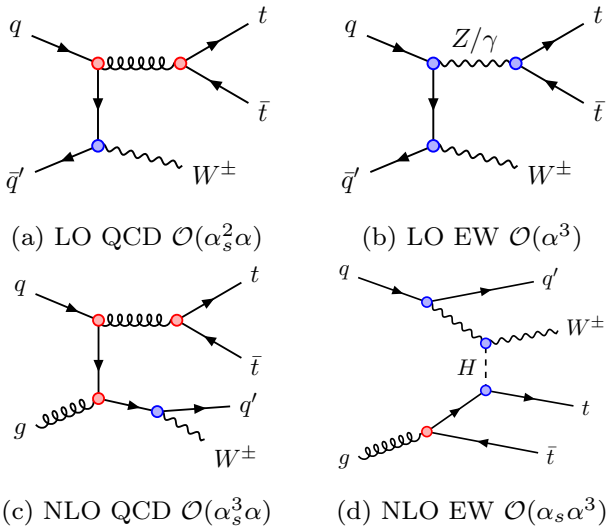
The complete NLO $t\bar{t}W$ cross–section can be decomposed into the following contributions:

$$\begin{aligned} \sigma^{NLO}_{QCD+EW} &= \sigma^{NLO}_{QCD} + \delta\sigma_{EW}, \text{ with} \\ \sigma^{NLO}_{QCD} &= \underbrace{\mathcal{O}(\alpha_s^2\alpha)}_{\text{LO QCD}} + \underbrace{\mathcal{O}(\alpha_s^3\alpha)}_{\text{NLO QCD}}, \\ \delta\sigma_{EW} &= \underbrace{\mathcal{O}(\alpha_s^2\alpha^2)}_{\text{NLO EW}} + \underbrace{\mathcal{O}(\alpha^3) + \mathcal{O}(\alpha_s\alpha^3)}_{\text{tree–level EW contributions}} + \underbrace{\mathcal{O}(\alpha^4)}_{\text{negligible}}. \end{aligned} \tag{3.9}$$

Figure 3.10 shows illustrative Feynman diagrams that contribute to $t\bar{t}W$ production at LO and NLO for both QCD and EW production.

The CMS and ATLAS Collaborations have performed inclusive and differential (in the case of ATLAS) cross–section measurements of the $t\bar{t}W$ process [234, 235] using the full Run 2 dataset, and finding some tension with the SM calculation ($\text{NLO}_{\text{QCD+EW}}$) [236]. While the exact reason cannot be pin–pointed yet, several advances in the precision of theoretical predictions aim to include more effects, naively thought to be negligible, which close the gap between experiment and theory. The most prominent example are the EW contributions to the $t\bar{t}W$ process. While these are now properly accounted for in all MC simulations, they were once thought to be small following a simple power count in the order of the coupling strength constants [237, 238]. These EW contributions include the $\mathcal{O}(\alpha_s\alpha^3)$ terms which are conducive to the opening of tW scattering diagrams (see Figure 3.10 (d)). Furthermore, they modify some kinematic distributions and lead to about a 10% increase of the inclusive cross–section prediction at NLO in QCD ($\mathcal{O}(\alpha_s^2\alpha) + \mathcal{O}(\alpha_s^3\alpha)$). On the other hand, the EW corrections to the LO process $\mathcal{O}(\alpha_s^2\alpha^2)$ (also referred to as "NLO EW" terms) lead to a small decrease (about -4%) of the total cross–section.

Increasing the precision of the theoretical calculation can be tackled in several ways, such as performing QCD calculations including the full off–shell effects [239], or merging NLO calculations with additional parton emissions with the FxFx or MePs@NLO algorithms as mentioned in Section 2.2.3. These



(a) LO QCD $\mathcal{O}(\alpha_s^2\alpha)$ (b) LO EW $\mathcal{O}(\alpha^3)$

(c) NLO QCD $\mathcal{O}(\alpha_s^3\alpha)$ (d) NLO EW $\mathcal{O}(\alpha_s\alpha^3)$

Figure 3.10: Examples of Feynman diagrams of $t\bar{t}W$ production at LO (a,b) and NLO (c,d) with one extra parton. The diagrams show QCD and EW $t\bar{t}W$ production. The red circles at the vertices correspond to QCD couplings and the blue circles correspond to EW couplings.

multi–leg calculations include $\alpha_s^4\alpha$ terms and their effect, in combination with the EW corrections, seem to be slightly favoured by data, suggesting the need for the full NNLO result [240]. In particular, they can increase the inclusive $t\bar{t}W$ cross–section and modify the jet multiplicity, as well as the kinematics of the jets and the $t\bar{t}$ system [241] (see Appendix B for further details). In fact, the latest reference cross–section prediction from Ref. [236] makes use of these multi–leg setups, and includes the full EW corrections ($\delta\sigma_{EW}$). This helps to reduce the discrepancy with the measured cross–sections, but is still at about a 20% deficit.

3.10 Top quark effective couplings

This section builds upon the introduction to the SMEFT framework in Section 1.6.1. It assesses the sensitivity of the previously discussed production processes to the effective couplings that are considered in Chapter 4. The total number of possible dimension–6 operators is too large for them to be included in a single fit (see Refs. [64, 242, 243]). Therefore, a subset of these operators is typically used. In particular, this section describes those that affect the couplings of the top and bottom quark couplings with vector, tensor or scalar Lorentz structures. Moreover, the Warsaw basis [70] is used and a set of recommendations by the LHC Top Working Group are followed [244].

3.10.1 Two–fermion operators

Two–fermion operators affect the EW couplings between the top and bottom quarks and bosons. They are defined as:

Left and right–handed couplings of the top and bottom quark to the Z

$$\underline{O^1_{\varphi Q}} \equiv \tfrac{1}{2}\left(\bar{q}\gamma^\mu q\right)\left(\varphi^\dagger i\overleftrightarrow{D}_\mu\varphi\right),$$
$$O^{(3)}_{\varphi Q} \equiv \tfrac{1}{2}\left(\bar{q}\tau^I\gamma^\mu q\right)\left(\varphi^\dagger i\overleftrightarrow{D}^I_\mu\varphi\right),$$
$$O^-_{\varphi Q} \equiv O^1_{\varphi Q} - O^{(3)}_{\varphi Q},$$
$$O_{\varphi u} \equiv \tfrac{1}{2}\left(\bar{u}\gamma^\mu u\right)\left(\varphi^\dagger i\overleftrightarrow{D}_\mu\varphi\right),$$
$$O_{\varphi d} \equiv \tfrac{1}{2}\left(\bar{d}\gamma^\mu d\right)\left(\varphi^\dagger i\overleftrightarrow{D}_\mu\varphi\right),$$

EW dipole operators ($\mathrm{x} = u, d$)

$$\underline{O_{\mathrm{x}B}} \equiv \left(\bar{q}\sigma^{\mu\nu}\mathrm{x}\right)\left(\epsilon\varphi^* B_{\mu\nu}\right),$$
$$O_{\mathrm{x}W} \equiv \left(\bar{q}\tau^I\sigma^{\mu\nu}\mathrm{x}\right)\left(\epsilon\varphi^* W^I_{\mu\nu}\right),$$
$$O_{\mathrm{x}Z} \equiv -\sin\theta_W O_{\mathrm{x}B} + \cos\theta_W O_{\mathrm{x}W},$$

Chromo–magnetic dipole operators

$$O_{uG} \equiv \tfrac{1}{2}\left(\bar{q}\sigma^{\mu\nu}\lambda^a u\right)\left(\epsilon\varphi^* G^a_{\mu\nu}\right),$$
$$O_{dG} \equiv \tfrac{1}{2}\left(\bar{q}\sigma^{\mu\nu}\lambda^a d\right)\left(\epsilon\varphi^* G^a_{\mu\nu}\right),$$

Top/bottom Yukawa

$$O_{u\varphi} \equiv \left(\bar{q}u\right)\left(\epsilon\varphi^*\varphi^\dagger\varphi\right),$$
$$O_{d\varphi} \equiv \left(\bar{q}d\right)\left(\epsilon\varphi^*\varphi^\dagger\varphi\right),$$

Charged current interaction

$$O_{\varphi ud} \equiv \tfrac{1}{2}\left(\bar{u}\gamma^\mu d\right)\left(\varphi^T\epsilon i D_\mu\varphi\right),$$

(3.10)

where q represents the left–handed quark $SU(2)_L$ doublet, u, d are the right–handed singlets[§], φ is the Higgs doublet, and $\epsilon \equiv i\tau^2$ is the antisymmetric $SU(2)_L$ tensor. The underlined operators $O^1_{\varphi Q}$, O_{uB} and O_{dB} are not used directly, rather a rotation of the basis is performed following Ref. [244], thus appearing in the linear combinations of $O^-_{\varphi Q}$ and $O_{\mathrm{x}Z}$.

The operators $O^1_{\varphi Q}$ and $O^{(3)}_{\varphi Q}$ modify the left–handed couplings of the Z bosons to up– and down–type quarks, while $O_{\varphi u}$ and $O_{\varphi d}$ modify the analogous right–handed couplings to the Z boson. The difference between the former two operators, defined as $O^-_{\varphi Q}$, probes the left–handed coupling to the top quark to the Z boson at leading order. The EW dipole operators give rise to tensor couplings of the photon and the Z boson to the up– and down–type quarks, respectively, and induce an anomalous dipole moment. In addition, the $O^{(3)}_{\varphi Q}$ and O_{uW} operators modify interactions between the W boson, up–type quarks and left–handed down–type quarks, while $O_{\varphi ud}$ and O_{dW} affect interactions between W bosons, up–type quarks and right–handed down–type quarks. Moreover, the chromo–magnetic dipole operators modify the coupling of quarks to gluons. Finally, the $O_{u\varphi}$ operator parametrises the top–Higgs coupling which affects many

[§]The three quark generations are included in the definition of q, u and d.

important observables such Higgs boson production and decay rates.

3.10.2 Four–fermion operators

Four–fermion operators, in particular those that couple two–light–two–heavy quarks, are especially relevant in $t\bar{t}$ production cross–section measurements where they can be precisely constrained [231,243,245]. However, some can also impact the $t\bar{t}+X$ associate production processes [231]. They are defined as:

LL and RR chiral structure	LR and RL chiral structure
$O_{Qq}^{1,8} \equiv \frac{1}{4}\left(\overline{Q}\gamma_\mu\lambda^a Q\right)\left(\bar{q}_i\gamma^\mu\lambda^a q_i\right),$	$O_{Qu}^{8} \equiv \frac{1}{4}\left(\overline{Q}\gamma^\mu\lambda^a Q\right)\left(\bar{u}_i\gamma_\mu\lambda^a u_i\right),$
$O_{Qq}^{3,8} \equiv \frac{1}{4}\left(\overline{Q}\gamma_\mu\lambda^a\tau^I Q\right)\left(\bar{q}_i\gamma^\mu\lambda^a\tau^I q_i\right),$	$O_{Qd}^{8} \equiv \frac{1}{4}\left(\overline{Q}\gamma^\mu\lambda^a Q\right)\left(\bar{d}_i\gamma_\mu\lambda^a d_i\right),$
$O_{tu}^{8} \equiv \frac{1}{4}\left(\bar{t}\gamma_\mu\lambda^a t\right)\left(\bar{u}_i\gamma^\mu\lambda^a u_i\right),$	$O_{tq}^{8} \equiv \frac{1}{4}\left(\bar{q}_i\gamma^\mu\lambda^a q_i\right)\left(\bar{t}\gamma_\mu\lambda^a t\right),$
$O_{td}^{8} \equiv \frac{1}{4}\left(\bar{t}\gamma_\mu\lambda^a t\right)\left(\bar{d}_i\gamma^\mu\lambda^a d_i\right),$	$O_{Qu}^{1} \equiv \left(\overline{Q}\gamma^\mu Q\right)\left(\bar{u}_i\gamma_\mu u_i\right),$
$O_{Qq}^{1,1} \equiv \left(\overline{Q}\gamma_\mu Q\right)\left(\bar{q}_i\gamma^\mu q_i\right),$	$O_{Qd}^{1} \equiv \left(\overline{Q}\gamma^\mu Q\right)\left(\bar{d}_i\gamma_\mu d_i\right),$
$O_{Qq}^{3,1} \equiv \left(\overline{Q}\gamma_\mu\tau^I Q\right)\left(\bar{q}_i\gamma^\mu\tau^I q_i\right),$	$O_{tq}^{1} \equiv \left(\bar{q}_i\gamma^\mu q_i\right)\left(\bar{t}\gamma_\mu t\right),$
$O_{tu}^{8} \equiv \frac{1}{4}\left(\bar{t}\gamma_\mu\lambda^a t\right)\left(\bar{u}_i\gamma^\mu\lambda^a u_i\right),$	
$O_{td}^{8} \equiv \frac{1}{4}\left(\bar{t}\gamma_\mu\lambda^a t\right)\left(\bar{d}_i\gamma^\mu\lambda^a d_i\right).$	

(3.11)

The four–fermion operators describe effective contact interactions between the incoming quarks and the top–antitop quark pair in $t\bar{t}$ and $t\bar{t}+X$ production. Hence, the third and first two quark families are separated in the operator definitions: $q_i \equiv (u^i, d^i)_L^{\mathrm{T}}$ and $u_i \equiv u_R^i$, $d_i \equiv d_R^i$ with $i = 1, 2$; $Q \equiv (t, b)_L^{\mathrm{T}}$ and $t \equiv t_R$ are the left–handed doublet and right–handed singlets, respectively. Additional four–fermion operators that couple four heavy quarks or leptons to quarks are not listed here, since they are not explored in this thesis, but their definitions may be found in Ref. [70].

3.10.3 Characterisation of the relevant Wilson coefficients in the top quark EW sector

Table 3.2 shows the sensitivity of the the $t\bar{t}+X$ and single–top associate production processes to the operator's Wilson coefficients, in the chosen basis, and with the appropriate substitutions, i.e. $u \to t$ and $d \to b$. Considering the listed processes in Section 3.8, the available observables are not sensitive enough to some of the operators in Eq. 3.10 and 3.11 in order to constrain them properly. Such is the case of the $O_{\varphi d}$ and O_{dZ} operators which modify down–type interactions with the Z boson. The bottom quark Yukawa operator $O_{b\varphi}$ is not included since none of the studied observables in this thesis are sensitive to it.

Parameter	$t\bar{t}$	$t\bar{t}H$	$t\bar{t}Z$	$t\bar{t}W$	$t\bar{t}\gamma$	tZq	$t\gamma q$	tHq	tW	single t	t decay
$C^{-}_{\varphi Q}$	–	–	Λ^{-2}	–	–	Λ^{-2}	–	–	–	–	–
$C^{(3)}_{\varphi Q}$	–	–	Λ^{-2}	–	–	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-2}
$C_{\varphi t}$	–	–	Λ^{-2}	–	–	Λ^{-2}	Λ^{-2}	–	–	–	–
C_{tW}	–	–	–	–	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-2}
C_{tZ}	–	–	Λ^{-2}	–	Λ^{-2}	Λ^{-2}	Λ^{-2}	–	–	–	–
$C_{t\varphi}$	–	Λ^{-2}	–	–	–	–	–	Λ^{-2}	–	–	–
$C_{\varphi tb}$	–	–	–	–	–	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-4}	Λ^{-4}	Λ^{-4}
C_{bW}	–	–	–	–	–	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-4}	Λ^{-4}	Λ^{-4}
C_{tG}	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-2}	–	–	–	Λ^{-2}	$[\Lambda^{-2}]$	$[\Lambda^{-2}]$
$C^{1,8}_{Qq}$	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-2}	–	–	–	–	–	–
$C^{3,8}_{Qq}$	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$	–	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$
C^{8}_{tu}, C^{8}_{td}	Λ^{-2}	Λ^{-2}	Λ^{-2}	–	Λ^{-2}	–	–	–	–	–	–
$C^{1,1}_{Qq}$	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$	–	–	–	–	–	–
$C^{3,1}_{Qq}$	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-2}	Λ^{-2}	Λ^{-2}	–	Λ^{-2}	Λ^{-2}
C^{1}_{tu}, C^{1}_{td}	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$	–	Λ^{-4} $[\Lambda^{-2}]$	–	–	–	–	–	–
C^{8}_{Qu}, C^{8}_{Qd}	Λ^{-2}	Λ^{-2}	Λ^{-2}	–	Λ^{-2}	–	–	–	–	–	–
C^{8}_{tq}	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-2}	Λ^{-2}	–	–	–	–	–	–
C^{1}_{Qu}, C^{1}_{Qd}	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$	–	Λ^{-4} $[\Lambda^{-2}]$	–	–	–	–	–	–
C^{1}_{tq}	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$	Λ^{-4} $[\Lambda^{-2}]$	–	–	–	–	–	–

Table 3.2: The sensitivity of each of $t\bar{t}$ and single–top associated production processes and their contributions via SM interference (Λ^{-2}) and via dimension–6 squared terms only (Λ^{-4}). The square brackets indicate that the Wilson coefficient contributes via SM interference but only at NLO in QCD. "single t" stands for t–channel and s–channel single–top quark production only. Information to complete this table is collected from Ref. [231] and the analysis described in Chapter 4.

4 Constraining EFT couplings in the top quark EW sector

This chapter describes a global EFT fit with the detailed processes and two–fermion operators of Table 3.2 using the latest LHC results, as well as legacy measurements from Tevatron and the LEP and SLC experiments. The fits are performed without (with) including the Λ^{-4} terms. These two fits are also referred to as the linear and quadratic fits. With this, results for six (eight) effective operator Wilson coefficients that characterise top quark couplings to the photon, the Z and Higgs bosons, and the charged tWb vertex are reported. These results have been published in Ref. [246]. Figure 4.1 summarises which of the Wilson coefficients enter each of the two fits as additional degrees of freedom. The coefficients that modify b quark couplings ($C_{\varphi tb}$ and C_{bW}) can only be constrained in the quadratic fit since their linear terms $X_i^{(1)}$ in Eq. 1.62 vanish in the limit $m_b \to 0$.

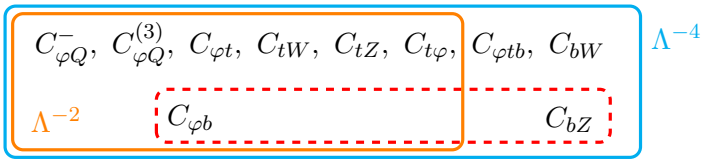


Figure 4.1: List of Wilson coefficients that enter these global EFT fits. There are a total of seven d.o.f. in the linear Λ^{-2} fit (orange) and ten d.o.f. in the quadratic $\Lambda^{-2} + \Lambda^{-4}$ fit (blue). The operators in the red box are included in the fits since EW–precision data are used, but are not reported as part of the final result due to the low sensitivity of the considered observables to them.

The choice of this basis omits several degrees of freedom: $\mathcal{CP}$–violating imaginary parts are not examined, and four–fermion operators are not considered as part of the main result. The impact of the inclusion in the fit of a subset of them, in particular the two–light–two–heavy quark operators (and also the O_{tG} operator), is studied as a robustness test of the fit. Other four–fermion operators that couple leptons to heavy quarks ($\ell\ell QQ$), or four heavy quarks ($QQQQ$), are not considered at all. Needless to say, these are studied in dedicated analyses [233, 247–249] and strong bounds are obtained from future lepton collider experiments [250, 251].

Section 4.1 describes the choice of measurements being used in the EFT fit. Section 4.2 details the fit and input EFT parameterisation setups that are used. The compatibility of the chosen observables with the SM is shown in Section 4.3, and their interplay with the Wilson coefficients is shown in Section 4.4. The results from the fit are given in Section 4.5 and a set of stress tests of the fit are presented in Section 4.6.

4.1 Measurements

Compared to previous EFT analysis in the top quark sector (Refs. [231, 249, 252–254] among others), the number of observables is extended. The latest differential measurements for $t\bar{t}Z$ and $t\bar{t}\gamma$ production are added alongside the other $t\bar{t}+X$ and single–top associate production processes. The precise measurements of the $Z \to b\bar{b}$ vertex by LEP and SLC [252] remain as powerful probes of some coefficients*, as well as some Tevatron legacy measurements. In addition, W boson helicity fractions are also included which puts the total number of observables n_{obs} to 30. These are all listed in Table 4.1.

In most of the cases only the largest available dataset is chosen, this corresponds to LHC data collected between 2015 and 2018 (Run 2) at $\sqrt{s} = 13$ TeV. This is done to avoid issues with statistical correlations between the different COM energy results; being the single–top s–channel measurement the exception, where both LHC and Tevatron data are included. However, since these are independent datasets, no correlations are expected. For results with various final–state topologies, depending on the top quark decay mode, measurements are chosen to be (nearly) orthogonal. Single–top measurements at $\sqrt{s} = 8$ TeV are used since inserting their 13 TeV counterparts as part of the fit has a negligible impact and ATLAS+CMS combinations are only available at 8 TeV. The latter is a convincing test that the choice of using a single COM result is adequate.

When available, experimental correlations among the observables are included in the fit. This is the case for the different bins of the boson's p_{T} in the $t\bar{t}Z$ and $t\bar{t}\gamma$ differential measurements, the two LEP/SLC observables and the three W helicity fractions. Nonetheless, similar sources of experimental and theoretical uncertainties can be shared by the other production processes considered in the fit. To account for these unknown correlations, an ansatz correlation matrix has been defined to further test the robustness of the fit (see Section 4.6.2).

A special note is necessary for some of the processes:

- $t\bar{t}\gamma$ production receives contributions from radiated photons from the top quarks, or its decay products, in $t\bar{t}$ events and from $tW\gamma$ production. The absence of the top–γ vertex dilutes the sensitivity of the EFT coefficients [255]. Photons emitted by decay products tend to be close to jets and are predominantly soft. In the ATLAS measurement [212], the fiducial region definition is such that these contributions are mitigated. Ultimately, the $tW\gamma$ and photons from top decay products are fully accounted for in the SM prediction, but are ignored in the EFT parameterisa-

*The observables R_b and A_{FBLR}^{bb}, measured in different experiments from LEP and SLC accelerators, are included since they are very sensitive to operators involving left–handed couplings of top and bottom quarks.

tion. In other words, the EFT parameterisation is obtained by simulating the $(2 \to 3)$ $pp \to t\bar{t}\gamma$ process and not the full $(2 \to 7)$ $pp \to e\nu_e b\mu\nu_\mu b\gamma$ process.

- $t\bar{t}H$ production is hard to distinguish from tHq production (due to the difficulty to disentangle the two processes experimentally). For this reason, the tHq contribution is included in the SM prediction and in the EFT parameterisation.

- $t\bar{t}W$ production receives large contributions from EW diagrams (see Section 3.9). The latter include tW scattering diagrams that are sensitive to the $O_{\varphi t}$ operator. Hence, the $t\bar{t}W$ EW diagram contributions are also parameterised.

Process	Observable	$\sqrt{s}$	$\int \mathcal{L}_{\text{int.}}$	Experiment	SM	Ref.
$pp \to t\bar{t}H + tHq$	σ	13 TeV	140 fb^{-1}	ATLAS	[227]	[256]
$pp \to t\bar{t}Z$	$d\sigma/dp_T^Z$ (7 bins)	13 TeV	140 fb^{-1}	ATLAS	[199]	[198]
$pp \to t\bar{t}\gamma + tW\gamma$	$d\sigma/dp_T^\gamma$ (11 bins)	13 TeV	140 fb^{-1}	ATLAS	[215, 216]	[212]
$pp \to tZq$	σ	13 TeV	77.4 fb^{-1}	CMS	[257]	[258]
$pp \to t\gamma q$	σ	13 TeV	36 fb^{-1}	CMS	[218]	[218]
$pp \to t\bar{t}W$	σ	13 TeV	36 fb^{-1}	CMS	[227, 237]	[259]
$pp \to t\bar{b}$ (s-ch)	σ	8 TeV	20 fb^{-1}	LHC	[260, 261]	[262]
$pp \to tW$	σ	8 TeV	20 fb^{-1}	LHC	[161]	[262]
$pp \to tq$ (t-ch)	σ	8 TeV	20 fb^{-1}	LHC	[260, 261]	[262]
$t \to Wb$	F_0, F_L	8 TeV	20 fb^{-1}	LHC	[170]	[168]
$p\bar{p} \to t\bar{b}$ (s-ch)	σ	1.96 TeV	9.7 fb^{-1}	Tevatron	[160]	[263]
$e^-e^+ \to b\bar{b}$	R_b , A_{FBLR}^{bb}	~ 91 GeV	202.1 pb^{-1}	LEP/SLC	–	[252]

Table 4.1: Measurements included in the EFT fit of the top quark electro-weak sector. For each measurement, the process, the observable, the center-of-mass energy, the integrated luminosity and the experiment/collider are given. The last two columns list the references for the SM predictions and measurements that are included in the fit. LHC refers to the combination of ATLAS and CMS measurements. In a similar way, Tevatron refers to the combination of CDF and D0 results, and LEP/SLC to several experiments from those two accelerators.

4.2 Fit setup

The dependence of the Wilson coefficients on the observables – the normalised $X_i^{(1)}/X_{\text{SM}}$ and $X_{ij}^{(2)}/X_{\text{SM}}$ terms from Eq. 1.62 – is parameterised using the MC generator aMC@NLO v.2.7.0 [98] using a fixed order calculation except for the top decay where an analytical NLO calculation is used [264]. Most of the predictions are derived at NLO. Out of all the available SMEFT models in

the market, SMEFT@NLO† [265] is the only one capable of such calculations and is therefore used. Other observables such as those from LEP/SLC or Tevatron are computed at LO. The dependence of the bottom quark operators (O_{bW}, $O_{\varphi tb}$, O_{bZ} and $O_{\varphi b}$) is computed at LO using the TEFT_EW model [266]. Both linear and quadratic terms are considered, as well as the interference between the different operators.

The SMEFT@NLO model uses the m_W, m_Z, G_F input electro-weak scheme while the TEFT_EW model uses the $\alpha_{EM}^{-1}(m_Z)$, m_Z, G_F scheme with

$$\alpha_{EM}^{-1}(m_Z) = 127.95, \qquad G_F = 1.16637 \times 10^{-5}\ \text{GeV}^{-2},$$

$$m_t = 173.3\ \text{GeV}, \quad m_H = 125\ \text{GeV}, \quad m_Z = 91.1876\ \text{GeV}, \quad m_W = 79.8244\ \text{GeV}.$$

The other fermion masses are taken to be zero. The values for the best SM predictions are taken from the references in Table 4.1 and in the cases where the prediction is superior to the one obtained with aMC@NLO (NLO_{QCD}), it is used to scale the X_{SM} term from Eq. 1.62, for each observable. The obtained parameterisations for the inclusive cross–section observables are given in Appendix A (Tables A.1, A.2 and A.3). The fit to data is performed using the open source HEPfit package [267,268]. HEPfit is a general tool designed to combine direct and indirect constraints in the SM [269], in EFT [250,270] or in particular BSM extensions [271,272]. Its flexibility allows to easily implement any BSM model or observable. The fit is performed as a Bayesian statistical analysis of the model, in which both theoretical and experimental uncertainties are included. HEPfit includes a Markov Chain Monte Carlo implementation provided by the Bayesian Analysis Toolkit [273] to explore the parameter space.

4.3 Compatibility with the SM predictions

The first step before performing any EFT fit is to compare the selected measurements with their SM prediction to assess their compatibility. All the previously detailed experimental correlations are included in this fit. The obtained chi–square value is $\chi^2_{\text{SM}}/(n_{\text{obs}} - 1) = 21.3/29$, corresponding to a p–value of 0.85. Overall, a good agreement of the SM predictions with the experimental results is seen. Figure 4.2 shows the individual contributions to the χ^2_{SM} for each of the observables. The largest discrepancies (about 1–2σ) come from some of the differential p_{T} bins from the $t\bar{t}Z$ and $t\bar{t}\gamma$ cross–sections results.

4.4 Interplay between measurements and operator coefficients

Figure 4.3 shows the sensitivity of the defined set of observables to each of the Wilson coefficients. Here, only inclusive calculations are shown. The sensitivity

†Either SMEFTatNLO_U2_2_U3_3_cG_4F_LO_UFO or v1.0.1 are used.

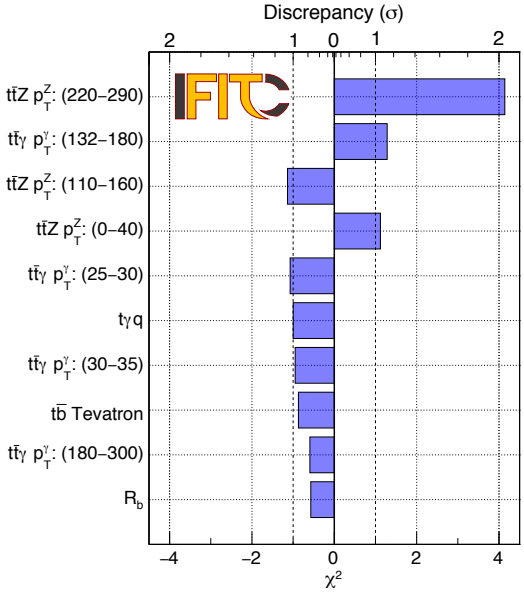


Figure 4.2: Contributions to the $\chi^2_{\rm SM}$ of some of the input measurements, together with their discrepancy with respect to the SM predictions. The full fit has 30 data points and a cumulative $\chi^2_{\rm SM} \sim 21.3$. The negative values indicate that the prediction is lower than the measurement. Both theoretical and experimental uncertainties are considered.

is defined as the variation of Eq. 1.62 with respect to the Wilson coefficient C_i considering both linear and quadratic terms. Likewise, the sensitivity of some rare processes that are not included in the fit is also shown (last 5 columns from the right). These correspond to those processes that have not been observed yet at the LHC (but experiments are close to do so) due to their low expected cross–sections. Nevertheless, their sensitivity is quite high for some operators, granting an exciting prospect for future EFT fits once observed. The $t\bar{t}$ cross–section measurements are not used in the fit since their sensitivity owes mainly to the C_{tG} and four–fermion operators which are not included in the baseline fit.

Figure 4.4 shows the linear plus quadratic terms for each of the $t\bar{t}Z$ and $t\bar{t}\gamma$ differential $p_{\rm T}$ bins for the most sensitive coefficients: C_{tZ} and $C_{\varphi t}$ for $t\bar{t}Z$, and C_{tZ} and C_{tW} for $t\bar{t}\gamma$. As expected, higher sensitivity to the Wilson coefficients comes from the high–$p_{\rm T}$ bins which further motivates the study of other processes differentially, just as larger datasets are recorded.

Figure 4.5 shows the individual 95% probability bounds on the eight operator coefficients considered in the fit with Λ^{-4} terms. The bounds from measurements in different processes, ordered from most to least constraining (going

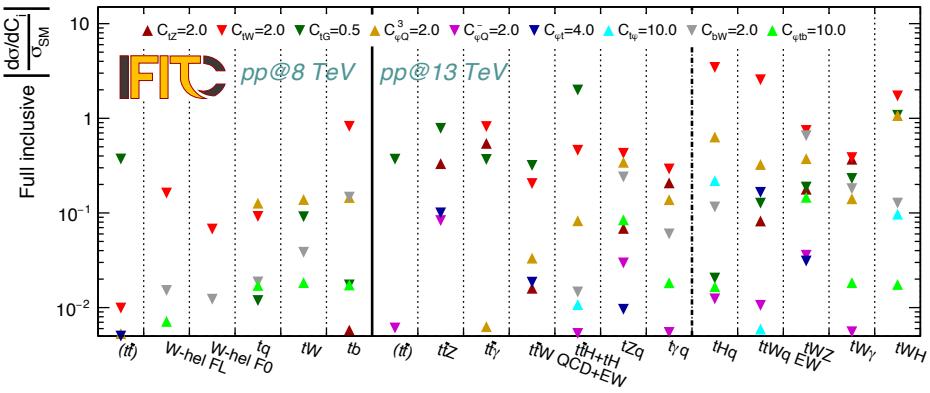
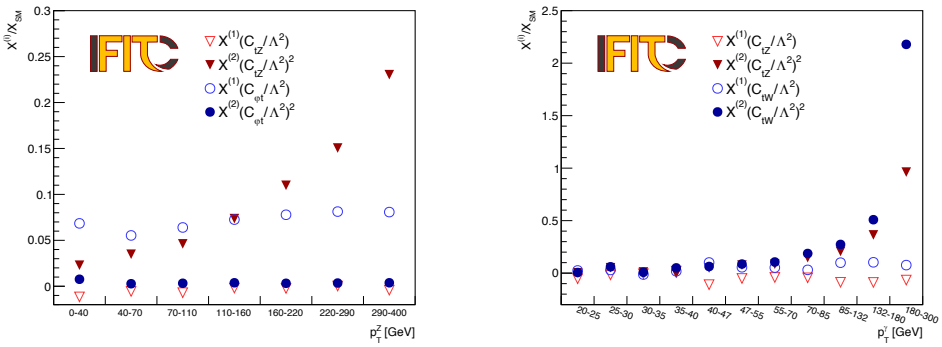


Figure 4.3: Sensitivity of LHC observables for the Wilson coefficients considered in this study. The solid black vertical line separates processes studied at 8 and 13 TeV, while the dashed one differentiates those for which measurements are or not already available.



(a) (b)

Figure 4.4: Linear (empty markers) and quadratic (filled markers) parameterisations for each of the bins in the differential (a) $t\bar{t}Z$ and (b) $t\bar{t}\gamma$ cross–sections. The most sensitive operators to these processes (C_{tZ}, $C_{\varphi t}$ and C_{tW}) are shown.

from left to right), are presented for each coefficient[‡]. For the $t\bar{t}Z$ and $t\bar{t}\gamma$ differential cross–sections, Figure 4.5 shows the improvement in sensitivity with respect to the inclusive measurement. This is depicted by the smaller darker bars, which correspond to the differential measurements for $t\bar{t}Z$ and $t\bar{t}\gamma$. Only the differential ones are included in the final fit setup.

[‡]Here, the smaller the bar is for a specific process, the more constraining power it has on the respective Wilson coefficient.

The power of the measurements to simultaneously constrain multiple Wilson coefficients is illustrated in Figure 4.6. The 95% probability bounds obtained in two-parameter fits to a single measurement are shown as areas of different colours. The red area shows the result of the global fit (including Λ^{-4} terms), where the result is marginalised over the remaining operator coefficients.

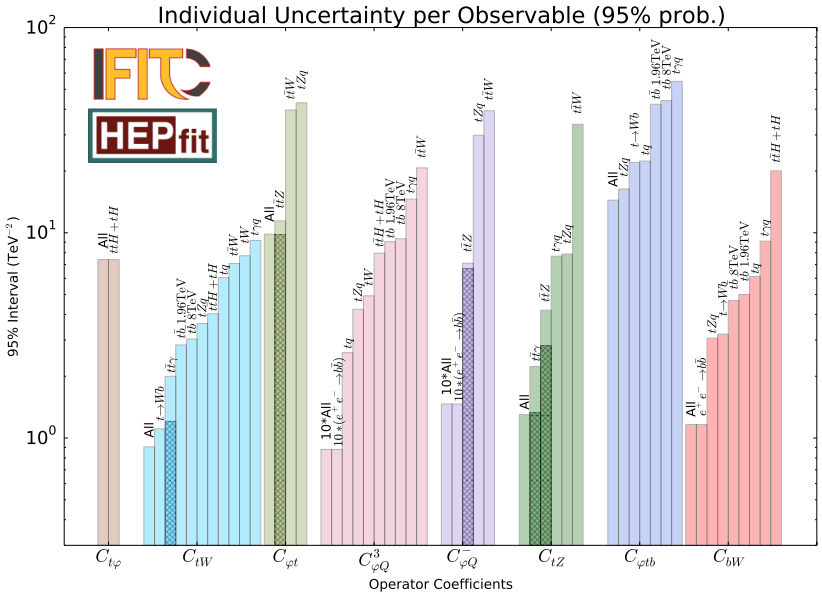


Figure 4.5: Individual constraints on the eight Wilson coefficient resulting from measurements in different processes. The y–axis corresponds to the full width of the 95% probability interval for fits including Λ^{-4} terms. The dark shading indicates the bound from the $t\bar{t}\gamma$ and $t\bar{t}Z$ differential cross–sections. Besides QCD production, in the case of $t\bar{t}W$ process, the contribution of the electro-weak $t\bar{t}Wq$ production is also included.

From these last two figures, the following can be extracted:

- $C_{t\varphi}$: Despite the large number of considered observables, $C_{t\varphi}$ is only constrained by the $t\bar{t}H$ cross–section measurement. Other fits including additional Higgs data can improve the constraints from the loop diagrams in $gg \to H$ production and the $H \to \gamma\gamma$ decay [274–278].

- $C_{\varphi t}$ and C_{tZ}: The $t\bar{t}Z$ production cross–section measurement is one of the most powerful observables for these two coefficients. For C_{tZ}, $t\bar{t}\gamma$ production provides an even more stringent constraint and, for both operators, the improvement due to the use of the differential measurements is significant.

- $C_{\varphi Q}^{(3)}$, $C_{\varphi Q}^{-}$ and C_{bW}: The constraints coming from the Z–pole data collected by the LEP/SLC experiments remain unrivalled for these coefficients. Their bounds are multiplied by a factor ten in Figure 4.5 to facilitate their comparison to the other observables. Withal, these only constrain linear combinations as seen by the thin orange lines in Figure 4.6 (left panels and bottom right panel). This degeneracy is disentangled when the $t\bar{t}Z$, tZq and single–top t–channel measurements are introduced.

- C_{tW}: The primary constraint on this operator arises from the top quark decay data, specifically the W boson helicity fractions. It is one of the coefficients that exhibits significant improvement when incorporating the $t\bar{t}\gamma$ differential cross–section measurement. In fact, the $t\bar{t}\gamma$ measurement provides complementary constraints on the linear combination of the two EW dipole operators (top and middle right panels of Figure 4.6).

- $C_{\varphi tb}$: Constraining this operator is difficult, with the recent tZq measurements providing a similar sensitivity in comparison to the W helicity fractions one. What is more, the tZq measurement is able to disentangle some directions in the $C_{\varphi tb}$– C_{tW} and $C_{\varphi tb}$– $C_{\varphi Q}^{(3)}$ planes (middle and bottom right panels of Figure 4.6).

4.5 Bounds of a global EFt fit

The main results of this study are the 68% and 95% probability intervals for the Wilson coefficients that modify the top quark electroweak couplings. Two sets of results are given for global fits with linear and quadratic terms, as shown in Figure 4.7, in orange and blue respectively. The results of the tests of the robustness of the fit, which are detailed in the next section (Section 4.6), are also given. The dotted light brown lines present the limits from a fit where the basis has been expanded to include the C_{tG} and seven four–fermion operators. Analogously, the dotted dark brown line presents the results including correlations between all observables in the baseline fit. Finally, the red lines show the envelope of all the stress tests that have been performed.

The limits coming from the linear and quadratic fits are quite similar for each of the Wilson coefficients. The main exception are the constraints on the C_{tZ} operator, that are reduced by a factor two when considering the Λ^{-4} terms. This improvement is due to the suppression of the linear term for the $t\bar{t}Z$ and $t\bar{t}\gamma$ processes [266] which can visually be seen in Figure 4.4 as the impact of the $X_i^{(2)}$ term rapidly increases as a function of the boson p_{T} while the $X_i^{(1)}$ term remains constant throughout. $C_{\varphi t}$ is slightly shifted from zero, mainly driven by the measured sixth p_{T} bin in $t\bar{t}Z$ production being lower than the SM prediction (see Figure 10 of Ref. [198]). $C_{\varphi Q}^{(3)}$ and $C_{\varphi Q}^{-}$ are also shifted in

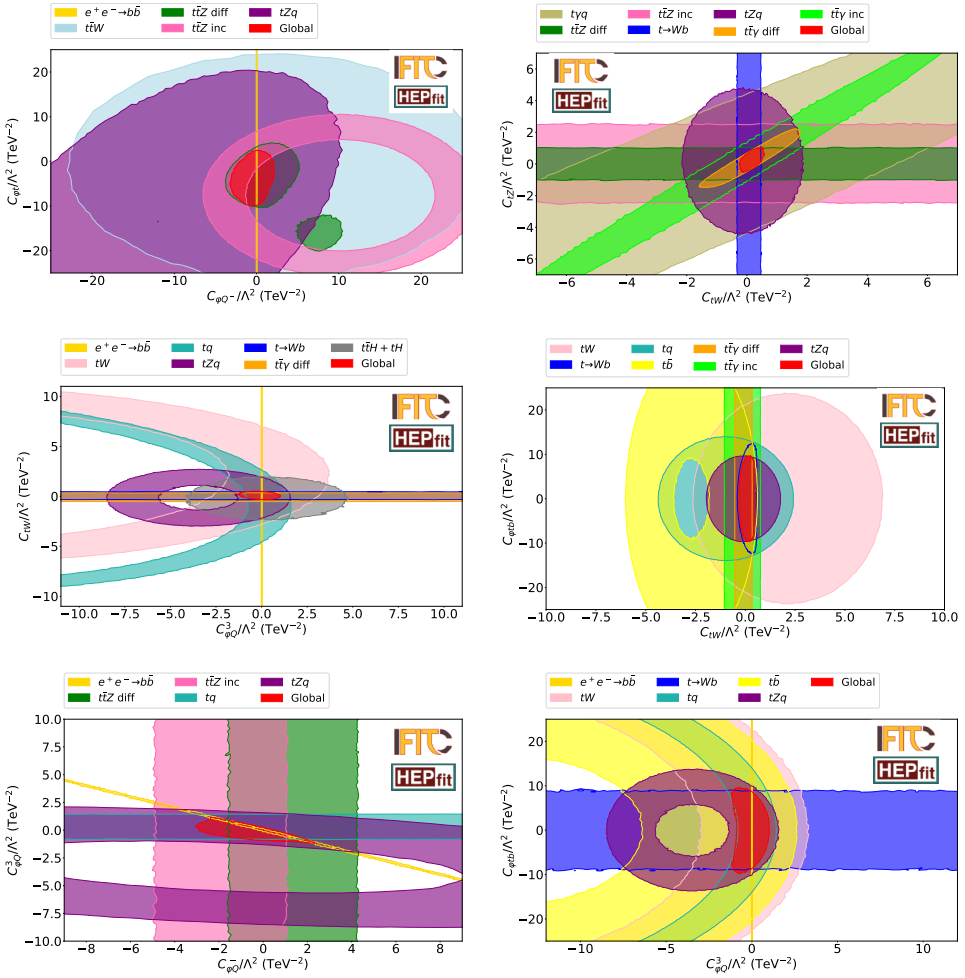


Figure 4.6: Two-dimensional 95% probability bounds on pairs of Wilson coefficients, $C^-_{\varphi Q}$ and $C_{\varphi t}$ in the upper leftmost panel, C_{tW} and C_{tZ} in the upper rightmost panel, $C^3_{\varphi Q}$ and C_{tW} in the middle leftmost panel, C_{tW} and $C_{\varphi tb}$ in the middle rightmost panel, $C^3_{\varphi Q}$ and $C^-_{\varphi Q}$ in the lower leftmost panel, and $C_{\varphi tb}$ and $C^3_{\varphi Q}$ in the lower rightmost panel. Bounds are presented for two-parameter fits to the most constraining measurements. The global fit results, marginalising over all other Wilson coefficients, are also shown (red area). All these fits include Λ^{-4} terms. Besides QCD production, in the case of $t\bar{t}W$ process, the contribution of the electro-weak $t\bar{t}Wq$ production is also included.

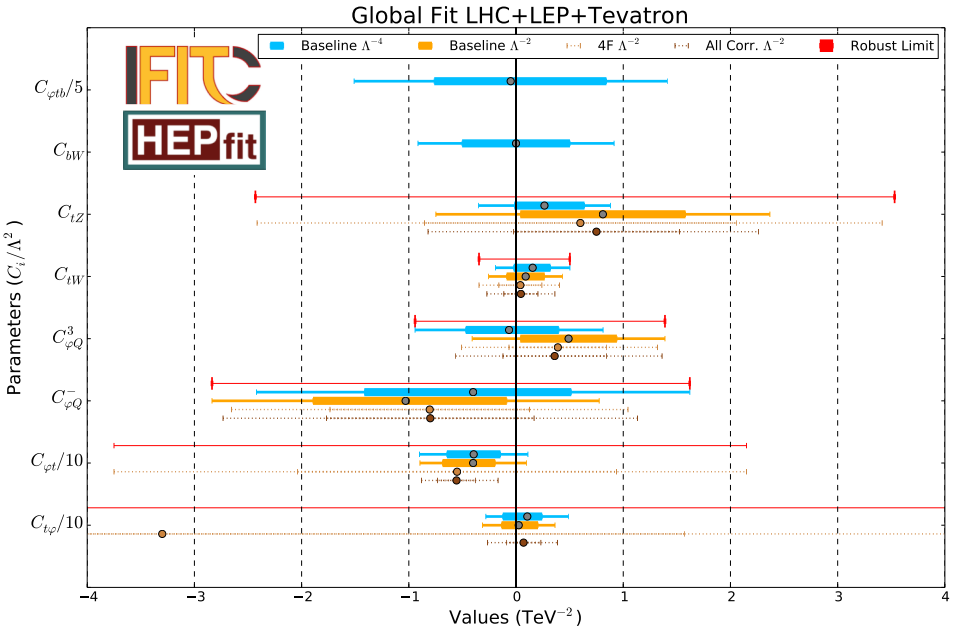


Figure 4.7: Results of the global fits with $O(\Lambda^{-2})$ (in orange) and $O(\Lambda^{-4})$ (in blue) terms. The two thin lines below each of them correspond to additional fits performed to test the robustness of the results: to account for the effects of the inclusion of additional operators and for the correlations between all the different measurements (as described in Section 4.6). The red markers correspond to the envelope of those additional fits plus another one that accounts for the theoretical uncertainties on the parameterisations.

opposite directions, steered by the interplay of the LEP/SLC observables with these coefficients (see bottom left panel of Figure 4.6). The dependence of the fit result on terms proportional to Λ^{-4} is significantly reduced compared to previous analyses [250]. This is expected as the larger datasets collected at the LHC, which are conducive to higher precision measurements, increase the dominance of the linear terms in constraining the EFT parameter space.

The obtained limits (95% probability intervals) for the fitted operator coefficients are presented in Table 4.2. For completeness, the limits obtained when fitting only one parameter and fixing the others to zero (individual fits) are also shown in this table. The obtained bounds for $C_{\varphi Q}^{(3)}$ and $C_{\varphi Q}^{-}$ in the individual fits are much better than for the global fits. For the rest of coefficients, the degradation of the limits in the global fit is minor.

C/Λ^2	Linear (95% probability)		Lin.+Quad. (95% probability)		(95% probability)
(TeV^{-2})	Individual	Global–Baseline	Individual	Global–Baseline	Global–Robust
$C_{t\varphi}$	[-3.17, 3.47]	[-3.13, 3.63]	[-3.05, 4.05]	[-2.82, 4.92]	[-121.82, 62.82]
$C_{\varphi Q}^{-}$	[-0.038, 0.079]	[-2.84, 0.78]	[-0.038, 0.079]	[-2.42, 1.62]	[-2.84, 1.62]
$C_{\varphi Q}^{(3)}$	[-0.019, 0.040]	[-0.41, 1.39]	[-0.019, 0.040]	[-0.94, 0.81]	[-0.94, 1.39]
$C_{\varphi t}$	[-6.6, 1.8]	[-8.96, 0.96]	[-8.6, 1.5]	[-9.01, 1.11]	[-37.50, 21.50]
C_{tW}	[-0.30, 0.38]	[-0.26, 0.44]	[-0.28, 0.32]	[-0.19, 0.50]	[-0.35, 0.50]
C_{tZ}	[-0.82, 2.21]	[-0.75, 2.37]	[-0.39, 0.57]	[-0.35, 0.88]	[-2.43, 3.53]
$C_{\varphi tb}$	–	–	[-6.61, 6.71]	[-7.55, 7.05]	–
C_{bW}	–	–	[-0.47, 0.47]	[-0.91, 0.91]	–

Table 4.2: Allowed ranges of the Wilson coefficients with a probability of 95% expressed in TeV^{-2} including only linear terms or linear and quadratic terms. The results of five fits, from left to right, are shown: individual with linear terms, global baseline with linear terms, individual with linear and quadratic terms, global baseline with linear and quadratic terms and global robust limits. The robust result accounts for the effects of the correlations between the observables, the inclusion of further operators and the theoretical uncertainties on the parameterisations.

4.6 Testing the robustness of the fit

A total of three additional tests are performed to check the validity of the obtained results. The results are stressed under the inclusion of further degrees of freedom coming from additional SMEFT Wilson coefficients, the inclusion of correlations among the different observables, and the effect of missing higher–orders (uncertainty on the scale choice) for the parameterisation between the observables and the Wilson coefficients.

4.6.1 Limitations of the basis

The set of chosen operators in the baseline fit omits the two–fermion chromo–magnetic dipole operator O_{tG} and the four–fermion operators. These are well constrained from $t\bar{t}$ cross–section measurements at the LHC and Tevatron but other $t\bar{t}+X$ processes can also have some sensitivity to them, as shown in Table 3.2. The impact of these additional d.o.f. is assessed by expanding the basis and including the following operator coefficients into the fit:

$$C_{tG},\ C^{8}_{tu},\ C^{8}_{td},\ C^{1,8}_{Qq},\ C^{3,8}_{Qq},\ C^{8}_{Qu},\ C^{8}_{Qd},\ C^{8}_{tq}. \tag{4.1}$$

They are parameterised at LO in QCD and quadratic terms are ignored since they do not contribute. With these additional floating parameters, the number of observables has to inevitably be increased to maintain enough freedom in the fit and constrain the new coefficients. The $t\bar{t}$ cross–section measurements coming from Tevatron [279] and LHC at $\sqrt{s} = 13$ TeV [280] are added. The target of this test is to quantify the degradation of the fit, not to find competitive limits on these operator coefficients. In order to do so, additional observables would have to be included such as the differential $pp \to t\bar{t}$, $t\bar{t}$ charge asymmetry [187, 281] or $t\bar{t}$ energy asymmetry measurements [282, 283].

Looking at the dotted light brown limits from Figure 4.7, this extension predominantly degrades the C_{tZ} limits (by a factor two) and the $C_{\varphi t}$ and $C_{t\varphi}$ limits in a much larger way. The main constraints for these coefficients come from $t\bar{t}+X$ measurements (see Figure 4.5) which are also sensitive to the new d.o.f. (see Table 3.2). As a result, it is expected for the limits on these operator coefficients to be degraded the most.

4.6.2 Correlations between measurements

Even though the selection of the observables aims to minimise the statistical correlations between them, correlations between experimental and theoretical systematics could be sizeable. In spite of the lack of these correlations from the experimental results, variations from zero to the ansatz described below are performed.

Experimental uncertainties reported by each of the measurements can share similarities that are classified in three categories: *statistical* when the same datasets are fully or partially used in both measurements, and *systematic* which are further distinguished into *mildly* or *highly correlated*. This distinction depends on the pair of considered observables and the examined systematic uncertainty. The combined correlation between two observables i and j is then given by

$$\bar{\rho}^{\alpha}_{ij} = \frac{\sum_{\alpha} \rho_{ij} u^{\alpha}_{i} u^{\alpha}_{j}}{\sqrt{\left(\sum_{\alpha} u^{\alpha}_{i} u^{\alpha}_{i}\right)\left(\sum_{\alpha} u^{\alpha}_{j} u^{\alpha}_{j}\right)}}, \tag{4.2}$$

where u_i^α is the uncertainty of the observable i, α runs over the three categories of uncertainties and ρ_{ij} is an ansatz which varies from 0 to 0.5 (depending on the source of the uncertainty).

Theoretical uncertainty correlations can be estimated considering similarities in the used PDF sets and in some production processes. For the observables measured at the LHC, the ansatz for this correlation is: 100% between the different bins of $t\bar{t}Z$ and $t\bar{t}\gamma$ differential cross–sections, and either 20% or 50% between the other 13 TeV measurements and some 13 TeV–8 TeV and 8 TeV–8 TeV pairs. The correlations of a given observable with F_L are of opposite sign to those with F_0, and a -100% corelation is assigned between the two W helicity fractions.

Looking at the dotted dark brown limits from Figure 4.7, the impact of the ansatz of the correlations is mild. Some small shifts are seen, for instance in the $C_{\varphi Q}^{(3)}$ and $C_{\varphi Q}^{-}$ coefficients of $\mathcal{O}(1\ \text{TeV}^{-2})$ and in the $C_{t\varphi}$ coefficient of $\mathcal{O}(10\ \text{TeV}^{-2})$. This shift in $C_{\varphi t}$ increases the tension with the SM but the impact on the p–value of the fit is small.

4.6.3 Theory uncertainties on the parameterisation

Missing higher–order corrections in α_s in the linear and quadratic terms can be estimated by considering the envelope of the scale variations provided by the MC generator. These are displayed in Tables A.1, A.2 and A.3. This effect does not have an important impact on the fit (maximum of 5%) given that most predictions are computed at NLO.

4.7 Summary and outlook

The work described in this part of thesis has presented a global EFT fit to the top EW couplings using the latest available LHC data at the time as well as some complementary Tevatron and LEP/SLC measurements. The baseline set of 30 observables is found to be compatible with the SM with a p–value close to unity. The global 95% intervals for the eight Wilson coefficients that modify the top quark electroweak couplings vary from ± 0.35 to $\pm 8\ \text{TeV}^{-2}$ and include the SM expectation ($C_i = 0$).

Results are reported for two fits (following the LHC Top Working Group recommendations [244]): one that includes only the parameterisations of the linear terms (Λ^{-2}) and another that also includes the quadratic terms ($\Lambda^{-2} + \Lambda^{-4}$). Differences between both fits is found to be small with the exception of a factor two improvement on C_{tZ} upon including the quadratic parameterisations. This is due to the suppression of the linear term in the most sensitive processes, $t\bar{t}Z$ and $t\bar{t}\gamma$, to this operator coefficient [266]. The addition of the recent differential measurements for these two processes have provided a signi-

ficant improvement on the constraints of several coefficients. This demonstrates the crucial role of high–precision measurements coming from high–energy collider experiments in future global EFT fits.

Each individual observable corresponds to the most precise available measurement in an effort to reduce unknown correlations with other less precise ones or other COM results. Further checks confirming the agnostic nature of the fit to the inclusion of several measurements at different COM have been performed, thus supporting this choice. Aside from the published experimental correlations for some of the included observables, a complete covariance matrix with plausible ansatz values is used to estimate the residual impact of correlations. The fit is found to be relatively robust to the assumed correlations, with minor changes in the results for three of the operator coefficients. The fact that single measurements are used for each observable limits the dependence with the correlations in comparison with previous results that report larger ones [284]. This reduction of the dataset will no longer be necessary once legacy combinations between ATLAS and CMS become available.

An extension of the baseline basis is also used as a test of the robustness of the fit. This extension includes operator coefficients to which the $t\bar{t}$ and $t\bar{t}+X$ cross–section measurements are sensitive. The C_{tZ}, $C_{\varphi t}$ and $C_{t\varphi}$ coefficients show the largest sensitivity to this extension as a result of the dependence of the main processes that constrain them (in the baseline fit) with the new operator coefficients. Finally, the result is stressed against the inclusion of higher–order corrections in α_s but is found to be very robust against it.

The disclosed results from this fit are in agreement with other, more comprehensive, analysis that expand on Higgs boson and top quark production data as well as including EW precision measurements [277, 278], once the differences in the operator basis and treatment of Λ^{-4} terms are accounted for. The main differences are driven by the $t\bar{t}Z$ and $t\bar{t}\gamma$ differential measurements which set stronger individual bounds on the C_{tZ} operator in this result. Others like $C_{\varphi t}$ and $C_{t\varphi}$ have tighter individual bounds in Refs. [277, 278] by virtue of the additional Higgs and EW data. While these differences are realised for the individual bounds, global bounds on top EW couplings remain dominated by top physics data. The interplay between Higgs, EW and top physics (and flavour [248]) definitely merits further exploration.

To all intents and purposes, the search for new observables that are able to probe certain operators or specific blind directions in the EFT parameter space is essential along with high–precision measurements. For this, the $t\bar{t}W$ process is well–suited. As shown in Table 3.2, the $t\bar{t}W$ process presents (almost) no dependence to the two–fermion operators and a limited one to the four–fermion operators. The W boson can only be emitted as ISR (see Table 3.1), making the final state identical to that of $t\bar{t}$ production which is (also) not sensitive to two–fermion operators other than O_{tG}. Moreover, it selects a left–handed

chiral structure of the initial quarks so ultimately only a selection of the four–fermion operators in Eq. 3.11 contribute at LO. Certain variables, like the charge asymmetries presented in Section 3.7.1, can turn into powerful probes of top quark properties and the SM by extension. These asymmetries can naturally be used to constrain the EFT parameter space, especially in the four–fermion sector [187, 283]. To this extent, the lepton charge asymmetry defined in Eq. 3.3 for the $t\bar{t}W$ process, that is predicted to be large in the SM [239, 285], is an ideal candidate to explore blind directions in the Wilson coefficient parameter space and constrain new physics effects in a model independent way. This is further discussed in Section 5.10.1.

4.7.1 Extrapolation of the EFT global fit to HL–LHC and future colliders

In the global strategic plans for particle physics, such as the *2020 update of the European strategy* [286], or the *Energy Frontier Topical Group on EW Physics* of the 2021 US Snowmass study [287], there are several proposals for new collider facilities that are expected to improve the precision of the measurements in the top and bottom quark sectors. The current LHC program, which includes the high-luminosity upgrade, can provide precise measurements for rare top quark production processes and can extend the classical hadron–collider top production processes well into the boosted regime [288]. An electron–positron collider, operated above the top-quark pair production threshold, is highly advantageous as it enables an very precise characterization of the top quark electroweak interactions [289]. Lastly, there is potential for a new hadron or lepton collider, which could operate at energy levels of 10 TeV, to open up a new kinematic regime and greatly enhance the sensitivity to four quark operators [290].

Taking the work presented in this chapter as a baseline, the impact of the HL–LHC, and several future electron–positron collider scenarios on the top and bottom quark sectors of the SMEFT are studied. The basis is extended to include the four–quark operators: $O_{tu}^{8}, O_{td}^{8}, O_{Qu}^{8}, O_{Qd}^{8}, O_{Qq}^{1,8}, O_{Qq}^{3,8}$ and O_{tq}^{8}.

For HL–LHC, the observable list from Table 4.1 is increased with the addition of differential $t\bar{t}$ production measurements from CMS and ATLAS, as shown in Table 4.3. The addition of these measurements, exploiting the $t\bar{t}$ boosted regime at large invariant mass, plays an important role to constrain these the four–fermion operators. To take maximal advantage of this potential, the range of the projections is extended further into the high–$m_{t\bar{t}}$ tail than the current Run 2 measurements. Projections for rare top quark production processes are modelled on the S2 scenario used to predict the precision of Higgs coupling measurements [291]. This scenario envisages that the statistical uncertainties scale as $\mathcal{L}_{\text{int.}}^{-1/2}$, and experimental uncertainties are reduced by a factor two. Theory and modelling uncertainties are also divided by two, which

assumes NNLO calculations will be achieved for rare associated production processes, and that MC modelling can significantly be improved in the next decade.

Process	Observable	$\sqrt{s}$	$\int \mathcal{L}_{\text{int.}}$	Experiment	SM	Ref.
$pp \to t\bar{t}$	$d\sigma/dm_{t\bar{t}}$ (15+3 bins)	13 TeV	140 fb^{-1}	CMS	[157]	[292]
$pp \to t\bar{t}$	$dA_C/dm_{t\bar{t}}$ (4+2 bins)	13 TeV	140 fb^{-1}	ATLAS	[157]	[293]
+ all from Table 4.1						

Table 4.3: Measurements included in the extended EFT fit of the top quark electroweak sector. For each measurement, the process, the observable, the center-of-mass energy, the integrated luminosity and the experiment/collider are given. The last two columns list the references for the predictions and measurements that are included in the fit. The numbers in brackets indicate the number of bins that are used in the fit. To exploit the potential of the HL–LHC $t\bar{t}$ production at large invariant mass, the last bin of the differential results is split into additional bins.

The 95% probability bounds from a fit, including only Λ^{-2} terms, to the current data are shown as the dark red bars in Figure 4.8. The light red bars present the extrapolation to the complete HL–LHC program, with an integrated luminosity of 3000 fb^{-1}.

Across the board, the HL–LHC program is expected to improve the bounds by a factor of two to four with respect to the current Run 2 limits, both for individual bounds and global fit results. Exceptions are the individual bounds on $C^{-}_{\varphi Q}$ and $C^{3}_{\varphi Q}$, that continue to strongly depend on the LEP/SLC legacy measurements. Marginalised four–quark operator limits remain an order of magnitude worse than the individual bounds after the HL-LHC. This is due to unresolved correlations between the coefficients, also observed in Refs. [231,254].

This S2 scenario is limited by theory and modelling uncertainties. Therefore an improvement on these fixed–order calculations for the relevant rare top quark production processes will directly improve the sensitivity. This will likely require, however, calculations at the N^{3}LO precision for $2 \to 3$ processes with top quarks in the final state. Being the LHC a top quark "factory", the increased rates of $t\bar{t}$ production at HL–LHC will allow to further exploit the boosted regime (perhaps even more than what is considered here) and improve the sensitivity to the four–quark operators. Moreover, this fit only considers linear terms; the inclusion of quadratic terms is expected to provide stricter limits [278].

Other four–fermion operators, involving two–quarks–two–leptons can also be studied at the LHC via the off–shell $pp \to t\bar{t}\ell^+\ell^-$ production [247, 294, 295]. What is more, future circular and linear electron–position colliders can also probe them with high precision via the $e^+e^- \to b\bar{b}$ and $e^+e^- \to t\bar{t}$ production

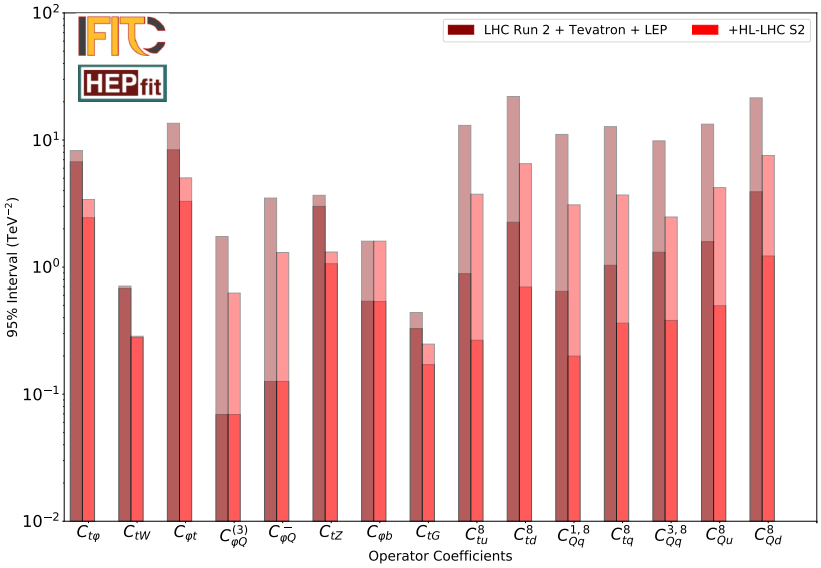


Figure 4.8: The 95% probability bounds on the Wilson coefficients for dimension-six operators that affect the top quark production and decay measurements listed in Table 4.3 after Run 2 of the LHC (in dark red) and prospects for the bounds expected after completion of the complete LHC program, including the high–luminosity stage (in light red). Only linear terms proportional to Λ^{-2} are taken into account in the dependence of the observables on the Wilson coefficients. The individual bounds obtained from a single-parameter fit are shown as solid bars, while the global or marginalised bounds obtained fitting all Wilson coefficients at once are indicated by the full bars (shaded region in each bar).

processes. The prospects of all these dimension–6 operators from simulations studies for ILC [296–298], CLIC [299, 300], FCC–*ee* [301] and CEPC [301] are also explored. All proposed energy stages and beam configurations are considered for each experiment. For $e^+e^- \to t\bar{t}$ production, a set of statistically optimal observables, defined at LO on the $e^+e^- \to t\bar{t} \to W^+bW^-\bar{b}$ differential distribution, are used [290]. For $e^+e^- \to b\bar{b}$ production, prospects for the R_b and A_{FBLR}^{bb} observables are used. With all, e^+e^- colliders can improve the bounds on the discussed two–fermion operators by up to two orders of magnitude. Moreover, the linear accelerators (ILC and CLIC) operating at COM energies above the $t\bar{t}$ production threshold, and even above the TeV scale, provide the tightest top quark sector SMEFT constraints on $e^+e^-t\bar{t}$ operators that can be achieved at any future collider (of the order of 10^{-3} TeV^{-2}). This does not include muon colliders, which have the potential to further improve these bounds. These studies are documented in Ref. [251].

