



BERGISCHE UNIVERSITÄT WUPPERTAL
FAKULTÄT FÜR
MATHEMATIK UND NATURWISSENSCHAFTEN
FACHGRUPPE PHYSIK

Search for flavour-changing
neutral current interactions
in the top-quark Higgs boson sector
in multi-lepton final states with the
ATLAS detector at the LHC at $\sqrt{s} = 13 \text{ TeV}$

Dissertation von Oliver Thielmann

Oktober 2023

Überarbeitete Version vom 23. Dezember 2023

Abstract

A search for flavour-changing neutral currents in the top-quark Higgs boson sector is presented, utilizing an integrated luminosity of $\mathcal{L}_{\text{int}} = 140 \text{ fb}^{-1}$ of proton-proton collisions recorded by the ATLAS experiment at a centre-of-mass energy of $\sqrt{s} = 13 \text{ TeV}$. As flavour-changing neutral currents are highly suppressed in the Standard Model of particle physics, these processes are sensitive to physics beyond the Standard Model. To differentiate between signal and background, advanced reconstruction methods are employed for creating kinematic distributions. Multivariate analysis techniques are utilised to optimize the differentiation between signal and background, leading to expected upper exclusion limits on the branching ratio $\mathcal{B}(t \rightarrow qH)$ in a binned profile likelihood fit including a full set of systematic uncertainties. The obtained expected exclusion limits amount to $\mathcal{B}(t \rightarrow uH) < 4.1 \cdot 10^{-4}$ and $\mathcal{B}(t \rightarrow cH) < 5.0 \cdot 10^{-4}$ for the branching ratio corresponding to the utH and ctH couplings respectively.

Contents

1. Introduction	1
2. The Standard Model of particle physics	3
2.1. Particles and interactions	3
2.2. Electroweak theory	5
2.3. Quantum chromodynamics	12
2.4. Experimental validation	14
2.5. Limitations	15
3. Flavour-changing neutral currents	19
3.1. FCNCs in the SM	19
3.2. FCNCs in Beyond-Standard-Model theories	20
3.3. FCNC tqH processes	23
4. The Large Hadron Collider and the ATLAS experiment	25
4.1. The Large Hadron Collider complex	25
4.1.1. The pre-accelerators	25
4.1.2. The Large Hadron Collider	26
4.2. The ATLAS Detector	28
4.2.1. Coordinate system	30
4.2.2. Inner detector	31
4.2.3. Calorimeters	33
4.2.4. Muon spectrometer	34
4.2.5. Trigger system	36
4.2.6. Luminosity measurement	36
5. Simulation of proton-proton collisions	39
6. Object definitions	45
6.1. Tracks and vertices	45
6.2. Electrons	47
6.3. Muons	48
6.4. Jets	50
6.5. Flavour tagging	52
6.6. Non-prompt lepton sources	55
6.7. Missing transverse momentum	58
6.8. Overlap removal	59

7. Data and simulated samples	61
7.1. Recorded data and trigger	61
7.2. Signal FCNC samples	61
7.3. Background samples	63
7.4. Building the non-prompt lepton templates	65
8. Event selection and region definition	67
8.1. Tri-lepton regions	68
8.2. Di-lepton same sign regions	73
9. Reconstruction	75
9.1. Neutrino weighting method	75
9.2. Reconstructed Z-boson mass	77
9.3. Neutrino-independent combinatorics estimator	78
9.4. Same-charged and opposite-charged leptons	81
9.5. Recursive jigsaw reconstruction	81
10. Multivariate analysis	85
10.1. Feed-forward neural networks	85
10.2. NeuroBayes	86
10.3. Applying the neural network	89
11. Statistical analysis techniques	93
11.1. Binned profile likelihood fits	93
11.2. Hypothesis testing	96
12. Systematic uncertainties	99
12.1. Experimental systematic uncertainties	99
12.2. Theoretical uncertainties	101
12.3. Implementation of systematic uncertainties	104
13. FCNC search results	107
13.1. Background-only fit	107
13.2. Realistic Asimov fit	115
13.3. Realistic Asimov fit results and exclusion limits	119
14. Conclusion and outlook	123
Appendices	126
A. b-tagging WP study	126