5 Search of the leptonic charge asymmetry in $t\bar{t}W$ production in the 3ℓ channel

In $t\bar{t}W$ production, the relative dominance of the $q\bar{q}'$ initial state leads to a larger rapidity–based charge asymmetry with respect to $t\bar{t}$ production [239,285]. Furthermore, the ISR W boson serves as a polariser of the initial $q\bar{q}'$ state, and in turn, of the final $t\bar{t}$ state. This polarisation further enhances the asymmetry between the decay products of the top and antitop quarks*. The prospects for an experimental observation of these asymmetries are greatest in the case of the charged leptons originating from their top quark parent decays due to the precision with which lepton kinematics can be reconstructed, and the power with which reducible background processes can be suppressed. This leptonic charge asymmetry, A_c^l, based on the absolute pseudorapidity differences of the leptons, is defined as in Eq. 3.3.

Ref. [285] gives a comparison of the A_c^l for $t\bar{t}$ and $t\bar{t}W$ production using a NLO+PS matched calculation in an inclusive phase–space at $\sqrt{s} = 13$ TeV. The leptonic charge asymmetry for $t\bar{t}W$ (about -13%) is larger with respect to $t\bar{t}$ production (less than 1%) at the expense of a smaller cross section of the process. It is not only sensitive to BSM physics, such as axigluons and SMEFT scenarios corresponding to four-fermion operators [67, 285], but also has the unique potential of discriminating between new physics signals with different chiral structure that would have indistinguishable effects on cross–section observables. What is more, the A_c^l is insensitive to the $t\bar{t}W$ production rate, hence allowing the study of this process in an independent way.

This chapter presents a search for the leptonic charge asymmetry (in the following just denoted as "charge asymmetry") in $t\bar{t}W$ production using pp collision data at $\sqrt{s} = 13$ TeV in the three light lepton (3ℓ) channel with the full Run 2 dataset, corresponding to an integrated luminosity of 139 fb^{-1}. The used data and MC simulated event samples are discussed in Sections 5.1 and 5.2, respectively. The object selection criteria and further event selection requirements to build the analysis regions are given in Section 5.3 and Section 5.4, respectively. Section 5.5 describes the matching of final–state leptons to their top quark parent using machine learning (ML) techniques. Sources of experimental and theoretical systematic uncertainties are detailed in Section 5.6, while the profile likelihood fitting technique used for the A_c^l extraction is described in Section 5.7. The results obtained at reconstruction level are shown in Section 5.8 and those at particle level (after an unfolding procedure) in Section 5.9. This work has been published in Ref. [302]†.

*The polarisation of the top quarks in the reference frame of the incoming quarks is of about 65%.

†Additional *internal* supporting documentation may be found in Ref. [303].

5.1 Data event samples

This analysis uses data coming from pp collisions at $\sqrt{s} = 13$ TeV delivered by the LHC and recorded by ATLAS during the 2015–2018 data–taking period. The bunch spacing for all this data is 25 ns with variable pile–up conditions as described in Section 2.2.5. After a set of data quality criteria are applied [304], the resulting data sample corresponds to a total integrated luminosity of 139 fb^{-1} [77, 305].

5.2 Simulated event samples

Simulated samples are obtained using different MC generators and PS programs to model the hard scatter, the UE and the hadronisation process, respectively (see Section 2.2.3). For some processes, the simulation is performed with alternative MC configurations to evaluate the uncertainty on the choice of the modelling (see Section 5.6).

All MC samples are generated using a 25 ns bunch spacing configuration. The effect of pile–up is modelled by overlaying the hard scattering event with simulated minimum–bias events generated with Pythia 8.186 [306], using the NNPDF2.3lo [307] set of PDFs and the A3 set of tuned MC parameters [308]. The simulated samples are reweighted to reproduce the observed distribution of the average number of collisions per bunch crossing in data. For all samples, the simulation of detector effects is performed with either a full ATLAS detector simulation based on the Geant4 [309] framework or a fast simulation (AtlFast–II) [310] using a parameterisation of the performance of the electromagnetic and hadronic calorimeters [311] and Geant4 for the other detector components [310]. Moreover, the decays of b– and c–hadrons are simulated using the EvtGen 1.2.0 program [312] for all but the Sherpa samples.

- $t\bar{t}W$ production: the signal process is simulated at NLO precision in QCD with Sherpa 2.2.10 and the NNPDF3.0nnlo PDF set [92]. This sample uses a multi–leg merged setup with up to one additional parton at NLO QCD accuracy and up to two additional partons at LO accuracy, with $\mu_Q = 30$ GeV. The choice of renormalisation and factorisation scales is $\mu_R = \mu_F = H_{\mathrm{T}}/2$, where H_{T} is defined as the scalar sum of the transverse masses $\sqrt{p_{\mathrm{T}}^2 + m^2}$ of all final state particles. The masses of the top quark and the W boson are set to 172.5 GeV and 80.4 GeV, respectively [16]. The corrections from the $\mathcal{O}(\alpha_s^2\alpha^2)$ and $\mathcal{O}(\alpha^3)$ terms to the nominal NLO QCD calculation are added through MC event weights derived using the virtual additive corrections in the formalism described in Ref. [313].

  An alternative sample is simulated using aMC@NLO 2.9.3, matched to Pythia 8.245, to model the $t\bar{t}W$ signal at NLO QCD. It uses the

NNPDF3.0NLO PDF set and the A14 set of tuned MC parameters [314]. Top quark decays are simulated at LO using the MadSpin program [315, 316]. This sample also uses a multi–leg merged setup with up to one additional parton at NLO in QCD accuracy, and up to two at LO accuracy, with $\mu_Q = 30$ GeV. Hence achieving a similar accuracy to that of the the nominal Sherpa sample. Further alternative $t\bar{t}W$ samples are simulated using Powheg, interfaced to either Pythia 8.245 or Herwig 7.2.1 at NLO QCD and with the NNPDF3.0NLO PDF set.

All $t\bar{t}W$ QCD samples are normalised to the nominal Sherpa 2.2.10 sample inclusive cross–section, i.e. $\sigma(t\bar{t}W) = 573.7$ fb (including the EW correction terms $\mathcal{O}(\alpha_s^2\alpha^2)$ and $\mathcal{O}(\alpha^3)$), but keeping acceptance effects, in order not to be sensitive to overall normalisation differences when comparing the different simulations.

A separate sample using Sherpa 2.2.10 is used to simulate the $\mathcal{O}(\alpha_s\alpha^3)$ EW contribution term. This sample shares the same configuration as the nominal sample[‡]. Since the contribution of this sample to the total event yields is negligible after the selection requirements, and thus does not impact the reconstructed A_c^l, it is treated as a background sample in the analysis.

- $t\bar{t}Z/\gamma^*$ production: the $t\bar{t}Z/\gamma^*$ sample is simulated at NLO QCD using aMC@NLO 2.8.1, matched to Pythia 8.244, with the NNPDF3.0NLO PDF set and the A14 tune. The mass of the Z boson is set to 91.2 GeV [16]. The $t\bar{t}\gamma^*$ contribution and Z/γ^* interference effects are taken into account, with the samples including events with invariant mass of the opposite–sign–same–flavour (OSSF) lepton pair down to $m_{\ell\ell} > 1$ GeV. It is normalised to the calculation at $\text{NLO}_{\text{QCD+EW}}$ accuracy reported in Ref. [227] for an on–shell Z boson, scaled to the leptonic contributions including off–shell $\gamma^* \to \ell^+\ell^-$ contributions with a correction estimated at one–loop levelin α_s. The resulting $t\bar{t}\ell^+\ell^-$ cross–section, with $m_{\ell\ell} > 1$ GeV, is 162 fb.

  An alternative $t\bar{t}Z$ aMC@NLO sample matched to Herwig 7.2.1, instead of Pythia, is also used. It uses the Herwig standard tuned parameters and the NNPDF3.0NLO PDF set. Uncertainties in additional jet modelling are derived by performing ISR α_s variations taken from the A14 tune [314]. As with $t\bar{t}W$, the alternative samples are normalised to the same cross–section as the nominal sample.

- $t\bar{t}$ and tW production: both the $t\bar{t}$ and tW samples are simulated at NLO QCD with Powheg, matched to Pythia 8.230, with the NNPDF3.0NLO

[‡]From the technical standpoint, this sample is generated by simulating the $t\bar{t}Wj$ process at LO.

PDF set and the A14 tune. The h_{damp} matching parameter is set to $h_{\text{damp}} = 1.5 \times m_t$ [317]. This parameter also regulates the high–p_{T} radiation against which the $t\bar{t}$ system recoils. The $t\bar{t}$ sample is normalised to the NNLO+NNLL calculation [156,157], and the tW sample is normalised to a NLO QCD calculation including NNLL soft-gluon corrections [161]. Alternative $t\bar{t}$ samples are simulated with the same ME configuration, but matched to Herwig 7.1.3 instead, with the standard Herwig tune, and the same PDF set as the nominal sample.

- $t\bar{t}H$ production: the $t\bar{t}H$ sample is simulated using the same MC configuration as the $t\bar{t}$ nominal sample but with $h_{\text{damp}} = 0.75 \times (2m_t + m_H)$. Two alternative samples are obtained varying the ME generator (aMC@NLO 2.6.0) or the PS (Herwig 7.2.1) with respect to the nominal ones. All samples where normalised to the cross–section calculated at $\text{NLO}_{\text{QCD+EW}}$ from Ref. [227].

- $t\bar{t}\gamma$ production: the $(2 \to 7)$ $pp \to e\nu_e b\mu\nu_\mu b\gamma$ sample is simulated with aMC@NLO 2.3.3, which is matched to Pythia 8.212 at LO QCD. This sample also uses the NNPDF3.0nlo PDF set and the A14 tune. This sample is only used to assign an extra uncertainty to additional photon emissions in the nominal $t\bar{t}$ prediction. This procedure is described in Section 5.6.2.

- tZq and tWZ production: the tZq sample uses aMC@NLO 2.3.3, which is matched to Pythia 8.245 at NLO QCD, the NNPDF3.0nlo PDF set, and the A14 tune. The tZq simulation includes off-shell Z boson decays into dilepton pairs in the range $m_{\ell\ell} > 5$ GeV. Single–top quark production in association with a W and a Z boson (tWZ) is also simulated with aMC@NLO 2.2.2, interfaced to Pythia 8.235 at NLO QCD, with the NNPDF3.0nlo PDF set and the A14 tune. The interference between $t\bar{t}Z$ and tWZ is removed using the DR1 scheme [318]. Both samples are normalised to their MC cross–section values.

- Z+jets and $Z\gamma$ production: the Z+jets samples are simulated at NLO QCD using Powheg, matched to Pythia 8.186, the NNPDF3.0nlo PDF set, the AZNLO set of tuned parameters [319], and uses Photos [320] for the FSR. An alternative Z+jets sample is simulated using Sherpa 2.2.11 and the NNPDF3.0nlo PDF set. Both of these samples include events in the range $m_{\ell\ell} > 10$ GeV and are normalised to NNLO predictions [321]. On the other hand, the $Z\gamma$ samples are simulated using the Sherpa 2.2.11 configuration at NLO QCD along with the NNPDF3.0nlo PDF set. Analogously to the $t\bar{t}\gamma$ sample, these are only used to assess an additional uncertainty in the Z+jets modelling.

- WZ, ZZ, WH and ZH production: diboson production includes final states with both three charged leptons and one neutrino (WZ+jets), or four charged leptons (ZZ+jets). They are simulated using Sherpa 2.2.2 in a multi–leg merged setup with up to one additional parton at NLO QCD, and two or three additional partons at LO accuracy. MC samples featuring Higgs boson production in association with a W or Z boson (WH and ZH) are generated at NLO QCD using Powheg interfaced to Pythia 8.230/8.235 for the PS, together with the NNPDF3.0nlo PDF set and the AZNLO MC tune, and they are normalised to the theoretical calculations in Ref. [227].

- Other minor processes: the production of the $t\bar{t}t\bar{t}$ process is simulated at NLO QCD with Sherpa 2.2.11, together with the NNPDF3.0nnlo PDF set. The production of three top quarks ($tt\bar{t}$) and the production of a $t\bar{t}$ pair with two W bosons ($t\bar{t}WW$) are simulated at LO using aMC@NLO 2.2.2 interfaced to Pythia 8.186 with the A14 MC tune and the NNPDF2.3lo PDF set. Fully leptonically decaying triboson processes (WWW, WWZ, WZZ and ZZZ) with up to six leptons in the final states are simulated with Sherpa 2.2.2 and the NNPDF3.0nlo PDF set. Final states with no additional partons are calculated at NLO QCD, whereas final states with up to three additional partons are calculated at LO.

5.3 Object selection

After the reconstruction techniques described in Section 2.5, the final–state objects such as the electrons, muons or jets must pass a *pre–selection* based on the final–state topology. Tables 5.1 and 5.2 show the pre–selection requirements for leptons and jets, respectively.

Additional isolation requirements are imposed on the leptons that are selected for the analysis regions[§]. In what follows, electrons and muons are referred to as *tight* or *loose* depending on whether they satisfy or not these isolation requirements, i.e. loose leptons satisfy the selection from Table 5.1, but fail the isolation requirements; whereas tight leptons pass all selection requirements. Furthermore, electron candidates originating from photon conversions ($\gamma^* \to e^+e^-$) are suppressed. Electrons can be identified as coming from material photon conversion candidates by checking for additional tracks close to the calorimeter energy clusters associated with them and the existence of displaced

[§]Non-prompt leptons are further rejected using a *Tight* isolation WP from the PLIV (Prompt Lepton Improved Veto) algorithm. This algorith uses a BDT discriminant that combines electromagnetic shower shapes and track information from the inner detector to distinguish prompt leptons from non–prompt ones. It is also referred to as "non-prompt lepton BDT" [234].

	Pre–selected electron	Pre–selected muon
Identification	Tight	Medium
Acceptance	$p_{\mathrm{T}} > 10$ GeV, $\lvert\eta_{\mathrm{clust}}\rvert < 2.47$ except $1.37 < \lvert\eta_{\mathrm{clust}}\rvert < 1.52$	$p_{\mathrm{T}} > 10$ GeV, $\lvert\eta\rvert < 2.5$
Impact parameter	$\lvert d_0\rvert/\sigma(d_0) < 5$ $\lvert z_0 \sin\theta\rvert < 5$ mm	$\lvert d_0\rvert/\sigma(d_0) < 3$ $\lvert z_0 \sin\theta\rvert < 5$ mm

Table 5.1: Summary of the electron and muon objects pre–selection requirements. These objects are used for the overlap removal and parts of the background estimation in the analysis. $|\eta_{\mathrm{clust}}|$ represents the pseudo–rapidity angular position at which the electron deposits its energy cluster in the ECAL. Electron candidates are rejected if it coincides with the ECAL transition region between the barrel and the end–cap regions. d_0 is the transverse impact parameter relative to the beam axis, $\sigma(d_0)$ is the uncertainty on d_0, and z_0 describes the longitudinal impact parameter relative to the reconstructed primary vertex. The requirements on these parameters improve the non–prompt lepton rejection.

	Selected jet
Algorithm	anti–k_t on PF objects, $R = 0.4$ (see Section 2.5.4)
Acceptance	$p_{\mathrm{T}} > 20$ GeV, $\lvert\eta\rvert < 2.5$
Pile–up rejection	JVT > 0.2 if $\lvert\eta\rvert < 2.4$ and $p_{\mathrm{T}} < 60$ GeV
	b–tagged jet
Acceptance	$p_{\mathrm{T}} > 20$ GeV, $\lvert\eta\rvert < 2.5$
Algorithm	DL1r algorithm [322] (77% WP)

Table 5.2: Summary of the jet and b–jet objects selection requirements.

conversion vertices. Such candidates are rejected following a set of requirements that are labelled as "e/γ ambiguity cuts" and are applied to both tight and loose electrons.

Associated scale factors are applied as multiplicative factors to the MC event weights to correct for the mis–modelling of efficiencies associated with the reconstruction, identification, isolation and trigger selection of electrons and muons, as well as the JVT and b–tagging selection for jets.

5.4 Event selection and definitions of signal and control regions

Given that this analysis targets a final state with three light charged leptons, events are selected with either single–lepton or dilepton (ee, $\mu\mu$ or $e\mu$) triggers. They have varying minimum p_{T} thresholds, between 12 and 26 GeV, depending on the lepton flavour, the trigger type, and the data-taking period [323, 324].

A logical OR between the triggers is applied. Events with anything other than exactly three tight leptons in the signal regions (SRs) are rejected, while the selection is agnostic to additional loose leptons. The p_{T} requirements of the leptons that fire the trigger have to be above the corresponding trigger threshold in order to achieve maximum trigger efficiency. In this respect, they have to be above 30, 20 and 15 GeV for the respective leading, sub–leading, and third lepton[¶]. Furthermore, a geometrical matching between these selected leptons, and the ones reconstructed by the trigger algorithms is required.

There are a total of four SRs. Events are classified in each of them depending on their jet and b–jet multiplicities, and the $E_{\mathrm{T}}^{\mathrm{miss}}$. The motivation behind this classification is two–fold. In first place, the $t\bar{t}W$ 3ℓ final–state topology contains two b–jet emissions from the hard process, with the possibility of an additional parton emission coming from gluon radiation or showering effects. In second place, this classification leads to optimal signal over background (S/B) ratios in all SRs, which is conducive to smaller uncertainties on the extracted A_c^l compared to other SR definitions. For these reasons, events are split in "lowN_{jets}" regions with two or three jets, and "highN_{jets}" with at least four jets, and these are further split depending on whether they have one b–jet or more.

All signal and control regions (CRs) follow a set of general requirements that are summarised in Table 5.3. These select events with three charged leptons, the aforementioned p_{T} requirements, as well as the sum of the lepton charges to be ± 1, and the invariant mass of the OSSF lepton pair $m_{\ell\ell}^{\mathrm{OSSF}}$ to be at least 30 GeV. The latter removes the contributions from low–mass lepton resonances (i.e. $J/\psi \to \ell^+\ell^-$). CRs follow an additional specific selection criteria to target the main backgrounds of the analysis, in order to constrain them in the simultaneous fit to extract the result. After the general selection described above, these are: $t\bar{t}Z$ production (irreducible), and non–prompt lepton sources from heavy–flavour (HF) hadron decays or γ–conversions. Non–prompt leptons are a source of reducible background. The number of background events of this type is looked to be reduced by applying specific selection criteria, such as tight lepton isolation requirements, or more efficient lepton reconstruction algorithms.

The number of Z boson candidates ($N_{Z\text{–cand.}}$) is defined by the number of OSSF lepton pairs found in the event that have an invariant mass in the range $[m_Z - 10\ \mathrm{GeV},\ m_Z + 10\ \mathrm{GeV}]$. Exactly one Z boson candidate is required for the CR for the $t\bar{t}Z$ background, but zero for all the other regions. The CRs targeting non-prompt electron/muons arising from HF hadron decays (CR–HF$_e$ and CR–HF$_\mu$) are separated by the flavour of the third lepton, which must fail the isolation requirements. This ensures that this third lepton originates from

[¶]These p_{T} requirements ensure a distinction between the two softest leptons, which is exploited in the control region definition.

a HF decay more than 99% of the time, by checking with the MC event record, and enriches the CRs with HF leptons. The CR for γ–conversions (CR–γ–conv) requires at least one of the leptons to be an electron candidate that fails the e/γ ambiguity cuts. These three sources of non–prompt leptons are simulated using events from $t\bar{t}$, tW and Z+jets MC samples, where only two prompt leptons are expected, plus additional non–prompt ones. Throughout the SRs and HF CRs, the contribution coming from the $t\bar{t}$ sample dominates, with over 90% of the total non–prompt lepton yields for all categories. This is slightly reduced to 80% for the CR–γ–conv. The contribution from electrons with misidentified electric charge has been studied in MC simulations and found to be negligible in the SRs. The classification of all these non–prompt lepton categories is performed using a set of tools within the ATLAS software environment. Details of their implementation can be found in Appendix C.

	General requirements			
N_ℓ ($\ell = e/\mu$)	$= 3$			
p_{T}^{ℓ} ($1^{\mathrm{st}}/2^{\mathrm{nd}}/3^{\mathrm{rd}}$)	≥ 30 GeV, ≥ 20 GeV, ≥ 15 GeV			
Sum of lepton charges	± 1			
$m_{\ell\ell}^{\mathrm{OSSF}}$	≥ 30 GeV			
	Region-specific requirements			
	SR–1b–**low**$N_{\mathbf{jets}}$	**SR**–1b–**high**$N_{\mathbf{jets}}$	**SR**–2b–**low**$N_{\mathbf{jets}}$	**SR**–2b–**high**$N_{\mathbf{jets}}$
N_{jets}	$[2, 3]$	≥ 4	$[2, 3]$	≥ 4
$N_{b\text{–jets}}$	$= 1$	$= 1$	≥ 2	≥ 2
$E_{\mathrm{T}}^{\mathrm{miss}}$	≥ 50 GeV	≥ 50 GeV	–	–
$N_{Z\text{–cand.}}$	$= 0$			
Lepton criteria	TTT			
e/γ ambiguity cuts	satisfy all			
	CR–$t\bar{t}Z$	**CR**–**HF**$_e$	**CR**–**HF**$_\mu$	**CR**–γ–**conv**
$\ell^{1\mathrm{st}/2\mathrm{nd}/3\mathrm{rd}}$	$\ell\ell\ell$	$\ell\ell e$	$\ell\ell\mu$	$\ell\ell e$, $\ell e\ell$, $e\ell\ell$
N_{jets}	≥ 4	≥ 2	≥ 2	≥ 2
$N_{b\text{–jets}}$	≥ 2	$= 1$	$= 1$	≥ 1
$E_{\mathrm{T}}^{\mathrm{miss}}$	–	< 50 GeV	< 50 GeV	< 50 GeV
$N_{Z\text{–cand.}}$	$= 1$	$= 0$	$= 0$	$= 0$
Lepton criteria	TTT	TTL	TTL	TTT
e/γ ambiguity cuts	satisfy all	satisfy all	satisfy all	fail ≥ 1

Table 5.3: Summary of the requirements applied to define the signal and control regions of the analysis. N_{jets} includes b-tagged and non–b–tagged jets. The labels T and L refer to tight leptons that satisfy all selection requirements described in Section 5.3 and loose leptons that fail to satisfy the isolation requirements, respectively.

Figures 5.1 and 5.2 show the compositions of the SRs and CRs, respectively, before the fit to data (*pre–fit*). In both, the contribution to the non–prompt lepton backgrounds ($\mathrm{HF}_{e,\mu}$ and γ–conversions) comes primarily from $t\bar{t}$ production, but also from Z+jets and tW production; and the "Other" category includes tWZ, $t\bar{t}WW$, $H + W/Z$, VVV, $tt\bar{t}$ and $t\bar{t}t\bar{t}$ production.

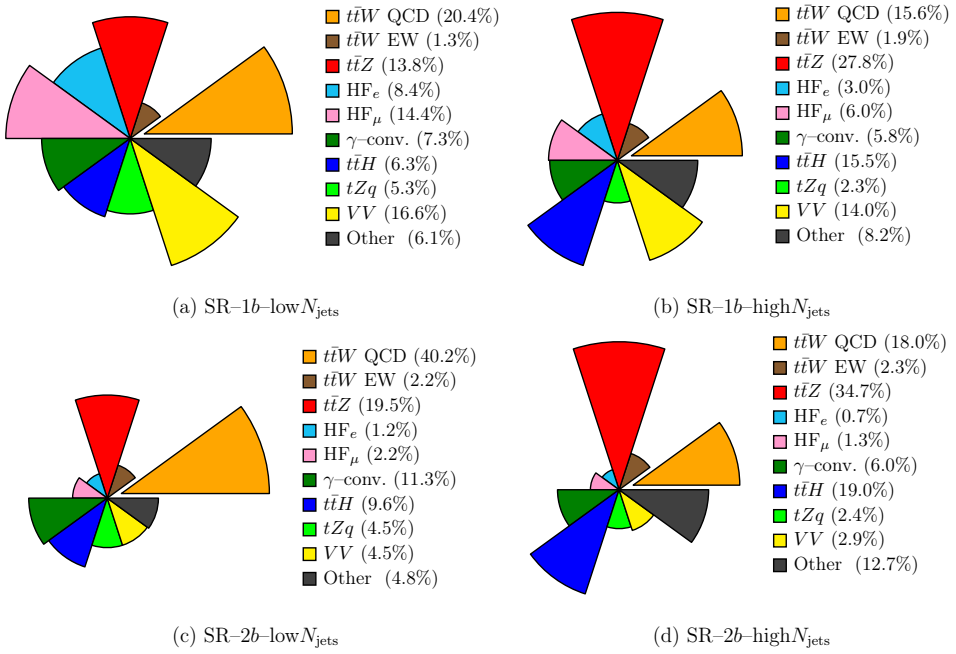
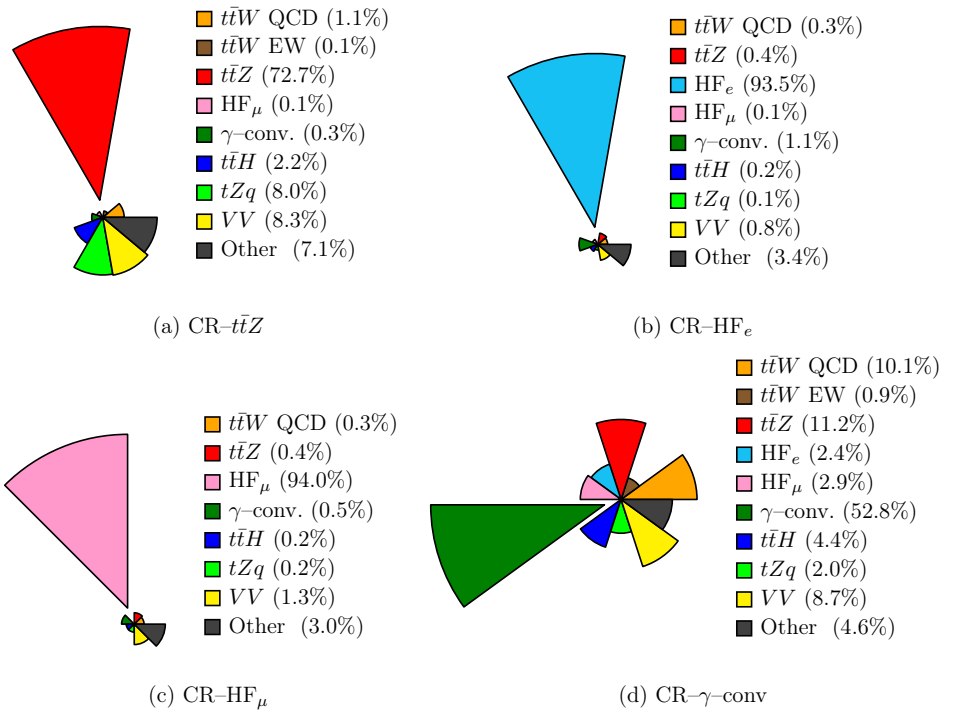


Figure 5.1: Expected contribution from the signal and background processes before the fit to data for the four SRs: (a) SR–1b–lowN_{jets}, (b) SR–1b–highN_{jets}, (c) SR–2b–lowN_{jets} and (d) SR–2b–highN_{jets}.

5.5 Lepton–top quark matching

One of the main challenges is to identify the leptons that originate from the top and antitop quarks in order to build the $\Delta|\eta_\ell|$ observable, and ultimately the A_c^l, as defined in Eq. 3.3. In a 3ℓ $t\bar{t}W$ event, the leptons coming from the $t\bar{t}$ pair have opposite electric charge signs, while the one originating from the ISR W boson has the same charge sign as one of the former two. This is true for any possible $t\bar{t}W$ 3ℓ final–state configuration. This ensures that the lepton with opposite sign to the other two, labelled as the *odd* lepton, always originates from a top (or antitop) quark. The remaining two leptons, of the same sign, are labelled as the *even* leptons. Therefore, the problem is reduced to identifying which one of the even leptons originates from the antitop (or top) quark.



(a) CR–$t\bar{t}Z$ (b) CR–HF$_e$

(c) CR–HF$_\mu$ (d) CR–γ–conv

Figure 5.2: Expected contribution from the signal and background processes before the fit to data for the four CRs: (a) CR–$t\bar{t}Z$, (b) CR–HF$_e$, (c) CR–HF$_\mu$, and (d) CR–γ–conv. The separated sections indicate the targetted process in the region. The absence of a background in a legend indicates that it contributes less than 0.1% to the total CR yields.

This problem is addressed using a BDT classifier algorithm that computes a discriminator value for each even lepton in each event. Large discriminator values correspond to large probabilities that a given lepton originated from a top or antitop quark decay. The lepton with the highest BDT discriminator score is selected to calculate $\Delta|\eta_\ell|$. This section describes the BDT training procedure and performance to achieve the final lepton–top quark association.

5.5.1 Parton level matching

Prior to the BDT training, the reconstructed even leptons have to be matched to the parton level objects* in order to classify them as *signal* leptons, that

*The set of information corresponding to an event before being showered and passed through the detector simulation is known as the parton level information. This is because it contains information from the ME generation that incorporates the knowledge of the original particle from which the final–state particles decay from, and no uncertainties due to particle reconstruction.

are identified as decaying from the top quark, or *background* leptons, which are not matched and are therefore classified as originating from the associated ISR W boson.

This matching procedure is based on the angular distance ΔR between the parton level and the reconstructed lepton objects, where the parton level lepton collection only contains the pair that comes from the $t\bar{t}$ decay. For a successful matching, $\Delta R_{\text{parton}-\text{reco.}} < 0.3$ and the leptons must have the same charge and flavour. In the case where more than one reconstructed lepton satisfies these requirements, the one with the smallest $\Delta R_{\text{parton}-\text{reco.}}$ distance is selected. Above 90% of the matched leptons have $\Delta R_{\text{parton}-\text{reco.}} < 0.1$, and the portion of $t\bar{t}W$ events where both parton level leptons are not matched is approximately 1%. These outlying events cannot be used in the training, and are thus removed from it.

5.5.2 Input BDT variables

Five input variables that discriminate between leptons from top quark or anti-top quark decays and leptons from the associated W boson decays are defined. They are the masses of the two systems formed from the lepton and the closest and second closest b–jets[†] ($m_{\ell b_0}$ and $m_{\ell b_1}$), the angular distances between the lepton and these b–jets ($\Delta R_{\ell b_0}$ and $\Delta R_{\ell b_1}$), and the lepton p_{T}. For events with only one b–tagged jet, the collection of jets that pass a looser DL1r WP than the default one (77% WP) is scanned, and the one with the highest tagging score is used. If no other jet passes any b–tagging WP, the variables are built using the closest untagged jets to the leptons.

Figure 5.3 shows the normalised distributions for the five input variables for the two even leptons in each event: the signal leptons (in blue), and the background leptons (in red). Figure 5.4 shows the correlation between the five BDT input variables which is in any case smaller than 52%. Moreover, the correlation between these input variables and the corresponding lepton pseudorapidity is very small.

5.5.3 $m_{\ell b_0}$–based lepton–top quark matching

The best discrimination is achieved by the $m_{\ell b_0}$ variable. In addition, the correct modelling of this variable is validated by comparing MC simulations to data in the analysis signal regions. This can be seen for $m_{\ell b_0}$, and the other input variables, in Figure 5.24 (a)–(e) after to the fit to real data described in Section 5.8.3. Given this, it is possible to exploit this variable to define a procedure of selecting the correct even lepton, without having to use the BDT. This method, denoted as "$m_{\ell b}$–matching", has two main advantages: in first

[†]The closest and second closest b–jets are defined by using the respective ΔR distances.

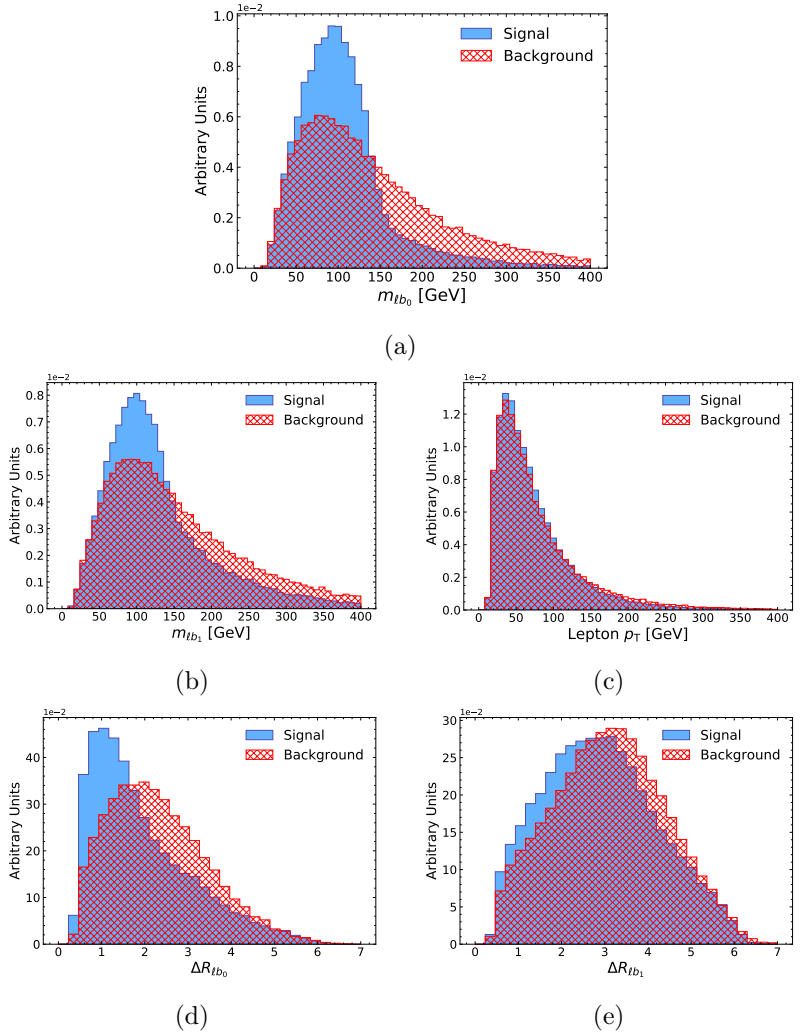


Figure 5.3: Normalised distributions comparing the leptons coming from top quarks (blue) against the leptons coming from ISR W bosons (red) in $t\bar{t}W$ events for the five input BDT variables: (a) $m_{\ell b_0}$, (b) $m_{\ell b_1}$, (c) lepton p_{T}, (d) $\Delta R_{\ell b_0}$, and (e) $\Delta R_{\ell b_1}$.

place, it is much simpler to understand since it avoids the inherent complexity added by the BDT; in second place, it greatly improves the reproducibility of the chosen phase–space, and is independent of the generator–specific MC event record. The $m_{\ell b_0}$ peaks at about 92 GeV for the signal lepton (top panel in Figure 5.3). Therefore, the even lepton whose $m_{\ell b_0}$ is closest to 92 GeV can be selected to be the one originating from the top quark. Parton level information is used to determine the efficiency of this lepton–top quark association method: the $m_{\ell b}$–matching has an efficiency of selecting the correct lepton of about 65%.

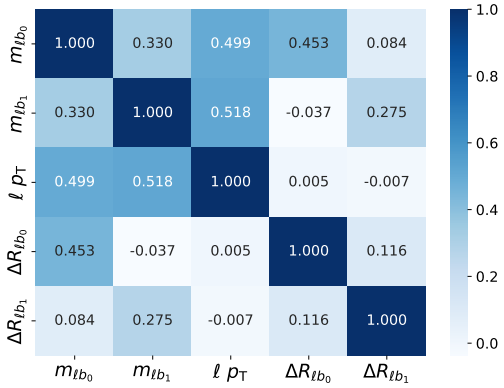


Figure 5.4: Correlation matrix between the five input BDT variables.

5.5.4 BDT–based lepton–top quark matching

The $m_{\ell b}$–matching efficiency can be improved upon considering a multivariate phase–space to decide which even lepton is the one that decays from the top quark; this is the job for a BDT. BDTs are one of the many ML techniques that are available. Broadly, ML algorithms are built upon a large models of data, known as training data, and are able to detect patterns or correlations among the input variables. This allows them to make reliable predictions or decisions when confronted with a new set of data. Nowadays, they are extensively being used in almost all physics analyses due to their great versatility.

For this analysis, the BDT classifier is implemented using the Scikit–learn package [325]. It is trained over the nominal $t\bar{t}W$ Sherpa sample using a k–fold cross validation, with five folds being used for the training and testing with a 80 : 20 ratio. The k–fold technique ensures that the performance of the BDT is independent from the chosen dataset. Furthermore, given that the same $t\bar{t}W$ sample is used to populate the analysis regions, it is necessary to perform this splitting to guarantee that the evaluated events are independent from those used for the training, in each of the five folds. Figure 5.5 shows a schematic view of the k–fold cross validation used for the training, testing and final evaluation of the events.

One of the most important metrics that comes out of ML algorithms is the feature importance. The feature importance is a metric used to identify which of the variables used by the ML algorithm provides the best discrimination between the signal and background. The left panel of Figure 5.6 shows the feature importance for the first fold. As anticipated, $m_{\ell b_0}$ leads the ranking, followed by $m_{\ell b_1}$. The right panel shows the BDT output on both the training (discontinuous line) and testing (continuous line) datasets, with signal leptons in blue and background leptons in red. The vertical error bars rep-

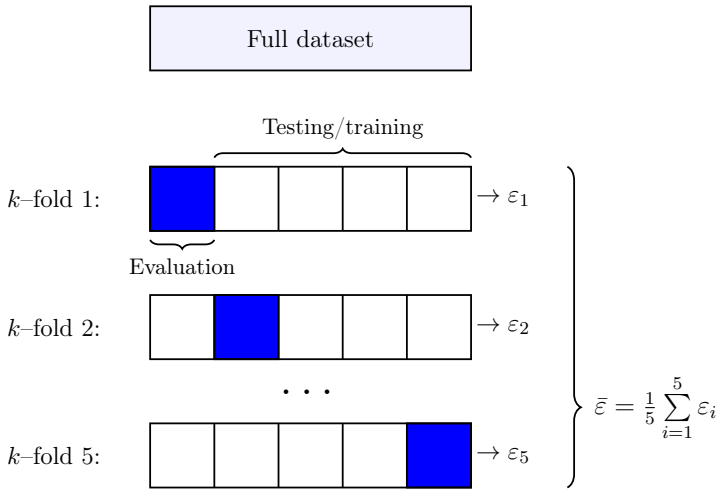
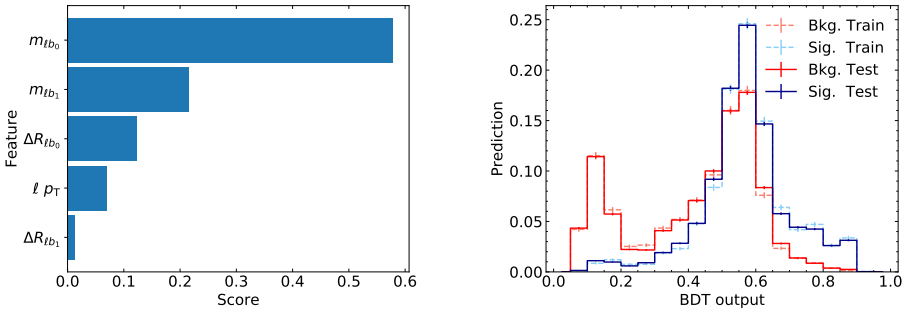


Figure 5.5: Schematic representation of the k–fold cross validation technique. The full sample is split in five parts so that the evaluated events are different from the ones used for the training and testing, in each fold. The efficiency ε_i of each classifier is evaluated for each of the folds, and the total efficiency $\bar{\varepsilon}$ is the average of all of them. In each of the folds, the white squares are split in a 80 : 20 ratio for the training and testing, respectively.

resent the statistical uncertainty of the MC sample. Good agreement is seen between both training and testing datasets, implying that the ML algorithm has not been overtrained. Two distinct peaks can be seen for signal (around 0.6) and background leptons (around 0.1). However, background leptons also peak around the signal peak. This contamination leads to a decrease in the efficiency of the BDT to select the correct lepton. Both figures show the results for the first fold, i.e. k–fold 1, but the results for the other four folds are analogous.

Moreover, the receiver operating characteristic (ROC) curve and the area under the curve (AUC) metrics are used to evaluate the efficiency of the ML algorithms. The AUC is the integral of the ROC curve which varies between zero and one. The closer the AUC is to one, the better the algorithm is at distinguishing between the signal and background classes. Figure 5.7 shows the ROC curves for the five folds. All of them are alike and have similar AUC integrals, advocating the independence of the BDT performance to the chosen dataset, and ensuring, again, that the model is not overtrained. The BDT–based lepton–top quark matching efficiency of selecting the correct even lepton is 71%. This is obtained by using the MC event record on the evaluated events, and is identical to the average AUC score from the testing datasets. The $\Delta|\eta_\ell|$ value obtained with the odd lepton and the even lepton with the highest BDT discriminant value is referred to as $\Delta|\eta_\ell^{\mathrm{BDT}}|$. Figure 5.24 (f)–(g)

shows the good data/MC agreement for the BDT discriminator values for both even leptons after the fit to real data described in Section 5.8.3. Appendix F shows the bidimensional distributions of the each of the input variables against the BDT discriminant score for the three leptons in the event.



(a) (b)

Figure 5.6: (a) Feature importance and (b) normalised BDT output for the signal (red) and background (blue) training (discontinuous) and testing (continuous) datasets for the first fold. The other 4 folds exhibit analogous results.

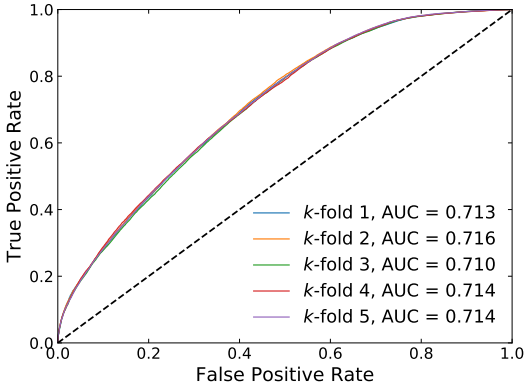


Figure 5.7: ROC curves and AUC scores for the five folds. The diagonal discontinuous line represents the case where the correct lepton is chosen randomly.

5.6 Experimental and theoretical systematic uncertainties

The predictions of the $t\bar{t}W$ signal and the SM backgrounds are affected by several sources of experimental and theoretical systematic uncertainties. These uncertainties are classified into the different categories that are described in the following.

5.6.1 Experimental uncertainties

Experimental uncertainties target detector–related uncertainties in the measurement of quantities such as luminosity, pile–up, or the several calibrations on the reconstruction and identification of the physics objects:

- Luminosity: the uncertainty in the combined 2015–2018 integrated luminosity is 1.7%* [305], obtained using the LUCID–2 detector for the primary luminosity measurements. This systematic affects all MC samples that are not normalised to data in the simultaneous fit to extract the A_c^l (see Section 5.8).

- Pile–up reweighting: the uncertainty in the reweighting of the MC pile–up distribution to match the distribution in data is evaluated by varying the reweighting factors and has a very small impact.

- Charged leptons (electrons and muons): uncertainties associated with the lepton selection arise from the trigger, reconstruction, identification and isolation efficiencies, and the lepton momentum scale and resolution. SFs are applied to MC to correct to data efficiencies (see Section 2.5.2) and their associated systematic uncertainties are coherently propagated to the analysis distributions. Overall, the impact of these uncertainties is small.

- Jets: uncertainties associated with the jet reconstruction and calibration arise from the JES calibration, JER and the JVT requirement. The JES and its uncertainties are derived by combining information from test-beam data, LHC collision data and simulation (see Section 2.5.4.1). These are decomposed in a set of 30 independent nuisance parameters, with contributions from pile–up, jet flavour composition, single–particle response, and effects of jets not contained within the calorimeter. Likewise, the JER is measured separately for data and MC using in-situ techniques [150]. Its uncertainty is evaluated using 13 components considering differences between MC and data in jet p_{T} and η. A SF is applied to correct for the JVT efficiency, and its associated uncertainty is also propagated through the fit. Uncertainties on the JER are the leading experimental ones in this result.

- b–tagging: likewise, variations in the SFs that are used to correct for tagging efficiencies are used to estimate the systematic uncertainty coming from flavour–tagging (see Section 2.5.5). The impact of the uncertainties on the b–tagging calibration is evaluated separately for b–jets, c–jets and light–flavour jets, and is found to be small.

*The final luminosity uncertainty for the full Run 2 was released after this analysis was finalised and was found to be 0.83% [77].

- $E_{\mathrm{T}}^{\mathrm{miss}}$: uncertainties are applied to scale and resolution of the soft term (see Section 2.5.6), as well as to the energy and momentum scales of electrons, muons and jets. Their impact is negligible.

5.6.2 Theoretical uncertainties

The signal and main background samples are simulated using different MC configurations. The comparison of two configurations allows for the estimation of the uncertainty in the choice of the ME+PS matching scheme, the PS or the modelling of certain aspects of the simulation. The impact of the missing orders in the pQCD calculation are assessed by varying the μ_R and μ_F scales (see Section 2.2.2). This is done by varying each of them, but not simultaneously, by a factor 2.0 or 0.5 from their nominal values. Uncertainties associated with the PDF sets are evaluated according to the PDF4LHC prescription [326]. They include internal variations of the nominal PDF sets that are added in quadrature, uncertainties due to the choice of the PDF set, as well as variations of the α_s parameter.

Table 5.4 summarises all the theoretical uncertainties on the signal and background samples in the analysis. The alternative samples used for each process are detailed in Section 5.2. In all cases, the alternative sample cross–sections are normalised to the same cross–section value as the nominal samples (which is the best theoretical prediction), so that the systematic uncertainties only cover differences between the acceptances and shapes of the kinematic variable distributions, but not the overall normalisations of the processes. For background samples that are not normalised to data, a fixed overall cross–section normalisation uncertainty "XS norm." is assigned. For the $t\bar{t}H$ and tZq processes, the respective theoretical cross–section uncertainties from Refs. [227] and [201] are used.

For the WW/WZ+jets backgrounds, a conservative normalisation uncertainty of 20% is used to account for differences in the WW/WZ+jets modelling for different b–jet multiplicities. This uncertainty is estimated from the level of agreement between data and MC simulation in several validation regions enriched in WW/WZ+jets and reflects the level of agreement between data and MC simulations in Ref. [327].

In addition, processes that contribute less than 2% to the total event yields in the SRs ($t\bar{t}W$ EW, tWZ, $t\bar{t}WW$, $H + W/Z$, VVV, $tt\bar{t}$ and $t\bar{t}t\bar{t}$) receive a 30% normalisation uncertainty. This is a conservative approach, based on the precision of the latest cross section measurement of $t\bar{t}t\bar{t}$ [233], which should cover the known theoretical uncertainties of the other rare backgrounds.

In $t\bar{t}$ production, an extra uncertainty associated with the photon radiation in $t\bar{t}$ events is applied by comparing the predictions from $t\bar{t}$ and $t\bar{t} + t\bar{t}\gamma$ production. The latter can only be done if the overlap between the photons radiated

within the PS in $t\bar{t}$ and the photons coming from the $2 \to 7$ $t\bar{t}\gamma$ MEs is removed. The same overlap removal "OLR" is applied to compare the nominal Z+jets sample to that of the combination of the alternative Z+jets + $Z\gamma$ samples.

5.6.3 Treatment of systematic uncertainties in the fit

The sources of systematic uncertainties described above are treated differently depending on how they modify the nominal prediction. Some systematic variations (JVT, pile–up, lepton SFs and b-tagging) are obtained by reweighting the nominal events. The *up* and *down*[†] weights are usually symmetric around 1.0 and therefore the systematic variation is also symmetric. Nevertheless, a bin–by–bin symmetrisation scheme, where $\text{up}_{\text{symm.}} = \text{down}_{\text{symm.}} = (\text{up} - \text{down})/2$, is applied to iron out punctual fluctuations. Since these *weight systematics* show a good level of symmetry from the start, the effect of this symmetrisation scheme, known as the *two–sided* scheme, is small. Other systematics modify the kinematics of the particle objects and cannot be applied as weights (JES, JER – see below –, $E_{\text{T}}^{\text{miss}}$, electron scale/resolution and muon scale/resolution). A parallel set of events, where these variations have been applied, are simulated and then compared to the nominal prediction. Similarly, the up and down variations should show a good level of symmetry but still the two–sided symmetrisation scheme is applied to ensure it. In contrast to the others, some $E_{\text{T}}^{\text{miss}}$ systematics provide only one variation with respect to the nominal. In this case, the unique variation is symmetrised to the other side of the nominal distribution.

For JER systematics, both the up and down variations are compared and the one that gives a larger variation with respect to the nominal distributions is symmetrised. This more conservative approach is needed since these systematics are usually not symmetric from the start, and the use of the two–sided scheme would greatly reduce the impact of these uncertainties on the fit.

Theoretical systematics are much easier to understand: alternative modelling systematics provide only one variation with respect to the nominal. This variation is set as the up variation, for example, and symmetrised to the other side in the fit. These are referred to as *two–point systematics*. Others, like cross–section normalisations or the luminosity uncertainty, are added as flat two–sided symmetric variations throughout the whole analysis.

A procedure known as *smoothing* is also applied to all weight systematics, the JER systematics, and the alternative $t\bar{t}$ and Z+jets MC modelling uncertainties. The smoothing procedure aims to remove statistical fluctuations from the systematic variations and contributes to the removal of artificial pulls or constraints, and of the double–counting of MC statistical uncertainties.

[†]Up and down variations consider a $\pm 1\sigma$ variation of the expected systematic uncertainty.

Additionally, a *pruning* procedure is applied to all systematics. This procedure rejects systematic uncertainties that have an effect smaller than 0.5% per bin on either the normalisation or the shape. This is done in order to avoid inestabilities of the fit due to very small uncertainties that have a negligible impact.

5.6.4 $t\bar{t}W$ signal modelling

Given the relevance of the modelling of the $t\bar{t}W$ process, the comparison between the nominal SHERPA sample and the FXFX one is shown in Figure 5.8 for the four SRs[‡]. Although both samples are normalised to the same cross–section, differences in the modelling, as well as acceptance effects, are conducive to a sizeable differences between both predictions. The alternative FXFX sample predicts lower expected yields throughout. An updated FXFX merging procedure has been developed and proposed by the authors [236]. Thus, an improved version of this MC sample is currently being simulated and is expected to reduce the disagreement and hence the associated modelling uncertainty in future studies.

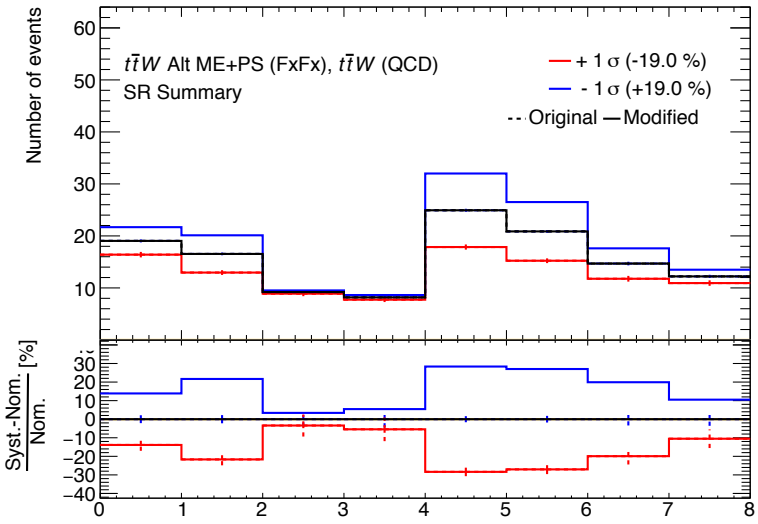


Figure 5.8: Two–point systematic variations for the aMC@NLO + PYTHIA FXFX generator setup for $t\bar{t}W$ production in the SRs. From left to right, each pair of bins show the SR–1b–lowN_{jets}, SR–1b–highN_{jets}, SR–2b–lowN_{jets} and SR–2b–highN_{jets}; and for each pair, the left one corresponds to $\Delta|\eta_\ell| \leq 0$ and the right one to $\Delta|\eta_\ell| > 0$. Uncertainty bands cover MC statistical uncertainties of the nominal SHERPA sample.

[‡]Modelling systematic variations for the other main backgrounds can be seen in Appendix D.

Process	Nominal ME+PS	Alt. matching scheme	Alternative PS	XS norm.	Weight variations
$t\bar{t}W$ QCD	Sherpa	aMC@NLO + Pythia FxFx		–	μ_R, μ_F, PDF and α_s
			Powheg + Pythia vs. Powheg + Herwig		
$t\bar{t}Z$	aMC@NLO + Pythia		aMC@NLO + Herwig	–	μ_R, μ_F, A14
$t\bar{t}H$	Powheg + Pythia	aMC@NLO + Pythia	Powheg + Herwig	$^{+5.8\%}_{-9.2\%}$	μ_R, μ_F
$t\bar{t}$	Powheg + Pythia	$t\bar{t}$ + $t\bar{t}\gamma$ (OLR)		–	μ_R, μ_F
			Powheg + Herwig		
Z+jets	Powheg + Pythia	Z+jets + $Z\gamma$ Sherpa (OLR)		–	μ_R, μ_F
tZq	aMC@NLO + Pythia	–		±14%	μ_R, μ_F
VV	Sherpa	–		±20%	μ_R, μ_F
Other backgrounds	–	–		±30%	–

Table 5.4: Summary of the theoretical uncertainties on the signal and background samples used in the analysis.

5.7 Binned maximum profile likelihood fitting

In this analysis, the best–fit value of the A_c^l to the selected data is extracted using a frequentist approach implemented in the ROOSTATS framework [328]. This uses the asymptotic approximation formulae described in Ref. [329]. The fitted data is distributed in some variable x that is used to construct a histogram $\mathbf{n} = (n_1, \ldots, n_N)$, with N being the total number of bins. The expected number of events in the i–th bin can be written as

$$\mathrm{E}[n_i] = \mathcal{N}_s s_i(\boldsymbol{\theta}) + \sum_{b \in \text{bkgs.}} \mathcal{N}_b b_{i,b}(\boldsymbol{\theta}), \tag{5.1}$$

where $\mathcal{N}_{s,b}$ are the normalisation factors (NFs) of the signal and background processes, with $\mathcal{N}_{s,b} = 1$ corresponding to the SM expected number of events, and b running over all the background processes described in Section 5.4. The signal and background events are distributed following their probability density functions $f(x, \boldsymbol{\theta})$ in each bin i, i.e.

$$s_i(\boldsymbol{\theta}) = s_{\text{tot}} \int_{\text{bin } i} f_s(x_i, \boldsymbol{\theta}) dx, \qquad b_i(\boldsymbol{\theta}) = b_{\text{tot}} \int_{\text{bin } i} f_b(x_i, \boldsymbol{\theta}) dx, \tag{5.2}$$

where $\boldsymbol{\theta}$ are the nuisance parameters (NPs), including systematic uncertainties, that may change the shapes of the probability density functions. s_{tot} and b_{tot} represent the total number of MC signal and background events, and their normalisations are fixed from MC simulation cross–sections or higher order calculations.

The likelihood is constructed as the product of the Poisson probabilities for all fitted bins, in all signal and control regions:

$$L(\mathcal{N}_s, \mathcal{N}_b, \boldsymbol{\theta}) = \prod_{r \in \text{regions}} \prod_{i=1}^{N} \frac{(\mathrm{E}[n_i])^{n_i}}{n_i!} e^{-\mathrm{E}[n_i]}. \tag{5.3}$$

Additional terms are included in the likelihood to constrain the NPs ($\boldsymbol{\theta}$). These are treated as Gaussian distributions with standard deviation σ and a floating mean value μ. The expected mean value is $\mathrm{E}[\mu] = \theta_0$. If the fitted NP has a best–fit mean value $\hat{\theta}$ that differs from zero, it is said to be *pulled*. Likewise, if the uncertainty on the fitted NP is smaller than the expected one, it is said to be *constrained*. The effect of the limited amount of available MC statistics is also included in the fit as additional NPs (referred to as "γ" parameters). Poisson constraints terms are used for these NPs. Similarly, they can also be pulled after comparing the fitted NP mean value with the expected one. These two additional types of constraints are added as multiplicative terms to Eq. 5.3.

The $\mathcal{N}_s$, $\mathcal{N}_b$, $\boldsymbol{\theta}$, and γ parameters are free floated in the fit to the selected data (n_i) which outputs the best–fit values ($\hat{\mathcal{N}}_s$, $\hat{\mathcal{N}}_b$, $\hat{\boldsymbol{\theta}}$, and $\hat{\gamma}$) that maximise

the likelihood. Typically, $\mathcal{N}_s$ is regarded as the parameter of interest (POI) of the fit. Nonetheless, in this analysis two signal NFs are defined on different sets of selected events and are reparameterised to extract the A_c^l as the POI of the fit (see Section 5.8).

5.8 Extraction of the charge asymmetry at reconstruction level

To extract the leptonic charge asymmetry from the reconstructed leptons, a simultaneous fit to the number of observed events in the SRs and CRs, as defined in Section 5.4, is performed. The fit is based on the binned maximum profile likelihood technique described in the previous section.

The NFs of the most relevant background processes, which are constrained by the dedicated CRs, are allowed to free float in the fit: namely, $\mathcal{N}_{t\bar{t}Z}$ for $t\bar{t}Z$ production, $\mathcal{N}^e_{\text{HF}}$ and $\mathcal{N}^\mu_{\text{HF}}$ for non–prompt electrons and muons from HF decays, and $\mathcal{N}^e_{\gamma\text{-conv}}$ for electrons from γ–conversions. The variables used for the binning in the CRs are the scalar sum of the p_{T} of the selected jets in the event, H_T, for the CR–$t\bar{t}Z$; and the the p_{T} of the third (softest) lepton for the CR–HF$_e$ and CR–HF$_\mu$. These variables are chosen because their distributions present a significant shape difference between the targetted process, and the other SM backgrounds*. The total number of event yields are used in the CR–γ–conv and SRs†.

Both SRs and CRs are separated into $\Delta|\eta_\ell^{\text{BDT}}| \leq 0$ ($\Delta\eta^-$) and $\Delta|\eta_\ell^{\text{BDT}}| > 0$ ($\Delta\eta^+$) regions. Consequently, separate NFs are defined for the four main background processes, in the $\Delta\eta^-$ and $\Delta\eta^+$ regions, and allowed to free float in the fit. This separation is done to avoid any possible bias from an assumption of SM asymmetries for these processes in data‡. For the $t\bar{t}W$ signal process, two free floating NFs, $\mathcal{N}_{\Delta\eta^-}$ and $\mathcal{N}_{\Delta\eta^+}$, are defined to model the normalisation of the signal yields across all $\Delta\eta^-$ and $\Delta\eta^+$ analysis regions, respectively. Therefore, the simultaneous fit to data in all (16) analysis regions outputs results for two parameters for the signal process and eight for the main background processes.

From the charge asymmetry expression in Eq. 3.3, $N^- \equiv N(\Delta|\eta_\ell| < 0)$ and $N^+ \equiv N(\Delta|\eta_\ell| > 0)$ are defined, where N^- and N^+ represent the number of signal events distributed in the $\Delta\eta^-$ and $\Delta\eta^+$ analysis regions, respectively. The normalisation of N^- and N^+ in the fit can trivially be linked to $\mathcal{N}_{\Delta\eta^-}$ and $\mathcal{N}_{\Delta\eta^+}$. Hence, given that the objective is to extract the A_c^l directly from the

*The chosen binning is $p_{\text{T}}^\ell = [15, 20, 25, 32, 50]$ GeV for CR–HF$_e$ and CR–HF$_\mu$; and $H_T = [150, 270, 390, 550, 1000]$ GeV for CR–$t\bar{t}Z$.

†Alternative binning options were tested for the SRs without any improvements on the final result. Thus, the strategy with the most simple setup was chosen.

‡The inclusive charge asymmetries at parton level for the simulated $t\bar{t}Z$ and $t\bar{t}$ processes are $A_c^l = -0.015$ and 0.004, respectively

fit, one of the signal NFs, $\mathcal{N}_{\Delta\eta^+}$, can be reparameterised using Eq. 3.3 as

$$\mathcal{N}_{\Delta\eta^+} = \frac{\mathcal{N}_{\Delta\eta^-} \times (1 + A_c^l)}{1 - A_c^l} \frac{N^-}{N^+}. \tag{5.4}$$

In doing this, the A_c^l (along with its uncertainties) can be directly extracted as the POI of the fit. The free floating parameter $\mathcal{N}_{\Delta\eta^-}$ is also extracted from the fit. This reparameterization is not strictly necessary. However, by including the A_c^l as a parameter in the fit, its uncertainty can be calculated without relying on assumptions through error propagation. This avoids the need to combine the results from both NFs ($\mathcal{N}_{\Delta\eta^-}$ and $\mathcal{N}_{\Delta\eta^+}$) after the fit in order to calculate the A_c^l.

5.8.1 Fit to an Asimov dataset in the signal and control regions

A first fit is performed, including all systematic uncertainties, to an *Asimov* dataset in the SRs and CRs. This pseudo–dataset is constructed from the expected SM yields of the samples in the regions. This study is meant as a self–consistency check of the fit configuration setup. None of the fitted NFs are expected to be pulled (centered at 1.0). The event yields in the SRs and CRs before the fit to data are given in Table 5.5. The indicated uncertainties consider statistical as well as all experimental and theoretical systematic uncertainties described in Section 5.6. The pre-fit SRs can be seen in Figure 5.9, and Figures 5.10 and 5.11 show the pre-fit CRs in the fit.

Figure 5.12 shows the NFs of the main background processes and the A_c^l and $\mathcal{N}_{\Delta\eta^-}$ parameters for the signal after the fit (post–fit) to the Asimov dataset. All NFs are centered at 1.0 and the extracted asymmetry is $A_c^l = -0.08 \pm 0.21\,(\text{stat.}) \pm 0.04\,(\text{syst.})$. Figure 5.13 shows the MC statistical uncertainties included as additional NPs in the fit for each bin in the analysis regions. They are centered at 1.0, as expected, and the uncertainties are in all cases below 5%. Figure 5.14 shows the constraints of the systematic NPs in the fit which are not pulled as expected. The largest constraints appear on the alternative $t\bar{t}$ PS modelling, especially for the HF_e background. These are due to sizeable shape differences, when comparing the two–point systematic to the nominal prediction, most notably in the high p_{T}^{ℓ} bins of the CR–HF_e. Figure 5.15 shows the correlations between the NFs and NPs included in the CR–only fit. Sources of significant correlations arise from alternative MC modelling NPs with their corresponding background NFs, and between the NFs that affect the same background for the $\Delta\eta^-$ and $\Delta\eta^+$ regions.

Figure 5.16 shows the first ranked 20 systematics with the largest impact on the fitted A_c^l value. The leading systematic uncertainty's impact is about 20 times smaller than the statistical uncertainty on the A_c^l.

Process	CR–$t\bar{t}Z$		CR–HF$_e$		CR–HF$_\mu$		CR–γ–conv	
	$\Delta\eta^-$	$\Delta\eta^+$	$\Delta\eta^-$	$\Delta\eta^+$	$\Delta\eta^-$	$\Delta\eta^+$	$\Delta\eta^-$	$\Delta\eta^+$
$t\bar{t}W$ (QCD)	1.8 ± 0.4	1.49 ± 0.19	1.18 ± 0.19	1.13 ± 0.18	1.72 ± 0.20	1.37 ± 0.28	4.1 ± 0.7	2.92 ± 0.18
$t\bar{t}W$ (EW)	0.18 ± 0.07	0.16 ± 0.06	0.10 ± 0.04	0.09 ± 0.04	0.09 ± 0.04	0.14 ± 0.05	0.23 ± 0.08	0.36 ± 0.12
$t\bar{t}Z$	107 ± 6	107 ± 6	1.42 ± 0.23	1.5 ± 0.4	2.20 ± 0.23	2.00 ± 0.14	4.04 ± 0.19	3.65 ± 0.32
HF$_e$	–	–	350 ± 40	362 ± 27	0.18 ± 0.11	0.20 ± 0.09	1.0 ± 0.6	0.67 ± 0.35
HF$_\mu$	0.14 ± 0.08	0.19 ± 0.09	0.20 ± 0.09	0.28 ± 0.10	520 ± 40	530 ± 50	0.9 ± 0.5	1.1 ± 0.9
γ–conv.	0.55 ± 0.14	0.41 ± 0.13	3.8 ± 2.5	4.7 ± 2.9	2.6 ± 2.4	3.3 ± 2.5	18.8 ± 1.4	17.5 ± 1.3
$t\bar{t}H$	3.3 ± 0.4	3.20 ± 0.32	0.87 ± 0.13	0.89 ± 0.11	1.18 ± 0.11	1.22 ± 0.22	1.48 ± 0.20	1.5 ± 0.4
tZq	12.6 ± 2.2	11.0 ± 1.9	0.48 ± 0.11	0.43 ± 0.09	0.95 ± 0.18	0.81 ± 0.15	0.68 ± 0.12	0.70 ± 0.13
WZ/ZZ+jets	12 ± 4	12 ± 4	3.0 ± 0.9	3.3 ± 1.0	7.2 ± 2.4	7.9 ± 2.5	3.1 ± 0.9	2.9 ± 0.8
Other	10.7 ± 3.3	10.2 ± 3.3	14 ± 4	13 ± 5	17 ± 7	17 ± 6	1.6 ± 0.8	1.5 ± 0.6
SM total	148 ± 10	146 ± 10	380 ± 40	387 ± 28	550 ± 40	560 ± 50	35.9 ± 2.4	32.9 ± 2.3
Data	156	176	315	373	551	592	34	40

Process	SR–1b–lowN_{jets}		SR–1b–highN_{jets}		SR–2b–lowN_{jets}		SR–2b–highN_{jets}	
	$\Delta\eta^-$	$\Delta\eta^+$	$\Delta\eta^-$	$\Delta\eta^+$	$\Delta\eta^-$	$\Delta\eta^+$	$\Delta\eta^-$	$\Delta\eta^+$
$t\bar{t}W$ (QCD)	19.0 ± 2.8	17 ± 4	9.2 ± 1.1	8.2 ± 1.1	25 ± 7	21 ± 6	14.7 ± 3.4	12.2 ± 1.9
$t\bar{t}W$ (EW)	1.06 ± 0.34	1.3 ± 0.4	1.05 ± 0.34	1.07 ± 0.34	1.2 ± 0.4	1.3 ± 0.4	1.8 ± 0.6	1.6 ± 0.5
$t\bar{t}Z$	12.0 ± 1.0	12.1 ± 1.1	15.5 ± 1.4	15.5 ± 1.1	11.4 ± 1.4	10.8 ± 1.4	26.2 ± 1.8	25.8 ± 1.7
HF$_e$	7.2 ± 1.2	7.5 ± 1.5	1.7 ± 0.7	1.6 ± 0.6	0.7 ± 0.5	0.6 ± 0.5	0.69 ± 0.35	0.37 ± 0.19
HF$_\mu$	12.5 ± 2.0	13 ± 4	3.2 ± 0.8	3.5 ± 1.3	1.35 ± 0.34	1.11 ± 0.33	1.0 ± 0.4	0.9 ± 0.5
γ–conv.	6.7 ± 0.9	6.1 ± 1.0	3.1 ± 0.5	3.4 ± 0.8	6.1 ± 0.8	6.9 ± 0.8	4.4 ± 0.7	4.6 ± 0.6
$t\bar{t}H$	5.5 ± 0.8	5.6 ± 0.8	8.6 ± 0.8	8.7 ± 0.9	5.5 ± 1.1	5.5 ± 1.0	14.1 ± 1.8	14.2 ± 1.7
tZq	5.1 ± 0.9	4.2 ± 0.7	1.40 ± 0.31	1.15 ± 0.27	2.8 ± 0.5	2.3 ± 0.4	1.92 ± 0.34	1.64 ± 0.30
WZ/ZZ+jets	15 ± 4	14 ± 4	8.0 ± 2.8	7.6 ± 2.5	2.9 ± 0.9	2.2 ± 0.7	2.2 ± 0.7	2.2 ± 0.7
Other	5.6 ± 2.0	5.1 ± 1.6	4.5 ± 2.4	4.7 ± 1.5	2.6 ± 1.1	2.9 ± 1.3	10 ± 6	9 ± 5
SM total	89 ± 6	85 ± 7	56 ± 6	56 ± 6	59 ± 9	55 ± 7	77 ± 8	73 ± 7
Data	94	89	50	69	84	81	89	81

Table 5.5: The predicted and observed numbers of events in the control and signal regions. The predictions are shown before the fit to data. The indicated uncertainties consider statistical as well as all experimental and theoretical systematic uncertainties. Background categories with event yields that contribute less than 0.01 to a region are shown as –.

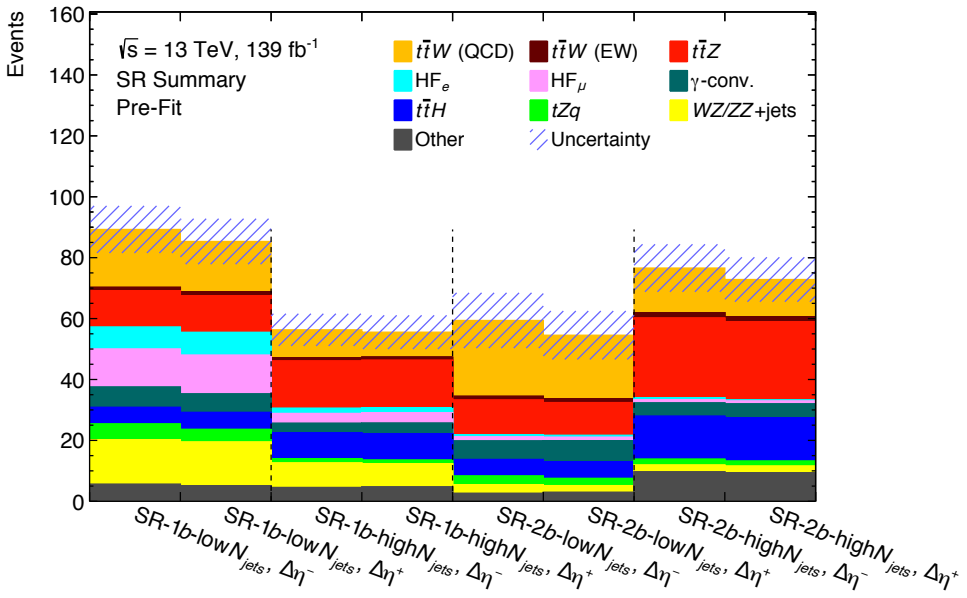


Figure 5.9: Pre–fit MC predictions for $\Delta|\eta_\ell^{\mathrm{BDT}}| \leq 0$ ($\Delta\eta^-$) and $\Delta|\eta_\ell^{\mathrm{BDT}}| > 0$ ($\Delta\eta^+$) in the four SRs. The error band includes the total pre–fit uncertainties.

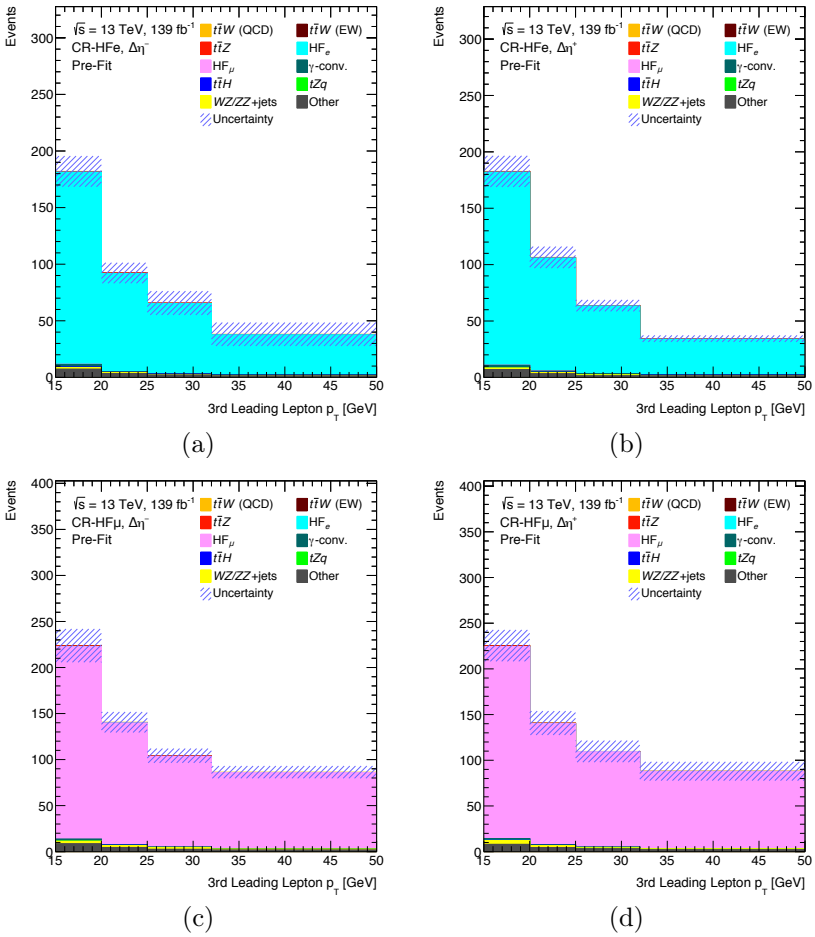


Figure 5.10: Pre–fit MC predictions in (a,b) CR–HF$_e$ and (c,d) CR–HF$_\mu$. The distributions show the p_{T} of the third lepton (electron or muon), which is the variable that is used for the binned likelihood fit. The regions are separated between $\Delta|\eta_\ell^{\mathrm{BDT}}| \leq 0$ ($\Delta\eta^-$) and $\Delta|\eta_\ell^{\mathrm{BDT}}| > 0$ ($\Delta\eta^+$). The error bands include the total pre–fit uncertainties. Events with the p_{T} of the third lepton above 50 GeV are included in the overflow bins.

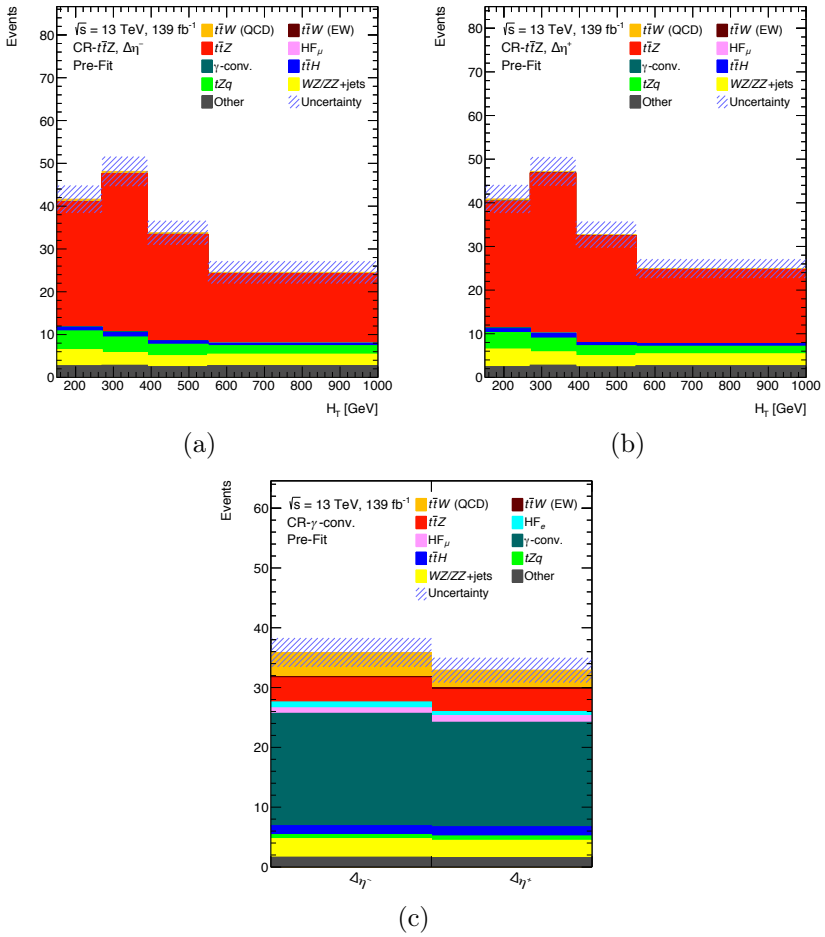


Figure 5.11: Pre–fit MC predictions in (a,b) CR–$t\bar{t}Z$ and (c) CR–γ–conv. The distributions are shown for the variables that are used for the binned likelihood fit: H_T for CR–$t\bar{t}Z$ and the total event yields for CR–γ–conv. The regions are separated between $\Delta|\eta_\ell^{\mathrm{BDT}}| \leq 0$ ($\Delta\eta^-$) and $\Delta|\eta_\ell^{\mathrm{BDT}}| > 0$ ($\Delta\eta^+$). The error bands include the total pre–fit uncertainties. Events with an H_T above 1 TeV are included in the overflow bins of (a) and (b).

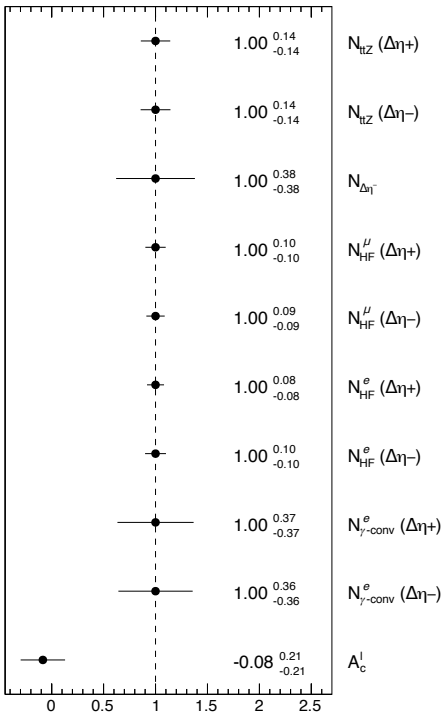


Figure 5.12: Normalisation factors for the major background processes, together with $\mathcal{N}_{\Delta\eta^-}$ for $t\bar{t}W$ and the A_c^l value extracted from the fit to Asimov data in the CRs and SRs. The normalisation factors, $\mathcal{N}_{t\bar{t}Z}$, $\mathcal{N}^e_{\gamma\text{-conv}}$, $\mathcal{N}^e_{\text{HF}}$ and $\mathcal{N}^\mu_{\text{HF}}$, are obtained separately for $\Delta|\eta_\ell^{\text{BDT}}| \leq 0$ ($\Delta\eta^-$) and $\Delta|\eta_\ell^{\text{BDT}}| > 0$ ($\Delta\eta^+$). The indicated uncertainties consider statistical as well as systematic uncertainties.

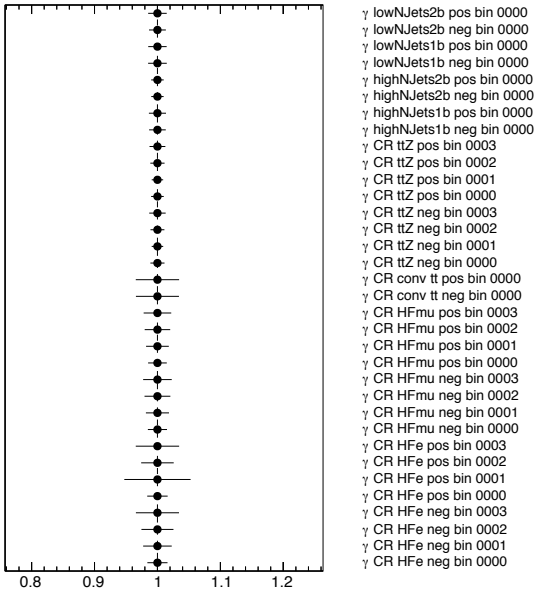


Figure 5.13: Post–fit MC statistical (γ) uncertainties after the fit to Asimov data for each of the bins throughout all analysis regions.

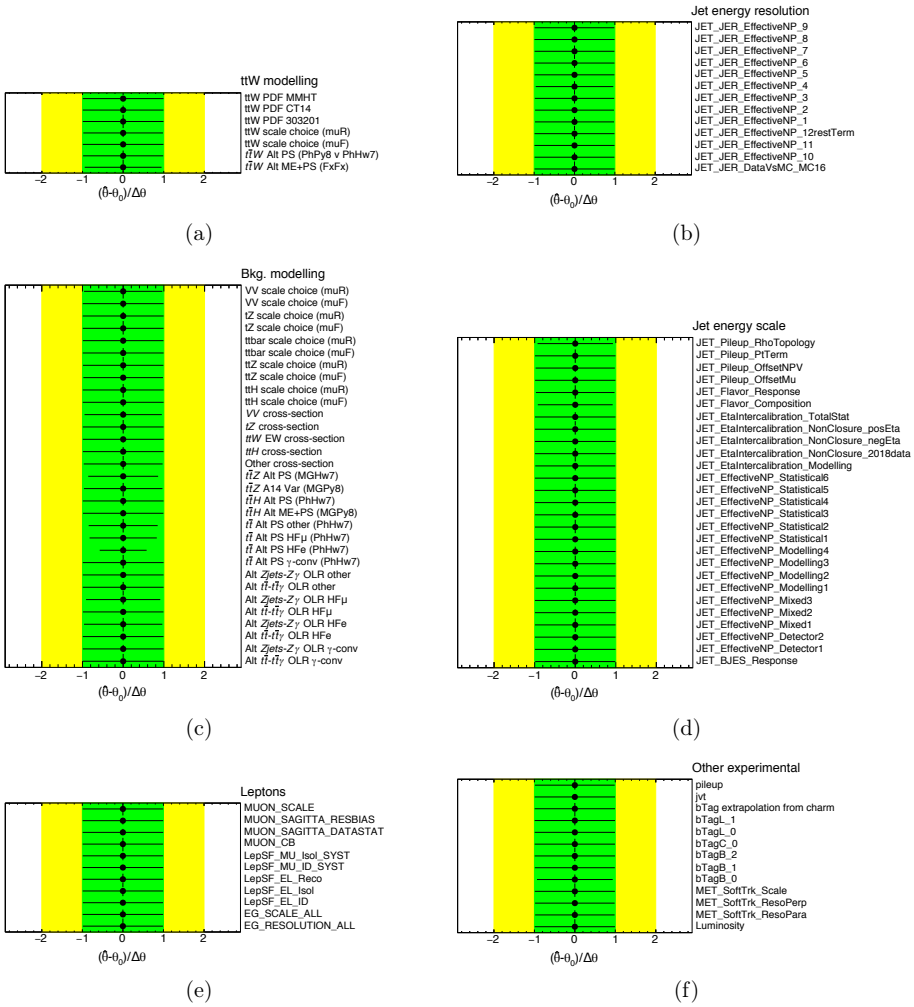


Figure 5.14: Post–fit systematic uncertainties after the fit to Asimov data in all analysis regions. The green (yellow) areas represent the $\pm 1(2)\sigma$ band on the pre–fit uncertainty $\Delta\theta$. The black points show the best–fit values $\hat{\theta}$ of the NPs, with the error bars representing the post–fit uncertainties. Each NP is shown relative to its nominal value, θ_0, and in units of its pre–fit uncertainty. Theory uncertainties are separated corresponding to (a) only signal or (c) background modelling; experimental uncertainties are separated in (b) JES, (d) JER, (e) leptons and (f) other experimental systematics, i.e. pile–up, JVT, b–tagging, $E_{\mathrm{T}}^{\mathrm{miss}}$ soft term and luminosity.

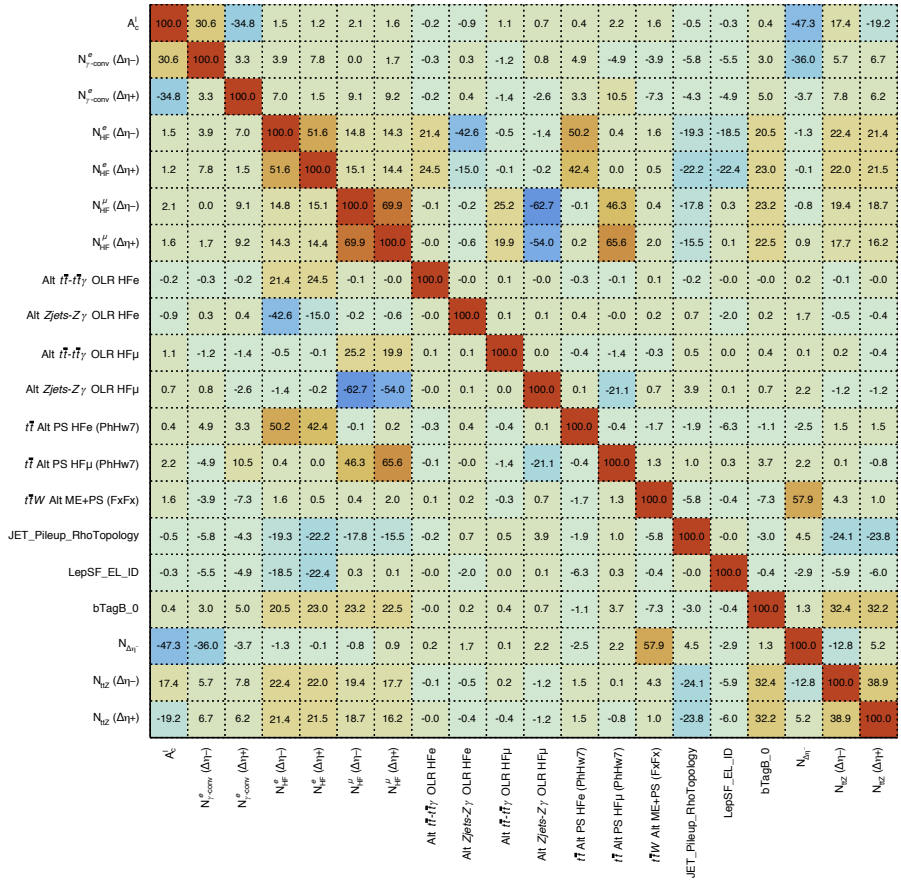


Figure 5.15: Correlation matrix between the NFs and the NPs after the fit to Asimov data in all analysis regions. Only NPs with correlations above 20% are displayed.

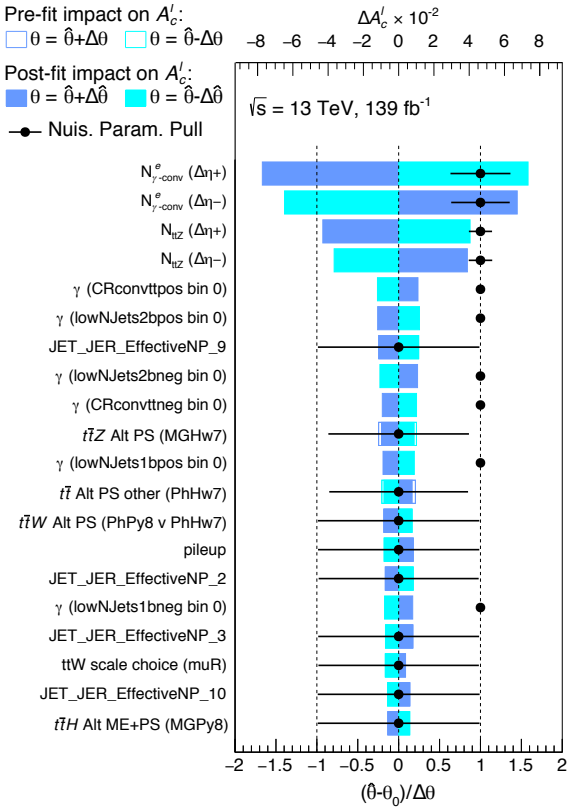


Figure 5.16: The most relevant (20) systematic uncertainties ranked by their impact on the A_c^l parameter at reconstruction level (top axis values). The impact ΔA_c^l of the uncertainties is shown before and after the fit to Asimov data in the signal and control regions. It is computed by comparing the nominal best–fit value of A_c^l, with the result of the fit when fixing the considered NP to its best–fit value, $\hat{\theta}$, shifted by its pre–fit (post–fit) uncertainties $\pm\Delta\theta$ ($\pm\Delta\hat{\theta}$). Pulls introduced by the fitting procedure are also shown (bottom axis values) and are identical to those in Figure 5.14. The top ranked systematic uncertainty is about 20 times smaller than the statistical uncertainty for the A_c^l.

5.8.2 Fit to real data in the control regions

Prior to injecting real data in the SRs, an intermediate fit (including all systematic uncertainties) is performed with real data but only using the bins in the CRs (referred to as the "CR–only" fit). Consequently, the signal parameters (A_c^l and $\mathcal{N}_{\Delta\eta^-}$) cannot be properly constrained due to the low signal contamination in the CRs, and are thus left out of this fit. The goal of this fit is to check that the enriched CRs, and therefore the MC modelling of the dominant backgrounds, show a good level of agreement with data, while keeping the signal yields to their SM expectation[§].

Figure 5.17 shows the NFs of the main background processes after the CR–only fit to data. The normalisation of the main backgrounds is found to be compatible with the MC prediction within 2σ. Figure 5.18 shows the post–fit MC statistical uncertainties. The size of the uncertainties remains unchanged with respect to the Asimov fit. However, the γ parameter associated with the second bin of the $\Delta\eta^+$ CR–HF$_e$ is slightly pulled. Figure 5.14 shows the post–fit pulls and constraints of the systematic NPs. The constraints are equivalent from those in the Asimov fit. Despite, some small pulls (less than 0.6σ) are seen for the HF$_e$ background alternative modelling for the $t\bar{t}$ PS and Z+jets two–point systematics. Figure 5.20 shows the correlations between the NFs and NPs included in the CR–only fit. Similar correlations as those in the Asimov fit are seen.

[§]Comparisons between data and the post–fit SM predictions in the CRs are given in Figure 5.21 for CR–HF$_e$ and CR–HF$_\mu$, and in Figure 5.22 for CR–$t\bar{t}Z$ and CR–γ–conv.

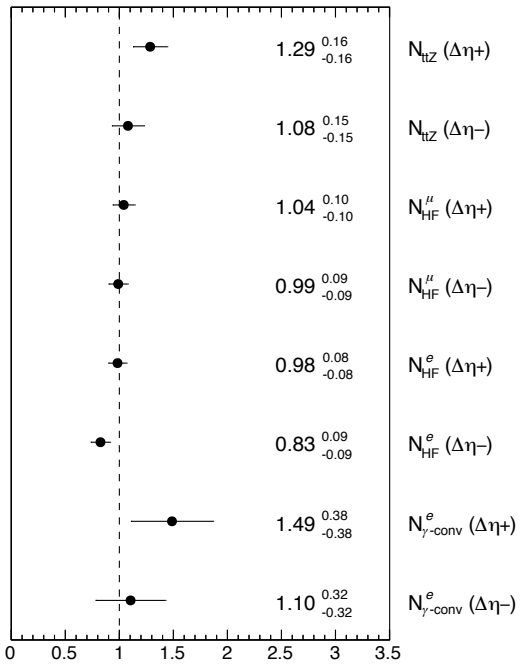


Figure 5.17: Normalisation factors for the major background processes, for the CR–only fit to data. The normalisation factors, $\mathcal{N}_{t\bar{t}Z}$, $\mathcal{N}^{e}_{\gamma\text{-conv}}$, $\mathcal{N}^{e}_{\text{HF}}$ and $\mathcal{N}^{\mu}_{\text{HF}}$, are obtained separately for $\Delta|\eta_{\ell}^{\text{BDT}}| \leq 0$ ($\Delta\eta^{-}$) and $\Delta|\eta_{\ell}^{\text{BDT}}| > 0$ ($\Delta\eta^{+}$). The indicated uncertainties consider statistical as well as systematic uncertainties.

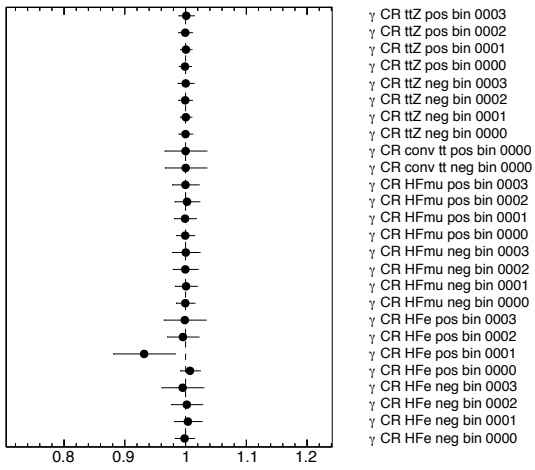


Figure 5.18: Post–fit MC statistical (γ) uncertainties after the CR–only fit to data for each of the bins throughout all analysis regions.

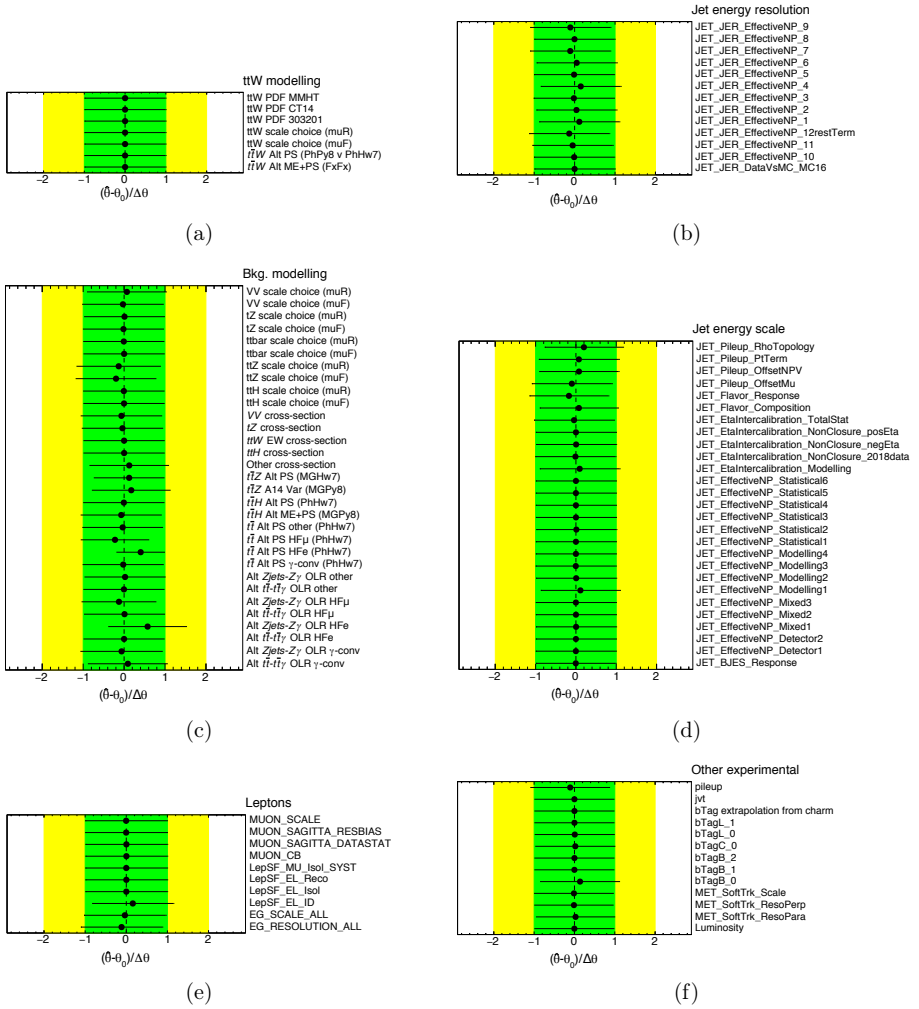


Figure 5.19: Post–fit systematic uncertainties after the CR–only fit to data in all analysis regions. The green (yellow) areas represent the $\pm 1(2)\sigma$ band on the pre–fit uncertainty $\Delta\theta$. The black points show the best–fit values $\hat{\theta}$ of the NPs, with the error bars representing the post–fit uncertainties. Each NP is shown relative to its nominal value, θ_0, and in units of its pre–fit uncertainty. Theory uncertainties are separated corresponding to (a) only signal or (c) background modelling; experimental uncertainties are separated in (b) JES, (d) JER, (e) leptons and (f) other experimental systematics, i.e. pile–up, JVT, b–tagging, $E_{\mathrm{T}}^{\mathrm{miss}}$ soft term and luminosity.

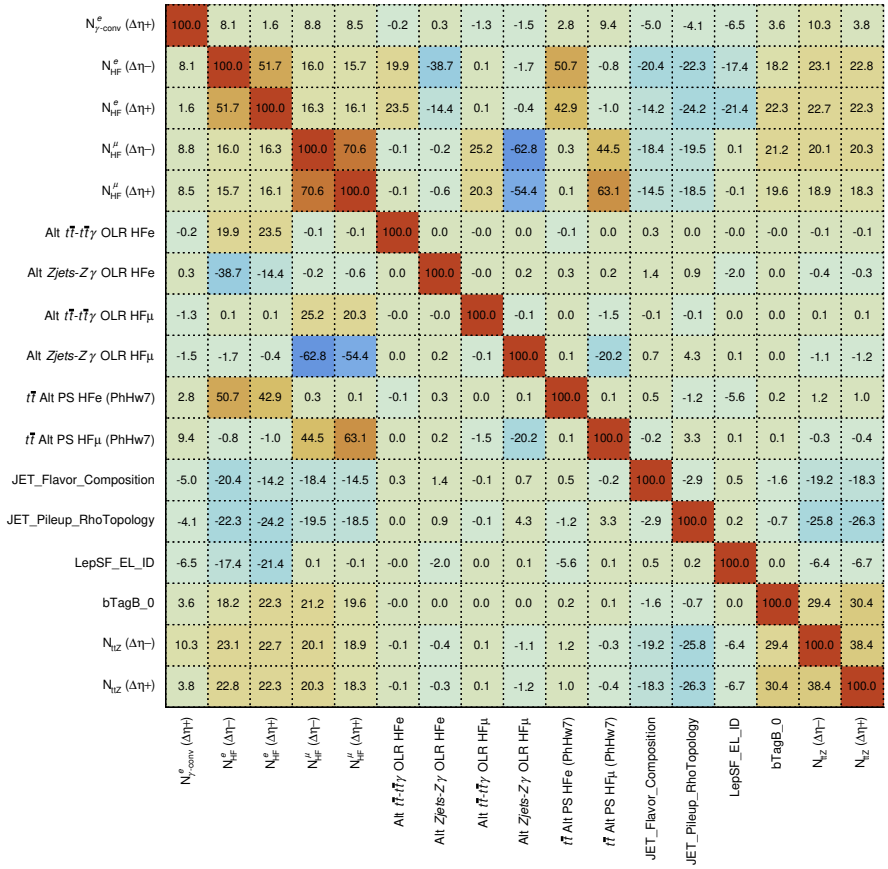


Figure 5.20: Correlation matrix between the NFs and the NPs in the CR–only fit to data in all analysis regions. Only NPs with correlations above 20% are displayed.

5.8.3 Fit to real data in the signal and control regions

Finally, a fit with real data in all analysis regions to extract the reconstructed A_c^l is performed. The event yields in the SRs and CRs after the fit to data are given in Table 5.6. The indicated uncertainties consider MC statistical as well as all experimental and theoretical systematic uncertainties described in Section 5.6. Comparisons between data and the post–fit SM predictions for the variables that are used for the binned likelihood fit are given in Figure 5.21 for CR–HF$_e$ and CR–HF$_\mu$, and in Figure 5.22 for CR–$t\bar{t}Z$ and CR–γ–conv. The data and the post–fit predictions for $\Delta\eta^-$ and $\Delta\eta^+$ in the four SRs are shown in Figure 5.23. In addition, Figure 5.24 shows the post–fit distributions of the five input BDT variables, the BDT scores of the two even leptons, and the N_{jets} and total charge of the leptons.

Process	CR–$t\bar{t}Z$		CR–HF$_e$		CR–HF$_\mu$		CR–γ–conv	
	$\Delta\eta^-$	$\Delta\eta^+$	$\Delta\eta^-$	$\Delta\eta^+$	$\Delta\eta^-$	$\Delta\eta^+$	$\Delta\eta^-$	$\Delta\eta^+$
$t\bar{t}W$ (QCD)	3.2 ± 0.7	2.2 ± 0.7	1.8 ± 0.5	1.7 ± 0.5	2.6 ± 0.8	1.8 ± 0.8	7.0 ± 1.3	4.4 ± 1.3
$t\bar{t}W$ (EW)	0.18 ± 0.06	0.16 ± 0.05	0.10 ± 0.03	0.09 ± 0.03	0.09 ± 0.03	0.14 ± 0.04	0.23 ± 0.07	0.36 ± 0.11
$t\bar{t}Z$	114 ± 13	138 ± 14	1.45 ± 0.27	1.7 ± 0.4	2.3 ± 0.4	2.55 ± 0.35	4.3 ± 0.6	4.6 ± 0.6
HF$_e$	–	–	290 ± 18	346 ± 20	0.15 ± 0.02	0.19 ± 0.02	0.59 ± 0.27	0.52 ± 0.17
HF$_\mu$	0.133 ± 0.012	0.201 ± 0.020	0.195 ± 0.018	0.277 ± 0.029	516 ± 25	556 ± 25	0.8 ± 0.4	1.3 ± 0.8
γ–conv.	0.40 ± 0.18	0.52 ± 0.16	2.8 ± 2.2	6 ± 4	1.9 ± 2.0	4.2 ± 3.4	14 ± 6	22 ± 7
$t\bar{t}H$	3.3 ± 0.4	3.23 ± 0.31	0.86 ± 0.13	0.87 ± 0.10	1.16 ± 0.11	1.19 ± 0.22	1.49 ± 0.20	1.6 ± 0.4
tZq	12.6 ± 2.2	11.0 ± 1.9	0.47 ± 0.10	0.42 ± 0.08	0.95 ± 0.17	0.79 ± 0.14	0.68 ± 0.11	0.70 ± 0.12
WZ/ZZ+jets	10.2 ± 2.9	10.6 ± 3.1	2.6 ± 0.7	2.8 ± 0.7	6.3 ± 1.7	6.7 ± 1.8	2.6 ± 0.7	2.5 ± 0.6
Other	10.8 ± 3.2	10.0 ± 2.9	14 ± 4	13 ± 5	18 ± 7	18 ± 6	1.7 ± 0.8	1.7 ± 0.6
SM total	155 ± 12	175 ± 13	315 ± 18	373 ± 19	550 ± 23	591 ± 24	33 ± 6	40 ± 6
Data	156	176	315	373	551	592	34	40

Process	SR–1b–lowN_{jets}		SR–1b–highN_{jets}		SR–2b–lowN_{jets}		SR–2b–highN_{jets}	
	$\Delta\eta^-$	$\Delta\eta^+$	$\Delta\eta^-$	$\Delta\eta^+$	$\Delta\eta^-$	$\Delta\eta^+$	$\Delta\eta^-$	$\Delta\eta^+$
$t\bar{t}W$ (QCD)	32 ± 6	27 ± 6	14 ± 4	12.1 ± 3.4	46 ± 9	36 ± 8	26 ± 6	19 ± 5
$t\bar{t}W$ (EW)	1.04 ± 0.32	1.3 ± 0.4	1.04 ± 0.32	1.05 ± 0.32	1.2 ± 0.4	1.3 ± 0.4	1.8 ± 0.5	1.6 ± 0.5
$t\bar{t}Z$	12.4 ± 2.0	15.0 ± 2.2	16.0 ± 2.2	19.6 ± 2.3	12.3 ± 2.3	14.3 ± 2.6	27.6 ± 3.3	33.2 ± 3.5
HF$_e$	6.4 ± 1.0	6.8 ± 0.8	1.5 ± 0.5	1.7 ± 0.4	0.40 ± 0.20	0.79 ± 0.35	0.45 ± 0.14	0.39 ± 0.14
HF$_\mu$	12.5 ± 1.5	13.6 ± 2.5	3.1 ± 0.6	3.6 ± 0.9	1.30 ± 0.23	1.19 ± 0.19	1.04 ± 0.29	0.9 ± 0.5
γ–conv.	4.9 ± 2.3	7.7 ± 2.6	2.3 ± 1.1	4.3 ± 1.6	4.6 ± 2.1	8.8 ± 2.9	3.3 ± 1.5	5.9 ± 1.9
$t\bar{t}H$	5.4 ± 0.8	5.5 ± 0.8	8.4 ± 0.8	8.6 ± 0.8	5.5 ± 1.1	5.6 ± 1.0	14.3 ± 1.7	14.4 ± 1.7
tZq	5.0 ± 0.9	4.1 ± 0.7	1.38 ± 0.27	1.16 ± 0.24	2.8 ± 0.5	2.3 ± 0.4	1.93 ± 0.33	1.65 ± 0.29
WZ/ZZ+jets	12.6 ± 3.0	12.3 ± 3.0	6.7 ± 2.0	6.5 ± 1.8	2.5 ± 0.7	1.9 ± 0.5	1.9 ± 0.6	1.9 ± 0.5
Other	6.0 ± 2.1	5.2 ± 1.6	3.6 ± 1.8	4.6 ± 1.4	2.9 ± 1.2	3.3 ± 1.3	8 ± 4	8 ± 4
SM total	99 ± 6	98 ± 6	58 ± 4	63 ± 4	80 ± 8	75 ± 7	85 ± 6	86 ± 5
Data	94	89	50	69	84	81	89	81

Table 5.6: The predicted and observed numbers of events in the control and signal regions. The predictions are shown after the fit to data. The indicated uncertainties consider statistical as well as all experimental and theoretical systematic uncertainties. Background categories with event yields that contribute less than 0.01 to a region are shown as –.

The normalisation factors for the major background processes, $\mathcal{N}_{t\bar{t}Z}$, $\mathcal{N}^e_{\text{HF}}$, $\mathcal{N}^\mu_{\text{HF}}$, and $\mathcal{N}^e_{\gamma\text{-conv}}$ (all obtained separately for $\Delta\eta^-$ and $\Delta\eta^+$), together with $\mathcal{N}_{\Delta\eta^-}$ and the A_c^l value for the $t\bar{t}W$ signal, are given in Figure 5.25. The

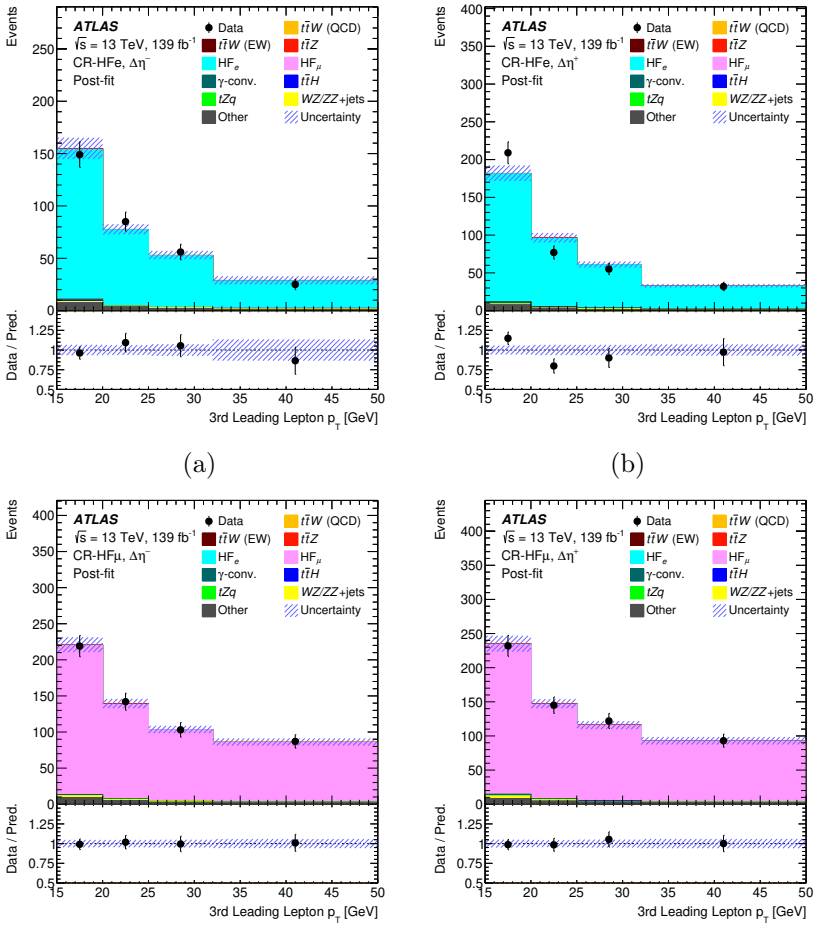


Figure 5.21: Comparison between data and the post–fit predictions in (a,b) CR–HF_e and (c,d) CR–HF_μ. The distributions show the p_T of the third lepton (electron or muon), which is the variable that is used for the binned likelihood fit. The regions are separated between $\Delta|\eta_\ell^{\mathrm{BDT}}| \leq 0$ ($\Delta\eta^-$) and $\Delta|\eta_\ell^{\mathrm{BDT}}| > 0$ ($\Delta\eta^+$). The error bands include the total uncertainties in the post–fit predictions. The ratios of the data to the total post–fit predictions are shown in the lower panels. Events with the p_T of the third lepton above 50 GeV are included in the overflow bins.

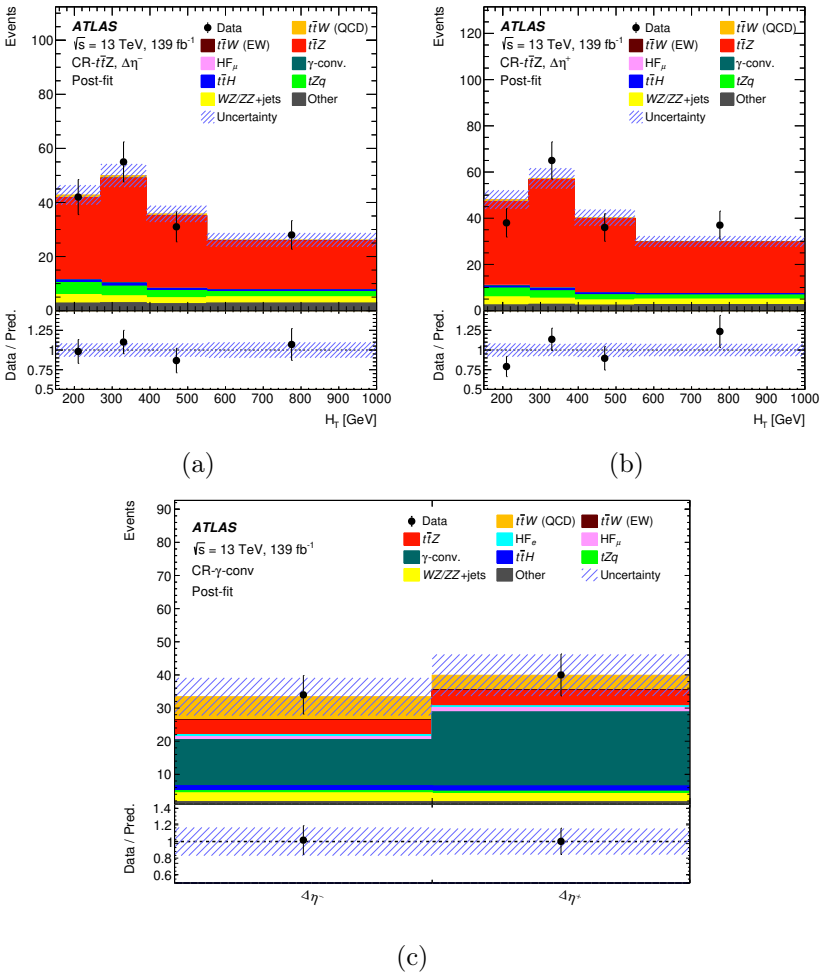


Figure 5.22: Comparison between data and the post–fit predictions in (a,b) CR–$t\bar{t}Z$ and (c) CR–γ–conv. The distributions are shown for the variables that are used for the binned likelihood fit: H_T for CR–$t\bar{t}Z$ and the total event yields for CR–γ–conv. The regions are separated between $\Delta|\eta_\ell^{\mathrm{BDT}}| \leq 0$ ($\Delta\eta^-$) and $\Delta|\eta_\ell^{\mathrm{BDT}}| > 0$ ($\Delta\eta^+$). The error bands include the total uncertainties in the post–fit predictions. The ratios of the data to the total post–fit predictions are shown in the lower panels. Events with an H_T above 1 TeV are included in the overflow bins of (a) and (b).

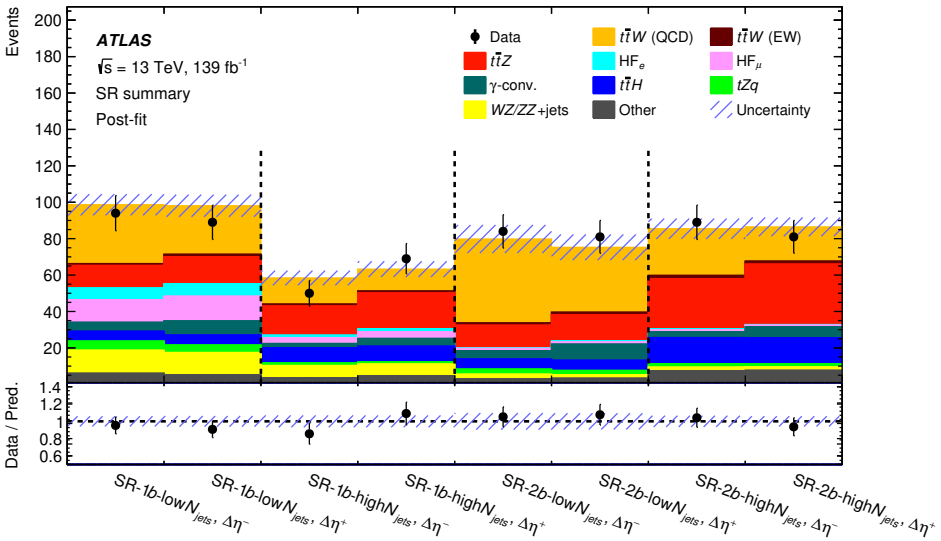


Figure 5.23: Comparison between data and the post–fit predictions in the four SRs for $\Delta|\eta_\ell^{\mathrm{BDT}}| \leq 0$ ($\Delta\eta^-$) and $\Delta|\eta_\ell^{\mathrm{BDT}}| > 0$ ($\Delta\eta^+$). The error band includes the total uncertainties of the post–fit predictions. The ratio of the data to the total post–fit predictions is shown in the lower panel.

normalisation factor for the $t\bar{t}W$ process is found to be (within its uncertainty) compatible with the latest ATLAS and CMS $t\bar{t}W$ cross–section measurements[¶] [234, 235]. Tests using MC simulations were also performed to validate that the extracted A_c^l value is not biased by the absolute normalisation of the $t\bar{t}W$ process.

The normalisation factors for some of the background processes (in particular $\mathcal{N}_{t\bar{t}Z}$ and $\mathcal{N}^e_{\gamma\text{-conv}}$) show small differences (within 1σ) between $\Delta\eta^-$ and $\Delta\eta^+$. In the nominal fit, due to the independent normalisation factors for $\Delta\eta^-$ and $\Delta\eta^+$ in the CRs, the background asymmetries are measured and are, in any case, less than 1.7σ significant. To account for the possibility that these measured asymmetries are due to statistical fluctuations, an alternative fit is performed. This fit follows a two–step procedure. First, NFs for the backgrounds are obtained from a fit to real data using only the CRs. Then, these NFs are used to scale the expected background yields in the SRs, and a new fit is performed with real data in the CRs and this new pseudo–data in the SRs. Two configurations are tested for the backgrounds: one with the nominal splitted NF scheme, and one where only one NF is assigned to each of the

[¶]Note that the $t\bar{t}W$ signal sample is normalised to the SHERPA cross–section prediction, which is about 18% smaller than the reference cross–section in Ref. [236]. By performing an *ad hoc* scaling, the obtained NF is in agreement.

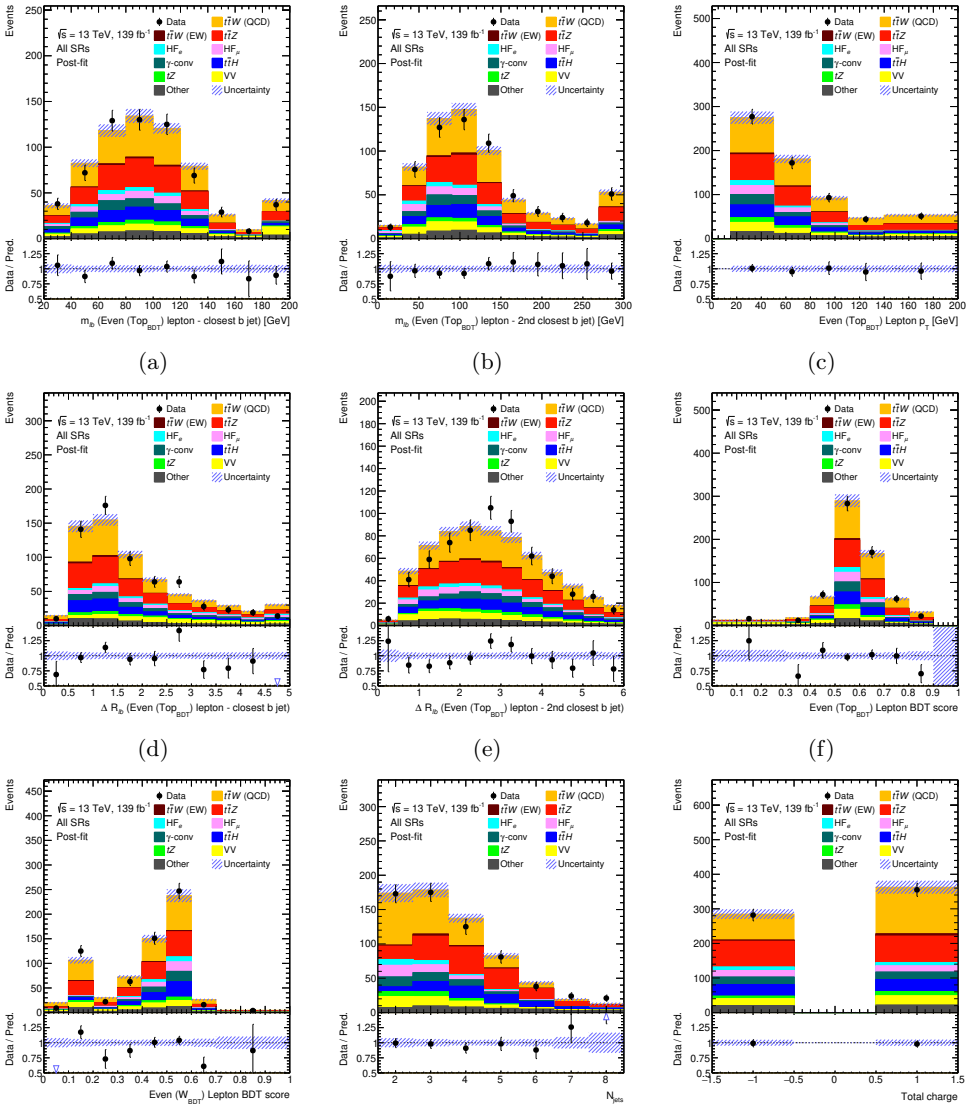


Figure 5.24: Comparison between data and the post–fit predictions, combining the four SRs, for the BDT input variables of the selected even lepton: (a) $m_{\ell b_0}$, (b) $m_{\ell b_1}$, (c) lepton p_{T}, (d) $\Delta R_{\ell b_0}$, (e) $\Delta R_{\ell b_1}$. Then, BDT discriminator values for the two even leptons: (f) selected lepton, (g) background lepton. Finally, (h) N_{jets} and (i) total charge of the leptons. The error band includes the total uncertainties of the post–fit predictions. The ratio of the data to the total post–fit predictions is shown in the lower panel.

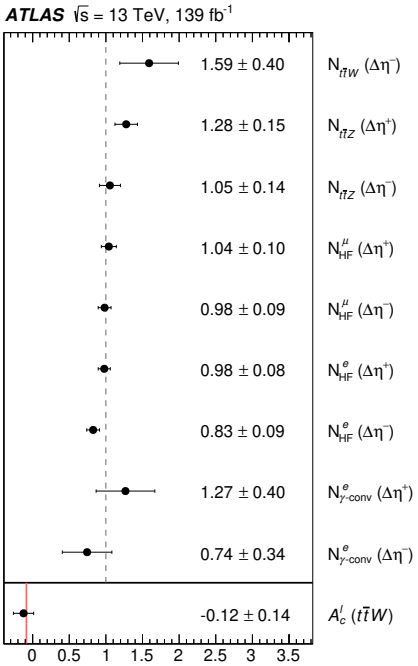


Figure 5.25: Normalisation factors for the major background processes, together with $\mathcal{N}_{\Delta\eta^-}$ for $t\bar{t}W$ and the A_c^l value extracted from the fit to data in the CRs and SRs. The normalisation factors, $\mathcal{N}_{t\bar{t}Z}$, $\mathcal{N}^e_{\gamma\text{-conv}}$, $\mathcal{N}^e_{\text{HF}}$ and $\mathcal{N}^\mu_{\text{HF}}$, are obtained separately for $\Delta|\eta_\ell^{\text{BDT}}| \leq 0$ ($\Delta\eta^-$) and $\Delta|\eta_\ell^{\text{BDT}}| > 0$ ($\Delta\eta^+$). The indicated uncertainties consider statistical as well as systematic uncertainties. The solid vertical line in the last entry shows the A_c^l SM expectation, calculated using the $t\bar{t}W$ Sherpa simulation.

backgrounds (thus fixing their asymmetries to the SM expectation). The difference between the results of these two fit configurations is assigned as an extra systematic uncertainty in the extracted A_c^l value. This uncertainty (denoted as $\Delta\eta^\pm$ CR–dependency) is found to be 0.046, and is the leading systematic uncertainty. Since this uncertainty is calculated after the fit, it is added in quadrature to the total A_c^l uncertainty (it is not part of the uncertainty bands of the relevant figures). Future iterations of the analysis, with larger datasets, will allow for a better understanding of the reported background asymmetries, meaning that this extra uncertainty will no longer be needed or conclusively be supported.

The leptonic charge asymmetry in $t\bar{t}W$ at reconstruction level is found to be

$$A_c^l(t\bar{t}W) = -0.12 \pm 0.14\,(\text{stat.}) \pm 0.05\,(\text{syst.}).$$

This is consistent with the SM expectation of

$$A_c^l(t\bar{t}W)_{\text{SM}} = -0.084\,^{+0.005}_{-0.003}\,(\text{scale}) \pm 0.006\,(\text{MC stat.}),$$

calculated using the nominal $t\bar{t}W$ SHERPA simulation. The contributions from the most relevant uncertainties are summarised in Table 5.7. The uncertainties are symmetrised and grouped into several type–related categories, and are shown together with the total systematic and statistical uncertainties. The result is severely limited by the statistical uncertainty of the data. The dominant systematic uncertainties are the $\Delta\eta^{\pm}$ CR–dependency, the JER, as well as the modelling uncertainties of the $t\bar{t}W$ and $t\bar{t}Z$ MC processes.

	$\Delta A_c^l(t\bar{t}W)$
Experimental uncertainties	
Jet energy resolution	0.013
Pile-up	0.007
b-tagging	0.005
Leptons	0.004
$E_{\mathrm{T}}^{\mathrm{miss}}$	0.004
Jet energy scale	0.003
Luminosity	0.001
Theoretical uncertainties	
$t\bar{t}W$ modelling	0.013
$t\bar{t}Z$ modelling	0.010
$\mathrm{HF}_{e/\mu}$ modelling	0.006
$t\bar{t}H$ modelling	0.005
Other uncertainties	
$\Delta\eta^{\pm}$ CR–dependency	0.046
MC statistical uncertainty	0.019
Data statistical uncertainty	0.136
Total uncertainty	0.145

Table 5.7: List of the most relevant systematic and statistical uncertainties in the extracted leptonic charge asymmetry $A_c^l(t\bar{t}W)$ from the simultaneous fit at reconstruction level. For this table, the uncertainties are symmetrised and grouped into categories. The sum in quadrature of the individual uncertainties is not necessarily equal to the total uncertainty due to correlations introduced by the fit.

Figure 5.26 shows the best–fit values, the pulls, and constraints of the NPs that are included in the fit. The alternative signal modelling NP is slightly pulled (about -0.5σ from its pre–fit value). Given that the fitted value for $\mathcal{N}_{\Delta\eta^-}$ is bigger than one, and that the alternative $t\bar{t}W$ sample predicts lower yields than the nominal one, this pull helps to push the $t\bar{t}W$ event yields in the direction of data. Post–fit MC statistical uncertainties in the CRs are analogous to those in the CR–only fit (see Figure 5.18), and are not pulled for the SR bins.

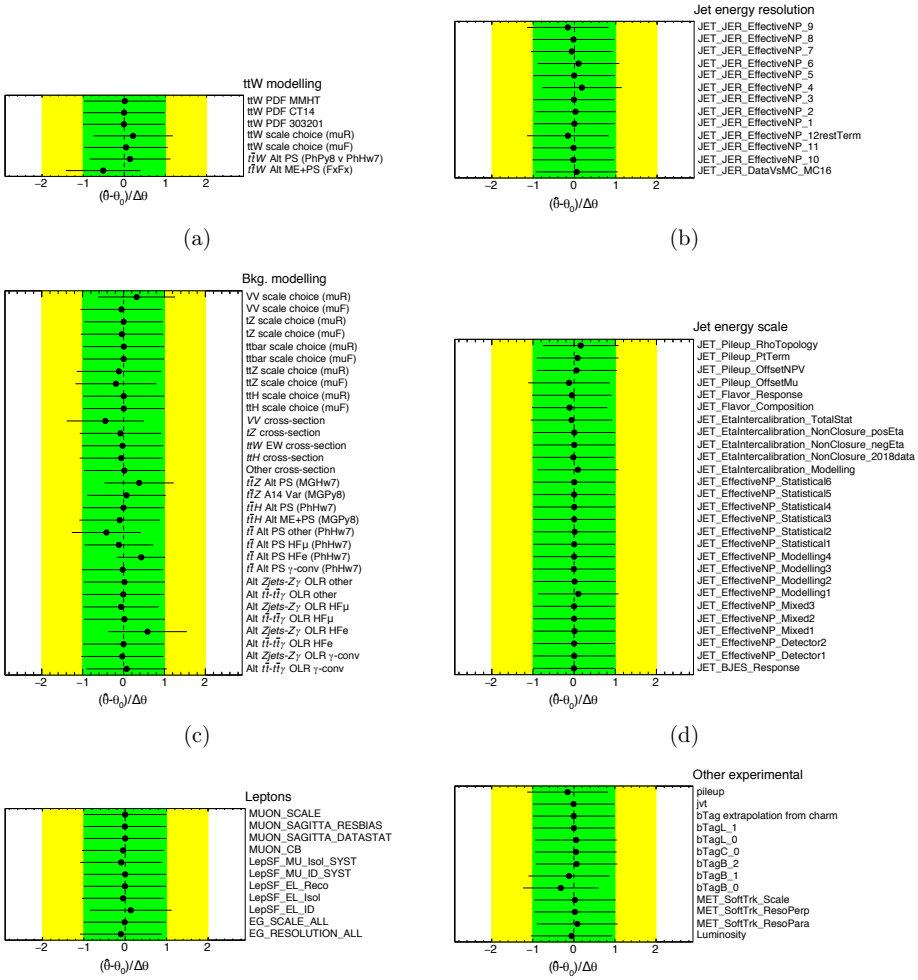


Figure 5.26: Post–fit systematic uncertainties after the fit to data in all analysis regions. The green (yellow) areas represent the $\pm 1(2)\sigma$ band on the pre–fit uncertainty $\Delta\theta$. The black points show the best–fit values $\hat{\theta}$ of the NPs, with the error bars representing the post–fit uncertainties. Each NP is shown relative to its nominal value, θ_0, and in units of its pre–fit uncertainty. NPs are grouped in a similar way as in Table 5.7: theory uncertainties are separated corresponding to (a) only signal or (c) background modelling; experimental uncertainties are separated in (b) JES, (d) JER, (e) leptons and (f) other experimental systematics, i.e. pile–up, JVT, b–tagging, $E_{\mathrm{T}}^{\mathrm{miss}}$ soft term and luminosity.