B. Monte Carlo Samples	128
C. Comparison of the left- and right handed prod-FCNC processes	132
D. Composition of the non-prompt templates	136
D.1. Composition in the 3ℓ regions	136
D.2. Composition in the 2ℓ SS regions	138
E. Pre-fit event yields	141
F. Plots related to the background-only fit	145
G. Validation plots of kinematic distributions	153
G.1. Validation plots for the 2ℓ SS regions	153
G.2. Validation plots for the 3ℓ regions	159
H. Plots related to the realistic Asimov Fit	165
Bibliography	173

1. Introduction

The Standard Model of particle physics offers a mathematical description of elementary particles and their interactions. This model, developed over 50 years ago, continues to provide accurate predictions for every interaction observed at present in the most powerful particle accelerator, the Large Hadron Collider. Interestingly, the Standard Model predicted the existence of elementary particles decades before their actual discovery. However impressive its achievements over the past half century, there are indications suggesting that the Standard Model is not the ultimate, most fundamental theory, but rather a “low-energy approximation”. The search for evidence of where the Standard Model breaks down is a significant aspect of the analyses carried out using the data recorded by the detectors at the Large Hadron Collider at CERN near Geneva. This thesis is one such analysis.

According to the Standard Model, only the $W^\pm$ boson is capable of altering the quark flavour, i.e. turning an up-type quark into a down-type quark and vice versa. This process is known as a flavour-changing charged current. Several extensions of the Standard Model introduce a new process, a flavour-changing neutral current, where the (electrically) neutral current is, for example, a Z boson, a photon, a gluon or a Higgs boson. Flavour-changing neutral currents in the top-quark sector are strongly suppressed in the Standard Model as they only appear in higher-order loop diagrams composed of flavour-changing charged currents. As a result, any measurement of the Standard Model’s prediction of flavour-changing neutral currents in the top-quark sector is virtually impossible with today’s experiments and technology. The discovery of such processes would revolutionise the world of particle physics. This thesis searches for a flavour-changing neutral current in the top-quark Higgs boson sector, with the Higgs boson being the neutral current.

The concept of searching for rare processes involves defining a selection of objects that is enriched in events of the rare process. In this analysis either two same-charged leptons or three charged leptons (electrons or muons) are required. Subsequently, a variable is constructed to differentiate as effectively as possible between the rare process of interest and the multitude of background processes. For this variable, the expectation (i.e. the prediction of the Standard Model) is compared to the measured data. To accomplish this, Chapter 2 provides an overview of the Standard Model, including its mathematical formulation. The flavour-changing neutral currents are then comprehensively discussed in Chapter 3. The Large Hadron Collider and the ATLAS experiment are described in Chapter 4, both of which are crucial for the acquisition of the data analysed in this thesis. In order to be able to compare the recorded data with an expectation, the simulation of proton-proton collisions is

discussed in Chapter 5 and the list of processes and samples used is given in Chapter 7. Because certain background templates require specific objects for their definition, the list of object definitions is provided beforehand in Chapter 6. Subsequently, using the defined objects, Chapter 8 defines an event selection and regions in which the rare flavour-changing neutral current process is enriched. Since many particles involved in the scattering process of proton-proton collisions decay further before being detected and recorded, a reconstruction of the events is carried out in Chapter 9. This reconstruction is important for creating variables that distinguish between rare signal events and background events. Afterwards, the discriminative power of the reconstructed variables, but also other kinematic distributions, are combined into a single variable using artificial neural networks in Chapter 10. After the description of the statistical analysis techniques in Chapter 11 and the sources of systematic uncertainties in Chapter 12, the results of this search are presented in Chapter 13, while Chapter 14 concludes this thesis with a summary.

2. The Standard Model of particle physics

The Standard Model (SM) [1, 2] of particle physics is a mathematical formulation, developed mainly in the 1960s and 1970s, which describes elementary particles and their interactions. Over the past decades, it has been tested and confirmed by many experiments and countless analyses, culminating in the discovery of the top quark in 1995 by the CDF and D0 collaborations at the Tevatron at Fermilab [3, 4] and the discovery of the Higgs boson in 2012 by the ATLAS and CMS collaborations at the LHC at CERN [5, 6]. This chapter discusses the SM with its particles and interactions, then demonstrates the enormous success of the SM in correctly predicting particles and processes over several orders of magnitude, and concludes with a section on why the SM is most likely not the most fundamental theory. In the following, electric charges are given in units of the elementary charge e , while the speed of light in vacuum c and the reduced Planck constant $\hbar$ are set to 1.