Figure 5.27 shows the correlations between the NFs and NPs included in the fit. The A_c^l is slightly correlated with the NFs that model the $t\bar{t}Z$ and γ–conversions backgrounds since these are quite abundant in the SRs. It is also anti–correlated with $\mathcal{N}_{\Delta\eta^-}$ through Eq. 3.3. Otherwise, correlations between the POI and the other NPs are small. The $t\bar{t}W$ alternative modelling is strongly correlated with $\mathcal{N}_{\Delta\eta^-}$ (about 65%). This correlation comes from the fact that both parameters modify the $t\bar{t}W$ yields in opposite directions (see Figure 5.8). Correlations between background NFs and the NPs are equivalent to those discussed in the CR–only fit.

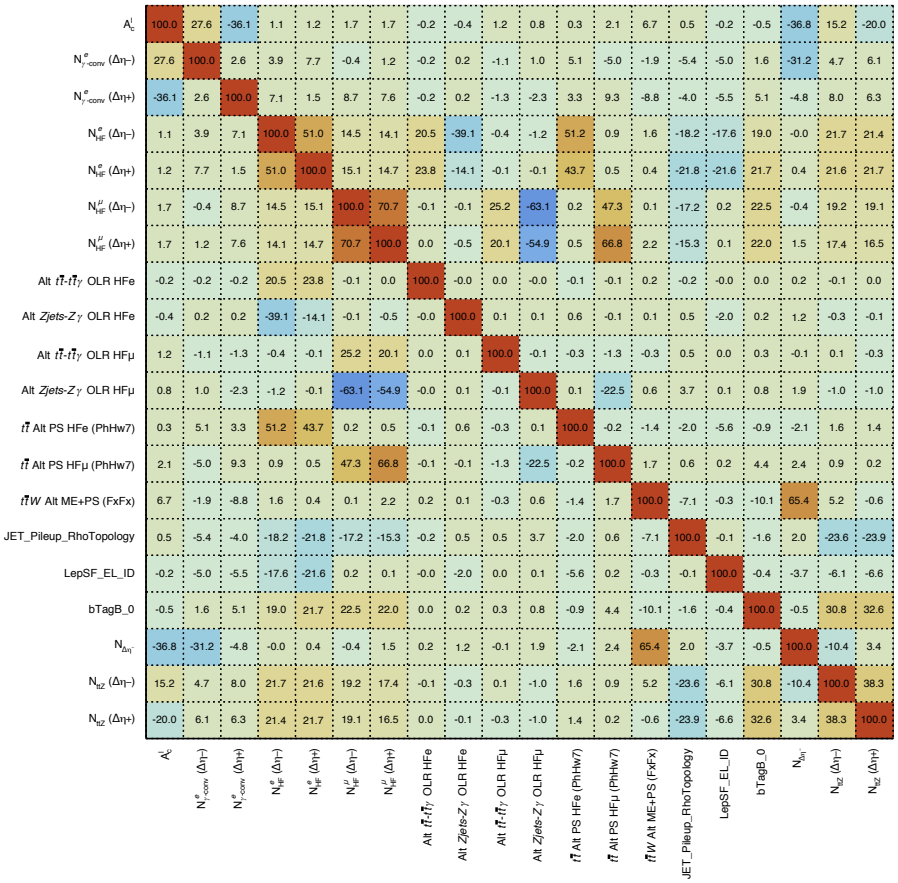


Figure 5.27: Correlation matrix between the NFs and the NPs in the fit to data in all analysis regions. Only NPs with correlations above 20% are displayed.

Figure 5.28 shows the first ranked 20 systematics with the largest impact on the fitted A_c^l value. The leading systematic corresponds to the alternative signal modelling NP which is expected given the interplay between this NP and $\mathcal{N}_{\Delta\eta^-}$ (see above), and ultimately the A_c^l. Nevertheless, its impact is about 10 times smaller than the statistical uncertainty on the A_c^l.

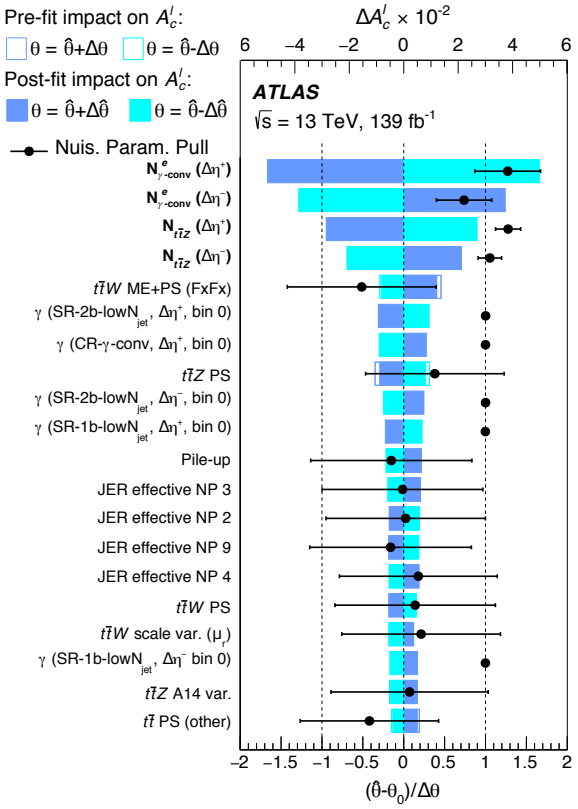


Figure 5.28: The most relevant (20) systematic uncertainties ranked by their impact on the A_c^l parameter at reconstruction level (top axis values). The impact ΔA_c^l of the uncertainties is shown before and after the fit to data in the signal and control regions. It is computed by comparing the nominal best–fit value of A_c^l, with the result of the fit when fixing the considered NP to its best–fit value, $\hat{\theta}$, shifted by its pre–fit (post–fit) uncertainties $\pm\Delta\theta$ ($\pm\Delta\hat{\theta}$). Pulls introduced by the fitting procedure are also shown (bottom axis values) and are identical to those in Figure 5.26. The entries shown in bold are the uncertainties in the free floating background normalisation parameters. The top ranked systematic uncertainty is about 10 times smaller than the statistical uncertainty for the A_c^l.

5.9 Unfolding and extraction of the leptonic charge asymmetry at particle level

To extract the value of the A_c^l at particle level in a defined fiducial phase–space, an unfolding procedure is performed to correct for resolution effects, as well as for signal efficiency and acceptance effects. The pertinent object definitions, unfolding procedure and results are presented in the following.

5.9.1 Particle level object definiton

Particle level objects in simulated events are defined using quasi–stable particles (with a mean lifetime greater than 30 ps) originating from pp collisions. They are selected after the parton shower and hadronisation, but before the interaction with the various detector components or consideration of pile–up effects.

Particle level leptons (electrons, muons and neutrinos) are required to originate from a W or a Z boson decay, including those from subsequent τ leptonic decays. In the case where they originate from a W boson, no match with the parent particle is required, i.e. it could be either from a prompt W boson, or from a W boson originating from a top quark decay. The four–momenta of the bare leptons are modified by adding the four–momenta of all radiated photons within a cone of size $\Delta R = 0.1$, excluding photons from hadron decays, to take into account FSR photons.

Particle level jets are reconstructed using the anti–k_t algorithm with a radius parameter of 0.4. It is applied to all stable particles, excluding the electrons, muons and photons used for the definitions of the selected charged leptons above, and neutrinos originating from W or Z boson decays. If b–hadrons with $p_{\mathrm{T}} > 5$ GeV are found in the MC event record, they are clustered into the stable–particle jets with their energies set to negligible positive values (denoted as "ghost–matching") [330]. Particle level jets containing one or more of these b–hadrons are considered as b–jets. The particle level $E_{\mathrm{T}}^{\mathrm{miss}}$ is defined as the vectorial sum of the transverse momenta of all selected neutrinos. Finally, particle level charged leptons within a cone size of $\Delta R = 0.4$ around a particle level jet are removed.

5.9.2 Particle level fiducial volume

The fiducial phase–space is defined to be close to the reconstruction level phase–space in order to avoid large acceptance effects. The following requirements on the particle level objects are imposed: exactly three charged light leptons (electrons or muons) with $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 2.5$ are required; the invariant mass of OSSF leptons must be larger than 25 GeV; no Z candidates (as described in Section 5.4) are allowed; and at least two jets with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$, one of them being identified as a b–jet, are required.

5.9.3 Unfolding procedure

An unfolding procedure is applied to the data events, distributed in $\Delta|\eta_\ell|$, and compared to the $t\bar{t}W$ SHERPA prediction in the fiducial volume. To construct the $\Delta|\eta_\ell|$ observable, a $m_{\ell b}$–matching procedure is applied to the particle level leptons in order to identify those coming from the top quark decay, given its advantages with respect to the multivariate method which are especially relevant here. The efficiency of the $m_{\ell b}$–matching at particle level is of 65%. Appendix E presents additional studies of the A_c^l at particle level using different lepton–top quark matching methods. Moreover, it also details the evolution of the predicted A_c^l as a function of the fiducial volume definition.

The unfolding procedure is based on the following formula:

$$N_i^{\text{folded}} = \underbrace{\frac{1}{\alpha_i}\sum_j \varepsilon_j M_{ij}}_{R_{ij}} N_j^{\text{fid}} \quad \text{with} \quad M_{ij} = \frac{N_{ij}^{(\text{reco}\,\cap\,\text{fid})}}{N_j^{(\text{reco}\,\cap\,\text{fid})}}, \quad \alpha_i = \frac{N_i^{(\text{reco}\,\cap\,\text{fid})}}{N_i^{\text{reco}}}, \quad \varepsilon_j = \frac{N_j^{(\text{reco}\,\cap\,\text{fid})}}{N_j^{\text{fid}}}, \tag{5.5}$$

with the number N_j^{fid} representing the content of bin j in the fiducial volume, N_i^{reco} being the events in bin i that satisfy the detector level selection and the symbol $\cap$ represents the logical intersection of the two regions. The N_i^{folded} represents the bin content (i) after the folding of the particle level bins (j) has been performed through the response matrix (R_{ij}). This matrix is constructed from the migration matrix (M_{ij}) and the acceptance and efficiency correction terms (α_i and ε_j) for each bin. The entries in the migration matrix represent the fractions of events at particle level in a y-axis bin that are reconstructed at detector level in an x-axis bin. They are normalised such that the sum of entries in each row is equal to one. The acceptance corrections account for events that are generated outside the fiducial volume but satisfy the selection at reconstruction level. The efficiency corrections account for events that are in the fiducial volume but fail to satisfy the reconstruction level selection.

Figure 5.29 shows an example of the migration matrix, efficiency and acceptance correction factors that are used for the SR–2b–lowN_{jets}, which is the SR with the highest $t\bar{t}W$ purity. The fraction of events that migrate between the $\Delta|\eta_\ell|$ bins (off–diagonal elements of M_{ij}) is around a reasonable 10% level. The acceptance corrections factors are small, at around 5%, and the efficiencies are about 11–12%. Moreover, no notable dependence with $\Delta|\eta_\ell|$ is observed. Other SRs exhibit similar migration matrices and acceptance levels, but efficiencies are lower (as low as 4%) depending on the selection criteria of each region (see Appendix G).

The unfolding procedure is based on a profile likelihood approach (similar to the one in Refs. [234,331]). With this approach, the unfolding problem is trans-

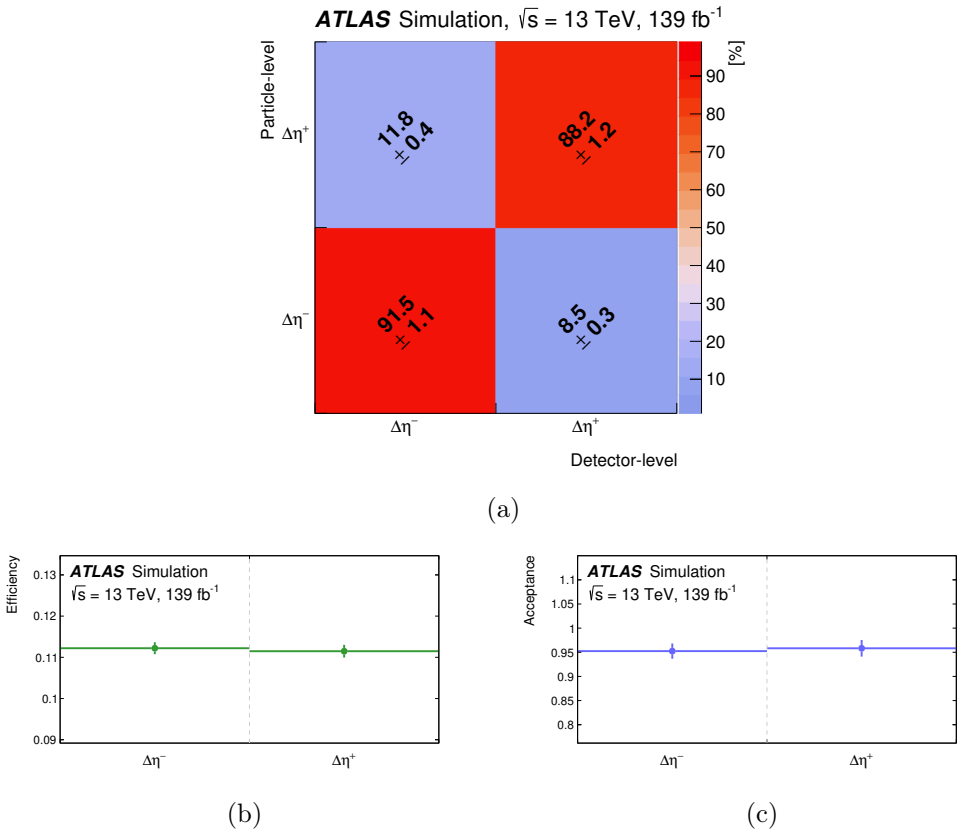


Figure 5.29: (a) The migration matrix, (b) the efficiency and (c) the acceptance corrections that are used as input for the unfolding of SR–2b–lowN_{jets}. The matrices are normalised such that the sum of any given row is 100%, although small differences may be present due to rounding. The error bars of the efficiency and acceptance correction terms represent the MC statistical uncertainties per bin based on the nominal $t\bar{t}W$ Sherpa sample.

formed into a standard problem of fitting normalisations of distributions. To achieve this, each bin in the particle level (signal) distribution is *folded* through the response matrix resulting in the same number of bins at reconstruction level. The particle level bins are treated as separate subsamples that are multiplied by their respective entries in the response matrix and freely floating parameters are assigned to each of these subsamples at reconstruction level. Therefore, the reconstruction level distributions are scaled by some factors, determined by fitting the data, and these factors are then used to scale the corresponding particle level bins which gives the desired unfolded result. The charge asymmetry is defined as the POI of the fit and follows an equivalent relation to Eq. 5.4 with the NFs of the unfolded bins.

In analogy to Section 5.8, the background NFs ($\mathcal{N}_{t\bar{t}Z}$, $\mathcal{N}^{e}_{\text{HF}}$, $\mathcal{N}^{\mu}_{\text{HF}}$ and $\mathcal{N}^{e}_{\gamma\text{-conv}}$) are split into $\Delta\eta^{-}$ and $\Delta\eta^{+}$ regions and are free floated in the fit. For the CRs, no response matrices are built. However, since the signal contamination is low, an approximation is made whereby the signal is treated as an additional background. The exception is the CR–γ–conv, where the signal contamination is sizeable (10%), and a response matrix is built. No regularisation is applied in the unfolding.

The collection of systematic uncertainties used in this fit is identical to the one described in Section 5.6. Systematic variations for the uncertainties affecting the signal process are obtained by computing and applying response matrices for all systematics on the particle level bins. On the other hand, variations for the uncertainties affecting the backgrounds are identical to those in the reconstruction level fit.

5.9.4 Injection tests

An injection test is performed in order to probe the stability of the unfolding procedure. This is done by injecting non–SM A_c^l values into the particle level prediction to obtain per–event *injection* weights. These are propagated to the reconstruction level distributions and treated as pseudo–data in the fit. The response matrices are not recomputed for each injected asymmetry value.

A total of seven pseudo–data points are obtained for several negative and positive variations of the A_c^l. The unfolding procedure is applied to each of the points to obtain a relation between the injected and extracted A_c^l value at particle level, and estimate the bias of the procedure. After the fit to real data, the observed A_c^l is substituted into the relation and the unfolding bias is found to be 0.004. Although this effect is well covered by the systematic uncertainties, it is added as an extra uncertainty to the unfolded A_c^l value.

Figure 5.30 shows the seven pseudo–data points, the fitted relation between the extracted and injected A_c^l values (blue line), and the case where both values would perfectly agree (discontinuous diagonal line), i.e. no bias. The full MC statistics of the $t\bar{t}W$ sample are used to estimate the uncertainties on each point. Here, the error bands cover the discrepancy of the fitted (blue) line with the discontinuous line within 2σ. Analogous results are obtained if the data statistics and systematic uncertainties are included in the injection test. Given that these uncertainties are much larger than the MC statistical ones, the discrepancies between both lines are also covered. Appendix H shows complementary tests to validate the unfolding procedure.

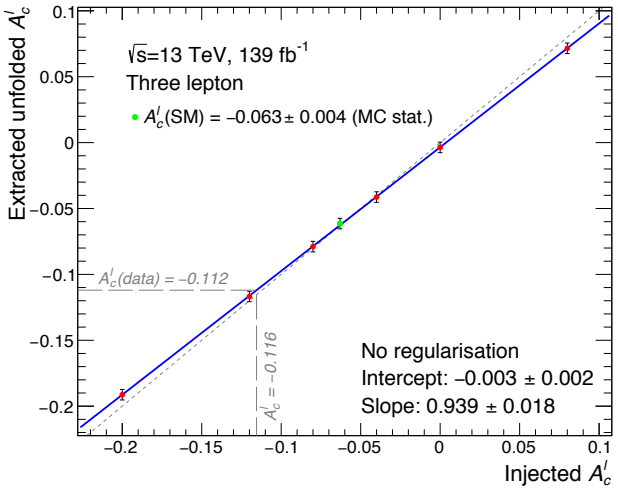


Figure 5.30: Results from the injection test with seven pseudo–data points (red), including the point where the SM A_c^l is also tested (green), computed with the $t\bar{t}W$ SHERPA sample. The error bars include MC statistical uncertainties. The blue line represents the best–fit line to the seven points and its fit parameters are also displayed. The diagonal discontinuous line portrays a perfect agreement between extracted and injected values of A_c^l. In the bottom left area (in grey), a schematic representation of how the unfolding bias is estimated, from the observed A_c^l value (see Section 5.9.5), is shown.

5.9.5 Extraction of the charge asymmetry at particle level

As in the reconstruction level fit, both Asimov and CR–only fits are performed before injecting data into the SRs. Given that the results and conclusions from these fits are analogous at particle level (PL), they are not presented here in detail*. The fitted SRs and CRs are identical to those in Section 5.8†. The unfolded charge asymmetry value, to a fit to data in all analysis regions, is found to be

$$A_c^l(t\bar{t}W)^{\mathrm{PL}} = -0.11 \pm 0.17\,(\mathrm{stat.}) \pm 0.05\,(\mathrm{syst.}),$$

with a SM expectation calculated using the nominal $t\bar{t}W$ SHERPA simulation of

$$A_c^l(t\bar{t}W)^{\mathrm{PL}}_{\mathrm{SM}} = -0.063\,^{+0.007}_{-0.004}\,(\mathrm{scale}) \pm 0.004\,(\mathrm{MC\ stat.}).$$

*The result from the unfolded Asimov fit is $A_c^l(t\bar{t}W)^{\mathrm{PL}} = -0.063 \pm 0.261\,(\mathrm{stat.}) \pm 0.056\,(\mathrm{syst.})$.

†Alternative binning options were tested for the unfolded A_c^l extraction without any improvements on the final result. Thus, the strategy with the most simple setup was chosen.

The nominal values for the background normalisations are the same as those in the reconstruction level fit (see Figure 5.25). The contributions from the most relevant uncertainties in the charge asymmetry at particle level are given in Table 5.8. The sources of systematic uncertainty are similar to the ones reported in Table 5.7, with the $\Delta\eta^{\pm}$ CR–dependency, the modelling of the $t\bar{t}W$ and $t\bar{t}Z$ MC processes, and the statistical uncertainty being the dominant ones. The statistical uncertainty is slightly increased relative to the reconstruction level result due to the unfolding procedure.

	$\Delta A_c^l(t\bar{t}W)^{\mathrm{PL}}$
Experimental uncertainties	
Leptons	0.014
Jet energy resolution	0.011
Pile–up	0.008
Jet energy scale	0.004
$E_{\mathrm{T}}^{\mathrm{miss}}$	0.002
Luminosity	0.001
Jet vertex tagger	0.001
MC modelling uncertainties	
$t\bar{t}W$ modelling	0.022
$t\bar{t}Z$ modelling	0.017
$\mathrm{HF}_{e/\mu}$ modelling	0.015
Others modelling	0.015
WZ/ZZ+jets modelling	0.014
$t\bar{t}H$ modelling	0.006
Other uncertainties	
Unfolding bias	0.004
$\Delta\eta^{\pm}$ CR–dependency	0.039
MC statistical uncertainty	0.027
Response matrix	0.009
Data statistical uncertainty	0.170
Total uncertainty	0.179

Table 5.8: List of the most relevant systematic and statistical uncertainties in the leptonic charge asymmetry at particle level $A_c^l(t\bar{t}W)^{\mathrm{PL}}$. For this table, the uncertainties are symmetrised and grouped into categories. The sum in quadrature of the individual uncertainties is not necessarily equal to the total uncertainty due to correlations introduced by the fit.

Figure 5.28 shows the first ranked 20 systematics with the largest impact on the fitted unfolded A_c^l value. The alternative signal modelling NP is, again, the leading systematic. The pulls, constraints, and correlations among NPs, and the defined NFs, are analogous to those discussed in Section 5.8.3.

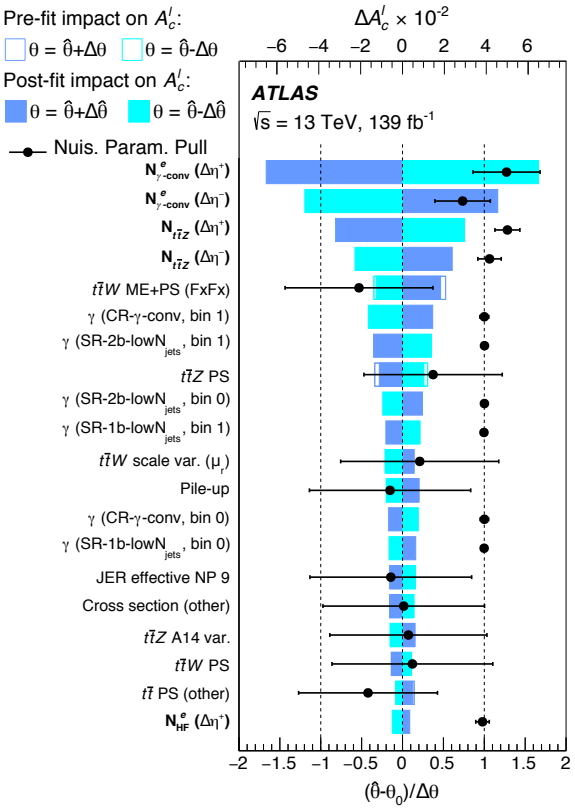


Figure 5.31: The most relevant (20) systematic uncertainties ranked by their impact on the A_c^l parameter at particle level (top axis values). The impact ΔA_c^l of the uncertainties is shown before and after the fit to data in the signal and control regions. It is computed by comparing the nominal best–fit value of A_c^l, with the result of the fit when fixing the considered NP to its best–fit value, $\hat{\theta}$, shifted by its pre–fit (post–fit) uncertainties $\pm\Delta\theta$ ($\pm\Delta\hat{\theta}$). Pulls introduced by the fitting procedure are also shown (bottom axis values). The entries shown in bold are the uncertainties in the free floating background normalisation parameters. The top ranked systematic uncertainty is about 10 times smaller than the statistical uncertainty for the unfolded A_c^l.

5.10 Summary and outlook

This chapter has presented a search for the leptonic charge asymmetry in $t\bar{t}W$ production using pp collision data at $\sqrt{s} = 13$ TeV with the full Run 2 dataset, corresponding to an integrated luminosity of 139 fb^{-1}. The leptonic charge asymmetry is defined as the absolute pseudorapidity difference between the two leptons associated with the decay of the $t\bar{t}$ pair. The search is performed in 3ℓ final states using reconstructed light leptons (electrons or muons), together with jets and b–jets. To correctly associate the leptons with either top quarks or antitop quarks, a technique based on a BDT is used.

The charge asymmetry at reconstruction level is obtained by performing a simultaneous profile likelihood fit to data using several signal and control regions optimised for either the $t\bar{t}W$ process or the major SM background processes ($t\bar{t}Z$, non-prompt leptons from HF decays and electrons from γ–conversions). The charge asymmetry is extracted together with the normalisations for these background processes and is found to be

$$A_c^l(t\bar{t}W) = -0.12 \pm 0.14\,(\text{stat.}) \pm 0.05\,(\text{syst.}),$$

with a SM expectation calculated using the nominal $t\bar{t}W$ Sherpa simulation of

$$A_c^l(t\bar{t}W)_{\text{SM}} = -0.084\,^{+0.005}_{-0.003}\,(\text{scale}) \pm 0.006\,(\text{MC stat.}).$$

An unfolding procedure is used to obtain the charge asymmetry at particle level in a specific fiducial volume. The particle level lepton–top quark matching is based on the invariant mass of the lepton and its closest b–jet. The unfolding is based on a profile likelihood approach, where the unfolding is performed together with fitting normalisations of the major background processes. The charge asymmetry at particle level yields

$$A_c^l(t\bar{t}W)^{\text{PL}} = -0.11 \pm 0.17\,(\text{stat.}) \pm 0.05\,(\text{syst.}),$$

with a SM expectation calculated using the nominal $t\bar{t}W$ Sherpa simulation of

$$A_c^l(t\bar{t}W)^{\text{PL}}_{\text{SM}} = -0.063\,^{+0.007}_{-0.004}\,(\text{scale}) \pm 0.004\,(\text{MC stat.}).$$

The most relevant systematic uncertainties affecting this search can be attributed to the $\Delta\eta^{\pm}$ CR–dependency of the fit, as well as the modelling uncertainties of the $t\bar{t}W$ and $t\bar{t}Z$ MC processes. However, both the reconstruction and particle level results are severely limited by the statistical uncertainties of the data. This motivates the study of the A_c^l once more data are available in the coming years. Considering a reduction of the statistical uncertainty proportional to $\mathcal{L}_{\text{int.}}^{-1/2}$, the statistical component of the uncertainty (based on the reconstruction level result) would be reduced to 0.096 for LHC Run 3 and 0.030 for

HL–LHC. In addition, after a plausible reduction of the total systematic uncertainty of about a half at HL–LHC, and the removal of the $\Delta\eta^{\pm}$ CR–dependency uncertainty following a better understanding of the background asymmetries, the total uncertainty on the reconstructed A_c^l would become 0.032. If the central value was found to be the same, the A_c^l would be about 3.8σ away from zero.

Further reductions on the A_c^l uncertainty are possible by increasing the S/B ratio in the SRs. In the 3ℓ channel, this could be studied with ML techniques that are able to discriminate the $t\bar{t}W$ signal against other backgrounds. The output scores from these ML techniques could be used to set additional requirements on the SR definitions, and increase the signal purity. During the course of this analysis, these studies were pursued using an event level BDT (with similar architecture to the one used for the lepton–top quark matching) that separated $t\bar{t}W$ events from all other background events. However, the improvements to the uncertainty were minimal and it was discarded due to the additional layer of complexity that it introduced, and its interplay with the existing BDT–based lepton–top quark matching. With a larger dataset, and more advanced ML techniques (e.g. that are able to output multiple scores for different processes), this methodology could prove to have a bigger impact in the reduction of the A_c^l uncertainty.

The A_c^l is sensitive to BSM models such as axigluons, some four–fermion operators in the SMEFT framework, and the chiral structure of possible new physics signals. Ref. [239] scrutinises similar asymmetries with SM calculations that include the $t\bar{t}W$ off–shell effects, or in the narrow–width approximation. With the reduction of the total A_c^l uncertainty, the (unfolded) result would have the potential to discriminate between these alternate models, and contribute to the better understanding of top quark physics.

5.10.1 EFT parameterisation and constraints with the A_c^l

The $t\bar{t}W$ A_c^l offers unique sensitivity to certain linear combinations of SMEFT coefficients that parameterise the effect of operators involving four quarks in the $t\bar{t}W$ production vertices. As discussed previously, it benefits from the particular characteristics of the $t\bar{t}W$ process and sensitivity to a reduced number of operators, and from its ability to distinguish between chiral structures to which cross–section observables are blind.

In this section, EFT parametrisations of the $t\bar{t}W$ A_c^l are obtained, and the presented measurement in this thesis is used to compute constraints for a set of three four–fermion operators: $O_{Qq}^{1,8}, O_{Qq}^{3,8}$ and O_{tq}^{8}. Parameterisations are computed at NLO accuracy (similarly to what is done in Chapter 4) for the A_c^l, and also for the $t\bar{t}W$ and $t\bar{t}$ processes' cross–sections. aMC@NLO v3.1.1 is used for the ME calculations. In order to compute the A_c^l, MadSpin is used to decay

the top quark pair leptonically*. Limits on the Wilson coefficients are obtained under the assumption that the expected and observed values of the observables are the same as those coming from the MC simulation. However, experimental uncertainties from the latest ATLAS public results are used. The parametrisation of the A_c^l is obtained from the parametrisations of the asymmetric part in the numerator and of the symmetric one in the denominator (see Eq. 3.3), where the denominator is, by definition, the $t\bar{t}W$ process cross–section.

Figure 5.32 shows the $t\bar{t}W$ cross–section (normalised to the SM prediction) and A_c^l dependence with the Wilson coefficients in the left and right panels, respectively. The solid curves show the allowed ranges after the constraints from the inclusive $t\bar{t}$ cross–section have been computed using the latest ATLAS measurement with a relative uncertainty of about 1.8% [332]. The shown $t\bar{t}$ allowed ranges are obtained using a linear fit for $O_{Qq}^{1,8}$ and O_{tq}^{8}, while a quadratic fit is used for $O_{Qq}^{3,8}$. These fits are chosen because they provide the best constraints with respect to their quadratic/linear counterparts. The grey region for the $t\bar{t}W$ cross–section observable indicates the relative total uncertainty of about 9% as reported by Ref. [234]. For the A_c^l, the uncertainty obtained in the work carried out within this thesis (± 0.15) spans the entire y–axis range and is therefore not shown. The displayed central value of the A_c^l is computed at parton level in the inclusive phase–space, with no selection requirements in the final state. These results show the potential of the A_c^l to probe the SMEFT parameter space, and its ability to distinguish the chiral structure of the $O_{Qq}^{1,8}$ (LL/RR) and O_{tq}^{8} (LR/RL) operators, which is not possible with cross–section observables.

Additional fits are performed for the three different possible 2D combinations of these operators, taking into account interference terms where the two operators are set to non–zero values. The 2D 68% CL bounds for each pair are shown in Figure 5.33. Assuming current relative experimental uncertainties (shown in the legend), the A_c^l is not able to provide any additional constraints on the Wilson coefficients. Notably, the inclusive $t\bar{t}W$ cross–section has a larger constraining power than the inclusive $t\bar{t}$ cross–section. However, current results that exploit the differential $t\bar{t}$ cross–sections and asymmetries [187,283,333,334] present very stringent limits on these coefficients and thus a better scenario for subsequent analysis with Run 3 data.

Nonetheless, these 2D bounds can be recomputed assuming the same scaling of uncertainties corresponding to the HL–LHC scenario presented in Section 5.10. These are shown in Figure 5.34. In particular for the $C_{Qq}^{1,8} - C_{tq}^{8}$ combination (bottom), the addition of the A_c^l provides a solution in a direction that is blind to cross–section observables, hence, reducing the allowed ranges of both coefficients to about ± 0.5.

*Decays are assumed to be SM–like.

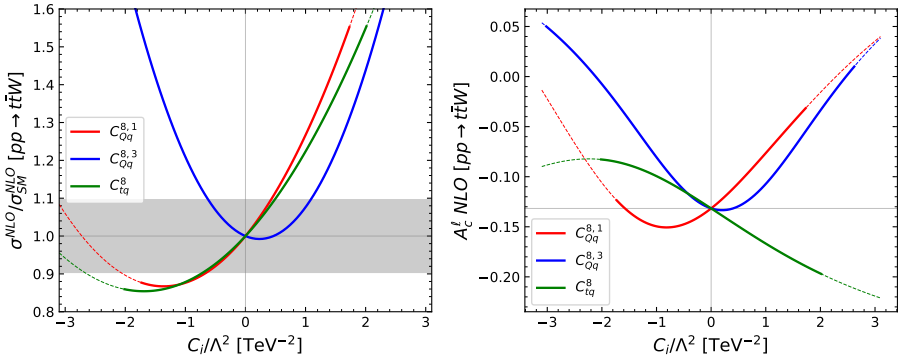


Figure 5.32: Left panel: $t\bar{t}W$ cross–section parameterisation (normalised to the SM prediction). The grey region indicates the relative total uncertainty as reported by Ref. [234]. Right panel: A_c^l parameterisation. In both, the solid curves show the allowed ranges after computing the constraints coming from the inclusive $t\bar{t}$ cross–section observable.

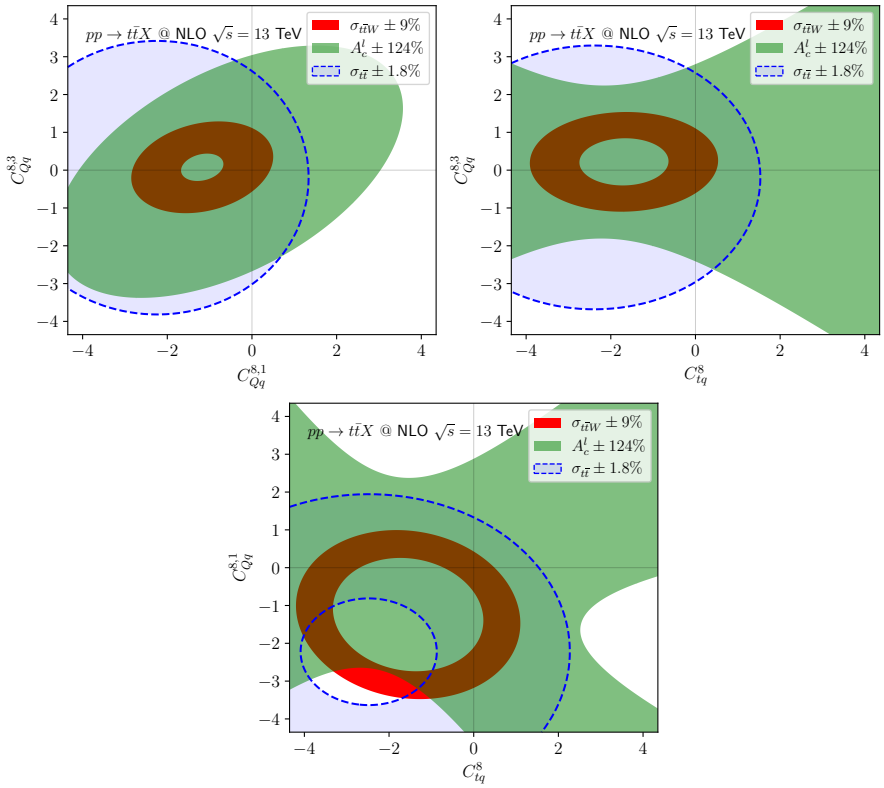


Figure 5.33: 2D 68% CL bounds for the three different possible combinations of the $t\bar{t}W$ and $t\bar{t}$ cross–sections, and A_c^l observables. The indicated uncertainties are the relative uncertainties taken from the corresponding public results.

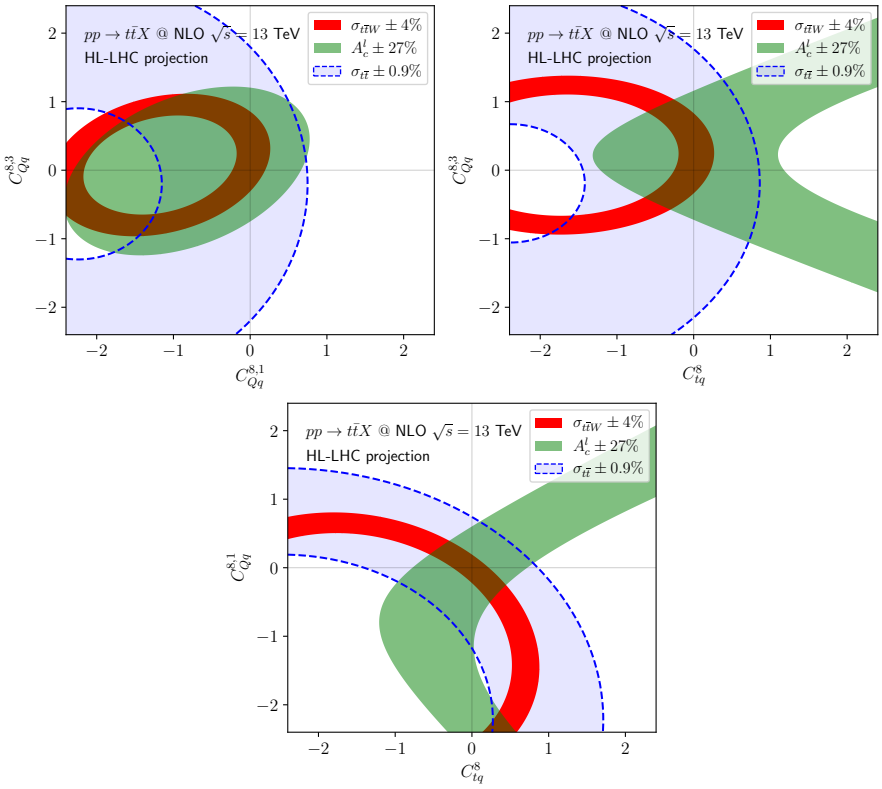


Figure 5.34: 2D 68% CL bounds for the three different possible combinations of the $t\bar{t}W$ and $t\bar{t}$ cross–sections, and A_c^l observables. The indicated uncertainties are the relative uncertainties taken from the corresponding public results and extrapolated to a HL–LHC scenario.

6 Study of the $\mathcal{CP}$ properties of the top quark Yukawa interaction in $t\bar{t}H$ and tH events

The study of the Yukawa interaction of the top quark with the Higgs boson is key to probe Higgs properties. Besides studying the magnitude and sign of such coupling (as summarised in Section 3.8.4), one can also study its $\mathcal{CP}$ properties. The SM predicts a scalar Higgs boson ($J^{\mathcal{CP}} = 0^{++}$) with a prescribed coupling to the top quark. However, the presence of a $J^{\mathcal{CP}} = 0^{+-}$ pseudoscalar admixture, which introduces a second coupling to the top quark, has not yet been excluded. The observation of this $\mathcal{CP}$–odd contribution would be a sign of BSM physics.

While the $\mathcal{CP}$–odd contributions have been studied by ATLAS and CMS in the couplings of the Higgs boson to gauge bosons, no signs of deviations from the SM $\mathcal{CP}$–even prediction have been seen [335–340]. In these HVV couplings, the $\mathcal{CP}$–odd terms are suppressed by powers of the new physics energy scale Λ^{-2} [243]. On the other hand, the study of the coupling of the Higgs boson to fermions is expected to yield higher sensitivities since the $\mathcal{CP}$–odd terms enter at the same order as the $\mathcal{CP}$–even ones.

The $\mathcal{CP}$ properties of the top quark Yukawa interaction can be studied in the production of a Higgs boson in association with top quarks at the LHC in $t\bar{t}H$ and tH ($tHjb$ and tWH) events. The $\mathcal{CP}$–odd component impacts the production rates of these processes and some kinematic distributions [341–344]. Moreover, this $\mathcal{CP}$–mixing modifies the rates of the loop–induced $H \to \gamma\gamma$ branching ratio and ggF production. To quantify its impact, the $\mathcal{CP}$–odd component to the top quark Yukawa is introduced using an EFT approach with the Higgs characterisation model [345], which is implemented in aMC@NLO. In this model, no new particles are assumed couple to the Higgs boson below the cut–off energy scale ($\Lambda = 1$ TeV).

6.1 MC samples using the Higgs characterisation model

The effective Lagrangian that modifies the top quark Yukawa coupling is

$$\mathscr{L} = -\frac{G_t}{\sqrt{2}} H\{\bar{t}\,\kappa_t[\cos(\alpha) + \mathrm{i}\sin(\alpha)\gamma_5]t\}, \tag{6.1}$$

where κ_t is the coupling strength factor and α is the $\mathcal{CP}$ mixing angle. Comparing this with the SM coupling in Eq. 3.7, a pure $\mathcal{CP}$–even coupling corresponds to $\kappa_t = 1$ and $\alpha = 0°$, while a pure $\mathcal{CP}$–odd coupling is realised when $\alpha = 90°$.

The simulation of the $t\bar{t}H$, $tHjb$ and tWH is done using arithmetical dynamical μ_R and μ_F scales such that $\mu_R = \mu_F = H_T/2$, where H_T runs over all outgoing partons from the ME calculation. Both the $t\bar{t}H$ and tWH samples are produced with the five–flavour (5F) scheme, while $tHjb$ uses the four–flavour

(4F) scheme. Figure 6.1 shows the cross–section dependence of the $t\bar{t}H$, $tHjb$ and tWH processes, as well as the $H \to \gamma\gamma$ decay and ggF production, with the $\mathcal{CP}$ mixing angle α and for $\kappa_t = 1$. They are normalised to the SM prediction.

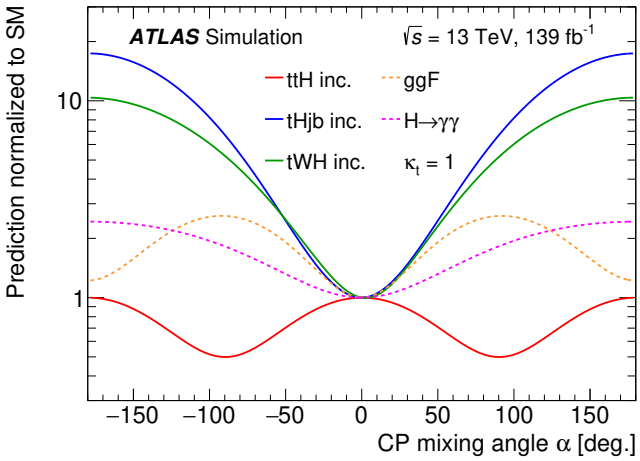


Figure 6.1: Dependence of the expected rates of the $t\bar{t}H$ (red), $tHjb$ (blue), tWH (green) processes, as well as the $H \to \gamma\gamma$ decay (magenta) and ggF production (orange) with the mixing angle for $\kappa_t = 1$. Normalised to the SM expectation.

For each process, the SM sample's ($\alpha = 0°$) cross–section is compared to that in the CERN YR4 [227], and a K–factor is computed. Samples in the YR4 use fixed scales, i.e. $\mu_{\text{fixed}} = \mu_R = \mu_F$, and the 5F scheme for all processes. Moreover, they are calculated at NLO in QCD accuracy for the $tHjb$ and tWH processes and at NLO+EW QCD accuracy for $t\bar{t}H$. Therefore, the $t\bar{t}H$, $tHjb$ and ggF processes are simulated with both dynamical and fixed scales to compute their ratio: K–factor $= \mu_{\text{fixed}}/\mu_{\text{dynamical}}$. The μ_{fixed} values are set equal to those in the YR4: $\mu_{\text{fixed}} = (m_t + m_H)/2 = 235$ GeV for $t\bar{t}H$, $\mu_{\text{fixed}} = (m_t + m_H)/4 = 74.4$ GeV for $tHbj$ (both in 4F and 5F schemes), and $\mu_{\text{fixed}} = m_H = 125$ GeV for ggF.

Table 6.1 shows the K–factors for the different processes and cross–section values for both scales. In addition, the $\mathcal{CP}$ mixing angle α is varied to study the dependence of the K–factor with it. Three points are used corresponding to $\alpha = 0°, 45°, 90°$ which are the SM pure $\mathcal{CP}$–even, $\mathcal{CP}$ maximum mixing, and pure $\mathcal{CP}$–odd cases, respectively. The results show to a good degree that the K–factor does not depend on the $\mathcal{CP}$ mixing angle since the maximum discrepancy that is observed between the K–factors is of about 6% for the $tHjb$ 5F process. As a result, the K–factors derived for the SM case are used to scale all samples with different α values.

$t\bar{t}H$ 5F cross–section [pb]			
Scale choice	$\alpha = 0^\circ$ (SM)	$\alpha = 45^\circ$	$\alpha = 90^\circ$
Aritm. dyn. $H_T/2$	0.457	0.325	0.197
Fixed scale	0.509	0.367	0.218
K–factor	**1.12**	**1.12**	**1.11**
$tHjb$ 4F cross–section [pb]			
Scale choice	$\alpha = 0^\circ$ (SM)	$\alpha = 45^\circ$	$\alpha = 90^\circ$
Aritm. dyn. $H_T/2$	0.061	0.098	0.271
Fixed scale	0.068	0.112	0.309
K–factor	**1.12**	**1.14**	**1.14**
$tHjb$ 5F cross–section [pb]			
Scale choice	$\alpha = 0^\circ$ (SM)	$\alpha = 45^\circ$	$\alpha = 90^\circ$
Aritm. dyn. $H_T/2$	0.058	0.096	0.262
Fixed scale	0.075	0.126	0.362
K–factor	**1.30**	**1.32**	**1.38**
Ratio (fixed5F/dyn4F)	1.23	1.29	1.33
$ggH2j$ 5F cross–section [pb]			
Scale choice	$\alpha = 0^\circ$ (SM)	$\alpha = 45^\circ$	$\alpha = 90^\circ$
Aritm. dyn. $H_T/2$	14.28	24.36	32.25
Fixed scale	14.80	24.14	32.87
K–factor	**1.04**	**0.99**	**1.01**

Table 6.1: K–factors for the $t\bar{t}H$ (top), $tHjb$ for both 4F and 5F schemes (middle), and $ggH2j$ (bottom) processes and their dependence with the $\mathcal{CP}$ mixing angle α.

Table 6.2 shows the NLO QCD cross–sections for the $t\bar{t}H$, $tHjb$, tWH and $ggH2j$ processes for different values of α and κ_t. They are normalised using the K–factors from Table 6.1 and the BR($H \to \gamma\gamma$) $\simeq 2 \times 10^{-3}$.

Normalised cross–sections × BR($H \to \gamma\gamma$) [fb]					
κ_t	α	$t\bar{t}H$	$tHjb$	tWH	$ggH2j$
1	0°(SM, $\mathcal{CP}$–even)	1.150	0.169	0.034	33.1
1	15°	1.113	0.177	0.038	–
1	30°	0.995	0.207	0.048	–
1	45° ($\mathcal{CP}$ max. mixing)	0.827	0.266	0.064	53.7
1	60°	0.666	0.382	0.087	–
1	75°	0.545	0.548	0.116	–
1	90° ($\mathcal{CP}$–odd)	0.500	0.753	0.150	74.4
−1	0°	–	1.980	0.307	–
0.5	0°	–	0.264	0.030	–
2	0°	–	0.666	0.190	–
2	45°	–	0.570	0.226	–

Table 6.2: NLO cross–sections × BR($H \to \gamma\gamma$) for the $t\bar{t}H$, $tHjb$, tWH and $ggH2j$ processes for different $\mathcal{CP}$ scenarios normalized with the K–factor (see text). The top part refers to MC samples in which κ_t value is fixed to 1 and α is varied, while the bottom part to samples with κ_t value different from 1 (for the tH processes).

These MC studies and samples serve as a validation of the Higgs characterisation model, as well as being used for the signal modelling in the analysis described in this section.

6.2 Analysis strategy and results

This section presents the first search for $\mathcal{CP}$–violation in the top quark Yukawa coupling using $t\bar{t}H$ and tH production modes in the diphoton decay channel ($H \to \gamma\gamma$). The analysis is based on the 139 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV recorded with the ATLAS detector at the LHC. This result has been published in Ref. [219].

6.2.1 Event selection and categorisation

Events are required to have two isolated photons (decaying from the Higgs boson) with p_{T} greater than 35 and 25 GeV. Both photons must satisfy the tight identification requirement [143]. In addition, events are separated into two $t\bar{t}H$–enriched regions: in first place, the "Lep" region ($1\ell, \geq 1j, \geq 1b$) includes events where a W boson (from a top quark decay) decays to a charged lepton and its associated neutrino; secondly, the "Had" region ($0\ell, \geq 3j, \geq 1b$) includes all other events with W bosons decaying *hadronically* to a $q\bar{q}$ pair producing two

additional jets with a minimum p_{T} of 25 GeV. An object level BDT, referred to as "Top Reco BDT", is trained with the $t\bar{t}H$ sample by using the XGBoost package [346] to extract a top quark candidate. Its goal is to discriminate random jet triplets from $t\bar{t}H$ $(\gamma\gamma)$ events from those originating from a top quark decay. The BDT uses information from the decay products of the top quark (W boson and b–jet), from the reconstructed jets or the final–state lepton for the Had and Lep regions, respectively. Figure 6.2 shows the reconstructed top quark candidate mass from the BDT which shows a good level of agreement with data.

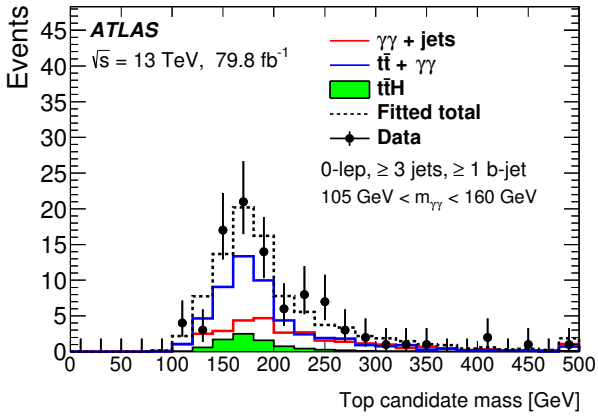


Figure 6.2: Distribution of the reconstructed top quark candidate in the diphoton mass window from a triplet of jets selected by the dedicated Top Reco BDT. A good agreement wit data is seen. Source: Ref. [224].

To improve the analysis sensitivity, selected events are categorized using partitions of a two–dimensional BDT space. Two BDTs are trained using the XGBoost package separately in both the Lep and Had channels:

- A background rejection BDT is trained to discriminate between $t\bar{t}H$ signal and background events [224]. It makes use of the four–momenta information of photons, jets, and leptons in the event. It exhibits good background rejection and $t\bar{t}H/tH$ acceptance, and weak dependence with $\mathcal{CP}$ mixing angle.

- Furthermore, a $\mathcal{CP}$ BDT is used to separate $\mathcal{CP}$–even from $\mathcal{CP}$–odd events. This BDT uses the four–momenta information of reconstructed top quark and Higgs candidates, angular distances between photons and jets, $\Delta R^{\gamma j}_{\mathrm{min}}$, as well as other event variables such as N_{jets}, $N_{b\text{–jets}}$, $E_{\mathrm{T}}^{\mathrm{miss}}$ or H_T. The leading variables for the classification in both channels are found to be: the Higgs candidate p_{T}, H_T, the second–smallest $\Delta R^{\gamma j}_{\mathrm{min}}$, and the invariant mass between both top quark candidates.

A schematic view of the hadronic (left) and leptonic (right) categories is shown in Figure 6.3. There are 20 categories in total: 12 in the Had region and 8 in the Lep region. Furthermore, Figure 6.4 shows the BDT discriminant distributions in data, as well as those expected from the $\mathcal{CP}$–even and the $\mathcal{CP}$–odd Higgs boson signals in the Had region.

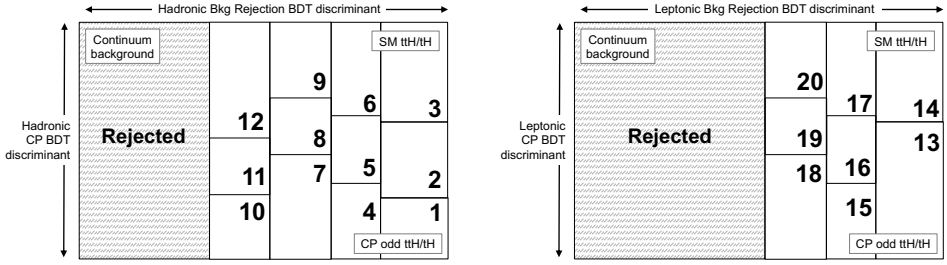


Figure 6.3: Schematic view of the hadronic (left) and leptonic (right) categories used in the analysis. The backgrounds are largest on the left of the figures. The pure $\mathcal{CP}$–even processes are largest on the top and the pure $\mathcal{CP}$–odd processes are largest on the bottom.

6.2.2 Fitting procedure and results

A simultaneous maximum–likelihood fit is performed to the diphoton invariant mass $m_{\gamma\gamma}$ spectra in all the categories. Signal and background shapes are modeled by analytic functions as discussed in Ref. [338]. Figure 6.5 shows the distributions of the reconstructed masses for the diphoton system and primary top quark. Figure 6.6 shows the yields for the $\mathcal{CP}$–even and $\mathcal{CP}$–odd signals, as well as the data calculated in the smallest $m_{\gamma\gamma}$ interval containing 90% of the signal shown in three groups of categories. The three categories are obtained by the combination of the appropriate event categories defined from the previous BDT categorisation. The last category shows a clear favour towards the $\mathcal{CP}$–even hypothesis.

Assuming a $\mathcal{CP}$–even coupling, and constraining all non–$t\bar{t}H$ Higgs boson processes to their SM predictions, the measured rate for $t\bar{t}H$ is found to be $\mu_{t\bar{t}H} = 1.43^{+0.33}_{-0.31}(\text{stat.})^{+0.21}_{-0.15}(\text{syst.})$ times the SM expectation. The background–only hypothesis is rejected with an observed (expected) significance of 5.2σ (4.4σ). Under the same assumptions, the CLs method [347] yields a 95% CL upper limit of 12 times the SM prediction for tH production cross–section; the same as the expected limit assuming the presence of a SM tH signal.

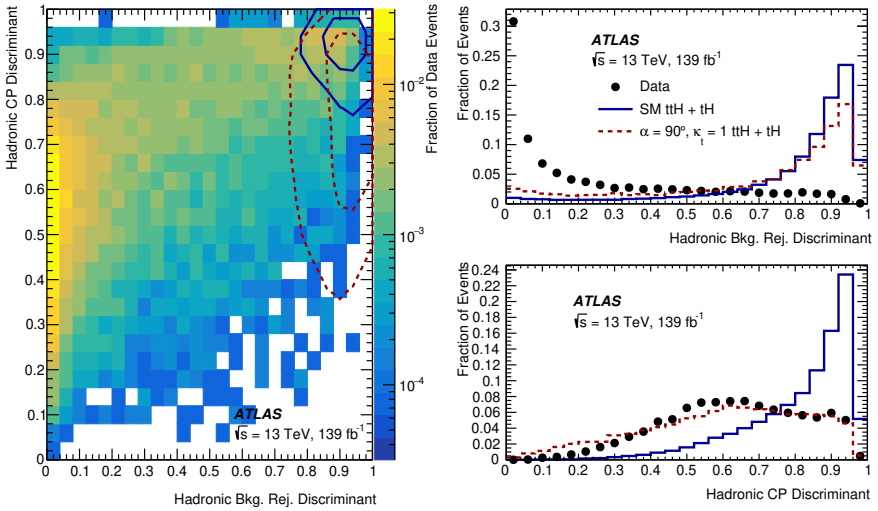


Figure 6.4: Left: two–dimensional BDT distribution for the hadronic channel showing the background rejection BDT (x–axis) and the CP BDT (y–axis). Inner (outer) contours capture 25% (50%) of $t\bar{t}H$ and tH signal events for $\mathcal{CP}$–even (blue) and $\mathcal{CP}$–odd (red) hypotheses. Right: Projections to the background rejection (top) and CP BDT (bottom) axes normalized to unity. The error bars include the statistical uncertainty on data.

From the simultaneous fit, the 1D ($\mathcal{CP}$ mixing angle α) and 2D ($\kappa_t \cos\alpha - \kappa_t \sin\alpha$ contours) limits on the $\mathcal{CP}$ properties of the top quark Yukawa coupling can be extracted. These are shown in Figure 6.7 in the left and right panels, respectively.

The corresponding Higgs boson coupling modifiers κ_γ and κ_g are taken from the Run 2 Higgs boson coupling combination [348] (without including the $t\bar{t}H$ process). With these, the observed (expected) exclusion limit on the $\mathcal{CP}$ mixing angle (without a prior constraint on κ_t) is $|\alpha| > 43°$ (63°) at 95% CL. The $\mathcal{CP}$–odd hypothesis is excluded at 3.9σ (2.5σ). In this analysis, the statistical uncertainty is the dominant one. A similar analysis in the diphoton decay channel by CMS excludes the $\mathcal{CP}$–odd hypothesis at 3.2σ [220].

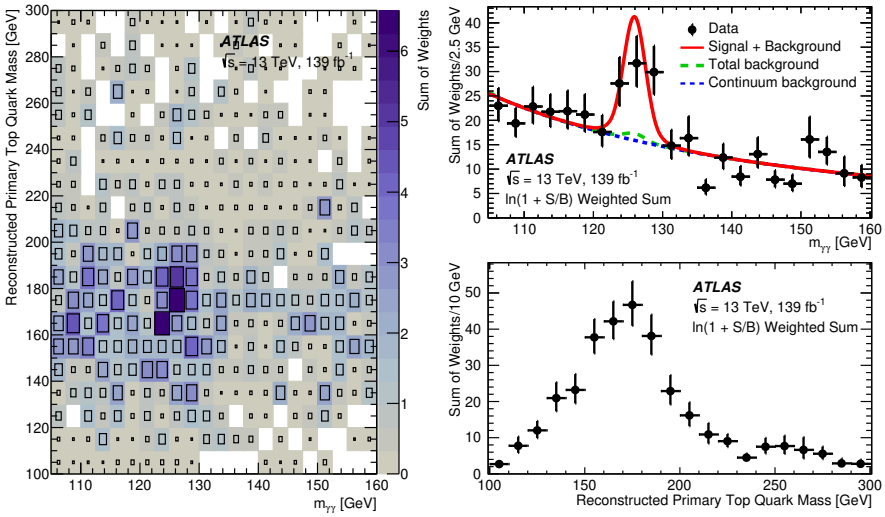


Figure 6.5: Left: Distribution of the reconstructed primary top mass versus Higgs mass in data events. Right: Projection to the Higgs boson mass (top) and primary top quark mass (bottom) axes. The events are weighted by $\ln(1+S/B)$ with S and B being the fitted signal and background yields in the smallest $m_{\gamma\gamma}$ interval containing 90% of the signal in each category. The error bars include the statistical uncertainty on data.

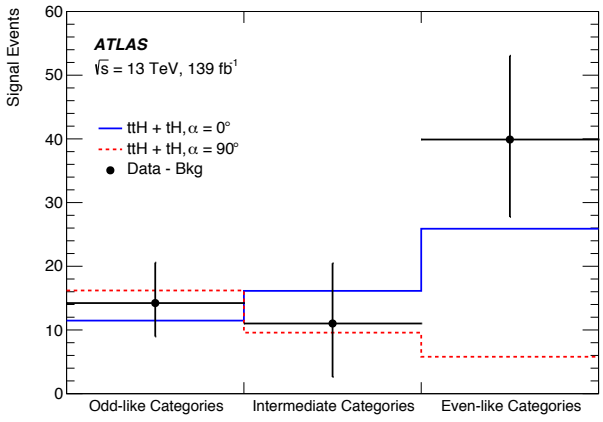


Figure 6.6: Signal yields to with the $\mathcal{CP}$-even (blue) and the $\mathcal{CP}$-odd (red) hypothesis compared to data. The data is shown with fitted background subtracted and the error bars includes the statistical uncertainty. The three categories (Odd–like, Intermediate and Even–like) are labelled corresponding to the CP BDT score ranges that they cover. In the last bin, the SM $\mathcal{CP}$–even coupling is favoured.

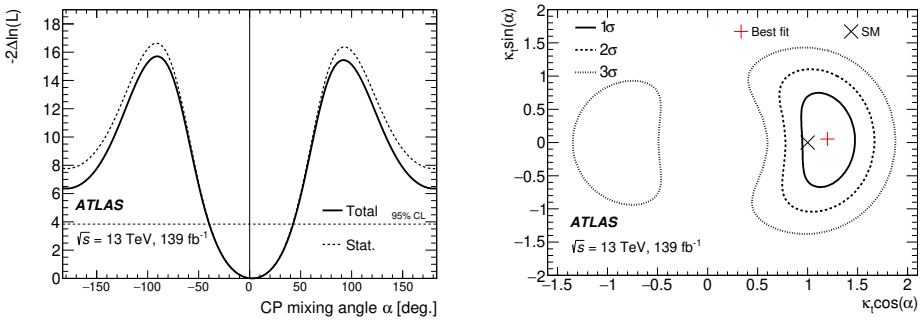


Figure 6.7: Left: One–dimensional likelihood distribution of the mixing angle α. The solid and dashed curves represent the results with total uncertainty and only statistical uncertainty considered, respectively. The horizontal dashed line represents the 95% CL. Right: two-dimensional $\kappa_t \cos\alpha - \kappa_t \sin\alpha$ likelihood contours results for the $\mathcal{CP}$ limit.

Conclusions

The SM of fundamental particles is a highly predictive theory that has driven many of the experimental, theoretical and accelerator science achievements through the last decades. It accurately describes the interactions between matter and force fields, which are treated with the same formalism, based on the local gauge invariance of the Lagrangian. The success of the SM can be partially attributed to the observation of the physical phenomena that it predicts in high–energy collider experiments. Such is the case of the discovery of the top quark or the Higgs boson, or the precise determination of the properties of these particles, and that of their couplings. In spite of this, the SM is an incomplete theory given the number of unresolved mysteries like the insufficient $\mathcal{CP}$–violating terms to explain the baryon asymmetry, or the origin of neutrino masses. These limitations motivate the study of models that extend the SM to probe higher (or lower) energy scales, and that may contain the underlying theories to explain all these physical phenomena. One of such extensions are the effective field theories, and in particular those that examine the couplings of the most massive particle of the SM, i.e. the top quark.

The top quark, its couplings and its properties serve as excellent probes of the SM. Due to its high mass, it is the only quark that decays before hadronisation, allowing for its study through its decay products. Moreover, it is the only quark with a Yukawa coupling to the Higgs boson of the order of unity, and has the potential to strongly couple to other heavy BSM particles. With all, the top quark is a crucial particle to study for both SM high–precision measurements and state–of–the–art BSM interpretations.

Progress and validation of the most advanced theories is closely tied up to the research and development of powerful collider machines and detectors that are able to capture the relevant physical processes. In this regard, the LHC provides one of the largest physics programmes spanning over a large variety of SM measurements and BSM searches; specialising in the production of top quarks in extended high–energy regimes and in association with other bosons. The work in this thesis presents several studies focused on the properties of the top quark and its EW couplings, both at the production level and by examining the angular properties of its decay products. In all of the discussed analysis, extensive work has been done to test and validate the MC simulation configurations to produce the samples that are used in them.

New physics effects can be encapsulated within additional degrees of freedom in the context of an EFT extension of the SM. Distinctively, the top quark has the potential to strongly couple to new physics objects above the threshold scale $\Lambda = 1$ TeV, which could be conducive to deviations of the SM predictions in some experimental observables. The search for such deviations typically comprises of a global fit in the EFT Wilson coefficient parameter space

to available observables measured at high–energy physics experiments, in order to constrain them. A global EFT fit in the top quark EW sector, using data from the LHC, Tevatron, and LEP/SLC that is sensitive to the corresponding Wilson coefficients ($C^{-}_{\phi Q}$, $C^{(3)}_{\phi Q}$, $C_{\varphi t}$, C_{tW}, C_{tZ}, $C_{t\varphi}$, $C_{\varphi tb}$, C_{bW}), is performed. No significant deviations from the SM are found. Nonetheless, in virtue of the latest differential measurements of $t\bar{t}Z$ and $t\bar{t}\gamma$ production and the novel parameterisation of the impact of the EFT operators on the various cross–section observables at NLO accuracy, some of the most stringent bounds on the set of Wilson coefficients are reported. These bounds have been tested against the introduction of new degrees of freedom and different sources of correlations among them in the fit, finding them to be remarkably robust. The results from this EFT fit have been published in Ref. [246].

Improvements on the constraining power of these global EFT fits is expected with the production of larger datasets at subsequent LHC and HL–LHC data–taking periods, and from the measurement of new observables that are able to disentangle degenerate directions in the Wilson coefficient parameter space.

$t\bar{t}W$ production is one of the most intriguing SM processes to date. Aside from the cross–section discrepancy between the theoretical prediction and the latest LHC measurements, the emittance of the W boson from the initial state grants it some unique characteristics. These can be probed using final–state observables that target angular properties, such as the central–forward asymmetry A_c^l based on the absolute pseudorapidity difference of the charged leptons decaying from the top quark pair. This asymmetry is enhanced in $t\bar{t}W$ production due to the absence of the symmetric gg initial state and the initial–state polarisation induced by the aforementioned ISR W boson (with respect to $t\bar{t}$ production). Additionally, it carries the discrimination power to distinguish between new BSM models independently from absolute cross–section normalisation factors.

The A_c^l is obtained at reconstruction level using 139 fb^{-1} of pp collision data collected by ATLAS at the LHC. It is measured in the 3ℓ channel and using a BDT to associate charged leptons to their top quark parents. The normalisation of the main backgrounds are constrained by dedicated control regions and are extracted simultaneously in the fit to data. In order to provide a direct comparison with theoretical predictions, the A_c^l is unfolded to particle level in a fiducial region which is chosen to be close to the reconstruction level region to minimise acceptance effects. The unfolding procedure is also validated after a set of closure and linearity tests. The results for the A_c^l, both at reconstruction and particle level, are in agreement with the SM prediction and severely dominated by the data statistics. The largest systematic uncertainties arise from the MC modelling of the signal process and the $t\bar{t}Z$ background. These results have been published in Ref. [302].

Addressing the accuracy of the theoretical calculations is paramount to improve the MC modelling of the events, which can help to reduce the magnitude of the uncertainties in any experimental measurement. While a complete NNLO calculation is inaccessible for many processes, the merging of samples with different parton multiplicities at NLO accuracy can provide significant improvements in the prediction of absolute cross–sections and kinematic distributions. Such is the case in $t\bar{t}W$ production, where sizeable differences in both are seen not only when comparing multi–leg merged setups with NLO–accurate inclusive setups, but also when varying the renormalisation and factorisation scales away from their nominal value. Furthermore, including the so–called EW contribution diagrams has a significant impact on $t\bar{t}W$ production. These studies have been documented in Ref. [241].

One of the limitations of the SM is the insufficient contributions to $\mathcal{CP}$ violation to explain, for example, the matter–antimatter asymmetry of the early universe. In the top quark sector, additional sources of $\mathcal{CP}$ violation can be probed using EFTs that can modify the chiral structure of the tWb vertex, or introduce $\mathcal{CP}$–violating terms in the top quark Yukawa coupling. The latter can be studied using MC simulations of the $t\bar{t}H$ and tH processes that provide a direct access to the top–Higgs vertex. They are parameterised with the $\mathcal{CP}$ mixing angle α, which exhibits a dependence with the inclusive cross–section of these processes. With this, a hypothesis test can be built to extract exclusion limits on the $\mathcal{CP}$ mixing angle. This analysis is performed using 139 fb^{-1} of pp collision data collected by ATLAS in a set of event categories targeting 0ℓ and 1ℓ selections. After a simultaneous fit to data in the diphoton mass spectra in all categories, the $\mathcal{CP}$–odd hypothesis ($\alpha = 90°$) is excluded at 3.9σ and $|\alpha| > 43°$ is excluded at 95% CL. These results are published in Ref. [219].

In conclusion, the top quark and its EW couplings remain as one of the cornerstones of high–energy physics programmes in the coming decades. The work on this thesis has provided a significant contribution to the study of the top quark EW sector, which has been examined in both SM properties and BSM EFT modifications of these couplings, finding a reasonable agreement with the SM predictions throughout. Nevertheless, new data coming from the LHC will allow for improvements in the precision of these measurements, possible discoveries of new physics effects, as well as the fine–tunning and development of current MC generators and reconstruction, and calibration techniques at the detectors.

The quest to crack the intricate mysteries of nature *is* on.

A EFT Wilson coefficient correlation matrices and parameterisations

This appendix shows auxiliary correlation matrices (Figures A.1 to A.6) and parameterisation tables (Tables A.1 to A.3) for the global EFT fit analysis described in Chapter 4.

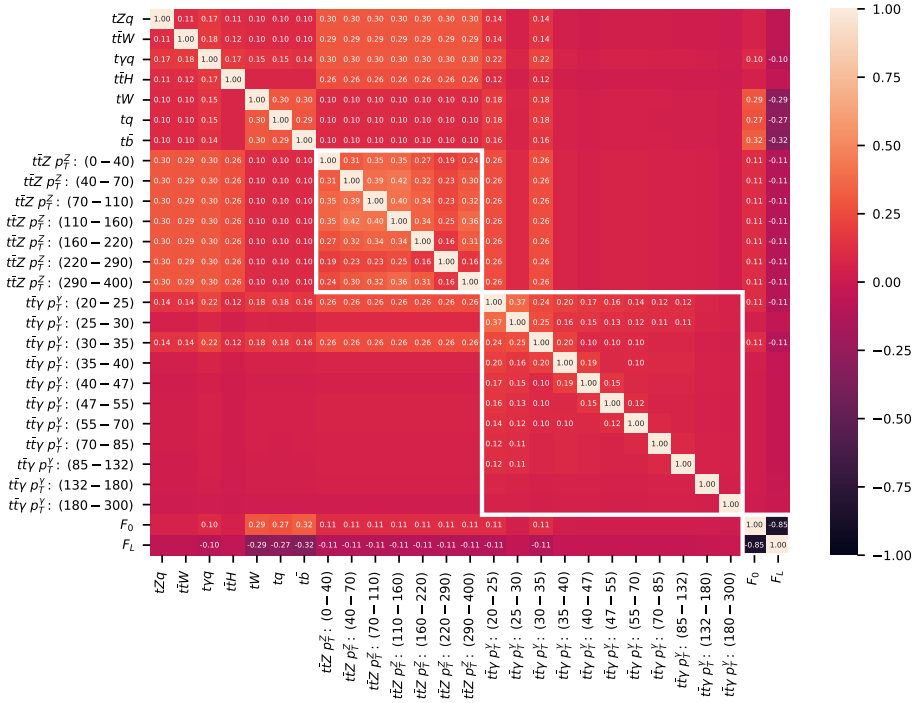


Figure A.1: Experimental correlation matrix used. The boxes in white correspond to the correlations published by the experiments for $t\bar{t}Z$, $t\bar{t}\gamma$ and W boson helicity fractions [168, 198, 212]. The rest of the entries correspond to the ansatz, as described in Section 4.6.2. Cells are filled if the correlation is higher than 10% in absolute value.

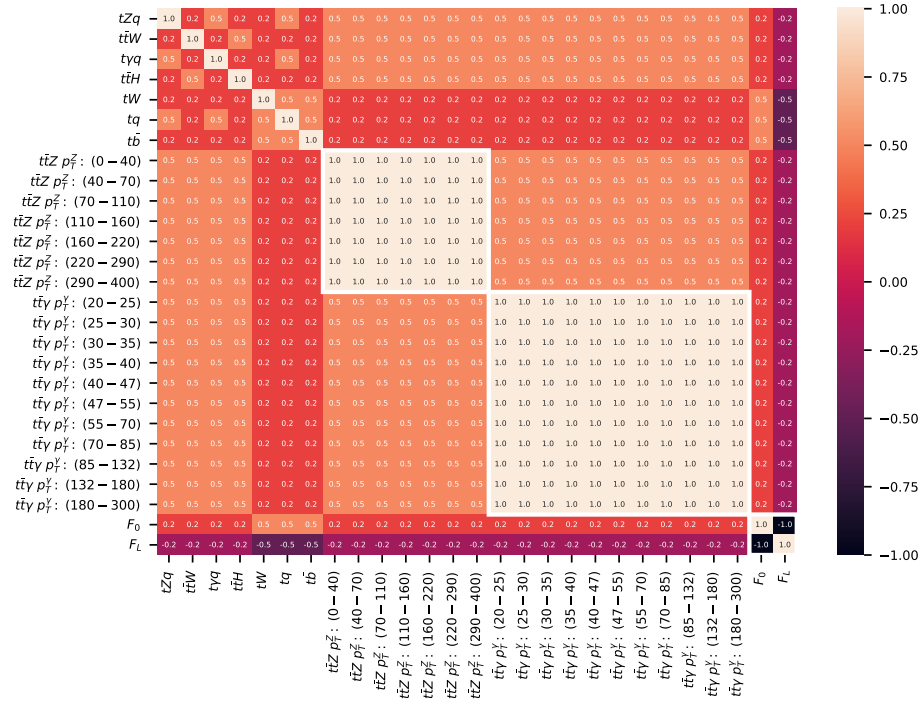


Figure A.2: Theoretical correlation matrix used. The boxes in white correspond to the correlations between the differential bins in $t\bar{t}Z$ and $t\bar{t}\gamma$ (a 100% correlation is assumed among them) and the W boson helicity fractions (a -100% correlation is assumed among them). The rest of the entries correspond to the ansatz, as described in Section 4.6.2. Cells are filled if the correlation is higher than 10% in absolute value.

Figure A.3: Correlation matrix between the different EFT operators obtained in the baseline linear (Λ^{-2}) fit. Cells are filled if the correlation is higher than 10% in absolute value. The operator $O_{\varphi b}$, that modifies only the bottom quark electro-weak couplings, is taken into account in the fit but limits on its coefficients are not reported since the obtained values are not competitive using only the observables considered in the fit.

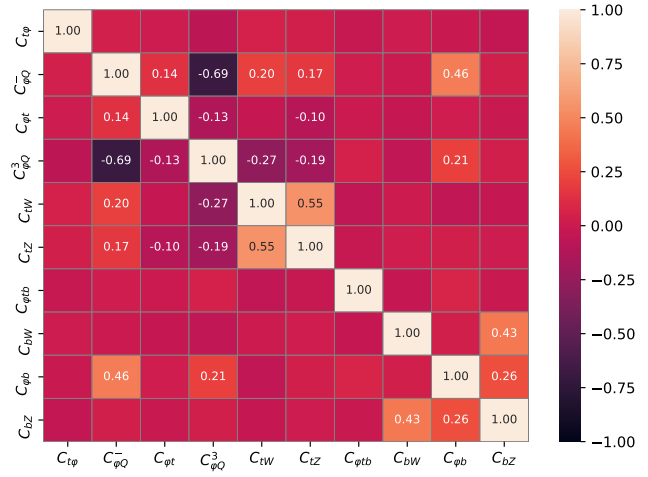


Figure A.4: Correlation matrix between the different EFT operators obtained in the baseline quadratic (Λ^{-4}) fit. Cells are filled if the correlation is higher than 10% in absolute value. The operators that modify only the bottom quark electro-weak couplings, $O_{\varphi d}$ and O_{dZ}, are taken into account in the fit but limits on their coefficients are not reported since the obtained values are not competitive using only the observables considered in the fit.

Figure A.5: Correlation matrix between the different EFT operators obtained in the linear (Λ^{-2}) "All correlations" robustness test fit described in Section 4.6.2. Cells are filled if the correlation is higher than 10% in absolute value. The operator $O_{\varphi b}$, that modifies only the bottom quark electro-weak couplings, is taken into account in the fit but limits on its coefficients are not reported since the obtained values are not competitive using only the observables considered in the fit.

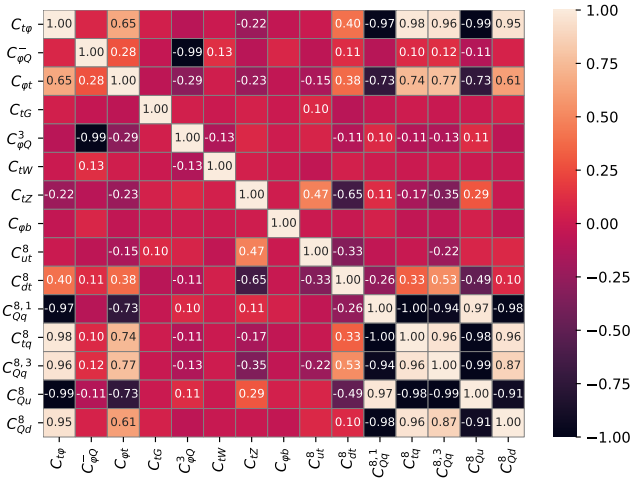


Figure A.6: Correlation matrix between the different EFT operators obtained in the linear (Λ^{-2}) "4F" robustness test fit described in Section 4.6.1. Cells are filled if the correlation is higher than 10% in absolute value. The operator $O_{\varphi b}$, that modifies only the bottom quark electro-weak couplings, is taken into account in the fit but limits on its coefficients are not reported since the obtained values are not competitive using only the observables considered in the fit.

13 TeV	Param. [fb]	$\mathcal{O}_{tZ}$	$\mathcal{O}_{tW}$	$\mathcal{O}_{\varphi Q}^{(3)}$	$\mathcal{O}_{\varphi Q}^{(-)}$	$\mathcal{O}_{\varphi t}$	$\mathcal{O}_{tG}$	$\mathcal{O}_{t\varphi}$	$\mathcal{O}_{bW}$	$\mathcal{O}_{\varphi tb}$
$t\bar{t}Z$	$\sigma_{SM,NLO}$	$740.8^{+9.7\%}_{-11\%}$								
	$\frac{\sigma^{(1)}_{i,NLO}}{\sigma_{SM,NLO}} \times \%$	$-0.4^{+38\%}_{-15\%}$	–	–	$-10.3^{+1.1\%}_{-0.1\%}$	$6.6^{+1.3\%}_{-0.2\%}$	$38.9^{+0.9\%}_{-0.9\%}$	–	–	–
	$\frac{\sigma^{(2)}_{i,NLO}}{\sigma_{SM,NLO}} \times \%$	$8.4^{+0.8\%}_{-1.4\%}$	–	–	$0.5^{+0.2\%}_{-2.2\%}$	$0.4^{+1.6\%}_{-1.2\%}$	$39.4^{+10\%}_{-6.5\%}$	–	–	–
tZq	$\sigma_{SM,NLO}$	$807.7^{+2.2\%}_{-1.7\%}$								
	$\frac{\sigma^{(1)}_{i,NLO}}{\sigma_{SM,NLO}} \times \%$	$-0.7^{+3.1\%}_{-1.5\%}$	$2.1^{+0.5\%}_{-0.2\%}$	$21.8^{+0.3\%}_{-0.1\%}$	$2.3^{+1.0\%}_{-0.8\%}$	$0.6^{+1.9\%}_{-0.1\%}$	$-0.5^{+14\%}_{-17\%}$	–	–	–
	$\frac{\sigma^{(2)}_{i,NLO}}{\sigma_{SM,NLO}} \times \%$	$1.9^{+1.2\%}_{-0.2\%}$	$10.1^{+1.6\%}_{-0.5\%}$	$3.1^{+1.0\%}_{-1.0\%}$	$0.2^{+7.0\%}_{-2.3\%}$	$0.1^{+6.3\%}_{-7.6\%}$	$0.7^{+27\%}_{-25\%}$	–	$6.0^{+2.3\%}_{-0.4\%}$	$0.4^{+1.4\%}_{-0.5\%}$
$t\bar{t}\gamma$	$\sigma_{SM,NLO}$	$1986.0^{+10.6\%}_{-11.5\%}$								
	$\frac{\sigma^{(1)}_{i,NLO}}{\sigma_{SM,NLO}} \times \%$	$-4.2^{+8.3\%}_{-6.3\%}$	$5.5^{+6.6\%}_{-3.0\%}$	–	–	–	$29.1^{+1.3\%}_{-0.4\%}$	–	–	–
	$\frac{\sigma^{(2)}_{i,NLO}}{\sigma_{SM,NLO}} \times \%$	$14.6^{+0.2\%}_{-0.5\%}$	$19.1^{+0.6\%}_{-0.3\%}$	–	–	–	$7.7^{+14\%}_{-3.0\%}$	–	–	–
$t\gamma q$	$\sigma_{SM,NLO}$	$869.5^{+2.0\%}_{-1.9\%}$								
	$\frac{\sigma^{(1)}_{i,NLO}}{\sigma_{SM,NLO}} \times \%$	$-0.6^{+8.5\%}_{-13\%}$	$2.9^{+2.9\%}_{-2.2\%}$	$12.1^{+0.1\%}_{-0.1\%}$	–	$0.1^{+13\%}_{-17\%}$	$0.5^{+20\%}_{-17\%}$	–	–	–
	$\frac{\sigma^{(2)}_{i,NLO}}{\sigma_{SM,NLO}} \times \%$	$5.3^{+0.2\%}_{-0.1\%}$	$6.6^{+1.2\%}_{-0.2\%}$	$0.4^{+0.3\%}_{-2.4\%}$	–	$< 10^{-3}$	$-0.7^{+76\%}_{-14\%}$	–	$1.5^{+1.3\%}_{-0.1\%}$	$0.1^{+1.1\%}_{-0.7\%}$

Table A.1: Parameterisations for some of the considered processes in the global fit at 13 TeV. Cells with – indicate that they contribute very little and are not included in the fit. Uncertainties are estimated using μ_R and μ_F scale variations.

13 TeV	Param. [fb]	$\mathcal{O}_{tZ}$	$\mathcal{O}_{tW}$	$\mathcal{O}_{\varphi Q}^{(3)}$	$\mathcal{O}_{\varphi Q}^{(-)}$	$\mathcal{O}_{\varphi t}$	$\mathcal{O}_{tG}$	$\mathcal{O}_{t\varphi}$	$\mathcal{O}_{bW}$	$\mathcal{O}_{\varphi tb}$
$t\bar{t}W$	$\sigma_{SM,NLO}$					$543^{+11\%}_{-10\%}$				
	$\frac{\sigma^{(1)}_{i,NLO}}{\sigma_{SM,NLO}} \times \%$	–	–	–	–	–	$29.1^{+0.7\%}_{-0.2\%}$	–	–	–
	$\frac{\sigma^{(2)}_{i,NLO}}{\sigma_{SM,NLO}} \times \%$	–	–	–	–	–	$4.3^{+33\%}_{-19\%}$	–	–	–
$t\bar{t}W$ EW	$\sigma_{SM,NLO}$					$47.5^{+19.5\%}_{-15.1\%}$				
	$\frac{\sigma^{(1)}_{i,NLO}}{\sigma_{SM,NLO}} \times \%$	$-1.0^{+7.9\%}_{-5.4\%}$	$25.2^{+8.1\%}_{-6.0\%}$	$7.6^{+0.1\%}_{-0.3\%}$	$-7.6^{+0.9\%}_{-0.9\%}$	$3.7^{+3.1\%}_{-2.6\%}$	$10.3^{+0.9\%}_{-1.0\%}$	–	–	–
	$\frac{\sigma^{(2)}_{i,NLO}}{\sigma_{SM,NLO}} \times \%$	$2.3^{+10\%}_{-7.1\%}$	$57.5^{+4.3\%}_{-3.1\%}$	$6.2^{+3.5\%}_{-2.5\%}$	$1.6^{+2.2\%}_{-2.2\%}$	$1.6^{+3.2\%}_{-2.4\%}$	$2.4^{+7.9\%}_{-8.8\%}$	–	–	–
$t\bar{t}H$	$\sigma_{SM,NLO}$					$459^{+6.3\%}_{-9.3\%}$				
	$\frac{\sigma^{(1)}_{i,NLO}}{\sigma_{SM,NLO}} \times \%$	–	–	–	–	–	$99.8^{+0.6\%}_{-0.2\%}$	$-12.9^{+1.8\%}_{-1.0\%}$	–	–
	$\frac{\sigma^{(2)}_{i,NLO}}{\sigma_{SM,NLO}} \times \%$	–	–	–	–	–	$130^{+1.2\%}_{-1.7\%}$	$0.4^{+4.3\%}_{-1.8\%}$	–	–
tHq	$\sigma_{SM,(N)LO}$				$71.2^{+1.9\%}_{-1.6\%}$ (NLO)				$66.6^{+2.5\%}_{-4.3\%}$ (LO)	
	$\frac{\sigma^{(1)}_{i,(N)LO}}{\sigma_{SM,(N)LO}} \times \%$	–	$29.6^{+1.6\%}_{-1.5\%}$	$-11.6^{+3.3\%}_{-2.3\%}$	–	–	$-0.2^{+32\%}_{-38\%}$	$-0.7^{+16\%}_{-18\%}$	$0.1^{+0.1\%}_{-0.1\%}$	$0.01^{+43\%}_{-45\%}$
	$\frac{\sigma^{(2)}_{i,(N)LO}}{\sigma_{SM,(N)LO}} \times \%$	–	$78.8^{+0.6\%}_{-0.1\%}$	$18.8^{+0.6\%}_{-0.2\%}$	–	–	$-1.9^{+23\%}_{-22\%}$	$1.1^{+0.6\%}_{-0.4\%}$	$2.9^{+0.1\%}_{-0.9\%}$	$0.1^{+1.7\%}_{-1.9\%}$

Table A.2: Continuation of Table A.1. Parameterisations for some of the considered processes in the global fit at 13 TeV. Cells with – indicate that they contribute very little and are not included in the fit. Uncertainties are estimated using μ_R and μ_F scale variations.

8 TeV	Param. [pb]	$\mathcal{O}_{tZ}$	$\mathcal{O}_{tW}$	$\mathcal{O}_{\varphi Q}^{(3)}$	$\mathcal{O}_{\varphi Q}^{(-)}$	$\mathcal{O}_{\varphi t}$	$\mathcal{O}_{tG}$	$\mathcal{O}_{t\varphi}$	$\mathcal{O}_{bW}$	$\mathcal{O}_{\varphi tb}$
tq (t-ch)	$\sigma_{SM,NLO}$					$82.6^{+2.6\%}_{-2.2\%}$				
	$\frac{\sigma^{(1)}_{i,NLO}}{\sigma_{SM,NLO}} \times \%$	–	$3.1^{+0.5\%}_{-0.2\%}$	$11.7^{+0.4\%}_{-0.5\%}$	–	–	–	–	–	–
	$\frac{\sigma^{(2)}_{i,NLO}}{\sigma_{SM,NLO}} \times \%$	–	$1.5^{+2.0\%}_{-1.2\%}$	$0.2^{+2.3\%}_{-6.8\%}$	–	–	–	–	$0.5^{+2.4\%}_{-1.2\%}$	$0.1^{+1.8\%}_{-1.5\%}$
tW	$\sigma_{SM,NLO}$					$22.4^{+6.8\%}_{-6.7\%}$				
	$\frac{\sigma^{(1)}_{i,NLO}}{\sigma_{SM,LO}} \times \%$	–	$-8.8^{+0.3\%}_{-0.4\%}$	$12.1^{+0.3\%}_{-0.2\%}$	–	–	$7.5^{+1.7\%}_{-1.6\%}$	–	–	–
	$\frac{\sigma^{(2)}_{i,NLO}}{\sigma_{SM,LO}} \times \%$	–	$2.2^{+0.7\%}_{-0.5\%}$	$0.4^{+7.0\%}_{-1.6\%}$	–	–	–	–	$1.0^{+1.3\%}_{-2.7\%}$	$0.1^{+0.7\%}_{-0.4\%}$
tb (s-ch)	$\sigma_{SM,NLO}$					$4843.0^{+3.6\%}_{-3.3\%}$				
	$\frac{\sigma^{(1)}_{i,NLO}}{\sigma_{SM,NLO}} \times \%$	–	$46.9^{+0.1\%}_{-0.1\%}$	$12.3^{+0.1\%}_{-0.1\%}$	–	–	$-0.1^{+23\%}_{-18\%}$	–	–	–
	$\frac{\sigma^{(2)}_{i,NLO}}{\sigma_{SM,NLO}} \times \%$	–	$8.8^{+0.2\%}_{-0.5\%}$	$0.6^{+4.0\%}_{-3.2\%}$	–	–	$1.8^{+9.8\%}_{-2.9\%}$	–	$3.7^{+0.3\%}_{-0.1\%}$	$0.1^{+0.5\%}_{-0.1\%}$
tb (s-ch) @ 1.96 TeV	$\sigma_{SM,LO}$					$0.7^{+9.3\%}_{-8.1\%}$				
	$\frac{\sigma^{(1)}_{i,LO}}{\sigma_{SM,LO}} \times \%$	–	$46.1^{+0.1\%}_{-0.1\%}$	$12.1^{+0.1\%}_{-0.1\%}$	–	–	$< 10^{-3}$	–	–	–
	$\frac{\sigma^{(2)}_{i,LO}}{\sigma_{SM,LO}} \times \%$	–	$7.3^{+0.9\%}_{-0.5\%}$	$0.4^{+1.5\%}_{-2.2\%}$	–	–	$1.2^{+4.2\%}_{-6.3\%}$	–	$3.1^{+0.5\%}_{-0.3\%}$	$0.1^{+0.1\%}_{-0.1\%}$

Table A.3: Parameterisations for each of the considered processes in the global fit at 8 TeV. Cells with – indicate that they contribute very little and are not included in the fit. Uncertainties are estimated using μ_R and μ_F scale variations.

B $t\bar{t}W$ QCD and EW modelling with multi–leg setups

In this appendix, MC samples are generated to merge the $t\bar{t}W + 1j$ ($\mathcal{O}(\alpha_s^4\alpha)$) MEs at NLO accuracy with the standard $t\bar{t}W + 0j$@NLO calculation – what is known as a multi–leg setup. In addition, the effect of including the EW contributions to the $t\bar{t}W$ rate and kinematic distributions is also studied.

B.1 $t\bar{t}W$ QCD production

The cross–section for $t\bar{t}W$ production has been computed to NLO QCD+EW accuracy [349] (up to $\mathcal{O}(\alpha_s^3\alpha)$ and $\mathcal{O}(\alpha_s^2\alpha^2)$) and using a fixed scale $\mu = (2m_t + m_W)/2$ documented in the CERN YR4 [227]: $\sigma(t\bar{t}W) = 600.8^{+13\%}_{-12\%}$ fb, where the uncertainties are from variations of the μ_R and μ_F scales, PDFs and α_s. The top quark and the electroweak vector boson are treated as stable particles in the narrow–width approximation. Major efforts have been undertaken to improve this calculation. As already mentioned, complete–NLO corrections (including also the tree–level EW contributions) [237,238] have been computed. Furthermore, soft gluon resummation up to NNLL in perturbative QCD [350–352] with all $\delta\sigma_{\text{EW}}$ terms are available. Moreover, complete top quark off–shell effects and the NLO QCD corrections to top quark decays for fully leptonic decay modes [353] have been derived and these effects have been found to be rather small (less than 5%).

B.1.1 NLO inclusive and multi–leg merged setups

A first set of MC samples is generated using aMC@NLO v2.6.7 using a 5F scheme and two different configurations: one with NLO inclusive MEs and NLO matching to the PS (referred to as "NLOinc"), and one with NLO multi–leg merging using the FxFx algorithm [116] (referred to as "FxFx") using NLO–accurate MEs for up to one additional jet and LO–accurate matrix elements for up to two additional jets ($t\bar{t}W + 0, 1j$@NLO $+2j$@LO). The inclusive configuration makes use of a dynamical functional form of $H_T/2$ for the μ_R and μ_F scales in the MEs and for the shower starting scale μ_{sh} (also referred to as resummation scale), where

$$H_T = \sum_i m_{T,i} = \sum_i \sqrt{m_i^2 + p_{T,i}^2}\,, \tag{B.1}$$

and where i runs over all outgoing partons from the ME calculation. The multi–leg configuration makes use of complex functional forms for the μ_R, μ_F and μ_{sh} scales which are chosen dynamically and depend on the kinematics of the event after the merging of the core process with the additional partons following the

FxFx merging prescription* [116, 119, 354].

In both samples, the decay of the top quarks (and its decay product W boson) and of the associated W boson is performed at LO with MadSpin [315, 316], preserving spin–correlations and polarization effects. The top quark mass is set to $m_t = 172.5$ GeV and the W boson mass to $m_W = 80.399$ GeV. Pythia 8.244 is used for the PS, along with the A14 tune. The decays of bottom and charm hadrons are simulated using the EvtGen v.1.7.0. The PDFs used are NNPDF3.0nlo and NNPDF2.3lo for the ME and PS, respectively. Within the multi–leg configurations, there are two additional parameters. The aforementioned merging scale parameter μ_Q (see Section 2.2.3), and the minimum jet p_{T} (`ptj`) of the additional parton in the ME generation.

Table B.1 summarizes these setups and shows their respective cross–sections. An alternative sample using Sherpa 2.2.8 MePs@NLO, with the same accuracy as the FxFx sample, is also shown. The higher jet multiplicity diagrams that enter at $\mathcal{O}(\alpha_s^4\alpha)$ in the multi–leg samples enhance the cross–section by about 10%.

ME	PS	ME PDF PS PDF	Tune	Matching/ Merging	Cross–section [fb]
aMC@NLO v2.6.7 NLOinc	Pythia 8.244	NNPDF3.0nlo NNPDF2.3lo	A14	MC@NLO	$545.7^{+10.0\%}_{-9.8\%}$(scale)
aMC@NLO v2.6.7 $+0,1j$@NLO $+2j$@LO				FxFx merging $\mu_Q = 30$ GeV	$614.2^{+11.8\%}_{-12.6\%}$(scale)
Sherpa 2.2.8 $+0,1j$@NLO $+2j$@LO	Sherpa 2.2.8	NNPDF3.0nlo NNPDF3.0nlo	Author's tune	MePs@NLO $\mu_Q = 30$ GeV	$589.2^{+15.5\%}_{-14.7\%}$(scale)

Table B.1: A summary of the $t\bar{t}W$ QCD production samples studied in this appendix. The matrix element generator, parton shower generator, tuned parameters, and the corresponding PDF sets used are given. Additional information regarding the matching and merging scheme used by the NLO+PS setup is provided. For aMC@NLO, the uncertainties cover the variations of the μ_R and μ_F scales in the ME by a factor 2 using 7–point variations. For the Sherpa sample, the same procedure is applied but scale uncertainties include variations in both the ME and the PS. The PDF uncertainties are below 2%.

To estimate the effect of missing higher order corrections, the NLOinc configuration is tested with two different functional forms of the dynamical scale: an arithmetic mean $\mu_{a,0} = H_T/2 = \frac{1}{2}\sum_i m_{T,i}$, and a geometrical mean $\mu_{g,0} = \sqrt[3]{m_{T,t} \cdot m_{T,\bar{t}} \cdot m_{T,W}}$; and a fixed scale $\mu_{m,0} = (2m_t + m_W)/2$. The dependence can be seen in the left panel of Figure B.1. In each point, the μ_R

*They depend on the phase–space configuration and are related to the clustering scales of the additional partons and on the core process.

and μ_F scales are independently varied by a factor 2 around its central value using 7–point variations[†]. As a result, predictions of the cross–sections with variations up to 8 times around μ_0 are shown. For reference, the latest ATLAS and CMS $t\bar{t}W$ cross–section measurements [234, 235] are also given. Higher cross–sections are obtained using the fixed or geometrical mean scales. Furthermore, the central cross–sections, $\mu_{m,0}$ and $\mu_{g,0}$, are close to the predicted cross–section of the FxFx multi–leg setup in Table B.1; and lower values of these scales predict higher cross–sections that increase their compatibility with the measured values. The estimate of the uncertainties due to missing higher orders by considering variations of these scales by factors 2, 4, and 8, are $\pm(10-12)\%$, $\pm(25-30)\%$, and $\pm(30-50)\%$ respectively.

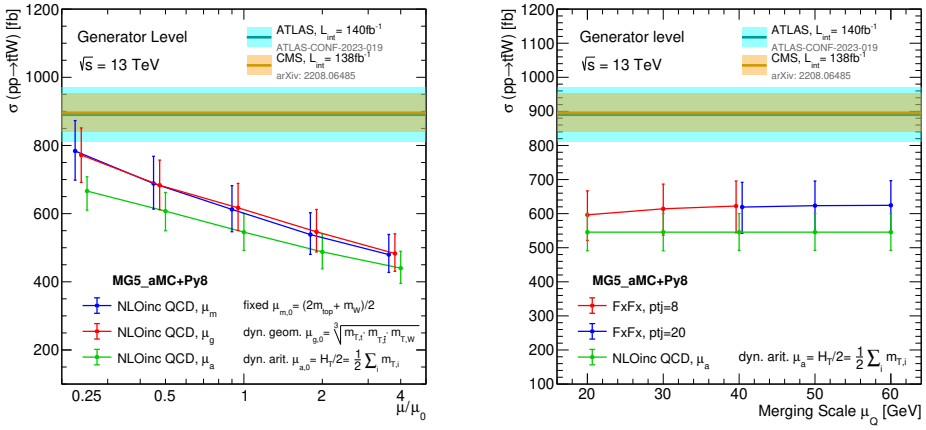


Figure B.1: Left: Dependence of the μ_R and μ_F scales with the aMC@NLO NLOinc sample cross–section, for three different functional forms. The scales are varied by factors 1/4, 1/2, 2 and 4 with respect to their nominal values μ_0. Right: Merging scale and `ptj` variations for the FxFx sample compared to the NLO inclusive sample with $\mu_a = H_T/2$. The latest ATLAS and CMS measurements of the $t\bar{t}W$ inclusive cross–section is also shown for reference in the cyan and orange bands, respectively. Uncertainty bands cover the variations of the μ_R and μ_F scales in the ME by a factor 2 around each point using 7–point variations. This figures are modified versions from those in Ref. [241]: the reference cross–section measurements have been updated to the latest values.

For the FxFx sample, different values of the merging scale and the minimum jet p_{T} of the additional parton are scanned following the recommendation of the authors to use values such that $2 \times \texttt{ptj} \leq \mu_Q$ [116]. A small dependence (below 5%) of the cross–section on these parameters is observed. This can be seen on

[†]The μ_R and μ_F scales are varied independently around the central value μ with the combinations $(\mu_R, \mu_F) = (\mu/2, \mu/2), (2\mu, 2\mu), (\mu, 2\mu), (2\mu, \mu), (\mu, \mu/2)$ and $(\mu/2, \mu)$ and taking an envelope of the predictions to estimate the error.

the right panel of Figure B.1 and in Table B.2. The benchmark rate is denoted by †. The effect of the merging scale on the cross–section indicates that opening the phase–space to the soft collinear emissions of the parton shower yields to slightly higher cross–section values for the studied range. In Ref. [240], similar studies using the same aMC@NLO and Pythia versions are conducted beyond the range that is shown here. Conversely, these show a small decrease of the cross–section for increasing $\mu_Q > 70$ GeV. To check for this effect, an additional point is generated for $\mu_Q = 150$ GeV, which outputs a cross–section of about 600 fb; in–line with the results in Ref. [240]‡.

ME	PS	ME PDF PS PDF	Tune	FxFx merging	Cross–section [fb]
MG5_aMC v2.6.7 $+0,1j$@NLO $+2j$LO	Pythia 8.244	NNPDF3.0nlo NNPDF2.3lo	A14	$\mu_Q = 20$ GeV ptj $= 8$ GeV	$596.5^{+11.8\%}_{-12.6\%}$(scale)
				$\mu_Q = 30$ GeV ptj $= 8$ GeV	$614.2^{+11.8\%}_{-12.6\%}$(scale)†
				$\mu_Q = 40$ GeV ptj $= 8$ GeV	$622.3^{+11.8\%}_{-12.6\%}$(scale)
				$\mu_Q = 40$ GeV ptj $= 20$ GeV	$619.3^{+11.7\%}_{-12.4\%}$(scale)
				$\mu_Q = 50$ GeV ptj $= 20$ GeV	$623.4^{+11.6\%}_{-12.4\%}$(scale)
				$\mu_Q = 60$ GeV ptj $= 20$ GeV	$624.4^{+11.7\%}_{-12.4\%}$(scale)

Table B.2: Simulated samples for $t\bar{t}W$ $+0,1j$@NLO $+2j$LO with FxFx merging. Different merging scale values from 20 to 60 GeV are scanned as well as a ptj variation for 8 and 20 GeV. The PDF uncertainty is 1.4% for all these samples.

For all the distributions shown in this appendix, the top quarks, associated W bosons, and their decay products are defined at parton level, i.e. before the event showering; while the jets and b–jets are defined at particle level.

B.1.2 Kinematic distributions with the multi–leg setups

In this section, the selected events are required to have a semileptonic decay of the $t\bar{t}$ system, as well as a leptonic decay of the associated W boson; being both charged leptons of the same sign§. Hence, the final state is expected to have two same–sign charged leptons and at least four jets, two of them originating

‡Simulations in Ref. [240] work in the 4F scheme, while the 5F scheme is used for the simulations this appendix. Cross–sections in the 4F scheme are expected to be about 6% higher than those in the 5F scheme [240].

§This requirement is applied at parton level using MC event record, but no explicit particle level event selection in terms of number of leptons, jets or b–jets is required. No additional requirements are applied to the leptons.

from b quarks. Kinematic cuts are applied to the jets and b–jets: $p_T > 25$ GeV and $|\eta| < 2.5$.

Figure B.2 shows the effect of the μ_R and μ_F scale variations for some kinematic distributions. It compares the NLOinc samples with $\mu_a = H_T/2$ and $\mu_a = H_T/8$, and the nominal FxFx multi–leg sample. The distributions are normalised to their respective cross–sections, and the ratio shows the difference with respect to the NLOinc $\mu_a = H_T/2$ sample. There are visible differences in the shapes of the distributions when comparing the $t\bar{t}W$ NLOinc $\mu_a = H_T/2$ sample, with the other setups. A higher jet multiplicity is observed in the multi–leg sample with a larger than 20% increase for events with $N_{\text{jets}} \geq 6$. Harder leading jet and $t\bar{t}$ system p_{T} distributions are also predicted by the multi–leg setup, as well as a different angular azimuthal separation between the two top quarks. Similarly to what is seen in the left panel of Figure B.1, the dependence on the scale choice is large for the NLOinc samples. In addition, the NLOinc $\mu_a = H_T/8$ sample shows a good agreement with the multi–leg sample. Therefore, selecting lower values of the central scale in the ME for NLO inclusive configurations, at least for the scale values compared here, mimics the behaviour of the higher–accuracy FxFx sample.

Figure B.3 presents the jet multiplicity and the leading jet p_{T} for different values of μ_Q for the FxFx multi–leg setup. The distributions are normalised to their respective cross–sections, and the ratio shows the difference with respect to benchmark FxFx sample. The choice of the value of the μ_Q parameter is rather small for these kinematic variables.

Figure B.4 shows comparisons of the two multi–leg samples displayed in Table B.1 (the aMC@NLO + Pythia 8 FxFx and Sherpa 2.2.8 samples) for some kinematic distributions[¶]. The distributions are normalised to their respective cross–sections, and the ratio shows the difference with respect to the FxFx sample. Overall, both samples agree within their scale variation uncertainties. This Sherpa sample has also been compared with previous versions (Sherpa 2.2.1) and the shapes of the distributions are found to be very similar.

[¶]The spectrum of the sixth jet in the event corresponds to the second additional jet and enters at LO accuracy in these MC samples.

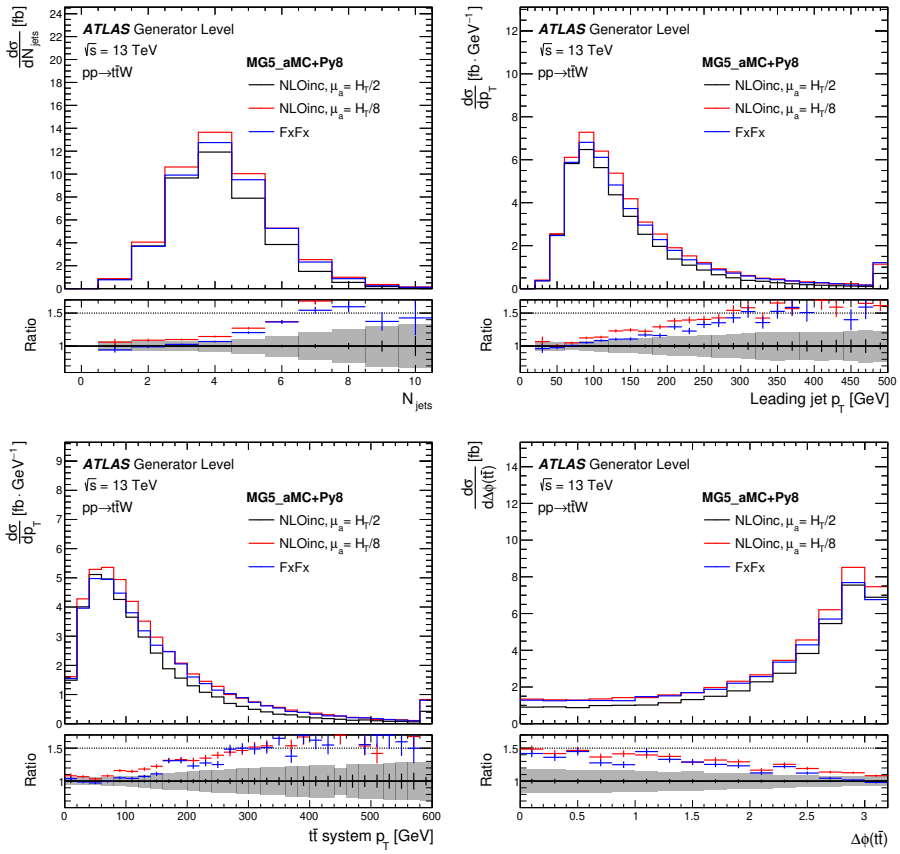


Figure B.2: Distributions of the jet multiplicity and leading jet p_{T} (top); p_{T} of the $t\bar{t}$ system and $\Delta\phi(t,\bar{t})$ (bottom) for both $t\bar{t}W$ NLO inclusive with $\mu_a = H_T/2$ and $\mu_a = H_T/8$ (black and red), and $t\bar{t}W$ FxFx multi–leg setup (blue). The distributions are normalised to their respective cross–sections. The bottom panels show the ratio with respect to the NLOinc $\mu_a = H_T/2$ sample. The vertical error lines represent the statistical errors and the shaded bands cover the variations of the μ_R and μ_F scales in the ME by a factor 2 using 7–point variations.

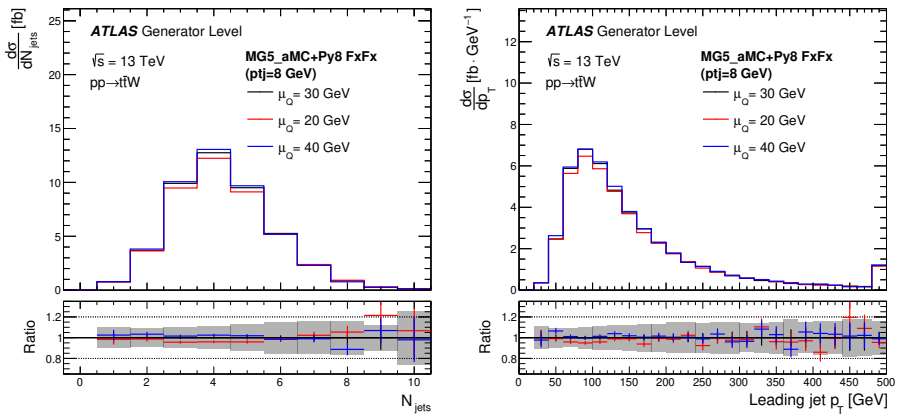


Figure B.3: Distributions of the jet multiplicity and leading jet p_{T} for $\mu_Q = 20, 30, 40$ GeV and a fixed `ptj`= 8 GeV. The distributions are normalised to their respective cross–sections. The bottom panels show the ratio with respect to the benchmark FxFx sample ($\mu_Q = 30$ GeV). The vertical error lines represent the statistical errors and the shaded bands cover the variations of the μ_R and μ_F scales in the ME by a factor 2 using 7–point variations.

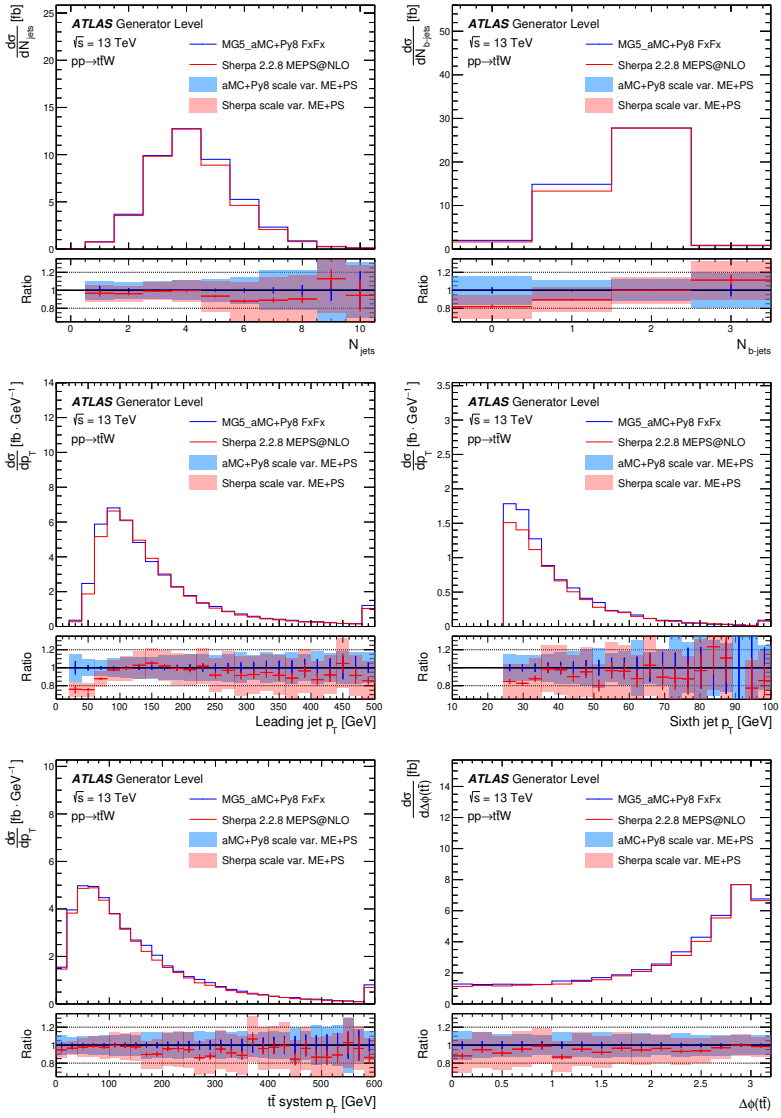


Figure B.4: Distributions of the jet multiplicity and b-jet multiplicity (top); leading jet and sixth jet p_{T} (middle); p_{T} of the $t\bar{t}$ system and $\Delta\phi(t,\bar{t})$ (bottom) for the FxFx multi–leg (blue) and the Sherpa 2.2.8 (red) samples. The distributions are normalised to their respective cross–sections. The bottom panels show the ratio with respect to the FxFx sample. The vertical error lines represent the statistical errors. For the FxFx sample, the blue shaded bands cover the variations of the μ_R and μ_F scales in the ME by a factor 2 using 7–point variations. For Sherpa, the red shaded band includes scale variations in both the ME and the PS.

B.2 $t\bar{t}W$ EW production

As discussed, contributions from $t\bar{t}W$ EW production increase the $t\bar{t}W$ cross–section by about 10% due to the opening of tW scattering diagrams [237, 238]. A separate sample, using aMC@NLO, is generated at NLO in QCD including the tree–level EW contribution terms $\mathcal{O}(\alpha^3) + \mathcal{O}(\alpha_s\alpha^3)$. It uses the same configuration as the NLOinc QCD production sample shown in Table B.1. The predicted cross–section is $49.1^{+19.1\%}_{-14.8\%}$(scale) fb.

In a similar way to $t\bar{t}W$ QCD production, the μ_R and μ_F scales are varied for a fixed (μ_m) and a dynamical (μ_a) functional form, as shown in Figure B.5. A 10% increase in the cross–section is seen throughout the whole range with respect to the NLO QCD inclusive samples described in Section B.1.1.

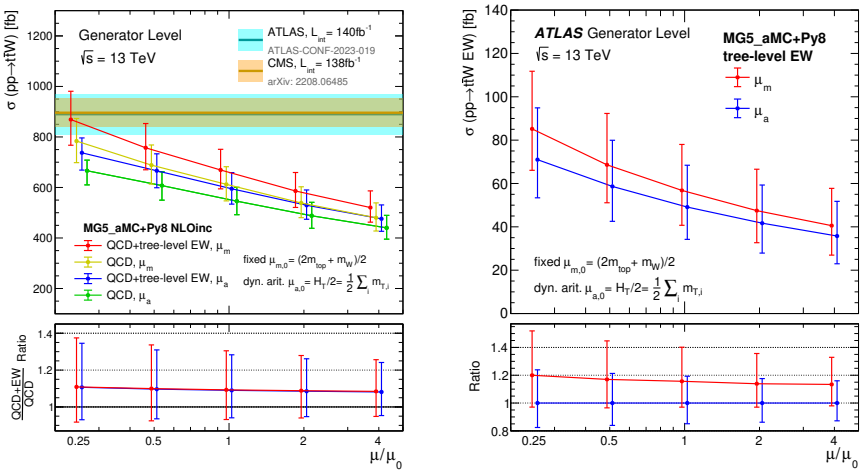


Figure B.5: Left: Dependence of the μ_R and μ_F scales with the aMC@NLO NLOinc sample cross–section, for three different functional forms. The tree–level EW contribution is added on top of the NLO QCD one where indicated. The scales are varied by factors 1/4, 1/2, 2 and 4 with respect to their nominal values μ_0. This figure is a modified version from that in Ref. [241]: the reference cross–section measurements have been updated to the latest values. Right: Same scale variations but only comparing the tree–level EW cross–section predictions. Uncertainty bands cover the variations of the μ_R and μ_F scales in the ME by a factor 2 around each point using 7–point variations.

The effect of the tree–level EW contributions is studied for some kinematical distributions in Figure B.6. The event selection follows the same requirements as in Section B.1.1, with the exception that forward jets (in the $2.5 < |\eta| < 4.5$ region) are also included. The EW contributions have a 20% effect for events with $N_{\text{central jets}} \geq 6$, as well as in the forward jets multiplicity, and in the high pseudorapidity region. This can be understood since the additional

jet produced in the tW scattering is expected to be in the forward region. Furthermore, it advocates that inclusive scaling factors are not sufficient and that EW contribution should be included in the simulations explicitly where possible.

B.3 Summary and outlook

For $t\bar{t}W$ production, both NLO inclusive and multi–leg merged (up to 1 jet at NLO and 2 jets at LO) configurations are studied for two different MC event generators: aMC@NLO interfaced to Pythia and Sherpa 2.2.8. In comparing these event generators with the same accuracy, no significant differences are observed. The largest difference is seen for the multiplicity of b–jets. This can be partially attributed to the different treatment in the decays of bottom and charm hadrons[‖]. When using a multi–leg configuration, a higher jet multiplicity and harder jets are seen with respect to the NLO inclusive one. Detailed studies varying the functional form of the μ_R and μ_F scales are also performed, and a strong dependence of the cross–section with the scale choice is seen for the NLO inclusive samples.

In addition, the tree–level $t\bar{t}W$ EW contributions to this process are studied. An overall increase of 10% in the $t\bar{t}W$ cross–section is observed for aMC@NLO, being enhanced in some regions of the kinematic distributions. Notably, the predicted cross–section calculated with a lower scale, combined with the EW contribution, agrees better with the latest experimental measurements.

Developments on multi–leg modelling have occurred since the realisation of the studies presented in this appendix. In fact, the MC generator versions used in the analysis described in Chapter 5 supersede those shown here. In particular for the FxFx sample, recent MC studies have shown a mis–modelling in the p_{T} spectra of the leading jet due to the incorrect treatment of extra emissions attached to an EW vertex in the final state [236]. Improvements on this treatment have been implemented in the latest aMC@NLO (>v3)+ Pythia 8 FxFx versions. With them, the cross–section showcases a bigger stability with the choice of μ_Q than the one reported in this appendix and in Ref. [240]. The most accurate $t\bar{t}W$ cross–section prediction to date, combining the FxFx multi–leg calculation and the complete set of EW contributions $\delta\sigma_{\mathrm{EW}}$, is calculated to be

$$\sigma = 722.4^{+9.7\%}_{-10.8\%}(\text{scale}) \pm 1.0\%(\text{PDF})\ \text{fb},$$

where the scale uncertainties are estimated from the μ_R and μ_F scale variations [236].

[‖] EvtGen has not been used for the Sherpa samples.

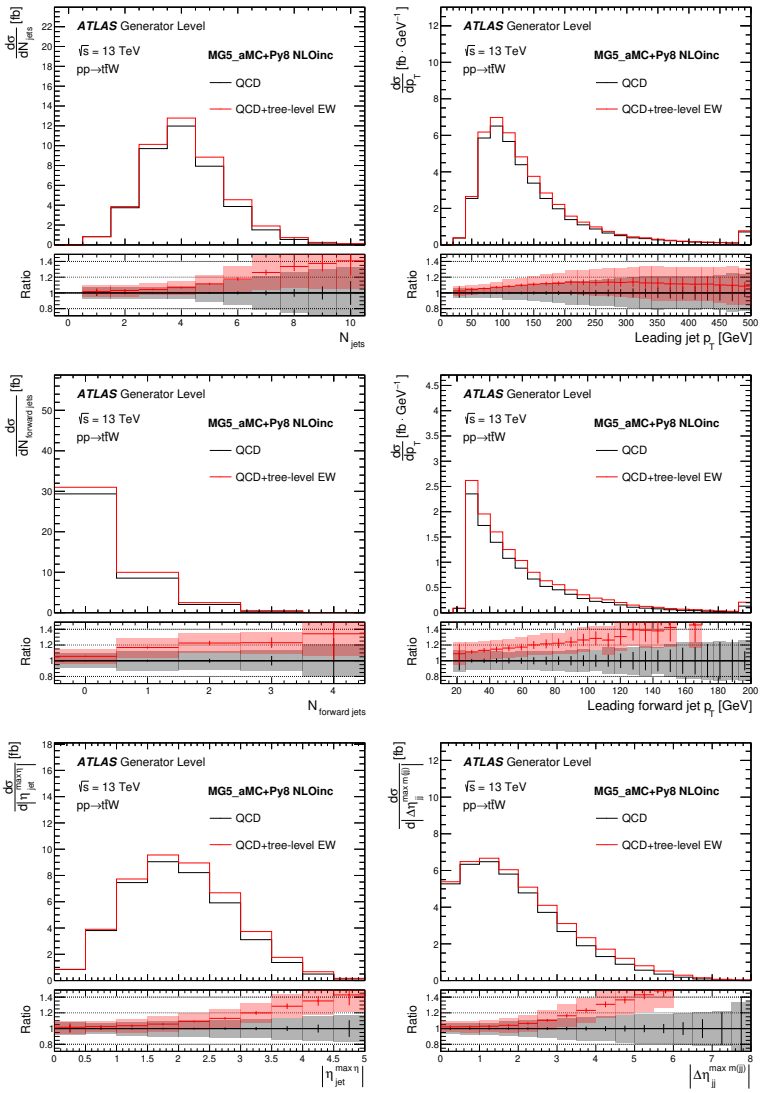


Figure B.6: Distributions for jet multiplicity and leading jet p_{T} (top); forward jet multiplicity and leading forward jet p_{T} (middle); $|\eta|$ of the most forward jet and $|\Delta\eta|$ of the di–jet pair with maximum invariant mass (bottom) for the aMC@NLO $t\bar{t}W$ QCD NLO inclusive (black) and the effect of the tree–level EW sample (red). The distributions are normalised to their respective cross–sections. The bottom panels show the ratio with respect to the NLOinc QCD sample. The vertical error lines represent the statistical errors and the shaded bands cover the variations of the μ_R and μ_F scales in the ME by a factor 2 using 7–point variations.

C Lepton truth classification

Within the ATLAS software, several tools are used to identify and classify prompt leptons, as well as non–prompt leptons, from different sources using the MC event record. One of such tools, the `IFFTruthClassifier`*, is used for the classification of the non–prompt lepton categories described in Section 5.4. Table C.1 shows the classification scheme for this tool.

Origin	Category
Unknown	0
KnownUnknown	1
Prompt (isolated) e	2
Charge–flip e	3
Prompt μ	4
Prompt γ conversion	5
e reconstructed as μ	6
τ decay	7
b–hadron decay	8
c–hadron decay	9
Light flavour decay	10

Table C.1: Correspondence between `IFFTruthClassifier` category and particle origin.

Figure C.1 shows the origin of the leading and third lepton for the combination of all SRs (top), the CR–HF$_e$ (middle), and the CR–HF$_\mu$ (bottom) based on this classification. The event yields correspond to the expected (pre–fit) yields. In the HF regions, the third lepton is the non–prompt one (categories 8 and 9) for more than 99% of the events, while the leading (and sub–leading) lepton originates from the top quark pair (categories 2 and 4); this behaviour is also present in the SRs. Throughout all analysis regions, the contributions coming from charge–flip electrons (category 3) is negligible. As already discussed, the nominal $t\bar{t}W$ Sherpa sample does not have the MC event record properly stored, due to this missing information, leptons are assigned mostly to category 1.

*The `IFFTruthClassifier` uses the definitions from a more comprehensive classification tool, but provides a more simplified and physics–oriented classification scheme.

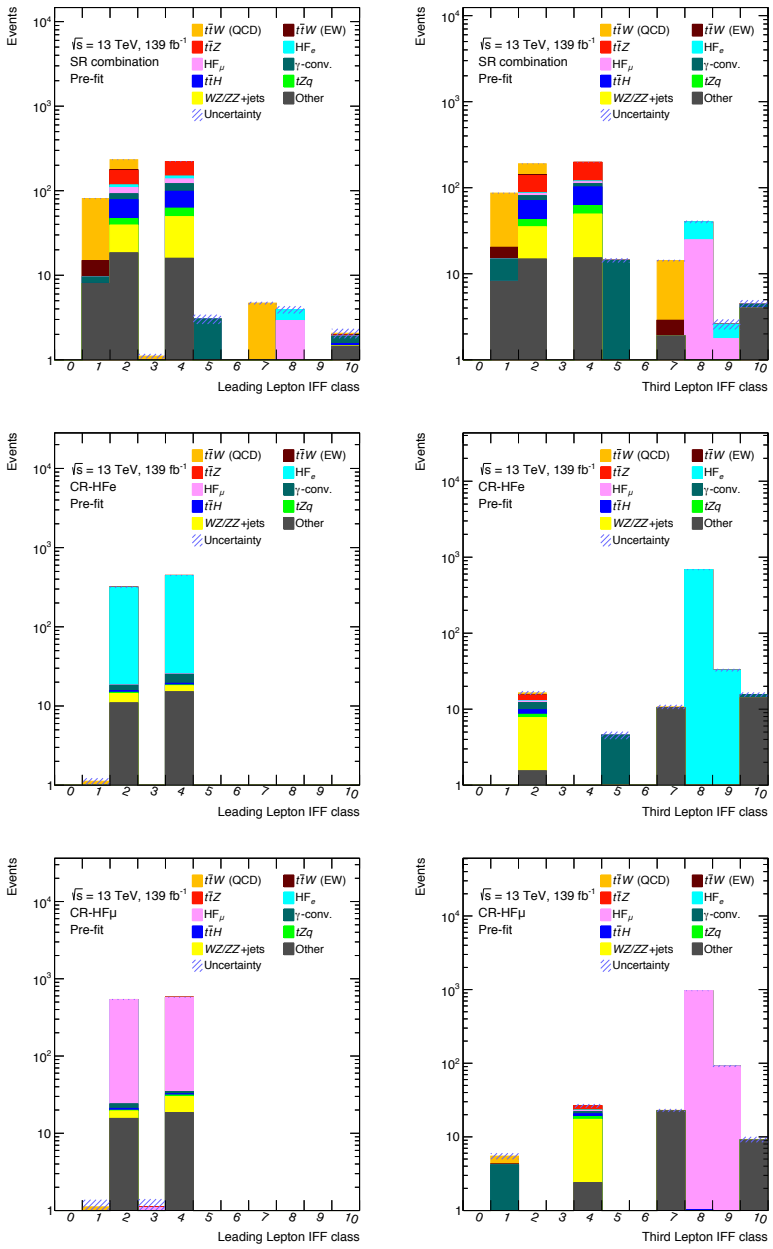


Figure C.1: Lepton origins based on the `IFFTruthClassifier` scheme for the leading (left) and third (right) leptons, for the combination of the SRs (top), CR–HF_e (middle), and CR–HF_μ (bottom). Uncertainty bands cover statistical errors.

D Theory modelling systematic variations

This appendix shows the theory modelling systematic variations for the main backgrounds of the A_c^l analysis described in Chapter 5.

D.1 $t\bar{t}Z$ modelling

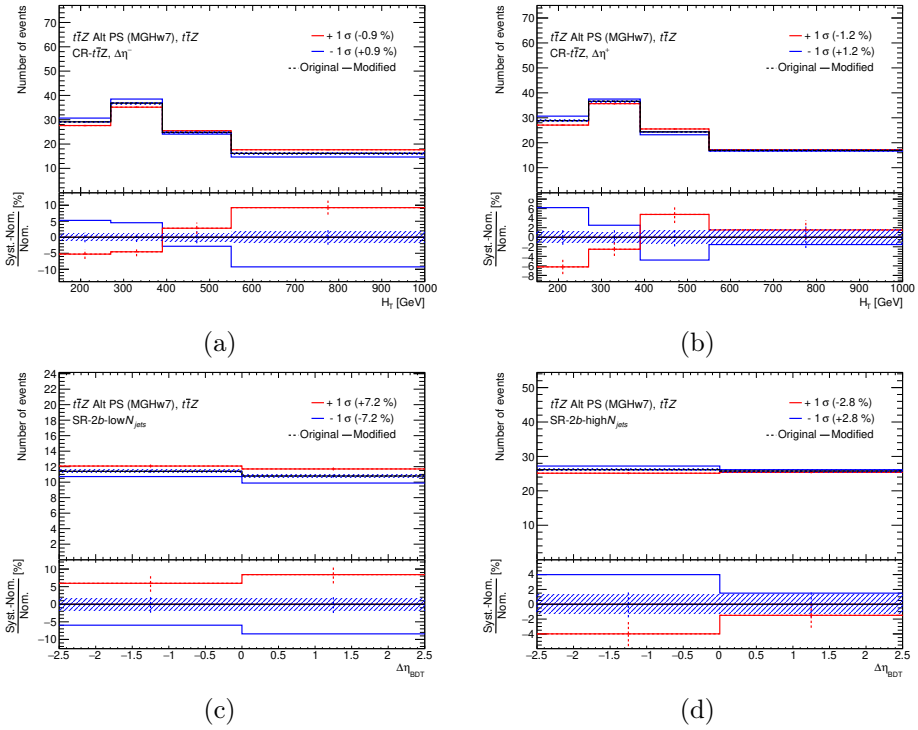


Figure D.1: Alternative $t\bar{t}Z$ PS modelling for (a) CR–$t\bar{t}Z$ $\Delta\eta^-$, (b) CR–$t\bar{t}Z$ $\Delta\eta^+$, (c) SR–$2b$–lowN_{jets}, and (d) SR–$2b$–highN_{jets}. Uncertainty bands cover MC statistical uncertainties of the nominal sample.

D.2 $t\bar{t}$ modelling

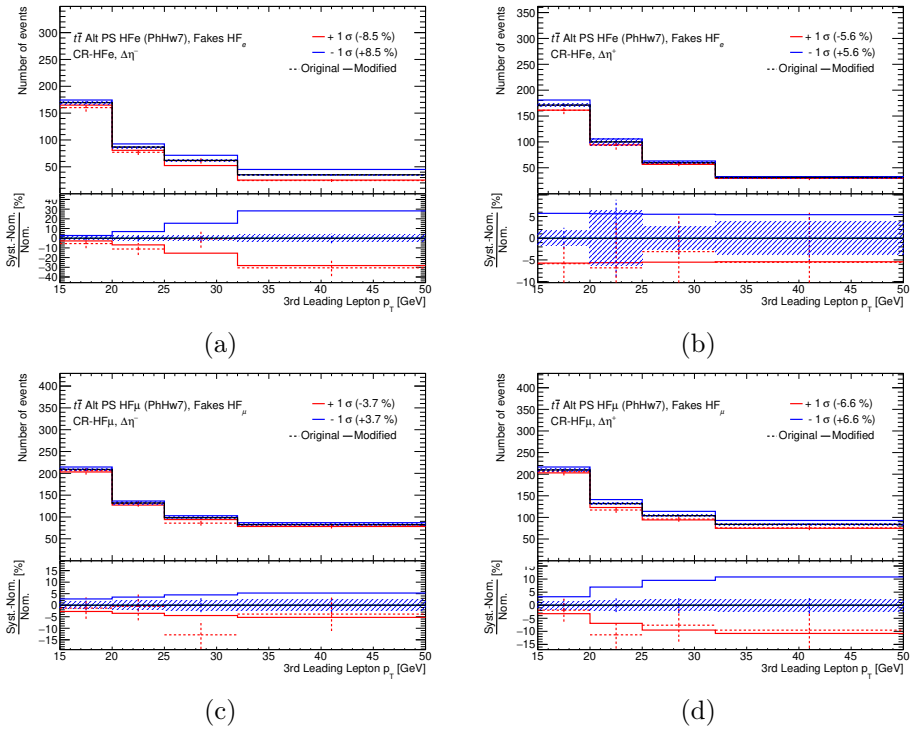


Figure D.2: Alternative $t\bar{t}$ PS modelling for (a) CR–HF$_e$ $\Delta\eta^-$, (b) CR–HF$_e$ $\Delta\eta^+$, (c) CR–HF$_\mu$, and (d) CR–HF$_\mu$. Uncertainty bands cover MC statistical uncertainties of the nominal sample.

D.3 $t\bar{t}H$ modelling

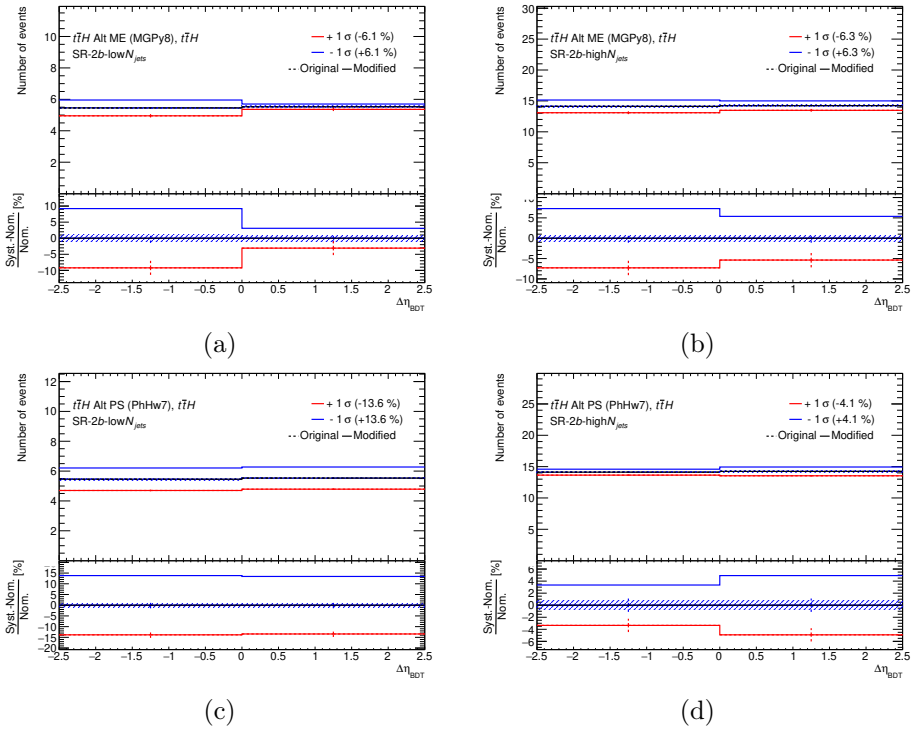


Figure D.3: Alternative $t\bar{t}H$ ME+PS matching for (a) SR–2b–lowN_{jets} and (b) SR–2b–highN_{jets}; and alternative $t\bar{t}H$ PS modelling for (c) SR–2b–lowN_{jets} and (d) SR–2b–highN_{jets}. Uncertainty bands cover MC statistical uncertainties of the nominal sample.

D.4 Z+jets modelling

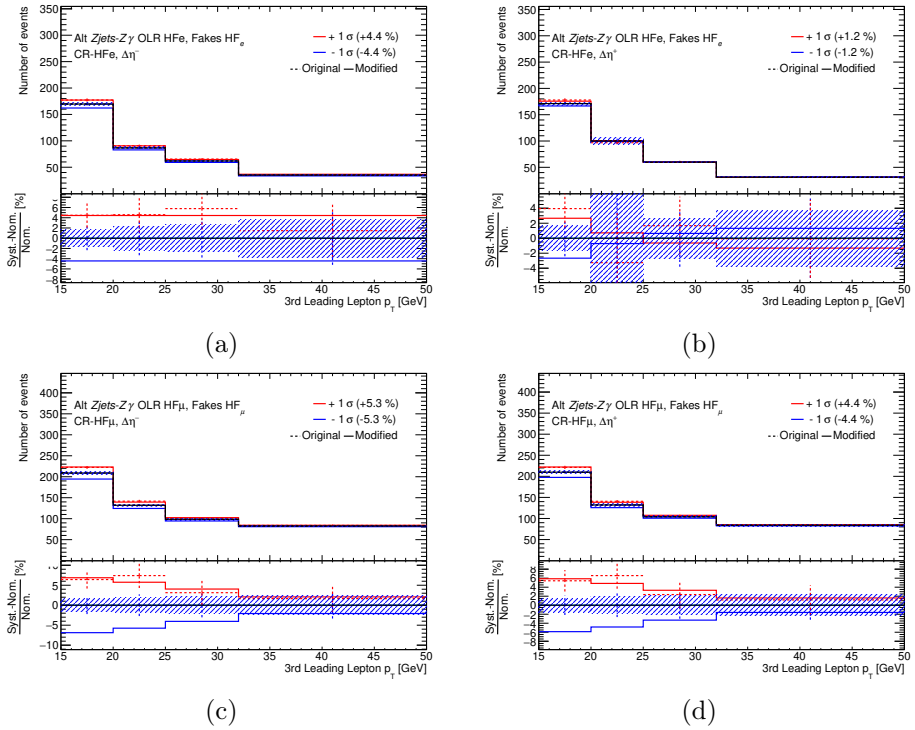


Figure D.4: Alternative Z+jets modelling for (a) CR–HF$_e$ $\Delta\eta^-$, (b) CR–HF$_e$ $\Delta\eta^+$, (c) CR–HF$_\mu$, and (d) CR–HF$_\mu$. Uncertainty bands cover MC statistical uncertainties of the nominal sample.

E Study of the A_c^l at parton, particle and reconstruction level

In this appendix, the A_c^l is studied at parton, particle and reconstruction level with the nominal $t\bar{t}W$ SHERPA sample.

E.1 Dependence of the A_c^l with the SRs and CR–γ–conv phase–space requirements and the lepton–top quark matching at particle and reconstruction level

In this section, the A_c^l is given for each SR and the CR–γ–conv of the analysis at both at particle and reconstruction level for the $t\bar{t}W$ SHERPA nominal sample. Table E.1 shows the different values of the A_c^l with different lepton–top quark association methods.

At reconstruction level (top), the first column presents the A_c^l for each region using the BDT for the lepton–top quark matching; these are the asymmetries that can be extracted from the analysis SRs. The second column makes use of the $m_{\ell b}$–matching algorithm instead of the BDT. At particle level (bottom), the first column shows the A_c^l where only fiducial space cuts have been applied. The second column is the denominator of the migration matrices (see Equation 5.5). It shows the A_c^l for events that pass both reconstruction and particle level (fiducial space) cuts. In order to make the separation for each of the regions, the additional cuts on jet and b–jet multiplicity of the SR definitions are applied.

Region	Reco. level$^{(\text{reco})}$ BDT	Reco. level$^{(\text{reco})}$ $m_{\ell b}$
SR–2b–lowN_{jets}	-0.089 ± 0.009	-0.073 ± 0.009
SR–2b–highN_{jets}	-0.092 ± 0.017	-0.058 ± 0.017
SR–1b–lowN_{jets}	-0.070 ± 0.011	-0.065 ± 0.011
SR–1b–highN_{jets}	-0.059 ± 0.021	-0.040 ± 0.021
CR–γ–conv	-0.163 ± 0.027	-0.134 ± 0.027
Combination	-0.084 ± 0.006	-0.067 ± 0.006

Region	Particle level$^{(\text{fid})}$ $m_{\ell b}$	Particle level$^{(\text{reco} \cap \text{fid})}$ $m_{\ell b}$
SR–2b–lowN_{jets}	-0.071 ± 0.004	-0.066 ± 0.009
SR–2b–highN_{jets}	-0.038 ± 0.008	-0.043 ± 0.018
SR–1b–lowN_{jets}	-0.105 ± 0.010	-0.073 ± 0.011
SR–1b–highN_{jets}	-0.028 ± 0.020	-0.047 ± 0.022
CR–γ–conv	–	-0.118 ± 0.027
Combination	-0.063 ± 0.004	-0.063 ± 0.006

Table E.1: A_c^l at reconstruction (top) and particle (bottom) level using both the BDT and the $m_{\ell b}$–matching algorithm (where applicable) for the lepton–top quark association. Uncertainties include MC statistical errors only.

There is a clear difference on the A_c^l value depending on which association lepton–top quark technique is used. In general, the more accurate BDT method gives larger (in absolute value) asymmetries, while those using the $m_{\ell b}$–matching are diluted in most of the regions due to the higher misidentification rates. For the combined result, this argument holds to the tee. All values using the $m_{\ell b}$–matching show a good agreement within 1σ of the MC statistical uncertainties with the exception of the SR–1b–lowN_{jets} where the agreement happens at the 2σ level.

Table E.2 shows the bottom columns of the previous table but with a different matching algorithm. In this case, each lepton has been matched to parton level objects. The matching has been done by looking at which of the particle level leptons is closest to each of the parton level lepton in ΔR. The closest particle level lepton for each of the two parton level leptons, with the same charge and flavour, is then set to be the one coming from the top or antitop quark, accordingly.

Region	Particle level$^{(\text{fid})}$ matching	Particle level$^{(\text{reco} \cap \text{fid})}$ matching
SR–2b–lowN_{jets}	-0.096 ± 0.005	-0.089 ± 0.009
SR–2b–highN_{jets}	-0.069 ± 0.009	-0.087 ± 0.020
SR–1b–lowN_{jets}	-0.100 ± 0.011	-0.080 ± 0.012
SR–1b–highN_{jets}	-0.082 ± 0.023	-0.079 ± 0.024
CR–γ–conv	–	-0.160 ± 0.029
Combination	-0.087 ± 0.004	-0.089 ± 0.007

Table E.2: A_c^l at particle level where a lepton matching to parton level objects has been performed. Uncertainties include MC statistical errors only.

All values in all regions agree within 1σ in Table E.2. When comparing the particle level predictions with the $m_{\ell b}$–matching algorithm and the parton level matching, a dilution of the A_c^l when using the $m_{\ell b}$–matching is clearly seen. From both tables, other conclusions can be extracted: the reconstruction level BDT asymmetries (Table E.1) are similar to the particle level$^{(\text{reco} \cap \text{fid})}$ matched ones (Table E.2). They are in fact compatible within 1σ for all regions. Additionally, the A_c^l ratio particle level$^{(\text{fid})}$: particle level$^{(\text{reco} \cap \text{fid})}$ in both tables is kept more or less constant for all SRs except the last one where the uncertainties are nevertheless quite large.

E.2 Value of the inclusive A_c^l using parton, particle and reconstruction level objects

The inclusive parton level A_c^l for the nominal $t\bar{t}W$ Sherpa sample is $A_c^l = -0.118 \pm 0.001$ (MC stat.). It is reduced when applying the fiducial particle level cuts detailed in Section 5.9.2 to $A_c^l = -0.087 \pm 0.004$ (MC stat.). Here, a match-

ing of particle level to parton level objects is performed – see Table E.2. The main dilution of the asymmetry between the full and the fiducial phase–space comes from the lepton p_{T} requirements. Moreover, when the lepton–top quark association is performed instead with the $m_{\ell b}$–matching, which has a 65% efficiency, the particle level A_c^l value is further reduced to $A_c^l = -0.063 \pm 0.004$ (MC stat.). This is the value to which the unfolding is performed. A similar value is obtained at reconstruction level when also using the $m_{\ell b}$–matching (see Table E.1). If the BDT is used, however, the reconstructed A_c^l value goes up to $A_c^l = -0.084 \pm 0.006$ (MC stat.) given the higher lepton–top quark matching efficiency.

Table E.3 presents the A_c^l at particle level using an inclusive phase–space for the different samples used for $t\bar{t}W$ modelling (top), as well as for the main background processes (bottom) in the analysis.

MC Generator	Particle level$^{(\mathrm{inc})}$
Sherpa (multileg)	-0.116 ± 0.001
aMC@NLO + Pythia FxFx (multileg)	-0.130 ± 0.001
Powheg + Pythia	-0.130 ± 0.001
Powheg + Herwig	-0.132 ± 0.001
$t\bar{t}Z$ aMC@NLO + Pythia	-0.014 ± 0.002
$t\bar{t}$ Powheg + Pythia	0.001 ± 0.001
$t\bar{t}H$ Powheg + Pythia	0.007 ± 0.001
tZq aMC@NLO + Pythia	0.037 ± 0.009

Table E.3: A_c^l at inclusive particle level for each MC generator and processes. Uncertainties include MC statistical errors only.

F Additional studies and performance of the BDT–based lepton–top quark matching

This appendix shows the bidimensional distributions of the each of the input variables ($m_{\ell b_0}$, $m_{\ell b_1}$, lepton p_{T}, $\Delta R_{\ell b_0}$ and $\Delta R_{\ell b_1}$) with the BDT discriminant score for the signal even lepton (Figure F.1), the background even lepton (Figure F.2), and the odd lepton (Figure F.3). The distributions show events from the nominal $t\bar{t}W$ SHERPA sample across all SRs, as used in the training.

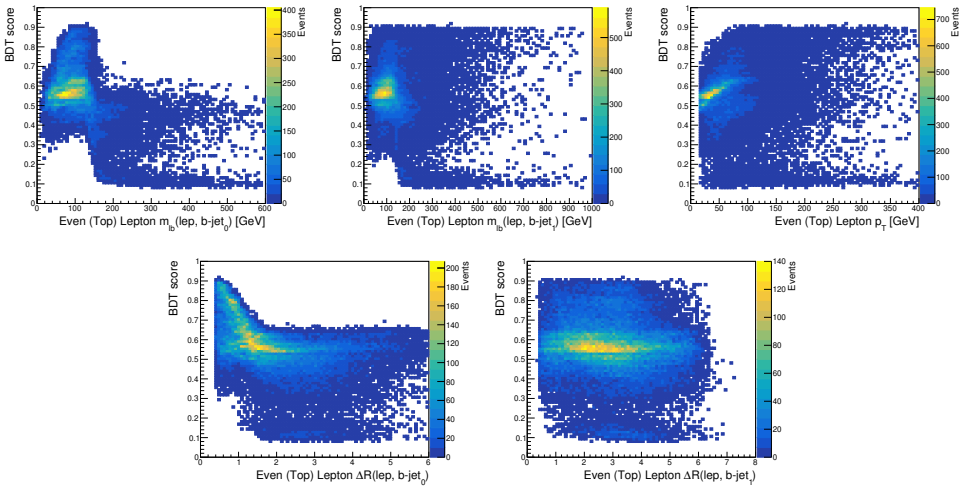


Figure F.1: From left to right, and top to bottom: $m_{\ell b_0}$, $m_{\ell b_1}$, lepton p_{T}, $\Delta R_{\ell b_0}$ and $\Delta R_{\ell b_1}$ distributions for the signal even lepton against the BDT discriminant score.

For the signal even leptons, a unique peak in the BDT score can be seen at around 0.6 (or higher), indicating that the BDT correctly identifies this lepton for most of the events. For the background even lepton, two peaks can be seen at around 0.1 and 0.6 for most of the input variables. This indicates that the BDT is not able to correctly identify the background lepton always, and it will classify it as the signal lepton some of the times. These are events with low lepton p_{T} and $m_{\ell b_0}$ around 70–90 GeV. This reduces the efficiency of the lepton–top quark matching.

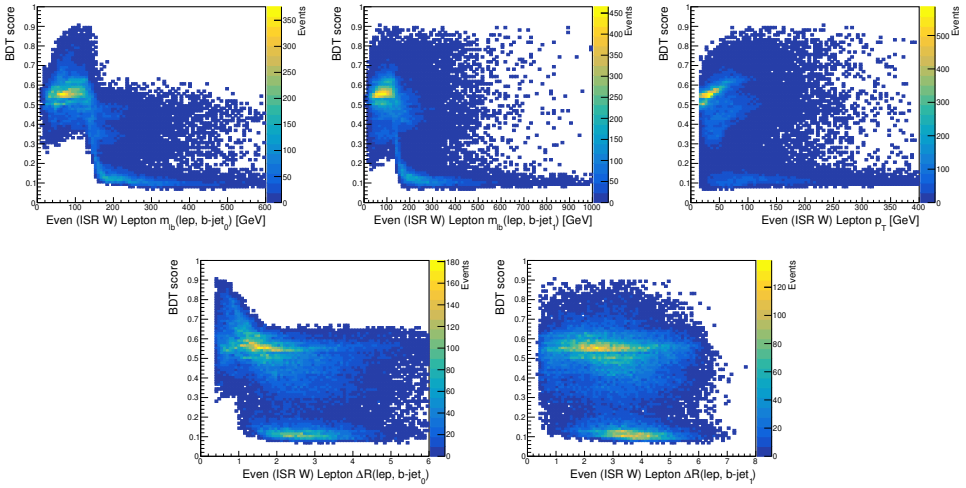


Figure F.2: From left to right, and top to bottom: $m_{\ell b_0}$, $m_{\ell b_1}$, lepton p_{T}, $\Delta R_{\ell b_0}$ and $\Delta R_{\ell b_1}$ distributions for the background even lepton against the BDT discriminant score.

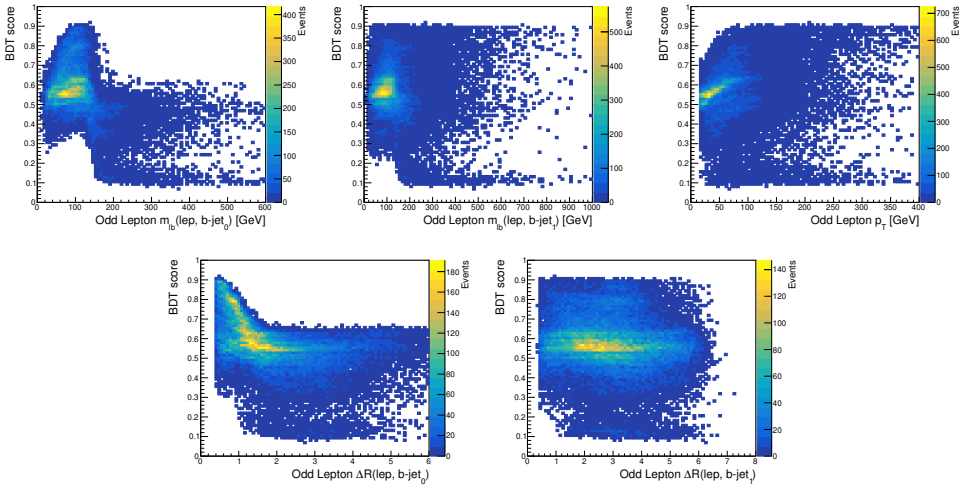


Figure F.3: From left to right, and top to bottom: $m_{\ell b_0}$, $m_{\ell b_1}$, lepton p_{T}, $\Delta R_{\ell b_0}$ and $\Delta R_{\ell b_1}$ distributions for the odd lepton against the BDT discriminant score.

G Inputs for the unfolded A_c^l extraction

This appendix shows the migration matrices, acceptance and efficiency factors in the SR–2b–highN_{jets} (Figure G.1), SR–1b–lowN_{jets} (Figure G.2), SR–1b–highN_{jets} (Figure G.3), and CR–γ–conv (Figure G.4) used for the extraction of the unfolded A_c^l in Section 5.9.

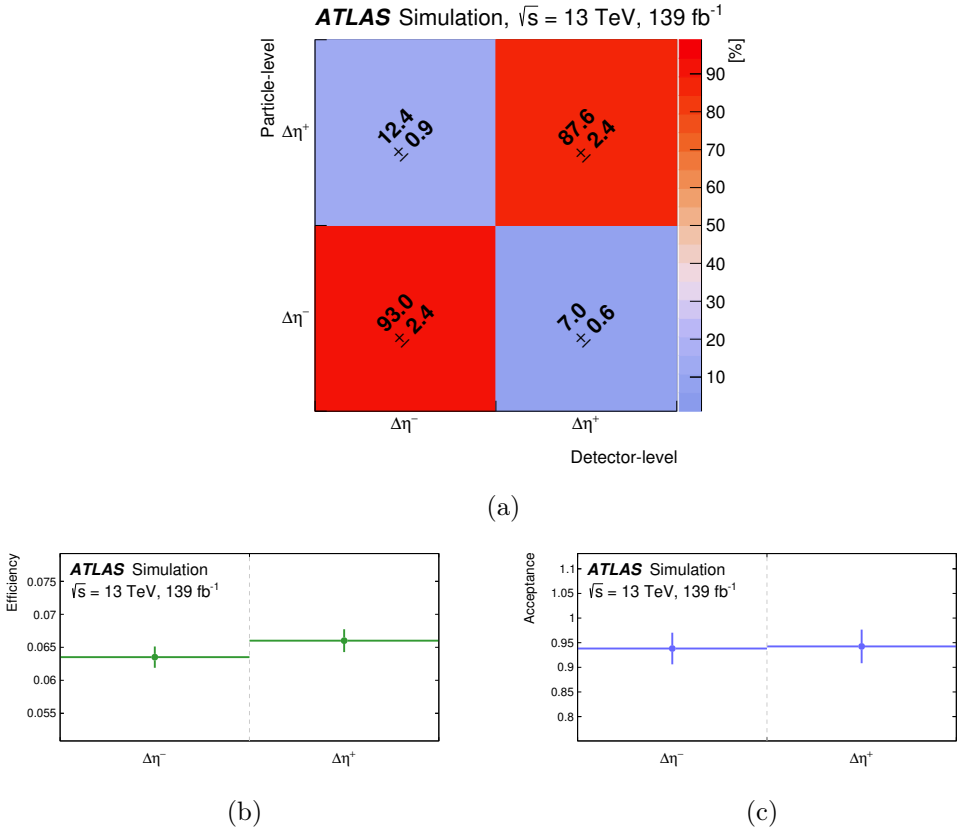


Figure G.1: (a) The migration matrix and (b) the efficiency and (c) the acceptance corrections that are used as input for the unfolding of SR–2b–highN_{jets}. The matrices are normalised such that the sum of any given row is 100%, although small differences may be present due to rounding. The error bars of the efficiency and acceptance correction terms represent the MC statistical uncertainties per bin based on the nominal $t\bar{t}W$ Sherpa sample.

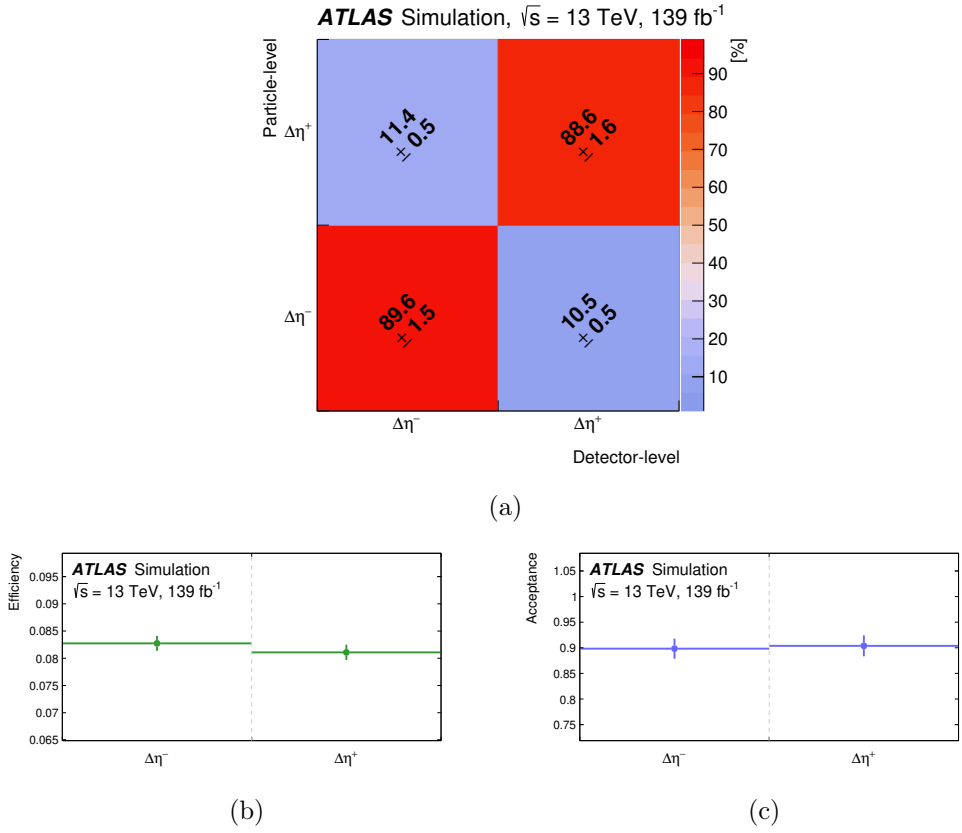


Figure G.2: (a) The migration matrix and (b) the efficiency and (c) the acceptance corrections that are used as input for the unfolding of SR–1b–lowN_{jets}. The matrices are normalised such that the sum of any given row is 100%, although small differences may be present due to rounding. The error bars of the efficiency and acceptance correction terms represent the MC statistical uncertainties per bin based on the nominal $t\bar{t}W$ SHERPA sample.

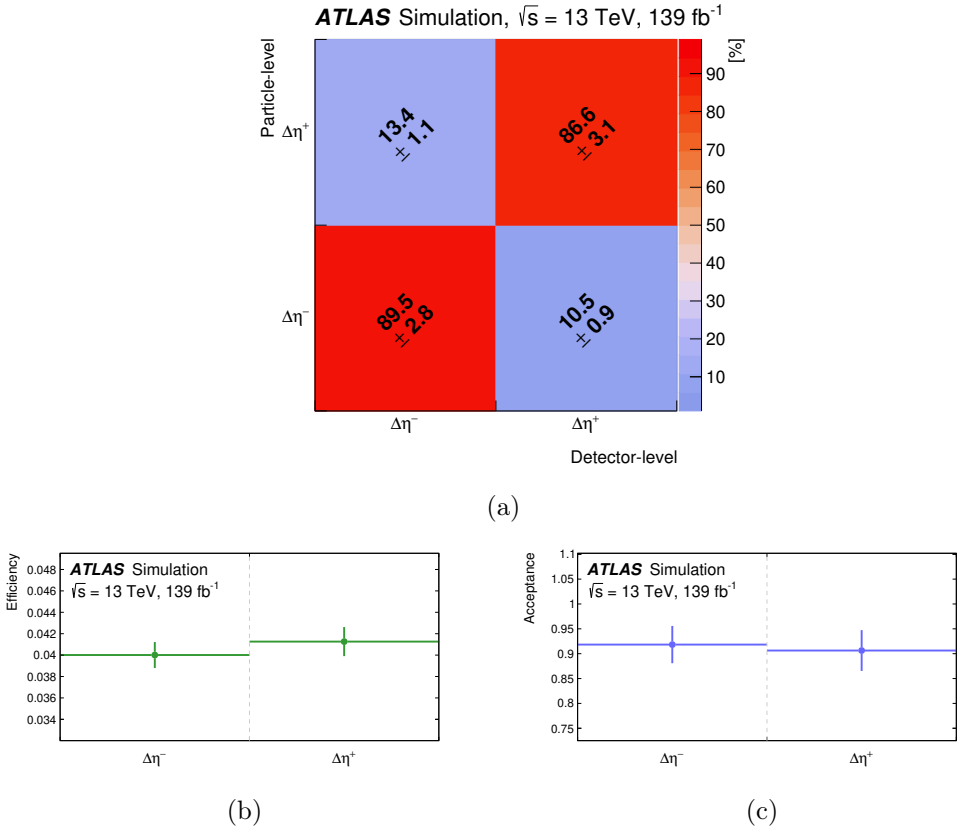


Figure G.3: (a) The migration matrix and (b) the efficiency and (c) the acceptance corrections that are used as input for the unfolding of SR–1b–highN_{jets}. The matrices are normalised such that the sum of any given row is 100%, although small differences may be present due to rounding. The error bars of the efficiency and acceptance correction terms represent the MC statistical uncertainties per bin based on the nominal $t\bar{t}W$ SHERPA sample.

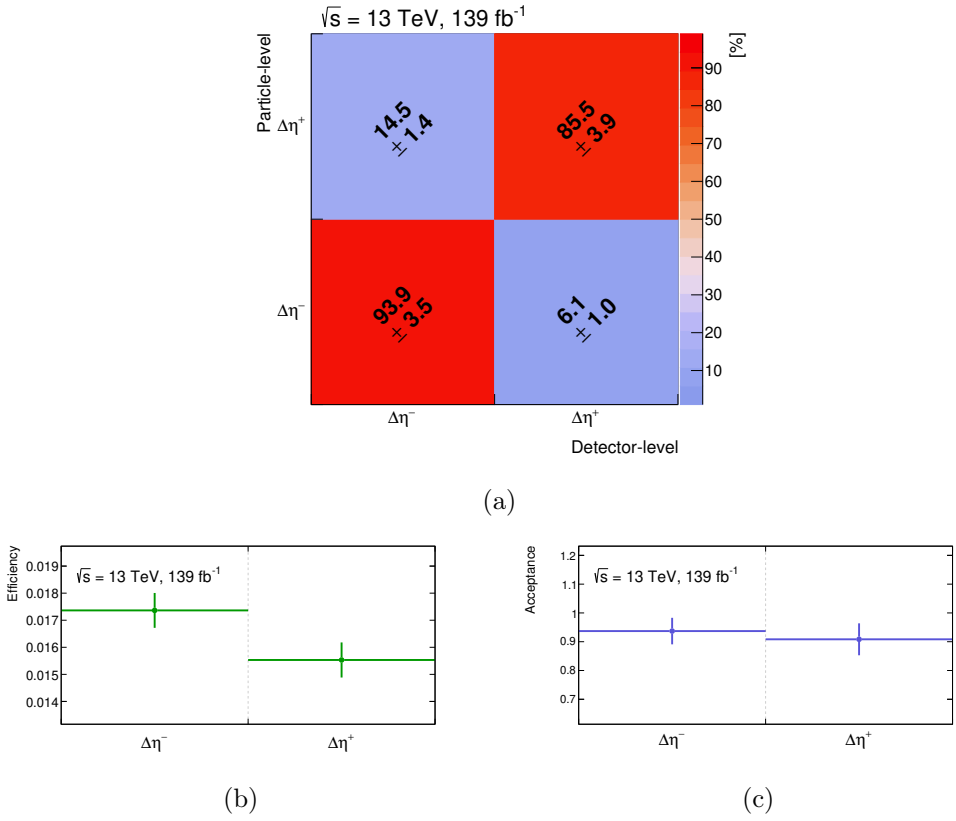


Figure G.4: (a) The migration matrix and (b) the efficiency and (c) the acceptance corrections that are used as input for the unfolding of CR–γ–conv. The matrices are normalised such that the sum of any given row is 100%, although small differences may be present due to rounding. The error bars of the efficiency and acceptance correction terms represent the MC statistical uncertainties per bin based on the nominal $t\bar{t}W$ SHERPA sample.

H Additional tests for the unfolded A_c^l extraction

In this appendix, additional tests are performed to stress and validate the unfolding procedure.

H.1 Reconstruction level injection test

An injection test is done at reconstruction level using the same pseudo–data points obtained after the reweighting described in Section 5.9.4. In this test the pseudo–data points are fitted in a detector level fit like the one presented in Section 5.8. The different input asymmetries at particle level result in different asymmetries at detector level. Table H.1 shows the injected values at particle level, and the resulting asymmetries at reconstruction level obtained via the reweighting. In Figure H.1, the reconstruction level injection test with only MC statistical uncertainties is shown, where the full raw MC statistics of the samples are used. An excellent agreement between the input and output reconstruction level asymmetries is seen.

Particle level A_c^l	Reconstruction level A_c^l
−0.20	−0.18
−0.12	−0.13
−0.08	−0.10
−0.06 (nominal)	−0.08
−0.04	−0.07
0.00	−0.04
0.08	0.02

Table H.1: Injected values at particle level and the resulting asymmetries at reconstruction level obtained via the reweighting.

H.2 Technical closure injection test

In the nominal unfolding injection test in Section 5.9.4, the response matrix is not changed for each pseudo–data point. Here, the response matrix is changed accordingly for each pseudo–data point which, by construction, should give a perfect agreement between the input and output unfolded result as in can be seen in Figure H.2 including only MC statistical uncertainties. This test is meant as a technical closure test to ensure that the unfolding mechanism works correctly.

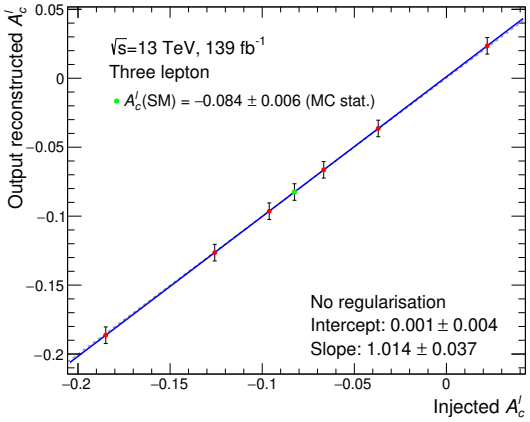


Figure H.1: Reconstruction level linearity test with seven pseudo–data points (red), including the point where the SM A_c^l is also tested (green), computed with the $t\bar{t}W$ SHERPA sample. The error bars include MC statistical uncertainties. The blue line represents the best–fit line to the seven points and its fit parameters are also displayed. The diagonal discontinuous line portrays a perfect agreement between output and injected values of A_c^l.

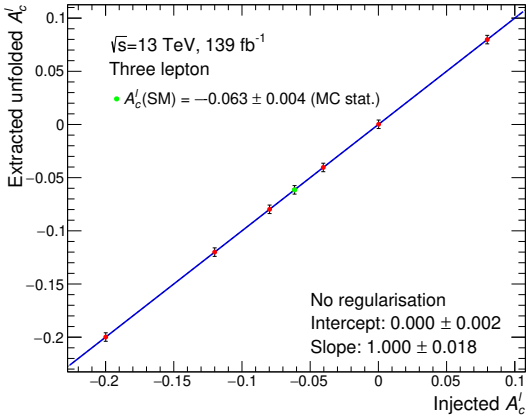


Figure H.2: Results from the technical closure injection test with seven pseudo–data points (red), including the point where the SM A_c^l is also tested (green), computed with the $t\bar{t}W$ SHERPA sample. The error bars include MC statistical uncertainties. The blue line represents the best–fit line to the seven points and its fit parameters are also displayed.

Resum

El treball d'aquesta tesi se centra en les propietats de producció de la partícula elemental (aquelles indivisibles) més massiva del Model Estàndard (SM)*: el quark *top*. En particular, s'examinen les interaccions electrofebles (EW) del quark top amb altres partícules del SM utilitzant les dades de col·lisions protó–protó (*pp*) principalment lliurades pel Gran Col·lisionador d'Hadrons (LHC) en el laboratori del CERN i enregistrades pel detector ATLAS entre 2015 i 2018.

El Model Estàndard de la física de partícules és un model matemàtic que descriu la natura a escala atòmic i subatòmic. És un model altament predictiu, capaç d'explicar la major part dels fenòmens observats, i que ens ha guiat al llarg dels últims 100 anys en la recerca de les partícules que el constitueixen. Tot i això, continuen existint múltiples interrogants que fan del SM una teoria incompleta. Per exemple, no és capaç d'incorporar la força gravitatòria, desentranyar l'origen de la matèria fosca, o explicar l'asimetria de matèria i antimatèria present a l'inici de l'univers. Per això, l'existència de nova física, més enllà del SM (BSM), és una de les principals línies d'investigació en la física moderna.

La recerca d'aquesta nova física és una de les principals motivacions de l'LHC, l'accelerador de partícules més potent del món. L'LHC està dissenyat per a col·lidir partícules a energies relativistes a un ritme de 40 milions de col·lisions per segon, el que permet estudiar la interacció, fins i tot de les partícules més pesants del SM, amb un gran nivell de detall. Un dels assoliments més importants de l'LHC va ser el descobriment del bosó de Higgs l'any 2012: l'última peça del trencaclosques del SM. D'altra banda, les mesures de precisió en el LHC també ens permeten posar a prova el SM i restringir les modificacions dels acoblaments del SM introduïdes per models de nova física.

Tant en la recerca de nova física com en les mesures de precisió, el quark top és una de les partícules més rellevants. En aquesta tesi, s'estudien les interaccions electrofebles del quark top, en ambdós àmbits, amb un enfocament especial en nou observables que puguen proporcionar informació complementària per acotar aquests acoblaments BSM del quark top.

De la mateixa manera que es fan molts esforços perquè les dades recollides siguen de màxima qualitat, les simulacions Monte Carlo (MC) també són una part fonamental per aconseguir resultats experimentals d'alt nivell. El desenvolupament de noves tècniques, amb major precisió, i que ens ajuden a reduir les incerteses teòriques de les nostres mesures, és crític en els pròxims anàlisis. En aquesta tesi s'estudien aquestes noves eines, i s'apliquen a la simulació d'algunes de les mostres que s'utilitzen.

*Al llarg d'aquest resum, els acrònims usen les sigles derivades dels seus noms en anglés.

R.1 Marc teòric

El SM és una teoria quàntica de camps, la qual cosa significa que els seus objectes fonamentals, és a dir, les partícules elementals, són descrites per les vibracions localitzades dels seus camps quàntics subjacents. Tracta tant els camps de matèria com els de força amb el mateix formalisme, ja que les interaccions es consideren mitjançades per les mateixes partícules. La dinàmica i la cinemàtica d'aquests camps es descriuen per la densitat Lagrangiana $\mathscr{L}_0$. A més, el SM és una teoria *gauge*, el que significa que la Lagrangiana és invariant sota transformacions locals. El grup de transformacions gauge sota el qual una Lagrangiana és invariant es diu grup de simetria, i el SM està definit pel grup de simetria $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$. $SU(3)_C$ és el grup del color, i està associat a la interacció forta descrita per la cromodinàmica quantica (QCD) [8, 44, 45]. D'altra banda, el grup $SU(2)_L \otimes U(1)_Y$ està associat a la interacció electrofeble [5–7, 23], que unifica les forces electromagnètiques i febles descrites per l'electrodinàmica quantica (QED).

Aquestes forçes son descrites pel que fa a l'intercanvi de camps gauge de spin–1: huit gluons sense massa per a la interacció forta, un fotó sense massa per a la interacció electromagnètica i tres bosons massius ($W^{\pm}$ i Z) per a la interacció feble. D'altra banda, el contingut de matèria en el SM es dóna per dues grups de fermions de spin–1/2: leptons (leptons carregats ℓ o neutrins ν) i quarks (tipus amunt q_u o avall q_d). Aquests fermions estan organitzats en tres generacions de *sabor* (columnes) i tipus (files), tal com es veu a la Figura R.3. Els camps de fermions tenen una propietat addicional, anomenada quiralitat, que els fa ser diferents en el seu comportament sota les transformacions de simetria. Això es tradueix en que poden ser a levogirs o a dextrogirs. En el SM, els camps levogirs es transformen com a doblets del grup $SU(2)_L$ i els camps dextrogirs com a singlets d'aquest. El contingut de fermions en cada una de les generacions es pot representar com a

$$\begin{bmatrix} \nu_\ell & q_u \\ \ell^- & q_d \end{bmatrix} \equiv \begin{pmatrix} \nu_\ell \\ \ell^- \end{pmatrix}_L, \quad \begin{pmatrix} q_u \\ q_d \end{pmatrix}_L, \quad \ell^-_R, \quad q_{uR}, \quad q_{dR}, \tag{R.1}$$

més les seues corresponents antipartícules[†]. Com que els neutrins no tenen massa en el SM, no tenen cap camp dextrogir necessari per a construir un terme de massa en la Lagrangiana.

Les interaccions febles dels fermions amb els bosons W i Z només són possibles amb camps levogirs, deixant els altres fora d'aquest tipus d'interaccions. Aquest fet es va comprovar en les mesures experimentals de la desintegració β dels nuclis com $n \to p\, e^-_L\, \overline{\nu}_{e,R}$ [18]. A més, aquest tipus d'interacció feble viola,

[†]Els camps d'antipartícules són el resultat d'aplicar la transformació de conjugació de càrrega, $\mathcal{C}$, sobre els camps de partícules associats. Tenen la mateixa massa, però càrrega física oposada a la seua partícula associada.

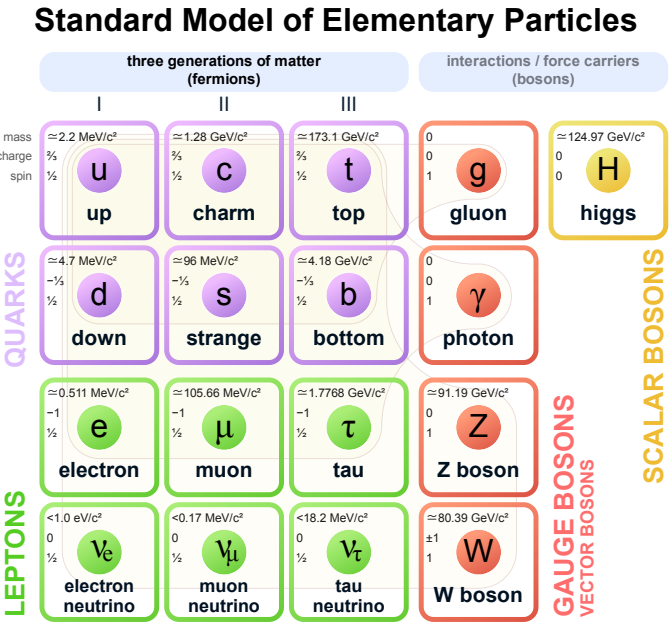


Figure R.2: Partícules fonamentals descrites pel Model Estàndard, incloent informació sobre les seues propietats. Font: Ref. [14].

de manera mínima, les simetries P (levogir ↔ dextrogir) i C (partícula ↔ antipartícula) individualment, però no conjuntament. És a dir, la transformació combinada CP segueix sent una simetria del sistema. D'altra banda, només els leptons més lleugers (els de la primera generació) són prou estables per a ser observats en la natura. Els més pesants eventualment es desintegraran en els més lleugers. Pel que fa els quarks, aquests tenen càrrega de color i no s'observen com a partícules lliures – sent aquest un dels principals resultats de la teoria QCD – sinó que romanen confinats i agrupats en partícules anomenades hadrons. Existeixen dos tipus d'hadrons: mesons (formats per un quark i un antiquark) i barions (formats per tres quarks).

Fins ara, el SM que s'ha descrit conté cauge sense massa per als bosons gauge i els fermions. No obstant això, les partícules que observem tenen massa, la qual cosa fa que aquest model s'allunye de la realitat. Per a donar massa a les partícules, i mantenir les simetries del SM, s'utilitza un mecanisme anomenat trencament espontani de simetria (SSB). Un dels ingredients necessaris és la introducció d'un nou camp escalar, que al seu torn predirà l'existència d'una nova partícula: el bosó de Higgs [9, 10]. De fet, el bosó de Higgs va ser predit en 1964 i no va ser fins al 2012 quan va ser descobert experimentalment [12, 13]. Per tant, les interaccions de les partícules del SM amb el bosó de Higgs són

les que donen lloc a la massa de les partícules elementals, i la intensitat de l'acoblament de les mateixes amb el bosó de Higgs depén de la massa de la partícula.

R.2 El quark top

El quark top és la partícula elemental més massiva del SM. És un quark de spin 1/2, tipus amunt (càrrega $+2/3e$), i en la tercera generació de partícules elementals. La seua descoberta va ser anunciada en 1995 per les col·laboracions CDF i D0 al col·lisionador Tevatron [30, 31]. La combinació de la seua alta massa, breu vida mitjana i la necessitat d'acceleradors de partícules d'alta energia per a la seua producció van contribuir al fet que el quark top fóra l'últim a ser descobert. Aquest esdeveniment no només va validar encara més el marc teòric del SM, sinó que també va obrir una via totalment nova per a les mesures i anàlisis relacionades amb el quark top, que són ara una peça fonamental de la física de partícules.

La seua alta massa li confereix propietats úniques. En primer lloc, és l'únic quark que es desintegra abans d'hadronitzar – donat que la seua vida mitjana, 5×10^{-25} s, és inferior a l'escala d'hadronització $\mathcal{O}(10^{-24}$ s) – el que permet el seu estudi a través dels seus productes de desintegració. Aquesta desintegració és, a més, molt peculiar ja que el quark top es descompon gairebé exclusivament en un quark *bottom* (b) i un bosó W proporcionant senyals molt característics als detectors.

Els bosons W, d'altra banda, també es desintegren en un leptó carregat amb el seu corresponent neutrí o en un parell quark–antiquark. De manera respectiva, a aquests canals de desintegració del quark top se'ls anomena leptònic o hadrònic. En segon lloc, la seua elevada massa es reflecteix en la interacció amb el bosó de Higgs, la més forta en el SM, i també fa que siga particularment sensible a efectes introduïts per nova física. Un exemple que cobrix ambdós aspectes és la seua sensibilitat a la violació de la simetria $\mathcal{CP}$, que pot ser estudiada en gran detall examinant tant el seu acoblament amb el bosó de Higgs com el vèrtex tWb.

El quark top és una de les partícules més produïdes en l'LHC. La seua producció principal és juntament amb un altre quark antitop ($t\bar{t}$), mitjançant la interacció forta; encara que també pot ser produït en solitari associat amb altres partícules, mitjançant la interacció electrofeble. Malgrat que la secció eficaç (la probabilitat d'interacció per unitat de superfície) d'aquest segon mode és unes tres vegades inferior al primer, aquest permet estudiar l'estructura del vèrtex tWb tant en la producció com en la desintegració del quark top. A més a més, en la producció en solitari el quark top està polaritzat.

R.3 Teoria de camps efectiva

Com s'ha dit abans, el SM no pot ser una teoria completa, ja que hi ha diversos fenòmens a escales d'energies superiors (i inferiors) que no és capaç d'explicar. Per a afrontar aquest problema, una de les extensions més populars del SM és la teoria de camps efectiva (EFT). Aquestes teories són àmpliament utilitzades per parametritzar els efectes de nova física en els observables mesurats pels experiments, i així poder interpretar possibles desviacions del SM. Amb aquestes tècniques, es poden aprendre propietats sobre la nova física d'altes energies, sense conéixer en detall la seua estructura subjacent.

En aquesta tesi es treballa amb el SMEFT [64–68], una EFT construïda amb els camps del SM. En el SMEFT, se suposa que tota la nova física es troba per damunt de l'escala d'alta energia Λ, on $\Lambda = 10^3$ GeV típicament. També se suposa que la nova física segueix les simetries i l'estructura del SM. Llavors, tots els termes compatibles es construeixen en la Lagrangiana efectiva, fent una expansió a cada ordre en la dimensió de la massa:

$$\mathscr{L}_{\text{SMEFT}} = \mathscr{L}_{\text{SM}} + \sum_{\forall i, \mathcal{D}\geq 5} \frac{C_i O_i^{(\mathcal{D})}}{\Lambda^{\mathcal{D}-4}}, \tag{R.2}$$

on $O_i^{(\mathcal{D})}$ són els operadors de dimensió–$\mathcal{D}$ construïts amb els camps del SM i C_i són les constants que regulen la intensitat dels acoblaments (també coneguts com a coeficients de Wilson). L'escala de nova física Λ apareix en el paràmetre d'expansió $(1/\Lambda^{\mathcal{D}-4})$ per suprimir els operadors de dimensió superior i per fer que els coeficients de Wilson siguen adimensionals.

Els operadors més rellevants són aquells de dimensió sis. Per establir límits experimentals sobre ells, habitualment es fa un ajust global en l'espai de paràmetres dels coeficients de Wilson. Desviacions dels coeficients de zero en l'ajust són indicacions de nova física en el corresponent vèrtex d'interacció efectiu. Per avaluar l'efecte de cada operador en un observable, X, les prediccions del SMEFT es poden escriure com una expansió en ordres de $1/\Lambda^2$:

$$X = X_{\text{SM}} + \frac{1}{\Lambda^2}\sum_i C_i X_i^{(1)} + \frac{1}{\Lambda^4}\sum_{ij} C_i C_j X_{ij}^{(2)} + \mathcal{O}(\Lambda^{-4}). \tag{R.3}$$

Els termes proporcionals a C_i/Λ^2 (lineals) inclouen interferències entre el SM i els operadors efectius de dimensió sis, mentre que aquells proporcionals a C_iC_j/Λ^4 (quadràtics) venen de l'amplitud quadrada entre els operadors de dimensió sis. L'efecte dels operadors de dimensió vuit, així com el de diagrames amb dos vèrtexs efectius, que serien proporcionals a Λ^{-4}, no es consideren en els estudis d'aquesta tesi.

R.4 CERN, l'LHC i el detector ATLAS

Amb l'objectiu d'examinar el SM i altres teories BSM amb gran detall, és crucial desenvolupar i construir màquines que puguen estudiar la física més rellevant i fer-ho amb la màxima precisió possible. En aquest sentit, l'LHC [73, 74], situat en el laboratori del CERN, és l'accelerador de partícules més gran i poderós construït fins ara. El potencial energètic que s'ha aconseguit en les seues col·lisions permet la producció i l'estudi exhaustiu de les partícules més massives del SM, i podria desvetllar evidències sobre partícules BSM més pesants. El quark top i altres partícules massives, com ara els bosons, es produeixen a l'LHC a una taxa formidable, convertint-lo en l'escenari ideal per estudiar-les amb una precisió sense precedents.

R.4.1 El Gran Col·lisionador d'Hadrons

L'LHC és un anell de 27 km de diàmetre situat a uns 100 metres sota terra a la frontera Franco-Suïssa, el que li confereix una protecció natural davant la radiació de fons d'altres fonts. Al llarg de l'anell, hi ha quatre punts d'interacció on es situen els detectors de partícules. En extrems oposats es troben els detectors ATLAS [72] i CMS [78], detectors de propòsit general, dissenyats per cobrir una àmplia gamma de cerques BSM i mesures de precisió del SM. En posseir característiques similars, els resultats que s'obtenen dels dos experiments poden ser contrastats i validats. Els altres dues detectors principals són ALICE [80], que es centra en la física d'hadrons pesants i l'estudi del plasma quark–glúó, i LHCb [79], que estudia la física del quark b detalladament.

R.4.2 El detector ATLAS

El detector ATLAS és el detector de partícules d'alta energia més gran, mai construït. Està compost per tres subsistemes distribuïts amb simetria cilíndrica al voltant del punt de col·lisió com es pot apreciar en la Figura R.4. Aquest punt marca l'origen del sistema de coordenades (x, y, z), on l'eix z apunta en la direcció del tub del feix i el pla x–y és transvers a ell. Aquest pla s'utilitza per definir algunes de les variables més importants com el moment transversal de les partícules p_{T}, o el moment transversal mancant de les col·lisions $E_{\mathrm{T}}^{\mathrm{miss}}$. També es pot treballar amb coordenades cilíndriques: l'angle azimutal ϕ es defineix al voltant del feix (en el pla x–y), i l'angle polar θ mesura la distància al feix. Tipicament, aquest últim no s'usa, sinó que es transforma en la pseudorapidesa definida com a $\eta = -\ln(\tan(\theta/2))$.

Les diferents capes del detector ATLAS permeten identificar les propietats de diferents tipus de partícules a mesura que viatgen a través d'ell:

- **Detector intern**: Localitzat en la part més interna del detector, el seu

objectiu és la mesura de l'origen i la trajectòria (traces) de les partícules carregades que el travessen. Per a açò, està submergit en un camp magnètic de 2 T creat per un solenoide que corba les trajectòries de les partícules, i permet la determinació de la seua càrrega i moment. Així mateix, el detector intern està format per altres subdetectores. De dins a fora es troben: els detectors de silici com l'IBL, el detector de píxels i l'SCT; i el TRT, que està format per tubs de deriva.

- **Calorímetres**: Després del detector intern, es troben els calorímetres, que estan dissenyats per a frenar les partícules produïdes en la col·lisió i mesurar la seua energia. Hi ha de dos tipus: el calorímetre electromagnètic, que mesura les energies dels electrons i fotons, i el calorímetre hadrònic, que mesura l'energia dels hadrons (*jets*) més pesants.

- **Cambra de muons**: Finalment, es troben les cambres de muons, que tenen com a objectiu mesurar el moment i les trajectòries dels muons, que típicament deixen poc senyal en els altres detectors. Els muons es corben a causa del camp magnètic d'uns 4 T produït pels imans superconductors toroidals, característics del disseny del detector ATLAS.

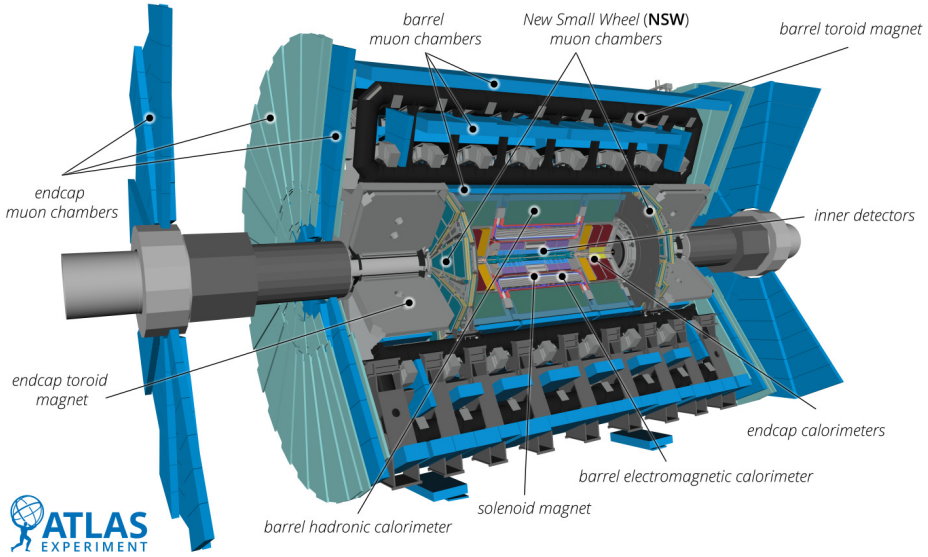


Figura R.4: Vista esquemàtica del detector ATLAS. Font: Ref. [127].

A causa de la gran quantitat de dades que es generen en cada col·lisió, s'utilitzen uns criteris de selecció anomenats *triggers* per reduir la quantitat de dades que s'emmagatzemen. ATLAS fa servir un sistema de dos nivells: el primer (L1), integrat en el mateix detector, i el segon (HLT), que fa servir

algoritmes de reconstrucció més sofisticats. Aquests són capaços de reduir el nombre d'esdeveniments des d'uns 40 milions per segon a només uns 1000 per segon, que es guarden per a la seua posterior anàlisi.

R.4.3 Reconstrucció d'objectes

Després dels triggers, els esdeveniments seleccionats es tornen a examinar i els diferents senyals recollits pels detectors es combinen per reconstruir els objectes físics que s'han produït en la col·lisió. En el cas d'ATLAS, es reconstrueixen els següents objectes:

- **Electrons**: Es reconstrueixen a partir de l'associació de les traçes del detector intern amb els dipòsits d'energia en el calorímetre electromagnètic. Els electrons candidats han de complir un conjunt de requisits que estan optimitzats per identificar electrons procedents de bosons vectorials. Les incerteses en el procés d'identificació, reconstrucció i en els criteris d'aïllament, així com l'eficiència i calibració d'aquests algoritmes, s'avaluen i es propaguen com a incerteses experimentals en els resultats finals de l'anàlisi.

- **Muons**: Es reconstrueixen a partir de les trajectòries en les cambres de muons i el detector intern, usant informació complementària dels calorímetres. De manera similar als electrons, s'apliquen una sèrie de talls als candidats a muons per assegurar la seua qualitat. Així mateix, les incerteses associades a la reconstrucció dels muons també es propaguen a les anàlisis.

- **Jets**: Els jets són col·leccions de partícules col·limades que provenen de l'hadronització dels quarks i gluons emesos en la col·lisió. Es reconstrueixen emprant els dipòsits del calorímetre hadrònic i les traçes del detector intern. Aquesta informació es combina en algoritmes de reconstrucció de jets com l'algoritme anti–k_t. També s'assignen incerteses en la reconstrucció dels jets. En particular, es consideren incerteses en la calibració dels jets i en l'estimació de l'escala i la resolució de l'energia dels jets.

- b**–jets**: La identificació de jets que provenen de quarks b és una part crucial en el procés de reconstrucció. Per a identificar b–jets és necessari fer ús d'algoritmes que analitzen les propietats d'aquestes partícules, com ara el paràmetre d'impacte o la presència de vèrtexs secundaris. En ATLAS, s'utilitzen algoritmes d'aprenentatge automàtic per a identificar els b–jets i també s'assignen incerteses en la reconstrucció i calibració d'aquests. Els b–jets són essencials per a la identificació de possibles candidats del quark top donada la seua característica desintegració.

- **Moment transvers mancant**: En el pla transversal al feix, la suma vectorial dels moments transversals de totes les partícules d'una col·lisió ha de ser zero a causa de la conservació del moment. L'energia de les partícules *invisibles* que s'escapen a la detecció, com és el cas dels neutrins, es veu reflectida en l'$E_{\mathrm{T}}^{\mathrm{miss}}$, que es defineix com la suma vectorial negativa del p_{T} de tots els objectes reconstruïts en l'esdeveniment. L'$E_{\mathrm{T}}^{\mathrm{miss}}$ és una variable important en la identificació dels candidats del quark top que es desintegren leptònicament, ja que es produeixen neutrins a l'estat final.

R.5 Ajust global EFT en el sector EW del quark top

El quark top té un paper molt important en moltes extensions del SM. Les seues propietats el fan particularment adequat per a la recerca experimental i s'ha desenvolupat un ric programa experimental al voltant del quark top des de la seua descoberta en 1995. A més a més, els experiments de l'LHC proporcionen noves mesures amb una precisió augmentada i en un règim cinemàtic ampliat, i han observat processos de producció associada poc freqüents que proporcionen una forma directa d'estudiar els acoblaments del quark top amb el fotó, el bosons Z i W, i el bosó de Higgs, a més del vèrtex tWb.

En aquesta secció, es descriu un ajust global fent servir el SMEFT i una sèrie d'observables de l'LHC, Tevatron, SLC i LEP que són sensibles als operadors efectius de dimensió sis en el sector EW del quark top. Els ajustos es presenten incloent i excloent els termes quadràtics Λ^{-4}, i un total de vuit (sis) límits en els coeficients de Wilson corresponents es donen per al cas d'incloure'ls (excloure'ls). La Figura R.5 resumeix quins dels coeficients de Wilson entren en cadascun dels dos ajustos com a graus de llibertat addicionals. Aquests resultats s'han publicat en la Ref. [246].

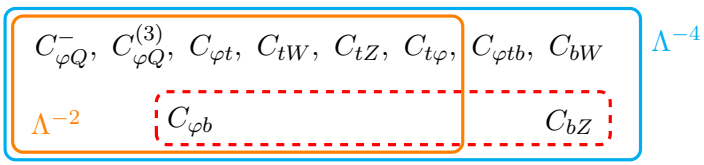


Figura R.5: Llista dels coeficients de Wilson que entren en els ajustos globals EFT. Hi ha un total de set graus de llibertat en l'ajust lineal de Λ^{-2} (taronja) i deu en l'ajust quadràtic $\Lambda^{-2} + \Lambda^{-4}$ (blau). Els operadors en el quadre vermell s'inclouen perquè s'utilitzen dades de precisió EW, però no es reporten com a part del resultat final a causa de la baixa sensibilitat dels observables considerats a ells.

Els operadors $O_{\varphi Q}^{(3)}$ i $O_{\varphi Q}^{-}$ modifiquen els acoblaments levogirs dels bosons Z als quarks top i bottom, mentre que $O_{\varphi t}$ i $O_{\varphi b}$ modifiquen els acoblaments dextrogirs analògics als bosons Z. Els operadors C_{tZ}, C_{bZ} i C_{tW} els quals representen els dipols electrofebles, modifiquen les interacciones entre el fotó

i els bosons W i Z amb els quarks de tipus amunt i avall. Finalment, $C_{\varphi tb}$ i C_{bW} afecten les interaccions dextrogires entre els bosons W i els quarks top i bottom, respectivament; i $C_{t\varphi}$ modifica l'acoblament del quark top amb el bosó de Higgs.

R.5.1 Configuració de l'ajust

S'utilitzen un total de 30 observables sensibles a aquests operadors, entre els quals cal destacar les mesures de seccions eficaces differencials dels processos $t\bar{t}Z$ i $t\bar{t}\gamma$ realitzades per ATLAS, així com els observables de LEP/SLC de la mesura del pol del bosó Z (producció $e^+e^- \to b\bar{b}$), que són extremadament sensibles a $C^{(3)}_{\varphi Q}$, $C^{-}_{\varphi Q}$ i C_{bW}. A més, les correlacions experimentals publicades en aquestes mesures s'inclouen en tots els ajustos que es presenten. Abans de realitzar qualsevol ajust EFT, s'avalua l'acord entre les prediccions del SM i les dades mesurades en cada observable, incloent-hi les correlacions experimentals. El valor recopilat de la chi–quadrada és $\chi^2_{\rm SM}/(n_{\rm obs} - 1) = 21.3/29$, el qual correspon a un *p–value* de 0.85. En general, s'observa una bona concordança entre les prediccions del SM i els resultats experimentals.

La dependència dels observables als coeficients de Wilson – els termes normalitzats $X^{(1)}_i/X_{\rm SM}$ i $X^{(2)}_{ij}/X_{\rm SM}$ de l'Eq. R.3 – es parametritza utilitzant el generador MC aMC@NLO [97,98] v.2.7.0 a segon order (NLO) de QCD per a la majoria dels observables. En concret, s'ha utilitzat el model `SMEFT@NLO` [265] que és l'únic capaç d'aquests càlculs. Altres observables, com els de LEP/SLC o Tevatron, es calculen a primer order (LO) de QCD, i la dependència dels operadors del quark bottom (O_{bW}, $O_{\varphi tb}$, O_{bZ} i $O_{\varphi b}$) es calcula a LO utilitzant el model `TEFT_EW`.

R.5.2 Resultats de l'ajust global EFT a les dades

L'ajust es realitza amb el programa de codi obert `HEPfit` [267,268], que fa servir una anàlisi estadística bayesiana del model, en la qual s'inclouen les incerteses teòriques i experimentals. La Figura R.6 mostra les cotes individuals al 95% d'interval de confiança sobre els vuit coeficients dels operadors considerats en l'ajust amb termes Λ^{-4}. Es presenten els límits obtinguts amb les diferents observables i processos considerats, ordenats de més a menys restrictius (de l'esquerra a la dreta), per a cada coeficient.

Els resultats principals d'aquesta anàlisi són els intervals de 68% i 95% de probabilitat per als coeficients de Wilson que modifiquen els acoblaments electrofebles del quark top. Els intervals amb el 95% de confiança de l'ajust global per als vuit coeficients de Wilson que modifiquen els acoblaments electroweak del quark top van de $\pm 0,35$ a ± 8 TeV^{-2} i inclouen l'expectativa del Model Estàndard ($C_i = 0$). Es presenten dos conjunts de resultats per als ajustos globals

amb termes lineals i quadràtics, com es mostra a la Figura R.7 en taronja i blau respectivament. Cal destacar l'ús de les mesures diferencials dels processos $t\bar{t}Z$ i $t\bar{t}\gamma$, que proporcionen millores significatives en les restriccions dels coeficients de Wilson (sobretot per a C_{tZ}), mostrant la necessitat de noves mesures d'alta precisió provinents d'experiments d'alta energia.

D'altra banda, es realitzen ajustos addicionals per avaluar la robustesa dels resultats. Les línies puntejades de color marró clar representen els límits d'un ajust on s'ha ampliat la base d'operadors per a incloure C_{tG} i un set addicional d'operadors amb quatre fermions. De manera anàloga, la línia puntejada de color marró fosc presenta els resultats que inclouen correlacions entre tots els observables en l'ajust bàsic, que han sigut estimats de manera apropiada. Finalment, les línies vermelles mostren l'envolvent de totes les proves d'estrés que s'han realitzat. Els resultats s'han trobat robustos en aquestes variacions.

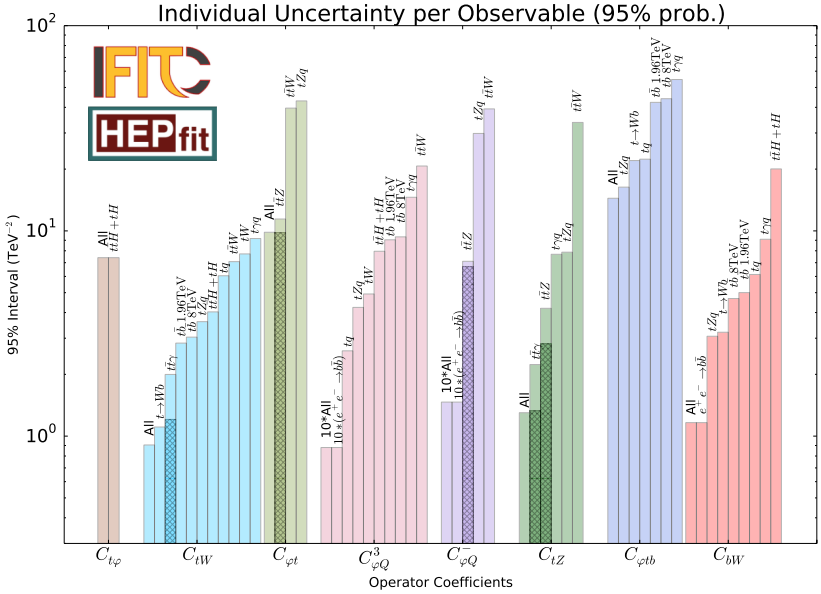


Figura R.6: Cotes individuals sobre els vuit coeficients de Wilson resultant de mesures en diferents processos incloent termes Λ^{-4}. L'ombreig fosc indica el resultat obtingut amb les mesures diferencials de $t\bar{t}\gamma$ i $t\bar{t}Z$ i mostra la millora respecte a la mesura inclusiva (que no s'inclou en l'ajust final).

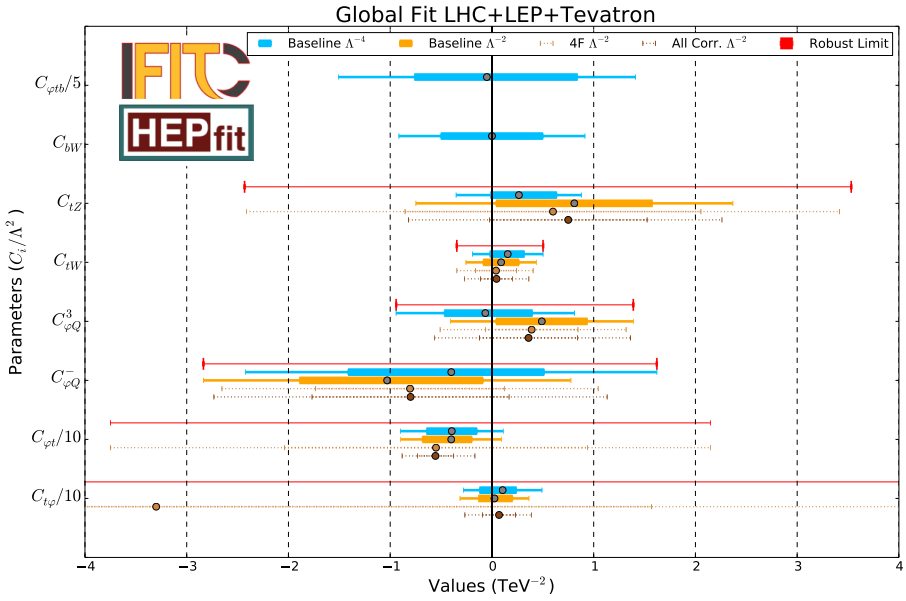


Figura R.7: Resultats dels ajustos globals amb termes $O(\Lambda^{-2})$ (en taronja) i $O(\Lambda^{-4})$ (en blau). Les dues línies primes sota cadascun d'ells corresponen a ajustos addicionals realitzats per provar la robustesa dels resultats: per a tenir en compte els efectes de la inclusió de més operadors, i per a tindre en compte les correlacions entre totes les diferents mesures. Els marcadors vermells corresponen a l'envolvent d'aquests ajustos addicionals més un altre que té en compte les incerteses teòriques en les parametritzacions.

R.5.3 Extrapolació a futurs acceleradors

Prenent el treball presentat en aquesta secció com a base, s'estudia l'impacte del HL–LHC[‡] i diversos escenaris futurs de col·lisionadors electró–positró en els sectors del quark top i bottom del SMEFT. En aquest estudi, diverses mesures diferencials de $t\bar{t}$ en l'LHC, així com les perspectives per a observables de producció $e^+e^- \to b\bar{b}$ i $e^+e^- \to t\bar{t}$ en col·lisionadors lineals i circulars com ara l'ILC, CLIC, FCC–*ee* o CEPC, s'utilitzen per obtindre cotes en els coeficients de Wilson del SMEFT. La base es va ampliar per incloure operadors que acoblen quatre fermions: tant operadors de quatre quarks com operadors de dos quarks i dos leptons. Les dades de col·lisions e^+e^- tenen una alta sensibilitat a aquests últims. Els resultats mostren una millora dels límits d'un factor de dos a quatre quan es fan les extrapolacions per al HL–LHC, i es veu com les perspectives dels futurs col·lisionadors e^+e^- milloren les cotes dels operadors amb dos fermions

[‡]HL–LHC serà la propera actualització de l'LHC. Es preveu que proporcione al voltant de 3000 fb^{-1} de dades de col·lisió – aproximadament 20 vegades més del que tenim ara.

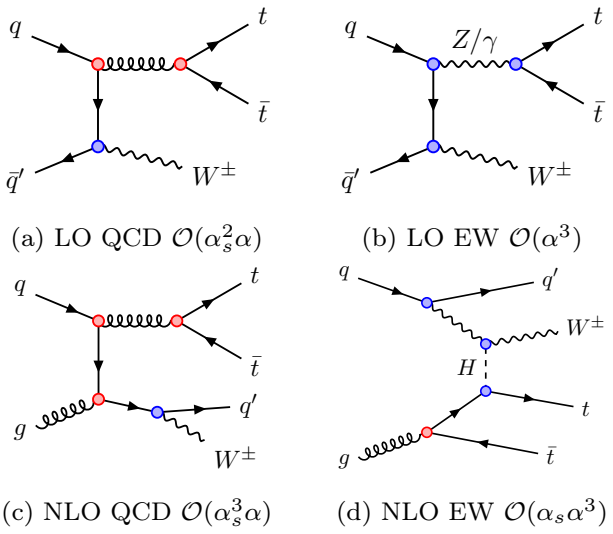
(aquells estudiats en aquesta secció) fins a dos ordres de magnitud. A més, els acceleradors lineals (ILC i CLIC), operant amb energies de centre de massa sobre del llindar de producció de $t\bar{t}$, proporcionen les millores cotes per als operadors $e^+e^-t\bar{t}$ (de l'ordre de 10^{-3} TeV^{-2}).

R.6 Recerca de l'asimetria de càrrega leptònica en producció d'esdeveniments $t\bar{t}W$

R.6.1 Producció $t\bar{t}W$

La producció $t\bar{t}W$ és un dels processos més únics que es poden estudiar a l'LHC. Els càlculs de secció eficaç són especialment complexos, ja que sorgeixen grans correccions a partir de potències superiors tant dels acoblaments forts com dels acoblaments electrofebles. Així, les mesures del procés $t\bar{t}W$ representen una prova sensible de les prediccions del sector QCD i del sector EW del SM, i de la seua mescla. Tant les mesures de secció eficaç inclusiva com la diferencial són molt rellevants, ja que poden proporcionar indicis indirectes de nova física BSM. A més a més, és un fons irreductible d'alguns processos rars del SM, com ara la producció $t\bar{t}H$ i $t\bar{t}t\bar{t}$.

La Figura R.8 mostra diagrames de Feynman il·lustratius que contribueixen a la producció de $t\bar{t}W$ a LO i NLO per a la producció QCD i EW.



(a) LO QCD $\mathcal{O}(\alpha_s^2\alpha)$ (b) LO EW $\mathcal{O}(\alpha^3)$

(c) NLO QCD $\mathcal{O}(\alpha_s^3\alpha)$ (d) NLO EW $\mathcal{O}(\alpha_s\alpha^3)$

Figura R.8: Exemples de diagrames de Feynman de la producció de $t\bar{t}W$ a LO (a,b) i NLO (c,d) amb una partícula addicional. Els diagrames mostren la producció $t\bar{t}W$ en QCD i EW. Els cercles vermells als vèrtexs corresponen als acoblaments de QCD i els cercles blaus corresponen als acoblaments EW.

En particular, les contribucions d'ordre $\mathcal{O}(\alpha_s\alpha^3)$, conegudes com a contribucions EW, augmenten la secció eficaç en aproximadament un 10% a causa de l'obertura de diagrames de dispersió tW, mentre que aquelles d'ordre $\mathcal{O}(\alpha_s^2\alpha^2)$ la disminueixen en aproximadament un 4%. A més a més, càlculs avançats que inclouen diagrames que consideren l'emissió de partons addicionals porten a un augment al voltant del 10% de la secció eficaç i, en combinació amb els efectes descrits anteriorment, es veu que modelen millor les dades. Això demana la necessitat d'un càlcul complet a tercer ordre (NNLO).

Tant ATLAS com CMS han mesurat seccions eficaces inclusives, i en el cas d'ATLAS també a nivell diferencial, del procés $t\bar{t}W$. Ambdós troben una taxa més alta en les observacions que en les prediccions del SM, fins i tot després d'incloure els termes esmentats anteriorment. Per aquesta raó, és interessant trobar altres observables que puguen provar la producció $t\bar{t}W$, i que no siguen sensibles a canvis de normalització en la secció eficaç, com es el cas de l'asimetria de càrrega.

R.6.2 Asimetria de càrrega leptònica

En la producció $t\bar{t}W$, el domini de l'estat inicial $q\bar{q}'$ porta a una asimetria de càrrega, basada en la pseudorapidesa, més gran respecte a la producció $t\bar{t}$. A més, el bosó W radiat d'un dels quark inicials (ISR), serveix com a polaritzador de l'estat inicial $q\bar{q}'$, i, així, dels quarks $t\bar{t}$ finals. Aquesta polarització es veu reforçada encara més en l'asimetria entre els productes de desintegració dels quarks top i antitop, com són els leptons carregats. Aquesta asimetria de càrrega leptònica, A_c^l, basada en les diferències absolutes de pseudorapidesa dels leptons $\Delta|\eta_\ell| = |\eta_{\ell^+}| - |\eta_{\ell^-}|$, es defineix com a

$$A_{c,\eta}^{\ell} = \frac{N(\Delta|\eta_\ell| > 0) - N(\Delta|\eta_\ell| < 0)}{N(\Delta|\eta_\ell| > 0) + N(\Delta|\eta_\ell| < 0)}. \tag{R.4}$$

La asimetria de càrrega leptònica per a $t\bar{t}W$ és més gran respecte a la producció de $t\bar{t}$ a costa d'una secció eficaç més xicoteta. No solament és sensible a la física BSM, com ara axigluons i escenaris SMEFT corresponents a operadors de quatre fermions, sinó que també té el potencial únic de discriminar entre senyals de nova física amb diferent estructura quiral que tindrien efectes indistinguibles en els observables de secció eficaç. Com s'ha dit abans, l'A_c^l és insensible a la normalització de la taxa de producció de $t\bar{t}W$, el qual permet l'estudi d'aquest procés de manera independent.

En aquesta secció, s'analitza l'asimetria de càrrega leptònica en la producció $t\bar{t}W$ al detector ATLAS. Aquesta anàlisi es basa en un conjunt de dades enregistrades en el període 2015–2018, conegut com a Run 2, amb una energia de col·lisió de 13 TeV. Els resultats es donen a nivell de reconstrucció, és a dir, utilitzant els objectes reconstruïts per ATLAS a partir de les partícules de les

col·lisions; i també a nivell de partícules, on els efectes de resolució del detector així com les efficiencies i l'acceptancia es tracten de corregir per a proporcionar una comparació més fàcil amb les prediccions teòriques. Aquests resultats s'han publicat en la Ref. [302].

R.6.3 Selecció d'esdeveniments

Aquesta anàlisi considera un estat final amb tres leptons carregats lleugers (electrons o muons). La presència d'aquests leptons permet una reducció natural dels processos de fons i, a més, són objectes que es reconstrueix amb alta precisió. Aquests leptons són preseleccionats pel trigger del detector ATLAS, seguint un conjunt de requisits cinemàtics. Per assegurar que tots els esdeveniments compleixen aquests requisits, els tres leptons han de tenir un p_{T} mínim de 30, 20 i 15 GeV, respectivament. Això permet la distinció del leptó menys energètic, que s'utilitza per al procediment d'estimació del fons.

Amb l'objectiu de reduir encara més el nombre d'esdeveniments dels processos de fons, s'imposen requisits addicionals: la suma total de les càrregues dels tres leptons ha de ser ± 1, i la massa invariant del parell de leptons amb signes contraris però del mateix sabor (OSSF), $m_{\ell\ell}^{\mathrm{OSSF}}$, ha de ser com a mínim de 30 GeV. A més, en les regions de senyal es requereix que aquesta massa invariant estiga fora del rang $[m_Z - 10\ \mathrm{GeV}, m_Z + 10\ \mathrm{GeV}]$ per a excloure esdeveniments amb leptons que vinguen del bosó Z.

Les regions de senyal es divideixen en regions anomenades "lowN_{jets}" amb dos o tres jets, i "highN_{jets}" amb almenys quatre jets, i aquestes es divideixen més depenent de si tenen un b–jet o més. Definint un total de quatre regions de senyal. Les regions de senyal amb exactament un b–jet també han de satisfer $E_{\mathrm{T}}^{\mathrm{miss}} \geq 50$ GeV. Per a poder extraure l'A_c^l, les quatre regions de senyal se separen addicionalment en regions $\Delta|\eta_\ell| \leq 0$ ($\Delta\eta^-$) i $\Delta|\eta_\ell| > 0$ ($\Delta\eta^+$). La regió de senyal més sensible al procés $t\bar{t}W$ és la SR–$2b$–lowN_{jets}.

R.6.3.1 Estimació del fons

Després de la selecció general descrita anteriorment, els principals fons són: la producció $t\bar{t}Z$ (irreduïble), i les fonts de leptons *falsos** procedents de desintegracions d'hadrons que es formen a partir d'un quark pesant (HF), o de fotons que produeixen un parell electró–positró ($\gamma^* \to e^+e^-$) en interaccionar amb el material del detector (γ–*conversions*). Els leptons falsos són una font important de fons reduïble, i en aquesta anàlisi provenen principalment del procés $t\bar{t}$ (i també de Z+jets i tW), on s'esperen només dos leptons reals i altres falsos.

*Els leptons falsos són objectes que semblen ser detectats pel detector com a leptons, però que en realitat són partícules que no ho són. Això pot ser degut a errors de mesura o a la presència d'altres partícules que imiten la signatura d'un leptó.

El nombre d'esdeveniments amb leptons falsos en les regions de senyal es tracta de reduir amb criteris de selecció específics i algoritmes de reconstrucció de leptons més eficaços.

Per a constrényer la normalització dels fons principals es realitza un ajust al nombre d'esdeveniments mesurats en les dades en regions de senyal i fons. Aquestes últimes segueixen un criteri de selecció específic addicional dirigit als principals fons de l'anàlisi, i serveixen a més per a verificar el correcte modelatge dels fons que representen. Es defineixen de la següent manera:

- CR–$t\bar{t}Z$: es seleccionen els esdeveniments amb almenys 4 jets, dos (o més) dels quals són b–jets, i un parell de leptons OSSF amb una massa invariant en el rang $[m_Z - 10\text{ GeV}, m_Z + 10\text{ GeV}]$. En la resta de regions de fons, també es requereix que cap parell de leptons tinga una massa invariant en aquest rang.

- CR–HF$_e$ i CR–HF$_\mu$: es defineixen dues regions, que se separen pel sabor del leptó menys energètic dels tres del esdeveniment (electró o muó). Aquest ha de fallar els requisits d'aïllament dels leptons. Això assegura que siga originari d'una desintegració HF més del 99% de les vegades.

- CR–γ–conv: s'utilitzen una sèrie de criteris especials per a identificar electrons que provenen de fotons convertits ($\gamma^* \to e^+e^-$). En aquesta regió, se seleccionen esdeveniments on hi haja un electró que satisfaça aquests criteris.

A més, estan separades en regions $\Delta\eta^-$ i $\Delta\eta^+$ per a permetre un millor modelatge de les asimetries del fons. Les Figures R.9 i R.10 mostren aquestes regions de control a nivell reconstrucció després de l'ajust a les dades. Com es pot observar, hi ha un bon acord entre la simulació i les dades, la qual cosa confirma un modelatge correcte d'aquestos processos del SM.

R.6.4 Associació de leptons i quarks top

Un dels principals reptes és identificar els leptons que provenen dels quarks top i antitop per a construir l'observable $\Delta|\eta_\ell|$ i, en última instància l'A_c^l, tal com es defineix a l'Eq. R.4. En un esdeveniment $t\bar{t}W$ amb tres leptons carregats (3ℓ), els que són provinents de la parella $t\bar{t}$ tenen signes de càrrega elèctrica oposats, mentre que el que prové del bosó W ISR té el mateix signe de càrrega que un dels dos anteriors. Això és cert per a qualsevol configuració possible dels estats finals 3ℓ. El qual garanteix que el leptó amb signe oposat als altres dos, anomenat leptó *imparell*, sempre prové d'un quark top (o antitop). Els altres dos leptons, del mateix signe, s'anomenen leptons *parells*. Per tant, el problema es redueix a identificar quin dels leptons parells prové del quark antitop (o top).

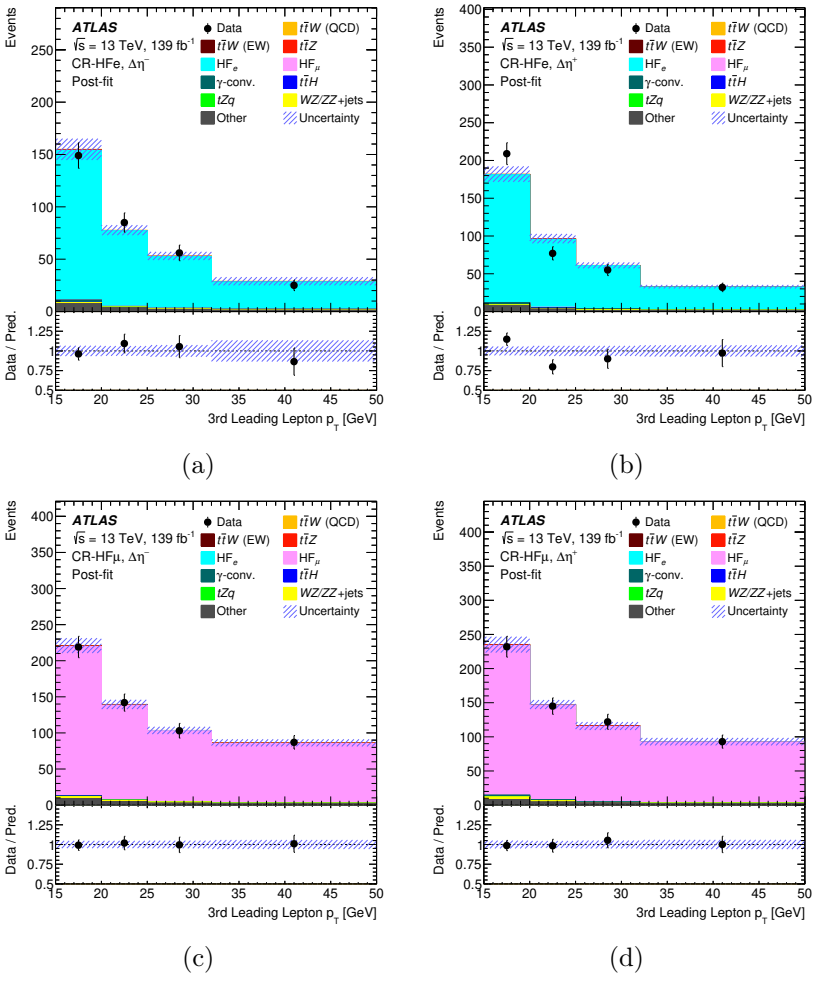


Figura R.9: Comparació entre les dades i les prediccions després de l'ajust en (a,b) CR–HF_e i (c,d) CR–HF_μ. Les distribucions mostren el p_{T} del leptó (electró o muó) menys energètic dels tres. Les regions estan separades entre $\Delta|\eta_\ell| \leq 0$ ($\Delta\eta^-$) i $\Delta|\eta_\ell| > 0$ ($\Delta\eta^+$). Les bandes d'error inclouen les incerteses totals en les prediccions després de l'ajust.

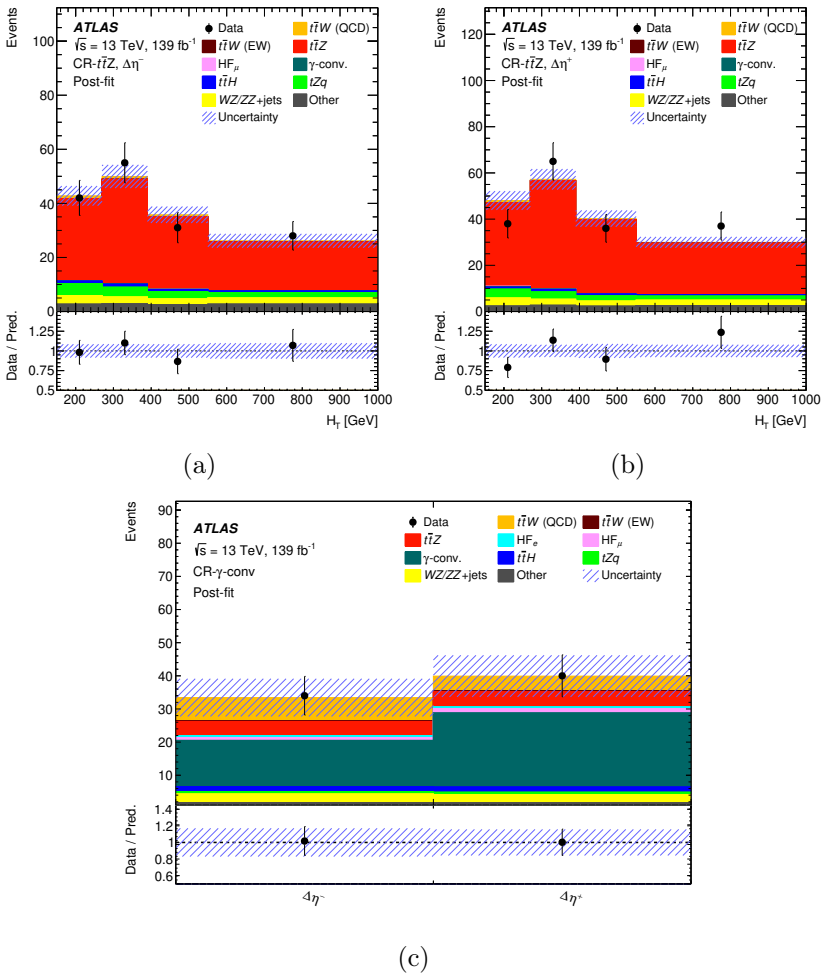


Figura R.10: Comparació entre les dades i les prediccions després de l'ajust en (a,b) CR–$t\bar{t}Z$ i (c) CR–γ–conv. Les distribucions mostren la suma del moment dels jets (H_T) per a CR–$t\bar{t}Z$, i el nombre total d'esdeveniments per a CR–γ–conv. Les regions estan separades entre $\Delta|\eta_\ell| \leq 0$ ($\Delta\eta^-$) i $\Delta|\eta_\ell| > 0$ ($\Delta\eta^+$). Les bandes d'error inclouen les incerteses totals en les prediccions després de l'ajust.

Aquest problema se soluciona mitjançant un algoritme d'aprenentatge automàtic (BDT) que calcula un valor discriminador per a cada leptó parell en cada esdeveniment. Els valors discriminadors alts corresponen a probabilitats altes que un determinat leptó haja originat d'una desintegració d'un quark top o antitop. El leptó amb la puntuació de discriminador BDT més alta s'escull per a calcular $\Delta|\eta_\ell|$. La BDT s'entrena amb esdeveniments $t\bar{t}W$ i, per a garantir que els esdeveniments avaluats siguen independents d'aquells utilitzats per a l'entrenament, s'utilitza la tècnica k–fold la qual garanteix que el rendiment del BDT siga independent de l'elecció del conjunt de dades. En aquesta tècnica, el conjunt de dades es divideix en $k = 5$ parts, i cada part es tracta com a conjunt de validació mentre que les altres $k - 1$ parts es tracten com a conjunt d'entrenament.

Per a l'entrenament, s'utilitza un conjunt de cinc variables d'entrada que són capaces de distingir entre els dos tipus de leptons parells. Aquestes variables són les masses dels dos sistemes formats pel leptó i els dos b–jets més propers ($m_{\ell b_0}$ i $m_{\ell b_1}$), així com les distàncies angulars entre el leptó i aquests b–jets ($\Delta R_{\ell b_0}$ i $\Delta R_{\ell b_1}$), i el p_{T} del leptó.

Amb aquesta configuració, el BDT és capaç d'identificar correctament el leptó del quark top aproximadament el 71% de les vegades. També és interessant mencionar que la variable que té la capacitat discriminatòria més gran és $m_{\ell b_0}$. De fet, si aquesta variable s'utilitza per realitzar la classificació, en lloc de la BDT, el rendiment es redueix a només un 65%. En aquest cas, el leptó seleccionat s'escull si el seu $m_{\ell b_0}$ està més a prop de 92 GeV que l'altre leptó. S'ha trobat que aquest valor és el valor de pic de la distribució de $m_{\ell b_0}$ per als leptons que provenen del quark top, com es mostra a la Figura R.11.

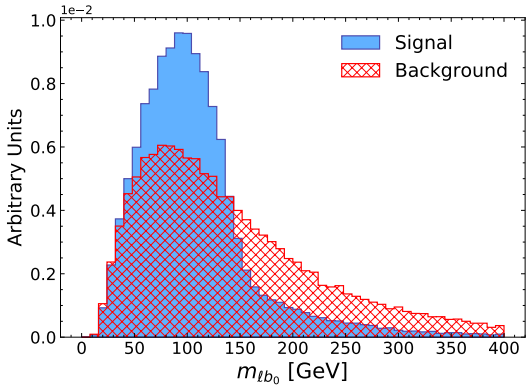


Figura R.11: Distribucions normalitzades que comparen els leptons que provenen dels quarks top (blau) amb els leptons que provenen dels bosons W ISR (vermell) en esdeveniments $t\bar{t}W$ per a la variable d'entrada de la BDT: $m_{\ell b_0}$.

R.6.5 Correcció a nivell de partícules

Per tal de comparar amb les prediccions teòriques en una regió fiducial, la distribució $\Delta|\eta_\ell|$ (després de la subtracció del fons) és corregida per l'acceptació del detector, així com per l'eficiència de reconstrucció. Aquesta correcció es fa mitjançant la tècnica de desplegament (*unfolding*) que consisteix a utilitzar la matriu de migració entre les regions de senyal reconstruïdes i a nivell de partícules. Tant la matriu com les altres correccions es calculen a partir de les dades de simulació MC del procés $t\bar{t}W$, fent servir el generador SHERPA [103]. Aquesta técnica ve descrita per l'expressió

$$N_i^{\text{plegat}} = \underbrace{\frac{1}{\alpha_i}\sum_j \varepsilon_j M_{ij}}_{R_{ij}} N_j^{\text{fid}} \quad \text{amb} \quad M_{ij} = \frac{N_{ij}^{(\text{reco}\,\cap\,\text{fid})}}{N_j^{(\text{reco}\,\cap\,\text{fid})}}, \quad \alpha_i = \frac{N_i^{(\text{reco}\,\cap\,\text{fid})}}{N_i^{\text{reco}}}, \quad \varepsilon_j = \frac{N_j^{(\text{reco}\,\cap\,\text{fid})}}{N_j^{\text{fid}}}, \tag{R.5}$$

on N_j^{fid} representa la distribució $\Delta|\eta_\ell|$ a nivell de partícules i N_j^{reco} a nivell de reconstrucció. El símbol $\cap$ representa la intersecció lògica de les dues regions. ε_j i α_i corregeixen l'eficiència i l'acceptació, respectivament, i M_{ij} és la matriu de migració. La Figura R.12 mostra aquestes tres correccions per a la regio e senyal SR–2b–lowN_{jets}. La matriu de migració és prou diagonal i no s'observa cap dependència amb $\Delta|\eta_\ell|$. Les altres regions de senyal mostren resultats pareguts. Per a validar aquest mètode, es fa servir un test de linealitat que confirma que l'A_c^l que s'obté a nivell de partícules és lineal amb l'A_c^l a nivell de reconstrucció. Aquest test es pot veure a la Figura R.13, i mostra que el mètode de desplegament no introdueix cap biaix significatiu.

R.6.6 Fonts d'incertesa

Les prediccions del senyal $t\bar{t}W$ i els fons del SM estan afectades per diverses fonts d'incerteses sistemàtiques experimentals i teòriques, així com per incerteses estadístiques degudes a la quantitat limitada de dades disponibles o d'esdeveniments seleccionats. Les incerteses experimentals se centren en les incerteses relacionades amb el detector en la mesura de quantitats com la lluminositat, o les diverses calibracions en la reconstrucció i identificació dels objectes físics. Les incerteses teòriques estan relacionades amb la modelització de processos com l'elecció del generador de MC. L'impacte d'aquestes incerteses en l'A_c^l es calcula per a cada mostra i es propaga a través de totes les regions de senyal i control. Després d'examinar cada incertesa, el seu efecte s'afegeix en quadratura a la incertesa total en l'A_c^l. En aquesta anàlisi, les incerteses estadístiques dominen severament la incertesa total.

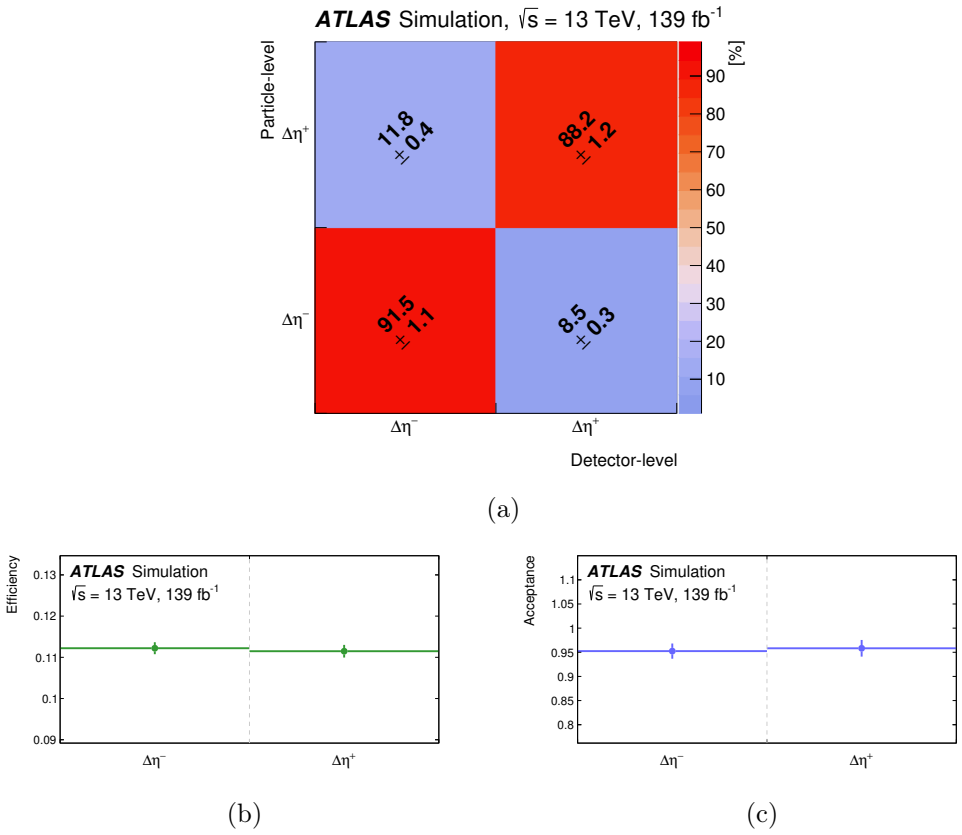


Figura R.12: (a) La matriu de migració, (b) l'eficiència i (c) les correccions d'acceptació per a la regió de senyal SR–2b–lowN_{jets}. La matriu està normalitzada per files. Les barres d'error de l'eficiència i les correccions d'acceptació representen les incerteses estadístiques de la simulació MC $t\bar{t}W$ SHERPA per a cada bin.

R.6.7 Resultats

Per extreure l'asimetria de càrrega leptònica a partir dels leptons reconstruïts, es realitza un ajust de màxima versemblança al nombre d'esdeveniments observats a les regions de senyal i control. Es defineixen factors de normalització separats per a les regions $\Delta\eta^-$ i $\Delta\eta^+$ per als principals fons i el senyal. Per als fons, aquesta separació es fa per evitar qualsevol possible biaix a partir d'una suposició d'asimetries del SM per a aquests processos en les dades. Per al senyal, es defineixen $\mathcal{N}_{\Delta\eta^-}$ i $\mathcal{N}_{\Delta\eta^+}$ i un d'ells, $\mathcal{N}_{\Delta\eta^+}$, es reparametritza utilitzant l'equació R.4 per permetre l'extracció de l'A_c^l directament de l'ajust com a paràmetre d'interés.

Les contribucions de $\Delta|\eta_\ell|$ en cadascuna de les regions de senyal a nivell reconstrucció es donen a la Figura R.14. S'observa bon acord entre la simulació

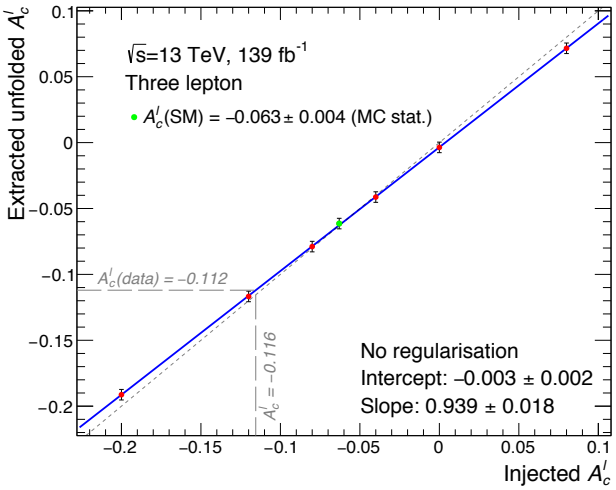


Figura R.13: Resultats de la prova de linealitat amb set punts de pseudo–dades (en vermell), incloent el punt on també es prova l'A_c^l del SM (en verd), calculat amb la mostra $t\bar{t}W$ SHERPA. Les barres d'error inclouen les incerteses estadístiques del MC. La línia blava representa la millor línia d'ajust als set punts i els seus paràmetres d'ajust també es mostren. La línia discontinua diagonal mostra una concordança perfecta entre els valors extrets i injectats de l'A_c^l. En la zona inferior esquerra, es mostra una representació esquemàtica de com s'estima el biaix de desplegament a partir del valor observat de l'A_c^l (vegeu Secció R.6.7).

i les dades.

El factor de normalització per al procés $t\bar{t}W$ es troba (dins de la seua incertesa) compatible amb les últimes mesures de secció eficaç de $t\bar{t}W$ d'ATLAS i CMS. L'asimetria de càrrega leptònica en $t\bar{t}W$ al nivell de reconstrucció es troba que és

$$A_c^l(t\bar{t}W) = -0.12 \pm 0.14\,(\text{est.}) \pm 0.05\,(\text{sist.}).$$

Això és consistent amb l'expectativa del SM de

$$A_c^l(t\bar{t}W)_{\text{SM}} = -0.084\,^{+0.005}_{-0.003}\,(\text{escala}) \pm 0.006\,(\text{est. MC}),$$

calculat utilitzant la simulació de $t\bar{t}W$ SHERPA.

Així mateix, l'asimetria de càrrega a nivell de partícules dona

$$A_c^l(t\bar{t}W)^{\text{PL}} = -0.11 \pm 0.17\,(\text{est.}) \pm 0.05\,(\text{sist.}),$$

amb una expectativa del SM calculada utilitzant la simulació de $t\bar{t}W$ SHERPA de

$$A_c^l(t\bar{t}W)^{\text{PL}}_{\text{SM}} = -0.063\,^{+0.007}_{-0.004}\,(\text{escala}) \pm 0.004\,(\text{est. MC}).$$

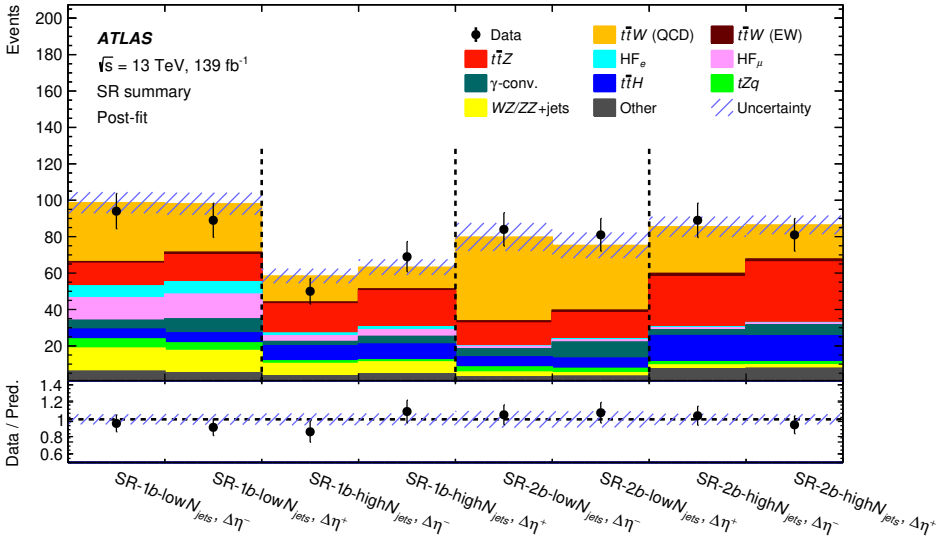


Figura R.14: Comparació entre les dades i les prediccions després de l'ajust en les quatre regions de senyal per a $\Delta|\eta_\ell| \leq 0$ ($\Delta\eta^-$) i $\Delta|\eta_\ell| > 0$ ($\Delta\eta^+$). Les bandes d'error inclouen les incerteses totals en les prediccions després de l'ajust.

Ambdós valors són compatibles amb les seues expectatives del SM i les incerteses estan dominades per la component estadística.

R.7 Estudi de la violació $\mathcal{CP}$ en esdeveniments de $t\bar{t}H$ i tH

L'estudi de l'acoblament Yukawa del quark top amb el bosó de Higgs és clau per a investigar les propietats del Higgs, com ara les propietats $\mathcal{CP}$ d'aquesta interacció. El Model Estàndard prediu un bosó de Higgs escalar ($J^{\mathcal{CP}} = 0^{++}$) amb una interacció prescrita amb el quark top. No obstant això, encara no s'ha exclòs la presència d'una mescla pseudoescalar $J^{\mathcal{CP}} = 0^{+-}$, que introduiria una segona interacció amb el quark top. L'observació d'aquesta contribució $\mathcal{CP}$–imparell seria una indicació de física més enllà del Model Estàndard.

Tant ATLAS com CMS han estudiat possibles contribucions $\mathcal{CP}$–imparell en els acoblaments del Higgs amb bosons vectorials o fermions, sent aquests últims els més sensibles, ja que el terme $\mathcal{CP}$–imparell entra en el mateix nivell que els $\mathcal{CP}$–parells. L'acoblament Yukawa del quark top pot ser modificat afegint un terme $\mathcal{CP}$–imparell a la Lagrangiana, que es pot escriure com a

$$\mathscr{L} = -\frac{G_t}{\sqrt{2}} H\{\bar{t}\,\kappa_t[\cos(\alpha) + \mathrm{i}\sin(\alpha)\gamma_5]t\}, \tag{R.6}$$

on κ_t és el factor de força de l'acoblament i α és l'angle de mescla $\mathcal{CP}$. El component $\mathcal{CP}$–imparell afecta les taxes de producció dels processos $t\bar{t}H$, tH,

així com a processos amb un bucle com són la desintegració $H \to \gamma\gamma$ i la producció ggF. Això es pot veure a la Figura R.15.

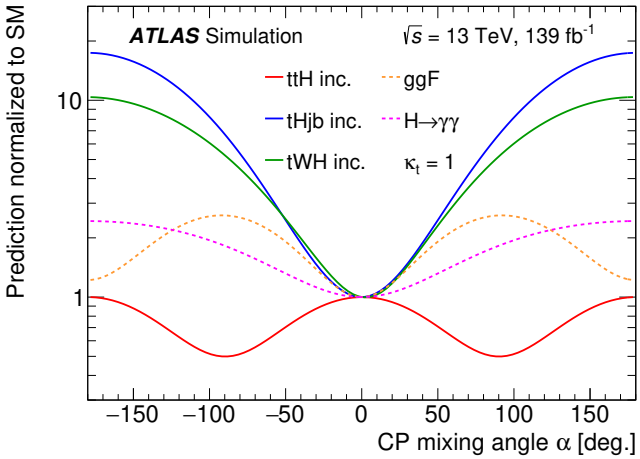


Figura R.15: Dependència de les taxes esperades dels processos $t\bar{t}H$ (vermell), $tHjb$ (blau), tWH (verd), així com la desintegració $H \to \gamma\gamma$ (magenta) i la producció ggF (taronja) amb l'angle de mescla per a $\kappa_t = 1$. Normalitzat a l'expectativa del SM.

Aquesta anàlisi presenta la primera cerca de violació $\mathcal{CP}$ en l'acoblament Yukawa del quark top utilitzant els modes de producció $t\bar{t}H$ i tH en el canal de desintegració a dos fotons. L'anàlisi es basa en les dades de col·lisió pp amb una energia de centre de massa de $\sqrt{s} = 13$ TeV enregistrades amb el detector ATLAS en el Run 2. Aquests resultats s'han publicat en la Ref. [219].

Les seccions eficaces de les mostres usades en aquesta anàlisi es corregeixen a les millors prediccions teòriques, que utilitzen diferents formes funcionals de les escales de renormalització i factorització. El factor de correcció també s'estudia en funció de l'angle de mescla $\mathcal{CP}$ ($\alpha = 0°, 45°, 90°$) i es troba que és compatible amb l'angle α triat.

Es requereixen esdeveniments amb dos fotons aïllats (de la desintegració del bosó de Higgs) amb un p_{T} superior a 35 i 25 GeV. A més, els esdeveniments se separen en dues regions enriquides en $t\bar{t}H$: en primer lloc, la regió "Lep" ($1\ell, \geq 1j, \geq 1b$) inclou esdeveniments on un bosó W (d'una desintegració del quark top) es desintegra en un leptó carregat i el seu neutrí associat; en segon lloc, la regió "Had" ($0\ell, \geq 3j, \geq 1b$) inclou la resta d'esdeveniments on els bosons W es desintegren en una parella $q\bar{q}$ i produint dos jets addicionals amb un p_{T} mínim de 25 GeV. Es fan servir tres BDT, amb el paquet XGBoost, per extreure el resultat d'aquesta anàlisi:

- Una BDT anomenada "Top Reco BDT" està entrenada amb la mostra $t\bar{t}H$ per extreure un candidat a quark top. El seu objectiu és discriminar

combinacions de jets aleatòries dels productes de desintegració del quark top. Fa servir informació dels productes de desintegració del quark top en ambdues categories.

- També es fa servir una BDT per a discriminar esdeveniments de senyal $t\bar{t}H$ d'aquells de fons. Exhibeix un bon rebuig de fons i una dependència feble amb l'angle de mescla $\mathcal{CP}$. Fa servir informació del quadrimoment dels fotons, jets i leptons a l'esdeveniment.

- Una última BDT s'utilitza per separar esdeveniments $\mathcal{CP}$–parells dels $\mathcal{CP}$–imparells. Fa servir informació del quadrimoment dels candidats a quark top i bosó Higgs reconstruïts, així com dels fotons, jets i leptons a l'esdeveniment.

Amb les dues últimes BDTs, es realitza una categorització per obtenir regions on se separen els esdeveniments de senyal i fons, utilitzant els valors discriminants de les BDTs. Aquesta categorització es fa per a les regions Lep i Had.

Després, es realitza un ajust simultani de màxima versemblança per a l'espectre de la massa invariant dels fotons $m_{\gamma\gamma}$ en totes les categories. A partir de l'ajust simultani, es poden extreure els límits unidimensionals (angle de mescla $\mathcal{CP}$ α) i bidimensionals (contorns $\kappa_t \cos\alpha - \kappa_t \sin\alpha$) sobre les propietats $\mathcal{CP}$ de l'acoblament Yukawa del quark top. Aquests es mostren a la Figura R.16 en els panells esquerre i dret, respectivament.

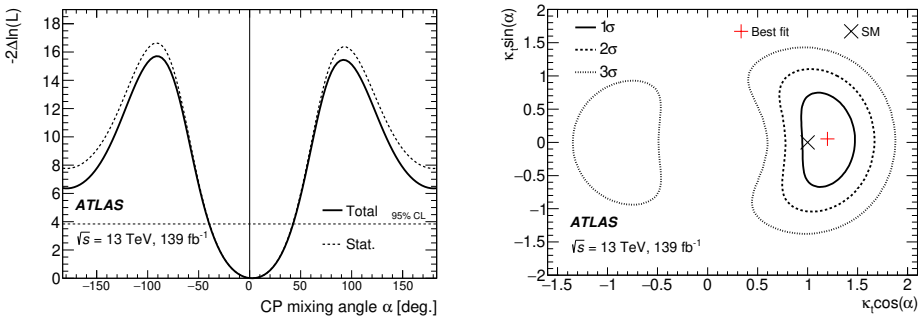


Figura R.16: Esquerra: Distribució de versemblança unidimensional de l'angle de mescla α. Les corbes sòlides i puntejades representen els resultats amb la incertesa total o només la incertesa estadística, respectivament. La línia horitzontal puntejada representa el 95% de nivell de confiança. Dreta: Contorns de versemblança bidimensionals $\kappa_t \cos\alpha - \kappa_t \sin\alpha$ dels resultats per al límit de $\mathcal{CP}$.

El límit d'exclusió observat (esperat) per a l'angle de mescla $\mathcal{CP}$ és $|\alpha| > 43°$ ($63°$) al 95% de nivell de confiança. La hipòtesi $\mathcal{CP}$–imparell queda exclosa a 3.9σ (2.5σ). La incertesa estadística domina en aquesta anàlisi.

R.8 Conclusions

El quark top, els seus acoblaments i les seues propietats, són excel·lents proves del SM. És l'únic quark amb un acoblament Yukawa al bosó de Higgs de l'ordre de la unitat, i té el potencial d'acoblar-se fortament amb altres partícules pesants. Tot plegat, el quark top és una partícula crucial per estudiar tant les mesures d'alta precisió del SM com búsquedes BSM o interpretacions en una EFT. El treball d'aquesta tesi s'ha centrat en l'estudi de les interaccions EW del quark top amb altres partícules del SM utilitzant les grans bases de dades recopilades pel detector ATLAS durant el Run 2 al LHC, així com bases de dades d'altres experiments.

En primer lloc, es realitza un ajust global EFT en el sector EW del quark top, utilitzant dades del LHC, Tevatron i LEP/SLC que són sensibles als corresponents coeficients de Wilson ($C_{\phi Q}^{-}$, $C_{\phi Q}^{(3)}$, $C_{\varphi t}$, C_{tW}, C_{tZ}, $C_{t\varphi}$, $C_{\varphi tb}$, C_{bW}), i no es troben desviacions significatives del SM. No obstant això, en virtut de les últimes mesures diferencials de producció de $t\bar{t}Z$ i $t\bar{t}\gamma$ i la nova parametrització amb precisió NLO de l'impacte dels operadors EFT en les diferents observables de secció eficaç, es presenten algunes de les limitacions més estrictes sobre el conjunt de coeficients de Wilson.

En segon lloc, s'estudia l'asimetria de càrrega leptònica en esdeveniments de $t\bar{t}W$ en el canal 3ℓ. En aquesta anàlisi, es fa servir d'una BDT per a associar els leptons carregats amb el quark top que els produeix. A més, la normalització dels principals fons es restringeixen mitjançant regions de control i s'extreuen simultàniament en l'ajust a les dades. Per proporcionar una comparació directa amb les prediccions teòriques, l'A_c^l es desplega a nivell de partícules en una regió fiducial que s'escull per a estar a prop de la regió de reconstrucció per minimitzar els efectes d'acceptació. Els resultats per l'A_c^l, tant a nivell de reconstrucció com a nivell de partícules, coincideixen amb la predicció del SM i estan severament dominats per la incertesa estasdítica. Per tant, podran millorar a mesura que es recullen més dades de l'LHC.

Finalment, s'estudia una de les limitacions del SM, com ara la insuficient contribució a la violació $\mathcal{CP}$ per a explicar, per exemple, l'asimetria matèria–antimatèria de l'univers primigeni. Es modifica l'acoblament Yukawa del quark top per a incloure termes de violació de $\mathcal{CP}$, que es parametritzen amb l'angle de mescla $\mathcal{CP}$ α. Se seleccionen esdeveniments dels processos $t\bar{t}H$ i tH, que proporcionen un accés directe a aquest vèrtex, i les secciones eficaces dels quals presenten una dependència amb α. Amb això, es pot construir una prova d'hipòtesi per extreure límits d'exclusió a l'angle de mescla $\mathcal{CP}$. Després d'un ajust simultani a les dades, la hipòtesi $\mathcal{CP}$–imparell queda exclosa a 3.9σ i $|\alpha| > 43°$ s'exclou al 95% de nivell de confiança. En aquesta anàlisi, la incertesa estadística també és predominant.

En resum, aquesta tesi proporciona una contribució significativa a l'estudi

de les interaccions EW del quark top, i a la cerca de nova física en aquest sector. Els resultats d'aquesta tesi són compatibles amb el SM i no es troben desviacions significatives. No obstant això, les noves dades procedents de l'LHC permetran millores en aquestes mesures de precisió, possibles descobriments d'efectes de nova física, així com el refinament i desenvolupament dels generadors de MC actuals, i les tècniques de reconstrucció i calibració als detectors.

List of Acronyms

ALICE A Large Ion Collider Experiment

ASCOT Apparatus with Super COnducting Toroids

ATLAS A Toroidal LHC ApparatuS

AUC Area Under the Curve

BDT Boosted Decision Tree

BR Branching Ratio

BSM Beyond the Standard Model

CC Charged Current

CDF Collider Detector at Fermilab

CEPC Circular Electron Positron Collider

CERN European Organization for Nuclear Research

CKM Cabibbo–Kobayashi–Maskawa

CLIC Compact Linear Collider

CL Confidence Level

CMS Compact Muon Solenoid

COM Centre–of–mass

CR Control Region

CSC Cathode Strip Chamber

CTEQ Coordinated Theoretical Experimental Project on QCD

D0 D–Zero

DESY Deutsches Elektronen-Synchrotron

DGLAP Dokshitzer–Gribov–Lipatov–Altarelli–Parisi

DM Dark Matter

d.o.f. degrees of freedom

DR Diagram Removal

DS Diagram Subtraction

EAGLE Experiment for Accurate Gamma, Lepton and Energy measurements

ECAL Electromagnetic Calorimeter

EFT Effective Field Theory

EM Electromagnetic

EW Electroweak

FASER ForwArd Search ExpeRiment

FCC Future Circular Collider

FCNC Flavour Changing Neutral Current

FCal Forward Calorimeter

FSR Final State Radiation

ggF gluon–gluon Fusion

GIM Glashow–Iliopoulos–Maiani

GPMC General Purpose Monte Carlo

HCAL Hadronic Calorimeter

HEC Hadronic End–cap Calorimeter

HF Heavy Flavour

HL–LHC High Luminosity LHC

HLT High Level Trigger

IBL Insertable B-Layer

ID Inner Detector

IFAE Institut de Física d'Altes Energies

IFIC Instituto de Física Corpuscular

ILC International Linear Collider

IP Interaction Point

IRC Infrared and Collinear

ISR Initial State Radiation

JER Jet Energy Resolution

JES Jet Energy Scale

JVT Jet Vertex Tagger

L1 Level 1

LAr Liquid Argon

LEP Large Electron-Positron Collider

LHCb Large Hadron Collider beauty

LHCf Large Hadron Collider forward

LHC Large Hadron Collider

LO Leading Order

LUCID LUminosity Cherenkov Integrating Detector

Linac4 Linear Accelerator 4

MC Monte Carlo

MDT Monitored Drift Tube

ME Matrix Element

MEC Matrix Element Correction

ML Machine Learning

MSTW Martin–Stirling–Thorne–Watt

MS Minimal Subtraction

MS Muon Spectrometer

MoMEDAL Monopole and Exotics Detector at the LHC

NC Neutral Current

NF Normalisation Factor

NLO Next-to-Leading Order

NNLL Next-to-Next-to-Leading Logarithm

NNLO Next-to-Next-to-Leading Order

NNPDF Neural Network Parton Distribution Function

NSW New Small Wheel

OLR Overlap Removal

OSSF Opposite–Sign–Same–Flavor

PDF Parton Distribution Function

PF Particle Flow

PIC Port d'Informació Científica

POI Parameter of Interest

pQCD perturbative Quantum Chromodynamics

PS Parton Shower

PV Primary Vertex

QCD Quantum Chromodynamics

QED Quantum Electrodynamics

ROC Receiver Operating Characteristic

RPC Resistive Plate Chamber

S/B Signal over Background

SCT SemiConductor Tracker

SF Scale Factor

SLAC Stanford Linear Accelerator Center

SLC SLAC Large Collider

SMEFT Standard Model Effective Field Theory

SM Standard Model

SND Scattering and Neutrino Detector

SR Signal Region

SSB Spontaneous Symmetry Breaking

TASSO Two-Arm Spectrometer Solenoid

TGC Thin Gap Chamber

TOTEM TOTal Elastic and diffractive cross section Measurement

TRT Transition Radiation Tracker

UAM Universidad Autónoma de Madrid

UE Underlying Event

UFO Universal FeynRules Output

$V - A$ Vector minus Axial–vector

WLCG Worldwide LHC Computing Grid

WP Working Point

YR4 Yellow Report 4

4F Four Flavour

5F Five Flavour

References

[1] E. Rutherford, *Collision of α particles with light atoms. IV. An anomalous effect in nitrogen*, *Phil. Mag. Ser. 6* **37** (1919) 581. (cit. on p. 3)

[2] P. A. M. Dirac, *A Theory of Electrons and Protons*, *Proc. Roy. Soc. Lond. A* **126** (1930) 360. (cit. on p. 3)

[3] E. Schrödinger, *The fundamental idea of wave mechanics*, *Resonance* **4** (1999) 92. (cit. on p. 3)

[4] S. L. Glashow, *The renormalizability of vector meson interactions*, *Nucl. Phys.* **10** (1959) 107. (cit. on p. 3)

[5] S. L. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl. Phys.* **22** (1961) 579. (cit. on p. 3 and 199)

[6] S. Weinberg, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (1967) 1264. (cit. on p. 3 and 199)

[7] A. Salam and J. C. Ward, *Weak and electromagnetic interactions*, *Nuovo Cim.* **11** (1959) 568. (cit. on p. 3 and 199)

[8] H. Fritzsch, M. Gell-Mann and H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, *Phys. Lett. B* **47** (1973) 365. (cit. on p. 3 and 199)

[9] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508. (cit. on p. 3 and 200)

[10] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321. (cit. on p. 3 and 200)

[11] J. J. Thomson, *Cathode rays*, *Phil. Mag. Ser. 5* **44** (1897) 293. (cit. on p. 3)

[12] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1 [arXiv:1207.7214]. (cit. on p. 3, 19, and 200)

[13] CMS Collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30 [arXiv:1207.7235]. (cit. on p. 3, 19, and 200)

[14] WikipediA, *The Standard Model*, https://en.wikipedia.org/wiki/Standard_Model. (cit. on p. 4 and 200)

[15] ATLAS Collaboration, *Measurement of the top quark mass in the $t\bar{t} \to$ lepton+jets channel from $\sqrt{s} = 8$ TeV ATLAS data and combination with previous results*, *Eur. Phys. J. C* **79** (2019) 290 [arXiv:1810.01772]. (cit. on p. 5)

[16] Particle Data Group Collaboration, *Review of Particle Physics*, *PTEP* **2022** (2022) 083C01. (cit. on p. 5, 7, 18, 24, 26, 56, 57, 58, 92, and 93)

[17] T. D. Lee and C.-N. Yang, *Question of Parity Conservation in Weak Interactions*, *Phys. Rev.* **104** (1956) 254. (cit. on p. 8)

[18] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes and R. P. Hudson, *Experimental Test of Parity Conservation in β Decay*, *Phys. Rev.* **105** (1957) 1413. (cit. on p. 8 and 199)

[19] Y. Nambu, *Quasiparticles and Gauge Invariance in the Theory of Superconductivity*, *Phys. Rev.* **117** (1960) 648. (cit. on p. 14)

[20] J. Goldstone, *Field Theories with Superconductor Solutions*, *Nuovo Cim.* **19** (1961) 154. (cit. on p. 14)

[21] J. Goldstone, A. Salam and S. Weinberg, *Broken Symmetries*, *Phys. Rev.* **127** (1962) 965. (cit. on p. 14)

[22] ATLAS Collaboration, *A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery*, *Nature* **607** (2022) 52 [arXiv:2207.00092]. (cit. on p. 16, 17, and 68)

[23] S. L. Glashow, J. Iliopoulos and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, *Phys. Rev. D* **2** (1970) 1285. (cit. on p. 18 and 199)

[24] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, *Phys. Rev. Lett.* **10** (1963) 531. (cit. on p. 18)

[25] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652. (cit. on p. 18)

[26] L.-L. Chau and W.-Y. Keung, *Comments on the Parametrization of the Kobayashi-Maskawa Matrix*, *Phys. Rev. Lett.* **53** (1984) 1802. (cit. on p. 18)

[27] J. H. Christenson, J. W. Cronin, V. L. Fitch and R. Turlay, *Evidence for the* 2π *Decay of the* K_2^0 *Meson*, *Phys. Rev. Lett.* **13** (1964) 138. (cit. on p. 18 and 53)

[28] S. W. Herb et al., *Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions*, *Phys. Rev. Lett.* **39** (1977) 252. (cit. on p. 18 and 53)

[29] W. R. Innes et al., *Observation of Structure in the* Υ *Region*, *Phys. Rev. Lett.* **39** (1977) 1240. (cit. on p. 18 and 53)

[30] CDF Collaboration, *Observation of top quark production in* $\bar{p}p$ *collisions*, *Phys. Rev. Lett.* **74** (1995) 2626 [arXiv:hep-ex/9503002]. (cit. on p. 18, 26, 53, and 201)

[31] D0 Collaboration, *Observation of the top quark*, *Phys. Rev. Lett.* **74** (1995) 2632 [arXiv:hep-ex/9503003]. (cit. on p. 18, 26, 53, and 201)

[32] H. M. Georgi, S. L. Glashow, M. E. Machacek and D. V. Nanopoulos, *Higgs Bosons from Two Gluon Annihilation in Proton Proton Collisions*, *Phys. Rev. Lett.* **40** (1978) 692. (cit. on p. 19)

[33] D. R. T. Jones and S. T. Petcov, *Heavy Higgs Bosons at LEP*, *Phys. Lett. B* **84** (1979) 440. (cit. on p. 19)

[34] S. L. Glashow, D. V. Nanopoulos and A. Yildiz, *Associated Production of Higgs Bosons and Z Particles*, *Phys. Rev. D* **18** (1978) 1724. (cit. on p. 19)

[35] R. Raitio and W. W. Wada, *Higgs Boson Production at Large Transverse Momentum in QCD*, *Phys. Rev. D* **19** (1979) 941. (cit. on p. 19)

[36] W. J. Marciano and F. E. Paige, *Associated production of Higgs bosons with t anti-t pairs*, *Phys. Rev. Lett.* **66** (1991) 2433. (cit. on p. 19)

[37] J. F. Gunion, *Associated top anti-top Higgs production as a large source of W H events: Implications for Higgs detection in the lepton neutrino gamma gamma final state*, *Phys. Lett. B* **261** (1991) 510. (cit. on p. 19)

[38] Z. Kunszt, *Associated Production of Heavy Higgs Boson with Top Quarks*, *Nucl. Phys. B* **247** (1984) 339. (cit. on p. 19)

[39] J. M. Butterworth, A. R. Davison, M. Rubin and G. P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, *Phys. Rev. Lett.* **100** (2008) 242001 [arXiv:0802.2470]. (cit. on p. 19)

[40] ATLAS Collaboration, *Constraints on Higgs boson production with large transverse momentum using* $H \to b\bar{b}$ *decays in the ATLAS detector*, *Phys. Rev. D* **105** (2022) 092003 [arXiv:2111.08340]. (cit. on p. 19)

[41] ATLAS Collaboration, *Measurement of the associated production of a Higgs boson decaying into b-quarks with a vector boson at high transverse momentum in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, *Phys. Lett. B* **816** (2021) 136204 [arXiv:2008.02508]. (cit. on p. 19)

[42] D. L. Rainwater, D. Zeppenfeld and K. Hagiwara, *Searching for $H \to \tau^+\tau^-$ in weak boson fusion at the CERN LHC*, *Phys. Rev. D* **59** (1998) 014037 [arXiv:hep-ph/9808468]. (cit. on p. 19)

[43] ATLAS Collaboration, *Measurements of Higgs boson production cross-sections in the $H \to \tau^+\tau^-$ decay channel in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, *JHEP* **08** (2022) 175 [arXiv:2201.08269]. (cit. on p. 19)

[44] Y. Ne'eman, *Derivation of strong interactions from a gauge invariance*, *Nucl. Phys.* **26** (1961) 222. (cit. on p. 20 and 199)

[45] M. Gell-Mann, *The Eightfold Way: A Theory of strong interaction symmetry*, . (cit. on p. 20 and 199)

[46] V. E. Barnes et al., *Observation of a Hyperon with Strangeness Minus Three*, *Phys. Rev. Lett.* **12** (1964) 204. (cit. on p. 20)

[47] O. W. Greenberg, *Spin and Unitary Spin Independence in a Paraquark Model of Baryons and Mesons*, *Phys. Rev. Lett.* **13** (1964) 598. (cit. on p. 20)

[48] M. Y. Han and Y. Nambu, *Three Triplet Model with Double SU(3) Symmetry*, *Phys. Rev.* **139** (1965) B1006. (cit. on p. 20)

[49] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, *Phys. Rev. Lett.* **30** (1973) 1343. (cit. on p. 20)

[50] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, *Phys. Rev. Lett.* **30** (1973) 1346. (cit. on p. 20)

[51] B. R. Stella and H.-J. Meyer, *Y(9.46 GeV) and the gluon discovery (a critical recollection of PLUTO results)*, *Eur. Phys. J. H* **36** (2011) 203 [arXiv:1008.1869]. (cit. on p. 21)

[52] A. Ali and G. Kramer, *Jets and QCD: A Historical Review of the Discovery of the Quark and Gluon Jets and its Impact on QCD*, *Eur. Phys. J. H* **36** (2011) 245 [arXiv:1012.2288]. (cit. on p. 21)

[53] P. F. de Salas, D. V. Forero, C. A. Ternes, M. Tortola and J. W. F. Valle, *Status of neutrino oscillations 2018: 3σ hint for normal mass ordering and improved CP sensitivity*, *Phys. Lett. B* **782** (2018) 633 [arXiv:1708.01186]. (cit. on p. 22)

[54] R. N. Mohapatra and G. Senjanovic, *Neutrino Mass and Spontaneous Parity Nonconservation*, *Phys. Rev. Lett.* **44** (1980) 912. (cit. on p. 22)

[55] J. Schechter and J. W. F. Valle, *Neutrino Masses in SU(2) x U(1) Theories*, *Phys. Rev. D* **22** (1980) 2227. (cit. on p. 22)

[56] J. Schechter and J. W. F. Valle, *Neutrino Decay and Spontaneous Violation of Lepton Number*, *Phys. Rev. D* **25** (1982) 774. (cit. on p. 22)

[57] E. Ma, *Pathways to naturally small neutrino masses*, *Phys. Rev. Lett.* **81** (1998) 1171 [arXiv:hep-ph/9805219]. (cit. on p. 22)

[58] E. Ma and O. Popov, *Pathways to Naturally Small Dirac Neutrino Masses*, *Phys. Lett. B* **764** (2017) 142 [arXiv:1609.02538]. (cit. on p. 22)

[59] S. Centelles Chuliá, E. Ma, R. Srivastava and J. W. F. Valle, *Dirac Neutrinos and Dark Matter Stability from Lepton Quarticity*, *Phys. Lett. B* **767** (2017) 209 [arXiv:1606.04543]. (cit. on p. 22)

[60] S. Centelles Chuliá, R. Srivastava and J. W. F. Valle, *Seesaw roadmap to neutrino mass and dark matter*, *Phys. Lett. B* **781** (2018) 122 [arXiv:1802.05722]. (cit. on p. 22)

[61] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, *Pisma Zh. Eksp. Teor. Fiz.* **5** (1967) 32. (cit. on p. 22)

[62] E. de Rafael, *Chiral Lagrangians and kaon CP violation*, in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 94): CP Violation and the limits of the Standard Model*, 1, 1995, arXiv:hep-ph/9502254. (cit. on p. 23)

[63] H. Georgi, *Heavy quark effective field theory*, in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 91): Perspectives in the Standard Model*, 8, 1991. (cit. on p. 23)

[64] W. Buchmuller and D. Wyler, *Effective Lagrangian Analysis of New Interactions and Flavor Conservation*, *Nucl. Phys. B* **268** (1986) 621. (cit. on p. 23, 70, and 202)

[65] C. N. Leung, S. T. Love and S. Rao, *Low-Energy Manifestations of a New Interaction Scale: Operator Analysis*, *Z. Phys. C* **31** (1986) 433. (cit. on p. 23 and 202)

[66] C. Degrande, N. Greiner, W. Kilian, O. Mattelaer, H. Mebane, T. Stelzer et al., *Effective Field Theory: A Modern Approach to Anomalous Couplings*, *Annals Phys.* **335** (2013) 21 [arXiv:1205.4231]. (cit. on p. 23 and 202)

[67] I. Brivio and M. Trott, *The Standard Model as an Effective Field Theory*, *Phys. Rept.* **793** (2019) 1 [arXiv:1706.08945]. (cit. on p. 23, 91, and 202)

[68] F. Maltoni, L. Mantani and K. Mimasu, *Top-quark electroweak interactions at high energy*, *JHEP* **10** (2019) 004 [arXiv:1904.05637]. (cit. on p. 23 and 202)

[69] E. Fermi, *Tentativo di una teoria dell'emissione dei raggi beta*, *Ric. Sci.* **4** (1933) 491. (cit. on p. 24)

[70] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085 [arXiv:1008.4884]. (cit. on p. 25, 70, and 72)

[71] ATLAS Collaboration, *ATLAS detector and physics performance : Technical Design Report, 1*, CERN-LHCC-99-014 , (1994). (cit. on p. 27 and 38)

[72] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003. (cit. on p. 27, 38, and 203)

[73] L. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001. (cit. on p. 27 and 203)

[74] M. Benedikt, P. Collier, V. Mertens, J. Poole and K. Schindl, *LHC Design Report. 3. The LHC injector chain*, CERN-2004-003-V-3 , (2004). (cit. on p. 27 and 203)

[75] CERN, *The CERN accelerator complex in 2019*, CERN-GRAPHICS-2019-002 , (2019). (cit. on p. 29)

[76] G. Apollinari, I. Béjar Alonso, O. Brüning, M. Lamont and L. Rossi, *High-Luminosity Large Hadron Collider (HL-LHC) : Preliminary Design Report*, CERN-2015-005 , (2015). (cit. on p. 29)

[77] ATLAS Collaboration, *Luminosity determination in pp collisions at* $\sqrt{s} = 13$ *TeV using the ATLAS detector at the LHC*, arXiv:2212.09379. (cit. on p. 29, 92, and 106)

[78] CMS Collaboration, *The CMS Experiment at the CERN LHC*, *JINST* **3** (2008) S08004. (cit. on p. 30 and 203)

[79] LHCb Collaboration, *The LHCb Detector at the LHC*, *JINST* **3** (2008) S08005. (cit. on p. 30 and 203)

[80] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, *JINST* **3** (2008) S08002. (cit. on p. 30 and 203)

[81] CMS, TOTEM Collaboration, *CMS-TOTEM Precision Proton Spectrometer*, CERN-LHCC-2014-021, TOTEM-TDR-003, CMS-TDR-13 , (2014). (cit. on p. 30)

[82] MoEDAL Collaboration, *Technical Design Report of the MoEDAL Experiment*, CERN-LHCC-2009-006, MoEDAL-TDR-001 , (2009). (cit. on p. 30)

[83] LHCf Collaboration, *LHCf experiment : Technical Design Report*, CERN-LHCC-2006-004, LHCF-TDR-001 , (2006). (cit. on p. 30)

[84] FASER Collaboration, *FASER: ForwArd Search ExpeRiment at the LHC*, arXiv:1901.04468. (cit. on p. 30)

[85] SND@LHC Collaboration, *SND@LHC: The Scattering and Neutrino Detector at the LHC*, arXiv:2210.02784. (cit. on p. 30)

[86] Worldwide LHC Computing Grid. https://wlcg.web.cern.ch. (cit. on p. 31)

[87] ATLAS Collaboration, *Computing activities at the Spanish Tier-1 and Tier-2s for the ATLAS experiment towards the LHC Run 3 and High-Luminosity periods*, *EPJ Web Conf.* **245** (2020) 07027. (cit. on p. 31)

[88] E. D. Bloom et al., *High-Energy Inelastic e p Scattering at 6-Degrees and 10-Degrees*, *Phys. Rev. Lett.* **23** (1969) 930. (cit. on p. 31)

[89] J. M. Campbell, J. W. Huston and W. J. Stirling, *Hard Interactions of Quarks and Gluons: A Primer for LHC Physics*, *Rept. Prog. Phys.* **70** (2007) 89 [arXiv:hep-ph/0611148]. (cit. on p. 32)

[90] T.-J. Hou et al., *New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC*, *Phys. Rev. D* **103** (2021) 014013 [arXiv:1912.10053]. (cit. on p. 32)

[91] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, *Parton distributions for the LHC*, *Eur. Phys. J. C* **63** (2009) 189 [arXiv:0901.0002]. (cit. on p. 32)

[92] NNPDF Collaboration, *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040 [arXiv:1410.8849]. (cit. on p. 32 and 92)

[93] L. Del Debbio, *Parton distributions in the LHC era*, *EPJ Web Conf.* **175** (2018) 01006. (cit. on p. 32)

[94] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [arXiv:1410.3012]. (cit. on p. 34 and 36)

[95] M. Bahr et al., *Herwig++ Physics and Manual*, *Eur. Phys. J. C* **58** (2008) 639 [arXiv:0803.0883]. (cit. on p. 34 and 36)

[96] E. Boos et al., *Generic User Process Interface for Event Generators*, in *2nd Les Houches Workshop on Physics at TeV Colliders*, 9, 2001, arXiv:hep-ph/0109068. (cit. on p. 35)

[97] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer and T. Stelzer, *MadGraph 5 : Going Beyond*, *JHEP* **06** (2011) 128 [arXiv:1106.0522]. (cit. on p. 35 and 207)

[98] 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 [arXiv:1405.0301]. (cit. on p. 35, 76, and 207)

[99] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **06** (2002) 029 [arXiv:hep-ph/0204244]. (cit. on p. 35)

[100] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040 [arXiv:hep-ph/0409146]. (cit. on p. 36)

[101] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070 [arXiv:0709.2092]. (cit. on p. 36)

[102] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043 [arXiv:1002.2581]. (cit. on p. 36)

[103] Sherpa Collaboration, *Event Generation with Sherpa 2.2*, *SciPost Phys.* **7** (2019) 034 [arXiv:1905.09127]. (cit. on p. 36 and 217)

[104] T. Gleisberg and S. Hoeche, *Comix, a new matrix element generator*, *JHEP* **12** (2008) 039 [arXiv:0808.3674]. (cit. on p. 36)

[105] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, *JHEP* **03** (2008) 038 [arXiv:0709.1027]. (cit. on p. 36)

[106] A. Denner, S. Dittmaier and L. Hofer, *Collier: a fortran-based Complex One-Loop LIbrary in Extended Regularizations*, *Comput. Phys. Commun.* **212** (2017) 220 [arXiv:1604.06792]. (cit. on p. 36)

[107] F. Buccioni, S. Pozzorini and M. Zoller, *On-the-fly reduction of open loops*, *Eur. Phys. J. C* **78** (2018) 70 [arXiv:1710.11452]. (cit. on p. 36)

[108] F. Buccioni, J.-N. Lang, J. M. Lindert, P. Maierhöfer, S. Pozzorini, H. Zhang et al., *OpenLoops 2*, *Eur. Phys. J. C* **79** (2019) 866 [arXiv:1907.13071]. (cit. on p. 36)

[109] F. Cascioli, P. Maierhofer and S. Pozzorini, *Scattering Amplitudes with Open Loops*, *Phys. Rev. Lett.* **108** (2012) 111601 [arXiv:1111.5206]. (cit. on p. 36)

[110] M. Bengtsson and T. Sjostrand, *Coherent Parton Showers Versus Matrix Elements: Implications of PETRA - PEP Data*, *Phys. Lett. B* **185** (1987) 435. (cit. on p. 36)

[111] A. Buckley et al., *General-purpose event generators for LHC physics*, *Phys. Rept.* **504** (2011) 145 [arXiv:1101.2599]. (cit. on p. 36)

[112] W. T. Giele, D. A. Kosower and P. Z. Skands, *Higher-Order Corrections to Timelike Jets*, *Phys. Rev. D* **84** (2011) 054003 [arXiv:1102.2126]. (cit. on p. 36)

[113] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, *Eur. Phys. J. C* **76** (2016) 196 [arXiv:1512.01178]. (cit. on p. 36)

[114] J. Bellm et al., *Herwig 7.1 Release Note*, arXiv:1705.06919. (cit. on p. 36)

[115] J. Bellm et al., *Herwig 7.2 release note*, *Eur. Phys. J. C* **80** (2020) 452 [arXiv:1912.06509]. (cit. on p. 36)

[116] R. Frederix and S. Frixione, *Merging meets matching in MC@NLO*, *JHEP* **12** (2012) 061 [arXiv:1209.6215]. (cit. on p. 37, 170, 171, and 172)

[117] S. Hoeche, F. Krauss, M. Schonherr and F. Siegert, *QCD matrix elements + parton showers: The NLO case*, *JHEP* **04** (2013) 027 [arXiv:1207.5030]. (cit. on p. 37)

[118] L. Lönnblad and S. Prestel, *Merging Multi-leg NLO Matrix Elements with Parton Showers*, *JHEP* **03** (2013) 166 [arXiv:1211.7278]. (cit. on p. 37)

[119] K. Hamilton, P. Nason and G. Zanderighi, *MINLO: Multi-Scale Improved NLO*, *JHEP* **10** (2012) 155 [arXiv:1206.3572]. (cit. on p. 37 and 171)

[120] ATLAS Collaboration, *Constituent-level pile-up mitigation techniques in ATLAS*, ATLAS-CONF-2017-065 , (2017). (cit. on p. 37)

[121] G. Soyez, *Pileup mitigation at the LHC: A theorist's view*, *Phys. Rept.* **803** (2019) 1 [arXiv:1801.09721]. (cit. on p. 37)

[122] ATLAS Collaboration, *Luminosity public results*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2. (cit. on p. 38)

[123] CERN, *General Meeting on LHC Physics and Detectors*, https://cds.cern.ch/record/236265 , (1992). (cit. on p. 38)

[124] ASCOT Collaboration, *The ASCOT detector at the LHC : expression of interest*, in *General Meeting on LHC Physics and Detectors*, 4, 1992, https://cds.cern.ch/record/1076511. (cit. on p. 38)

[125] EAGLE Collaboration, *EAGLE : Experiment for Accurate Gamma, Lepton and Energy measurements : expression of interest*, in *General Meeting on LHC Physics and Detectors*, 4, 1992, https://cds.cern.ch/record/247456. (cit. on p. 38)

[126] ATLAS Collaboration, *ATLAS : letter of intent for a general-purpose pp experiment at the large hadron collider at CERN*, CERN-LHCC-92-004 , (1992). (cit. on p. 39)

[127] ATLAS Collaboration, *ATLAS experiment schematic illustration*, ATLAS-PHOTO-2022-055 , (2022). (cit. on p. 40 and 204)

[128] ATLAS Collaboration, *ATLAS inner detector : Technical Design Report, 1*, CERN-LHCC-97-016 , (1997). (cit. on p. 40)

[129] ATLAS Collaboration, *Studies of radial distortions of the ATLAS Inner Detector*, ATL-PHYS-PUB-2018-003 , (2018). (cit. on p. 41)

[130] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, CERN-LHCC-2010-013 , (2010). (cit. on p. 42)

[131] ATLAS Collaboration, *Particle Identification Performance of the ATLAS Transition Radiation Tracker*, ATLAS-CONF-2011-128 , (2011). (cit. on p. 42)

[132] ATLAS Collaboration, *ATLAS liquid-argon calorimeter : Technical Design Report*, CERN-LHCC-96-041 , (1996). (cit. on p. 42)

[133] ATLAS Collaboration, *ATLAS tile calorimeter : Technical Design Report*, CERN-LHCC-96-042 , (1996). (cit. on p. 42)

[134] ATLAS Collaboration, *Computer generated image of the ATLAS calorimeter*, CERN-GE-0803015 , (2008). (cit. on p. 43)

[135] ATLAS Collaboration, *ATLAS muon spectrometer : Technical Design Report*, CERN-LHCC-97-022 , (1997). (cit. on p. 43)

[136] ATLAS Collaboration, *Computer generated image of the ATLAS Muons subsystem*, CERN-GE-0803017 , (2008). (cit. on p. 44)

[137] ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*, *JINST* **15** (2020) P10004 [arXiv:2007.12539]. (cit. on p. 44)

[138] ATLAS Collaboration, *Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2*, *Eur. Phys. J. C* **77** (2017) 673 [arXiv:1704.07983]. (cit. on p. 45)

[139] ATLAS Collaboration, *Performance of the reconstruction of large impact parameter tracks in the ATLAS inner detector*, ATL-PHYS-PUB-2017-014 , (2017). (cit. on p. 45)

[140] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, *Nucl. Instrum. Meth. A* **262** (1987) 444. (cit. on p. 45)

[141] ATLAS Collaboration, *Performance of the ATLAS Silicon Pattern Recognition Algorithm in Data and Simulation at $\sqrt{s} = 7$ TeV*, ATLAS-CONF-2010-072 , (2010). (cit. on p. 45)

[142] ATLAS Collaboration, *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC*, *Eur. Phys. J. C* **77** (2017) 332 [arXiv:1611.10235]. (cit. on p. 45)

[143] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, *JINST* **14** (2019) P12006 [arXiv:1908.00005]. (cit. on p. 46, 47, and 153)

[144] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **79** (2019) 639 [arXiv:1902.04655]. (cit. on p. 46)

[145] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data*, *JINST* **14** (2019) P03017 [arXiv:1812.03848]. (cit. on p. 46)

[146] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using 2012 LHC proton–proton collision data*, *Eur. Phys. J. C* **77** (2017) 195 [arXiv:1612.01456]. (cit. on p. 47)

[147] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **81** (2021) 578 [arXiv:2012.00578]. (cit. on p. 47, 48, and 49)

[148] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, *Eur. Phys. J. C* **77** (2017) 466 [arXiv:1703.10485]. (cit. on p. 49)

[149] M. Cacciari, G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, *JHEP* **04** (2008) 063 [arXiv:0802.1189]. (cit. on p. 49)

[150] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **81** (2021) 689 [arXiv:2007.02645]. (cit. on p. 50 and 106)

[151] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, *Eur. Phys. J. C* **76** (2016) 581 [arXiv:1510.03823]. (cit. on p. 50)

[152] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **79** (2019) 970 [arXiv:1907.05120]. (cit. on p. 51)

[153] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, ATL-PHYS-PUB-2017-013 , (2017). (cit. on p. 51)

[154] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **78** (2018) 903 [arXiv:1802.08168]. (cit. on p. 52)

[155] L. Di Lella and C. Rubbia, *The Discovery of the W and Z Particles*, *Adv. Ser. Direct. High Energy Phys.* **23** (2015) 137. (cit. on p. 53)

[156] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, *Comput. Phys. Commun.* **185** (2014) 2930 [arXiv:1112.5675]. (cit. on p. 54 and 94)

[157] M. Czakon, P. Fiedler and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$*, *Phys. Rev. Lett.* **110** (2013) 252004 [arXiv:1303.6254]. (cit. on p. 54, 89, and 94)

[158] ATLAS Collaboration, *Top working group cross-section summary plots June 2023*, ATL-PHYS-PUB-2023-014 . (cit. on p. 54, 55, and 64)

[159] J. Campbell, T. Neumann and Z. Sullivan, *Single-top-quark production in the t-channel at NNLO*, *JHEP* **02** (2021) 040 [arXiv:2012.01574]. (cit. on p. 55 and 67)

[160] N. Kidonakis, *NNLL resummation for s-channel single top quark production*, *Phys. Rev. D* **81** (2010) 054028 [arXiv:1001.5034]. (cit. on p. 55 and 76)

[161] N. Kidonakis and N. Yamanaka, *Higher-order corrections for tW production at high-energy hadron colliders*, *JHEP* **05** (2021) 278 [arXiv:2102.11300]. (cit. on p. 55, 76, and 94)

[162] K. Lannon, F. Margaroli and C. Neu, *Measurements of the Production, Decay and Properties of the Top Quark: A Review*, *Eur. Phys. J. C* **72** (2012) 2120 [arXiv:1201.5873]. (cit. on p. 56)

[163] ATLAS Collaboration, *Top quark mass and properties summary plots November 2022*, ATL-PHYS-PUB-2022-050 , (2022). (cit. on p. 57)

[164] A. H. Hoang, *The Top Mass: Interpretation and Theoretical Uncertainties*, in *7th International Workshop on Top Quark Physics*, 12, 2014, arXiv:1412.3649. (cit. on p. 57)

[165] S. Alekhin, A. Djouadi and S. Moch, *The top quark and Higgs boson masses and the stability of the electroweak vacuum*, *Phys. Lett. B* **716** (2012) 214 [arXiv:1207.0980]. (cit. on p. 57)

[166] J. Haller, A. Hoecker, R. Kogler, K. Mönig, T. Peiffer and J. Stelzer, *Update of the global electroweak fit and constraints on two-Higgs-doublet models*, *Eur. Phys. J. C* **78** (2018) 675 [arXiv:1803.01853]. (cit. on p. 57)

[167] Gfitter Group Collaboration, *The global electroweak fit at NNLO and prospects for the LHC and ILC*, *Eur. Phys. J. C* **74** (2014) 3046 [arXiv:1407.3792]. (cit. on p. 57)

[168] ATLAS, CMS Collaboration, *Combination of the W boson polarization measurements in top quark decays using ATLAS and CMS data at $\sqrt{s} = 8$ TeV*, *JHEP* **08** (2020) 051 [arXiv:2005.03799]. (cit. on p. 58, 76, and 163)

[169] ATLAS Collaboration, *Measurement of the polarisation of W bosons produced in top-quark decays using dilepton events at s=13 TeV with the ATLAS experiment*, *Phys. Lett. B* **843** (2023) 137829 [arXiv:2209.14903]. (cit. on p. 58)

[170] A. Czarnecki, J. G. Korner and J. H. Piclum, *Helicity fractions of W bosons from top quark decays at NNLO in QCD*, *Phys. Rev. D* **81** (2010) 111503 [arXiv:1005.2625]. (cit. on p. 58, 59, and 76)

[171] J. A. Aguilar-Saavedra and J. Bernabeu, *W polarisation beyond helicity fractions in top quark decays*, *Nucl. Phys. B* **840** (2010) 349 [arXiv:1005.5382]. (cit. on p. 58 and 60)

[172] J. A. Aguilar-Saavedra, J. Carvalho, N. F. Castro, F. Veloso and A. Onofre, *Probing anomalous Wtb couplings in top pair decays*, *Eur. Phys. J. C* **50** (2007) 519 [arXiv:hep-ph/0605190]. (cit. on p. 58)

[173] W. Bernreuther, D. Heisler and Z.-G. Si, *A set of top quark spin correlation and polarization observables for the LHC: Standard Model predictions and new physics contributions*, *JHEP* **12** (2015) 026 [arXiv:1508.05271]. (cit. on p. 58)

[174] G. Mahlon and S. J. Parke, *Spin Correlation Effects in Top Quark Pair Production at the LHC*, *Phys. Rev. D* **81** (2010) 074024 [arXiv:1001.3422]. (cit. on p. 58)

[175] A. Brandenburg, Z. G. Si and P. Uwer, *QCD corrected spin analyzing power of jets in decays of polarised top quarks*, *Phys. Lett. B* **539** (2002) 235 [arXiv:hep-ph/0205023]. (cit. on p. 59)

[176] CMS Collaboration, *Measurement of the top quark polarization and* $\mathrm{t\bar{t}}$ *spin correlations using dilepton final states in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, *Phys. Rev. D* **100** (2019) 072002 [arXiv:1907.03729]. (cit. on p. 59)

[177] ATLAS Collaboration, *Measurements of top-quark pair spin correlations in the* $e\mu$ *channel at* $\sqrt{s} = 13$ *TeV using pp collisions in the ATLAS detector*, *Eur. Phys. J. C* **80** (2020) 754 [arXiv:1903.07570]. (cit. on p. 59)

[178] G. Mahlon and S. J. Parke, *Improved spin basis for angular correlation studies in single top quark production at the Tevatron*, *Phys. Rev. D* **55** (1997) 7249 [arXiv:hep-ph/9611367]. (cit. on p. 59)

[179] G. Mahlon and S. J. Parke, *Single top quark production at the LHC: Understanding spin*, *Phys. Lett. B* **476** (2000) 323 [arXiv:hep-ph/9912458]. (cit. on p. 59)

[180] ATLAS Collaboration, *Measurement of the polarisation of single top quarks and antiquarks produced in the t-channel at* $\sqrt{s} = 13$ *TeV and bounds on the tWb dipole operator from the ATLAS experiment*, *JHEP* **11** (2022) 040 [arXiv:2202.11382]. (cit. on p. 60)

[181] E. L. Berger, J. Gao and H. X. Zhu, *Differential Distributions for t-channel Single Top-Quark Production and Decay at Next-to-Next-to-Leading Order in QCD*, *JHEP* **11** (2017) 158 [arXiv:1708.09405]. (cit. on p. 60)

[182] CDF, D0 Collaboration, *Combined Forward-Backward Asymmetry Measurements in Top-Antitop Quark Production at the Tevatron*, *Phys. Rev. Lett.* **120** (2018) 042001 [arXiv:1709.04894]. (cit. on p. 60)

[183] J. H. Kuhn and G. Rodrigo, *Charge asymmetry of heavy quarks at hadron colliders*, *Phys. Rev. D* **59** (1999) 054017 [arXiv:hep-ph/9807420]. (cit. on p. 61)

[184] J. H. Kuhn and G. Rodrigo, *Charge asymmetries of top quarks at hadron colliders revisited*, *JHEP* **01** (2012) 063 [arXiv:1109.6830]. (cit. on p. 61)

[185] W. Bernreuther and Z.-G. Si, *Top quark and leptonic charge asymmetries for the Tevatron and LHC*, *Phys. Rev. D* **86** (2012) 034026 [arXiv:1205.6580]. (cit. on p. 61 and 62)

[186] J. A. Aguilar-Saavedra, D. Amidei, A. Juste and M. Perez-Victoria, *Asymmetries in top quark pair production at hadron colliders*, *Rev. Mod. Phys.* **87** (2015) 421 [arXiv:1406.1798]. (cit. on p. 61)

[187] ATLAS Collaboration, *Evidence for the charge asymmetry in* $pp \to t\bar{t}$ *production at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, arXiv:2208.12095. (cit. on p. 61, 62, 85, 88, and 147)

[188] CMS Collaboration, *Measurement of the* $\mathrm{t\bar{t}}$ *charge asymmetry in events with highly Lorentz-boosted top quarks in pp collisions at* $\sqrt{s} = 13$ *TeV*, arXiv:2208.02751. (cit. on p. 61 and 62)

[189] M. Czakon, D. Heymes, A. Mitov, D. Pagani, I. Tsinikos and M. Zaro, *Top-quark charge asymmetry at the LHC and Tevatron through NNLO QCD and NLO EW*, *Phys. Rev. D* **98** (2018) 014003 [arXiv:1711.03945]. (cit. on p. 61)

[190] ATLAS, CMS Collaboration, *Combination of inclusive and differential* $\mathrm{t\bar{t}}$ *charge asymmetry measurements using ATLAS and CMS data at* $\sqrt{s} = 7$ *and 8 TeV*, *JHEP* **04** (2018) 033 [arXiv:1709.05327]. (cit. on p. 62)

[191] J. A. Aguilar-Saavedra, A. Juste and F. Rubbo, *Boosting the $t\bar{t}$ charge asymmetry*, *Phys. Lett. B* **707** (2012) 92 [arXiv:1109.3710]. (cit. on p. 62)

[192] CMS Collaboration, *Measurements of the $t\bar{t}$ charge asymmetry using the dilepton decay channel in pp collisions at $\sqrt{s} = 7$ TeV*, *JHEP* **04** (2014) 191 [arXiv:1402.3803]. (cit. on p. 62)

[193] ATLAS Collaboration, *Measurement of the charge asymmetry in dileptonic decays of top quark pairs in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, *JHEP* **05** (2015) 061 [arXiv:1501.07383]. (cit. on p. 62)

[194] C. Degrande, F. Maltoni, K. Mimasu, E. Vryonidou and C. Zhang, *Single-top associated production with a Z or H boson at the LHC: the SMEFT interpretation*, *JHEP* **10** (2018) 005 [arXiv:1804.07773]. (cit. on p. 63)

[195] ATLAS Collaboration, *Search for new phenomena in events with same-charge leptons and b-jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **12** (2018) 039 [arXiv:1807.11883]. (cit. on p. 66 and 69)

[196] ATLAS Collaboration, *Search for squarks and gluinos in final states with same-sign leptons and jets using 139 fb^{-1} of data collected with the ATLAS detector*, *JHEP* **06** (2020) 046 [arXiv:1909.08457]. (cit. on p. 66 and 69)

[197] CMS Collaboration, *Measurement of top quark pair production in association with a Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **03** (2020) 056 [arXiv:1907.11270]. (cit. on p. 66)

[198] ATLAS Collaboration, *Measurements of the inclusive and differential production cross sections of a top-quark–antiquark pair in association with a Z boson at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **81** (2021) 737 [arXiv:2103.12603]. (cit. on p. 66, 76, 81, and 163)

[199] A. Broggio, A. Ferroglia, R. Frederix, D. Pagani, B. D. Pecjak and I. Tsinikos, *Top-quark pair hadroproduction in association with a heavy boson at NLO+NNLL including EW corrections*, *JHEP* **08** (2019) 039 [arXiv:1907.04343]. (cit. on p. 66, 68, and 76)

[200] A. Kulesza, L. Motyka, D. Schwartländer, T. Stebel and V. Theeuwes, *Associated top quark pair production with a heavy boson: differential cross sections at NLO+NNLL accuracy*, *Eur. Phys. J. C* **80** (2020) 428 [arXiv:2001.03031]. (cit. on p. 66)

[201] ATLAS Collaboration, *Observation of the associated production of a top quark and a Z boson in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **07** (2020) 124 [arXiv:2002.07546]. (cit. on p. 66 and 107)

[202] CMS Collaboration, *Inclusive and differential cross section measurements of single top quark production in association with a Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **02** (2022) 107 [arXiv:2111.02860]. (cit. on p. 66)

[203] S. Frixione, E. Laenen, P. Motylinski, B. R. Webber and C. D. White, *Single-top hadroproduction in association with a W boson*, *JHEP* **07** (2008) 029 [arXiv:0805.3067]. (cit. on p. 67)

[204] ATLAS Collaboration, *Measurement of differential cross-sections of a single top quark produced in association with a W boson at $\sqrt{s} = 13$ TeV with ATLAS*, *Eur. Phys. J. C* **78** (2018) 186 [arXiv:1712.01602]. (cit. on p. 67)

[205] CMS Collaboration, *Measurement of inclusive and differential cross sections for single top quark production in association with a W boson in proton-proton collisions at $\sqrt{s} = 13$ TeV*, arXiv:2208.00924. (cit. on p. 67)

[206] M. Fael and T. Gehrmann, *Probing top quark electromagnetic dipole moments in single-top-plus-photon production*, *Phys. Rev. D* **88** (2013) 033003 [arXiv:1307.1349]. (cit. on p. 67)

[207] S. M. Etesami, S. Khatibi and M. Mohammadi Najafabadi, *Measuring anomalous WW γ and t $\bar{t}\gamma$ couplings using top+ γ production at the LHC*, *Eur. Phys. J. C* **76** (2016) 533 [arXiv:1606.02178]. (cit. on p. 67)

[208] A. O. Bouzas and F. Larios, *Electromagnetic dipole moments of the Top quark*, *Phys. Rev. D* **87** (2013) 074015 [arXiv:1212.6575]. (cit. on p. 67)

[209] M. Schulze and Y. Soreq, *Pinning down electroweak dipole operators of the top quark*, *Eur. Phys. J. C* **76** (2016) 466 [arXiv:1603.08911]. (cit. on p. 67)

[210] J. A. Aguilar-Saavedra, E. Álvarez, A. Juste and F. Rubbo, *Shedding light on the $t\bar{t}$ asymmetry: the photon handle*, *JHEP* **04** (2014) 188 [arXiv:1402.3598]. (cit. on p. 67)

[211] ATLAS Collaboration, *Measurements of inclusive and differential fiducial cross-sections of $t\bar{t}\gamma$ production in leptonic final states at $\sqrt{s}$ = 13 TeV in ATLAS*, *Eur. Phys. J. C* **79** (2019) 382 [arXiv:1812.01697]. (cit. on p. 67)

[212] ATLAS Collaboration, *Measurements of inclusive and differential cross-sections of combined $t\bar{t}\gamma$ and $tW\gamma$ production in the eμ channel at 13 TeV with the ATLAS detector*, *JHEP* **09** (2020) 049 [arXiv:2007.06946]. (cit. on p. 67, 75, 76, and 163)

[213] CMS Collaboration, *Measurement of the inclusive and differential $t\bar{t}\gamma$ cross sections in the single-lepton channel and EFT interpretation at $\sqrt{s}$ = 13 TeV*, *JHEP* **12** (2021) 180 [arXiv:2107.01508]. (cit. on p. 67)

[214] CMS Collaboration, *Measurement of the inclusive and differential $\mathrm{t\bar{t}}\gamma$ cross sections in the dilepton channel and effective field theory interpretation in proton-proton collisions at $\sqrt{s}$ =13 TeV*, *JHEP* **05** (2022) 091 [arXiv:2201.07301]. (cit. on p. 67)

[215] G. Bevilacqua, H. B. Hartanto, M. Kraus, T. Weber and M. Worek, *Precise predictions for $t\bar{t}\gamma/t\bar{t}$ cross section ratios at the LHC*, *JHEP* **01** (2019) 188 [arXiv:1809.08562]. (cit. on p. 67 and 76)

[216] G. Bevilacqua, H. B. Hartanto, M. Kraus, T. Weber and M. Worek, *Hard Photons in Hadroproduction of Top Quarks with Realistic Final States*, *JHEP* **10** (2018) 158 [arXiv:1803.09916]. (cit. on p. 67 and 76)

[217] ATLAS Collaboration, *Observation of single-top-quark production in association with a photon using the ATLAS detector*, arXiv:2302.01283. (cit. on p. 67)

[218] CMS Collaboration, *Evidence for the associated production of a single top quark and a photon in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, *Phys. Rev. Lett.* **121** (2018) 221802 [arXiv:1808.02913]. (cit. on p. 67 and 76)

[219] ATLAS Collaboration, *CP Properties of Higgs Boson Interactions with Top Quarks in the $t\bar{t}H$ and tH Processes Using $H \to \gamma\gamma$ with the ATLAS Detector*, *Phys. Rev. Lett.* **125** (2020) 061802 [arXiv:2004.04545]. (cit. on p. 68, 153, 161, and 221)

[220] CMS Collaboration, *Measurements of $\mathrm{t\bar{t}}H$ Production and the CP Structure of the Yukawa Interaction between the Higgs Boson and Top Quark in the Diphoton Decay Channel*, *Phys. Rev. Lett.* **125** (2020) 061801 [arXiv:2003.10866]. (cit. on p. 68 and 156)

[221] CMS Collaboration, *Analysis of the CP structure of the Yukawa coupling between the Higgs boson and τ leptons in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, *JHEP* **06** (2022) 012 [arXiv:2110.04836]. (cit. on p. 68)

[222] M. Miralles López, *Study of the $\mathcal{CP}$ properties of the top-quark Yukawa Interaction in $t\bar{t}H$ and tH events with $H \to \gamma\gamma$*, *PoS* **EPS-HEP2021** (2022) 570. (cit. on p. 68)

[223] CMS Collaboration, *Observation of $\mathrm{t\bar{t}}H$ production*, *Phys. Rev. Lett.* **120** (2018) 231801 [arXiv:1804.02610]. (cit. on p. 68)

[224] ATLAS Collaboration, *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, *Phys. Lett. B* **784** (2018) 173 [arXiv:1806.00425]. (cit. on p. 68 and 154)

[225] S. Biswas, E. Gabrielli and B. Mele, *Single top and Higgs associated production as a probe of the Htt coupling sign at the LHC*, *JHEP* **01** (2013) 088 [arXiv:1211.0499]. (cit. on p. 68)

[226] M. Farina, C. Grojean, F. Maltoni, E. Salvioni and A. Thamm, *Lifting degeneracies in Higgs couplings using single top production in association with a Higgs boson*, *JHEP* **05** (2013) 022 [arXiv:1211.3736]. (cit. on p. 68)

[227] LHC Higgs Cross Section Working Group Collaboration, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, arXiv:1610.07922. (cit. on p. 68, 76, 93, 94, 95, 107, 151, and 170)

[228] F. Demartin, F. Maltoni, K. Mawatari and M. Zaro, *Higgs production in association with a single top quark at the LHC*, *Eur. Phys. J. C* **75** (2015) 267 [arXiv:1504.00611]. (cit. on p. 68)

[229] CMS Collaboration, *Higgs boson production in association with top quarks in final states with electrons, muons, and hadronically decaying tau leptons at* $\sqrt{s} = 13$ TeV, CMS-PAS-HIG-19-008 , (2020). (cit. on p. 68)

[230] C. E. Gerber and C. Vellidis, *Review of Physics Results from the Tevatron: Top Quark Physics*, arXiv:1409.5038. (cit. on p. 68)

[231] I. Brivio, S. Bruggisser, F. Maltoni, R. Moutafis, T. Plehn, E. Vryonidou et al., *O new physics, where art thou? A global search in the top sector*, *JHEP* **02** (2020) 131 [arXiv:1910.03606]. (cit. on p. 68, 72, 73, 75, and 89)

[232] CMS Collaboration, *Search for production of four top quarks in final states with same-sign or multiple leptons in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, *Eur. Phys. J. C* **80** (2020) 75 [arXiv:1908.06463]. (cit. on p. 68)

[233] ATLAS Collaboration, *Observation of four-top-quark production in the multilepton final state with the ATLAS detector*, *Eur. Phys. J. C* **83** (2023) 496 [arXiv:2303.15061]. (cit. on p. 68, 74, and 107)

[234] ATLAS Collaboration, *Measurement of the total and differential cross-sections of* $t\bar{t}W$ *production in pp collisions at 13 TeV with the ATLAS detector*, ATLAS-CONF-2023-019 , (2023). (cit. on p. 69, 95, 131, 139, 147, 148, and 172)

[235] CMS Collaboration, *Measurement of the cross section of top quark-antiquark pair production in association with a W boson in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, arXiv:2208.06485. (cit. on p. 69, 131, and 172)

[236] R. Frederix and I. Tsinikos, *On improving NLO merging for* $\mathrm{t\bar{t}W}$ *production*, *JHEP* **11** (2021) 029 [arXiv:2108.07826]. (cit. on p. 69, 70, 109, 131, and 179)

[237] R. Frederix, D. Pagani and M. Zaro, *Large NLO corrections in* $t\bar{t}W^{\pm}$ *and* $t\bar{t}t\bar{t}$ *hadroproduction from supposedly subleading EW contributions*, *JHEP* **02** (2018) 031 [arXiv:1711.02116]. (cit. on p. 69, 76, 170, and 178)

[238] R. Frederix and I. Tsinikos, *Subleading EW corrections and spin-correlation effects in* $t\bar{t}W$ *multi-lepton signatures*, *Eur. Phys. J. C* **80** (2020) 803 [arXiv:2004.09552]. (cit. on p. 69, 170, and 178)

[239] G. Bevilacqua, H.-Y. Bi, H. B. Hartanto, M. Kraus, J. Nasufi and M. Worek, *NLO QCD corrections to off-shell* $t\bar{t}W^{\pm}$ *production at the LHC: correlations and asymmetries*, *Eur. Phys. J. C* **81** (2021) 675 [arXiv:2012.01363]. (cit. on p. 69, 88, 91, and 146)

[240] S. von Buddenbrock, R. Ruiz and B. Mellado, *Anatomy of inclusive $t\bar{t}W$ production at hadron colliders*, *Phys. Lett. B* **811** (2020) 135964 [arXiv:2009.00032]. (cit. on p. 70, 173, and 179)

[241] ATLAS Collaboration, *Modelling of rare top quark processes at $\sqrt{s}$ = 13 TeV in ATLAS*, ATL-PHYS-PUB-2020-024 , (2020). (cit. on p. 70, 161, 172, and 178)

[242] J. A. Aguilar-Saavedra, *A Minimal set of top anomalous couplings*, *Nucl. Phys. B* **812** (2009) 181 [arXiv:0811.3842]. (cit. on p. 70)

[243] C. Zhang and S. Willenbrock, *Effective-Field-Theory Approach to Top-Quark Production and Decay*, *Phys. Rev. D* **83** (2011) 034006 [arXiv:1008.3869]. (cit. on p. 70, 72, and 150)

[244] D. Barducci et al., *Interpreting top-quark LHC measurements in the standard-model effective field theory*, arXiv:1802.07237. (cit. on p. 70, 71, and 86)

[245] C. Degrande, J.-M. Gerard, C. Grojean, F. Maltoni and G. Servant, *Non-resonant New Physics in Top Pair Production at Hadron Colliders*, *JHEP* **03** (2011) 125 [arXiv:1010.6304]. (cit. on p. 72)

[246] V. Miralles, M. M. López, M. M. Llácer, A. Peñuelas, M. Perelló and M. Vos, *The top quark electro-weak couplings after LHC Run 2*, *JHEP* **02** (2022) 032 [arXiv:2107.13917]. (cit. on p. 74, 160, and 206)

[247] CMS Collaboration, *Search for new physics in top quark production with additional leptons in proton-proton collisions at $\sqrt{s}$ = 13 TeV using effective field theory*, *JHEP* **03** (2021) 095 [arXiv:2012.04120]. (cit. on p. 74 and 89)

[248] S. Bruggisser, R. Schäfer, D. van Dyk and S. Westhoff, *The Flavor of UV Physics*, *JHEP* **05** (2021) 257 [arXiv:2101.07273]. (cit. on p. 74 and 87)

[249] S. Bißmann, C. Grunwald, G. Hiller and K. Kröninger, *Top and Beauty synergies in SMEFT-fits at present and future colliders*, *JHEP* **06** (2021) 010 [arXiv:2012.10456]. (cit. on p. 74 and 75)

[250] G. Durieux, A. Irles, V. Miralles, A. Peñuelas, R. Pöschl, M. Perelló et al., *The electroweak couplings of the top and bottom quarks — Global fit and future prospects*, *JHEP* **12** (2019) 98 [arXiv:1907.10619]. (cit. on p. 74, 77, and 84)

[251] G. Durieux, A. G. Camacho, L. Mantani, V. Miralles, M. M. López, M. Llácer Moreno et al., *Snowmass White Paper: prospects for the measurement of top-quark couplings*, in *Snowmass 2021*, 5, 2022, arXiv:2205.02140. (cit. on p. 74 and 90)

[252] ALEPH, DELPHI, L3, OPAL, SLD, LEP Electroweak Working Group, SLD Electroweak Group, SLD Heavy Flavour Group Collaboration, *Precision electroweak measurements on the Z resonance*, *Phys. Rept.* **427** (2006) 257 [arXiv:hep-ex/0509008]. (cit. on p. 75 and 76)

[253] A. Buckley, C. Englert, J. Ferrando, D. J. Miller, L. Moore, M. Russell et al., *Global fit of top quark effective theory to data*, *Phys. Rev. D* **92** (2015) 091501 [arXiv:1506.08845]. (cit. on p. 75)

[254] N. P. Hartland, F. Maltoni, E. R. Nocera, J. Rojo, E. Slade, E. Vryonidou et al., *A Monte Carlo global analysis of the Standard Model Effective Field Theory: the top quark sector*, *JHEP* **04** (2019) 100 [arXiv:1901.05965]. (cit. on p. 75 and 89)

[255] G. Bevilacqua, H. B. Hartanto, M. Kraus, T. Weber and M. Worek, *Off-shell vs on-shell modelling of top quarks in photon associated production*, *JHEP* **03** (2020) 154 [arXiv:1912.09999]. (cit. on p. 75)

[256] ATLAS Collaboration, *A combination of measurements of Higgs boson production and decay using up to* 139 fb^{-1} *of proton–proton collision data at* $\sqrt{s}$ *= 13 TeV collected with the ATLAS experiment*, ATLAS-CONF-2020-027 , (2020). (cit. on p. 76)

[257] CMS Collaboration, *Measurement of the associated production of a single top quark and a Z boson in pp collisions at* $\sqrt{s}$ *= 13 TeV*, *Phys. Lett. B* **779** (2018) 358 [arXiv:1712.02825]. (cit. on p. 76)

[258] CMS Collaboration, *Observation of Single Top Quark Production in Association with a Z Boson in Proton-Proton Collisions at* $\sqrt{s}$ *=13 TeV*, *Phys. Rev. Lett.* **122** (2019) 132003 [arXiv:1812.05900]. (cit. on p. 76)

[259] CMS Collaboration, *Measurement of the cross section for top quark pair production in association with a W or Z boson in proton-proton collisions at* $\sqrt{s}$ *= 13 TeV*, *JHEP* **08** (2018) 011 [arXiv:1711.02547]. (cit. on p. 76)

[260] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer and M. Wiedermann, *HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR*, *Comput. Phys. Commun.* **182** (2011) 1034 [arXiv:1007.1327]. (cit. on p. 76)

[261] P. Kant, O. M. Kind, T. Kintscher, T. Lohse, T. Martini, S. Mölbitz et al., *HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions*, *Comput. Phys. Commun.* **191** (2015) 74 [arXiv:1406.4403]. (cit. on p. 76)

[262] ATLAS, CMS Collaboration, *Combinations of single-top-quark production cross-section measurements and* $|f_{LV}V_{tb}|$ *determinations at* $\sqrt{s}$ *= 7 and 8 TeV with the ATLAS and CMS experiments*, *JHEP* **05** (2019) 088 [arXiv:1902.07158]. (cit. on p. 76)

[263] CDF, D0 Collaboration, *Observation of s-channel production of single top quarks at the Tevatron*, *Phys. Rev. Lett.* **112** (2014) 231803 [arXiv:1402.5126]. (cit. on p. 76)

[264] C. Zhang, *Effective field theory approach to top-quark decay at next-to-leading order in QCD*, *Phys. Rev. D* **90** (2014) 014008 [arXiv:1404.1264]. (cit. on p. 76)

[265] 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 [arXiv:2008.11743]. (cit. on p. 77 and 207)

[266] O. Bessidskaia Bylund, F. Maltoni, I. Tsinikos, E. Vryonidou and C. Zhang, *Probing top quark neutral couplings in the Standard Model Effective Field Theory at NLO in QCD*, *JHEP* **05** (2016) 052 [arXiv:1601.08193]. (cit. on p. 77, 81, and 86)

[267] J. De Blas et al., ***HEPfit:*** *a code for the combination of indirect and direct constraints on high energy physics models*, *Eur. Phys. J. C* **80** (2020) 456 [arXiv:1910.14012]. (cit. on p. 77 and 207)

[268] HEPfit Collaboration, HEPfit*, a tool to combine indirect and direct constraints on high-energy physics*, http://hepfit.roma1.infn.it. (cit. on p. 77 and 207)

[269] J. de Blas, M. Ciuchini, E. Franco, S. Mishima, M. Pierini, L. Reina et al., *Electroweak precision observables and Higgs-boson signal strengths in the Standard Model and beyond: present and future*, *JHEP* **12** (2016) 135 [arXiv:1608.01509]. (cit. on p. 77)

[270] J. de Blas, O. Eberhardt and C. Krause, *Current and Future Constraints on Higgs Couplings in the Nonlinear Effective Theory*, *JHEP* **07** (2018) 048 [arXiv:1803.00939]. (cit. on p. 77)

[271] O. Eberhardt, A. P. n. Martínez and A. Pich, *Global fits in the Aligned Two-Higgs-Doublet model*, *JHEP* **05** (2021) 005 [arXiv:2012.09200]. (cit. on p. 77)

[272] O. Eberhardt, V. Miralles and A. Pich, *Constraints on coloured scalars from global fits*, *JHEP* **10** (2021) 123 [arXiv:2106.12235]. (cit. on p. 77)

[273] A. Caldwell, D. Kollar and K. Kroninger, *BAT: The Bayesian Analysis Toolkit*, *Comput. Phys. Commun.* **180** (2009) 2197 [arXiv:0808.2552]. (cit. on p. 77)

[274] E. Vryonidou and C. Zhang, *Dimension-six electroweak top-loop effects in Higgs production and decay*, *JHEP* **08** (2018) 036 [arXiv:1804.09766]. (cit. on p. 80)

[275] G. Durieux, J. Gu, E. Vryonidou and C. Zhang, *Probing top-quark couplings indirectly at Higgs factories*, *Chin. Phys. C* **42** (2018) 123107 [arXiv:1809.03520]. (cit. on p. 80)

[276] S. Jung, J. Lee, M. Perelló, J. Tian and M. Vos, *Higgs, top quark, and electroweak precision measurements at future e+e- colliders: A combined effective field theory analysis with renormalization mixing*, *Phys. Rev. D* **105** (2022) 016003 [arXiv:2006.14631]. (cit. on p. 80)

[277] 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 [arXiv:2012.02779]. (cit. on p. 80 and 87)

[278] SMEFiT Collaboration, *Combined SMEFT interpretation of Higgs, diboson, and top quark data from the LHC*, *JHEP* **11** (2021) 089 [arXiv:2105.00006]. (cit. on p. 80, 87, and 89)

[279] CDF, D0 Collaboration, *Combination of Measurements of the Top-Quark Pair Production Cross Section from the Tevatron Collider*, *Phys. Rev. D* **89** (2014) 072001 [arXiv:1309.7570]. (cit. on p. 85)

[280] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section and lepton differential distributions in $e\mu$ dilepton events from pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **80** (2020) 528 [arXiv:1910.08819]. (cit. on p. 85)

[281] M. P. Rosello and M. Vos, *Constraints on four-fermion interactions from the $t\bar{t}$ charge asymmetry at hadron colliders*, *Eur. Phys. J. C* **76** (2016) 200 [arXiv:1512.07542]. (cit. on p. 85)

[282] A. Basan, P. Berta, L. Masetti, E. Vryonidou and S. Westhoff, *Measuring the top energy asymmetry at the LHC: QCD and SMEFT interpretations*, *JHEP* **03** (2020) 184 [arXiv:2001.07225]. (cit. on p. 85)

[283] ATLAS Collaboration, *Measurement of the energy asymmetry in $t\bar{t}j$ production at 13 TeV with the ATLAS experiment and interpretation in the SMEFT framework*, *Eur. Phys. J. C* **82** (2022) 374 [arXiv:2110.05453]. (cit. on p. 85, 88, and 147)

[284] S. Bißmann, J. Erdmann, C. Grunwald, G. Hiller and K. Kröninger, *Constraining top-quark couplings combining top-quark and $\boldsymbol{B}$ decay observables*, *Eur. Phys. J. C* **80** (2020) 136 [arXiv:1909.13632]. (cit. on p. 87)

[285] F. Maltoni, M. L. Mangano, I. Tsinikos and M. Zaro, *Top-quark charge asymmetry and polarization in $t\bar{t}W^{\pm}$ production at the LHC*, *Phys. Lett. B* **736** (2014) 252 [arXiv:1406.3262]. (cit. on p. 88 and 91)

[286] European Strategy Group Collaboration, *2020 Update of the European Strategy for Particle Physics*. CERN Council, Geneva, 2020, 10.17181/ESU2020. (cit. on p. 88)

[287] K. Agashe et al., *Report of the Topical Group on Top quark physics and heavy flavor production for Snowmass 2021*, arXiv:2209.11267. (cit. on p. 88)

[288] P. Azzi et al., *Report from Working Group 1: Standard Model Physics at the HL-LHC and HE-LHC*, *CERN Yellow Rep. Monogr.* **7** (2019) 1 [arXiv:1902.04070]. (cit. on p. 88)

[289] M. S. Amjad et al., *A precise characterisation of the top quark electro-weak vertices at the ILC*, *Eur. Phys. J. C* **75** (2015) 512 [arXiv:1505.06020]. (cit. on p. 88)

[290] 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 [arXiv:1807.02121]. (cit. on p. 88 and 90)

[291] M. Cepeda et al., *Report from Working Group 2: Higgs Physics at the HL-LHC and HE-LHC*, *CERN Yellow Rep. Monogr.* **7** (2019) 221 [arXiv:1902.00134]. (cit. on p. 88)

[292] CMS Collaboration, *Measurement of differential $t\bar{t}$ production cross sections in the full kinematic range using lepton+jets events from proton-proton collisions at $\sqrt{s}$ = 13 TeV*, *Phys. Rev. D* **104** (2021) 092013 [arXiv:2108.02803]. (cit. on p. 89)

[293] ATLAS Collaboration, *Inclusive and differential measurement of the charge asymmetry in $t\bar{t}$ events at 13 TeV with the ATLAS detector*, ATLAS-CONF-2019-026 , (2019). (cit. on p. 89)

[294] G. Durieux, F. Maltoni and C. Zhang, *Global approach to top-quark flavor-changing interactions*, *Phys. Rev. D* **91** (2015) 074017 [arXiv:1412.7166]. (cit. on p. 89)

[295] M. Chala, J. Santiago and M. Spannowsky, *Constraining four-fermion operators using rare top decays*, *JHEP* **04** (2019) 014 [arXiv:1809.09624]. (cit. on p. 89)

[296] ILC International Development Team Collaboration, *The International Linear Collider: Report to Snowmass 2021*, arXiv:2203.07622. (cit. on p. 90)

[297] Y. Okugawa, A. Irles, V. Lohezic, S. Amjad, R. Yonamine, F. Richard et al., *Production and electroweak couplings of 3rd generation quarks at the ILC*, *PoS* **LeptonPhoton2019** (2019) 170. (cit. on p. 90)

[298] M. S. Amjad, M. Boronat, T. Frisson, I. Garcia, R. Poschl, E. Ros et al., *A precise determination of top quark electro-weak couplings at the ILC operating at $\sqrt{s}$ = 500 GeV*, arXiv:1307.8102. (cit. on p. 90)

[299] A. Robson and P. Roloff, *Updated CLIC luminosity staging baseline and Higgs coupling prospects*, arXiv:1812.01644. (cit. on p. 90)

[300] H. Abramowicz et al., *Higgs physics at the CLIC electron–positron linear collider*, *Eur. Phys. J. C* **77** (2017) 475 [arXiv:1608.07538]. (cit. on p. 90)

[301] G. Bernardi et al., *The Future Circular Collider: a Summary for the US 2021 Snowmass Process*, arXiv:2203.06520. (cit. on p. 90)

[302] ATLAS Collaboration, *Search for leptonic charge asymmetry in $t\bar{t}W$ production in final states with three leptons at $\sqrt{s}$ = 13 TeV*, *JHEP* **07** (2023) 033 [arXiv:2301.04245]. (cit. on p. 91, 160, and 212)

[303] ATLAS Collaboration, *Search for leptonic charge asymmetry in $t\bar{t}W$ production in final states with three leptons at $\sqrt{s}$ = 13 TeV*, ATL-COM-PHYS-2021-1042 , (2022). (cit. on p. 91)

[304] ATLAS Collaboration, *ATLAS data quality operations and performance for 2015–2018 data-taking*, *JINST* **15** (2020) P04003 [arXiv:1911.04632]. (cit. on p. 92)

[305] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, *JINST* **13** (2018) P07017. (cit. on p. 92 and 106)

[306] T. Sjostrand, S. Mrenna and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852 [arXiv:0710.3820]. (cit. on p. 92)

[307] R. D. Ball et al., *Parton distributions with LHC data*, *Nucl. Phys. B* **867** (2013) 244 [arXiv:1207.1303]. (cit. on p. 92)

[308] ATLAS Collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie-Landshoff diffractive model*, ATL-PHYS-PUB-2016-017 , (2016). (cit. on p. 92)

[309] GEANT4 Collaboration, *GEANT4–a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250. (cit. on p. 92)

[310] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, *Eur. Phys. J. C* **70** (2010) 823 [arXiv:1005.4568]. (cit. on p. 92)

[311] ATLAS Collaboration, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, ATL-PHYS-PUB-2010-013 , (2010). (cit. on p. 92)

[312] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth. A* **462** (2001) 152. (cit. on p. 92)

[313] S. Kallweit, J. M. Lindert, P. Maierhofer, S. Pozzorini and M. Schönherr, *NLO QCD+EW predictions for V + jets including off-shell vector-boson decays and multijet merging*, *JHEP* **04** (2016) 021 [arXiv:1511.08692]. (cit. on p. 92)

[314] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021 , (2014). (cit. on p. 93)

[315] S. Frixione, E. Laenen, P. Motylinski and B. R. Webber, *Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations*, *JHEP* **04** (2007) 081 [arXiv:hep-ph/0702198]. (cit. on p. 93 and 171)

[316] P. Artoisenet, R. Frederix, O. Mattelaer and R. Rietkerk, *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, *JHEP* **03** (2013) 015 [arXiv:1212.3460]. (cit. on p. 93 and 171)

[317] ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling for Top2016*, , (2016). (cit. on p. 94)

[318] F. Demartin, B. Maier, F. Maltoni, K. Mawatari and M. Zaro, *tWH associated production at the LHC*, *Eur. Phys. J. C* **77** (2017) 34 [arXiv:1607.05862]. (cit. on p. 94)

[319] ATLAS Collaboration, *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, *JHEP* **09** (2014) 145 [arXiv:1406.3660]. (cit. on p. 94)

[320] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A Precision tool for QED corrections in Z and W decays*, *Eur. Phys. J. C* **45** (2006) 97 [arXiv:hep-ph/0506026]. (cit. on p. 94)

[321] K. Melnikov and F. Petriello, *Electroweak gauge boson production at hadron colliders through $O(\alpha_s^2)$*, *Phys. Rev. D* **74** (2006) 114017 [arXiv:hep-ph/0609070]. (cit. on p. 94)

[322] ATLAS Collaboration, *ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset*, arXiv:2211.16345. (cit. on p. 96)

[323] ATLAS Collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*, *Eur. Phys. J. C* **80** (2020) 47 [arXiv:1909.00761]. (cit. on p. 96)

[324] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*, *JINST* **15** (2020) P09015 [arXiv:2004.13447]. (cit. on p. 96)

[325] F. Pedregosa et al., *Scikit-learn: Machine Learning in Python*, *J. Machine Learning Res.* **12** (2011) 2825 [arXiv:1201.0490]. (cit. on p. 103)

[326] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, *J. Phys. G* **43** (2016) 023001 [arXiv:1510.03865]. (cit. on p. 107)

[327] ATLAS Collaboration, *Measurement of $W^{\pm}Z$ production cross sections and gauge boson polarisation in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **79** (2019) 535 [arXiv:1902.05759]. (cit. on p. 107)

[328] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, *eConf* **C0303241** (2003) MOLT007 [arXiv:physics/0306116]. (cit. on p. 111)

[329] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J. C* **71** (2011) 1554 [arXiv:1007.1727]. (cit. on p. 111)

[330] M. Cacciari, G. P. Salam and G. Soyez, *The Catchment Area of Jets*, *JHEP* **04** (2008) 005 [arXiv:0802.1188]. (cit. on p. 138)

[331] ATLAS Collaboration, *Measurement of the charge asymmetry in top-quark pair production in association with a photon with the ATLAS experiment*, arXiv:2212.10552. (cit. on p. 139)

[332] ATLAS Collaboration, *Inclusive and differential cross-sections for dilepton $t\bar{t}$ production measured in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, arXiv:2303.15340. (cit. on p. 147)

[333] ATLAS Collaboration, *Measurements of differential cross-sections in top-quark pair events with a high transverse momentum top quark and limits on beyond the Standard Model contributions to top-quark pair production with the ATLAS detector at $\sqrt{s} = 13$ TeV*, *JHEP* **06** (2022) 063 [arXiv:2202.12134]. (cit. on p. 147)

[334] ATLAS Collaboration, *Differential $t\bar{t}$ cross-section measurements using boosted top quarks in the all-hadronic final state with 139 fb^{-1} of ATLAS data*, arXiv:2205.02817. (cit. on p. 147)

[335] ATLAS Collaboration, *Test of CP Invariance in vector-boson fusion production of the Higgs boson using the Optimal Observable method in the ditau decay channel with the ATLAS detector*, *Eur. Phys. J. C* **76** (2016) 658 [arXiv:1602.04516]. (cit. on p. 150)

[336] CMS Collaboration, *Combined search for anomalous pseudoscalar HVV couplings in VH($H \to b\bar{b}$) production and $H \to VV$ decay*, *Phys. Lett. B* **759** (2016) 672 [arXiv:1602.04305]. (cit. on p. 150)

[337] ATLAS Collaboration, *Measurement of the Higgs boson coupling properties in the $H \to ZZ^* \to 4\ell$ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **03** (2018) 095 [arXiv:1712.02304]. (cit. on p. 150)

[338] ATLAS Collaboration, *Measurements of Higgs boson properties in the diphoton decay channel with 36 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev. D* **98** (2018) 052005 [arXiv:1802.04146]. (cit. on p. 150 and 155)

[339] CMS Collaboration, *Measurements of the Higgs boson width and anomalous HVV couplings from on-shell and off-shell production in the four-lepton final state*, *Phys. Rev. D* **99** (2019) 112003 [arXiv:1901.00174]. (cit. on p. 150)

[340] CMS Collaboration, *Constraints on anomalous HVV couplings from the production of Higgs bosons decaying to τ lepton pairs*, *Phys. Rev. D* **100** (2019) 112002 [arXiv:1903.06973]. (cit. on p. 150)

[341] J. Ellis, D. S. Hwang, K. Sakurai and M. Takeuchi, *Disentangling Higgs-Top Couplings in Associated Production*, *JHEP* **04** (2014) 004 [arXiv:1312.5736]. (cit. on p. 150)

[342] F. Boudjema, R. M. Godbole, D. Guadagnoli and K. A. Mohan, *Lab-frame observables for probing the top-Higgs interaction*, *Phys. Rev. D* **92** (2015) 015019 [arXiv:1501.03157]. (cit. on p. 150)

[343] M. R. Buckley and D. Goncalves, *Boosting the Direct CP Measurement of the Higgs-Top Coupling*, *Phys. Rev. Lett.* **116** (2016) 091801 [arXiv:1507.07926]. (cit. on p. 150)

[344] S. Amor Dos Santos et al., *Probing the CP nature of the Higgs coupling in $t\bar{t}h$ events at the LHC*, *Phys. Rev. D* **96** (2017) 013004 [arXiv:1704.03565]. (cit. on p. 150)

[345] F. Demartin, F. Maltoni, K. Mawatari, B. Page and M. Zaro, *Higgs characterisation at NLO in QCD: CP properties of the top-quark Yukawa interaction*, *Eur. Phys. J. C* **74** (2014) 3065 [arXiv:1407.5089]. (cit. on p. 150)

[346] T. Chen and C. Guestrin, *XGBoost: A Scalable Tree Boosting System*, arXiv:1603.02754. (cit. on p. 154)

[347] A. L. Read, *Presentation of search results: The CL(s) technique*, *J. Phys. G* **28** (2002) 2693. (cit. on p. 155)

[348] ATLAS Collaboration, *Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment*, *Phys. Rev. D* **101** (2020) 012002 [arXiv:1909.02845]. (cit. on p. 156)

[349] S. Frixione, V. Hirschi, D. Pagani, H. S. Shao and M. Zaro, *Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons*, *JHEP* **06** (2015) 184 [arXiv:1504.03446]. (cit. on p. 170)

[350] H. T. Li, C. S. Li and S. A. Li, *Renormalization group improved predictions for $t\bar{t}W^{\pm}$ production at hadron colliders*, *Phys. Rev. D* **90** (2014) 094009 [arXiv:1409.1460]. (cit. on p. 170)

[351] A. Broggio, A. Ferroglia, G. Ossola and B. D. Pecjak, *Associated production of a top pair and a W boson at next-to-next-to-leading logarithmic accuracy*, *JHEP* **09** (2016) 089 [arXiv:1607.05303]. (cit. on p. 170)

[352] A. Kulesza, L. Motyka, D. Schwartländer, T. Stebel and V. Theeuwes, *Associated production of a top quark pair with a heavy electroweak gauge boson at NLO+NNLL accuracy*, *Eur. Phys. J. C* **79** (2019) 249 [arXiv:1812.08622]. (cit. on p. 170)

[353] G. Bevilacqua, H.-Y. Bi, H. B. Hartanto, M. Kraus and M. Worek, *The simplest of them all: $t\bar{t}W^{\pm}$ at NLO accuracy in QCD*, *JHEP* **08** (2020) 043 [arXiv:2005.09427]. (cit. on p. 170)

[354] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, *QCD matrix elements + parton showers*, *JHEP* **11** (2001) 063 [arXiv:hep-ph/0109231]. (cit. on p. 171)