Throughout this thesis, so-called Feynman diagrams [7] are used, which are a way of graphically representing the interactions of elementary particles. These graphical representations are subject to various rules and can be strictly translated into a mathematical formula to calculate the probability amplitude of a certain process.

2.1. Particles and interactions

The elementary particles, which are the constituents of the SM, are classified according to their intrinsic properties. First of all, the elementary particles can be divided according to their spin: particles with a spin of $1/2$ are called fermions, while particles with an integer spin are called bosons. Fermions are further divided according to their charge: Fermions with an integer charge are called leptons, while those with a charge of $1/3$ or $2/3$ are called quarks. Quarks also carry one of three so-called “colour charge”, typically denoted as red, green and blue. There are three charged leptons with a charge of -1 , called electron (e), muon (μ) and tau (τ), each of which comes with an electrically neutral neutrino partner (ν_e , ν_μ and ν_τ). The only difference between the charged leptons is their mass, with the electron being the lightest and the tau the heaviest. In the theory of the SM, the neutrinos are massless. The quarks are divided into up-type quarks (up, charm, top) with an electric charge of $+2/3$ and

down-type quarks (down, strange, bottom) with an electric charge of $-1/3$. Each fermion has its own antiparticle, which has the same mass as the particle but opposite quantum numbers, most notably opposite charges. A summary of the elementary particles in the SM, including their masses, can be seen in Figure 2.1. However, the figure does not include a separate list of all antiparticles.

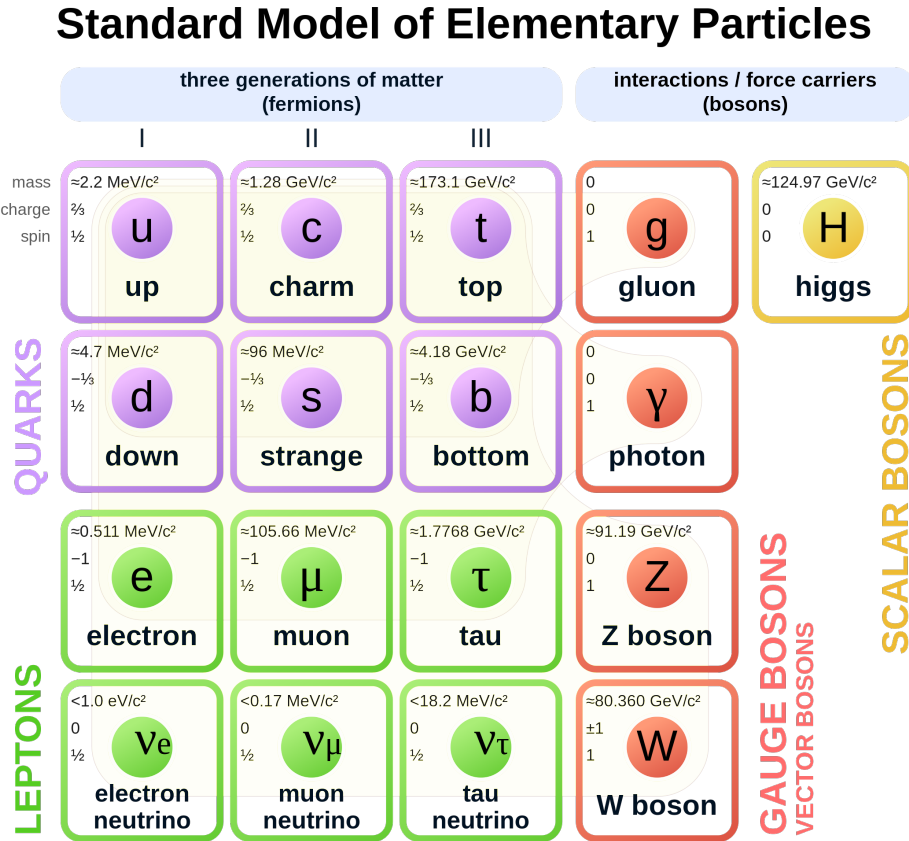


Figure 2.1: Particles of the Standard Model with their basic properties [8].

It can also be seen that the fermions can be further divided by their generation (I, II, III), where a higher generation indicates a higher mass. Only fermions of the first generation are stable and are the building blocks of all ordinary matter. Fermions of higher generations are short-lived and decay over time to first-generation fermions. To study them, higher generation fermions are usually created artificially in collider experiments.

To interact with each other, fermions exchange spin-1 gauge bosons. The electromagnetic force is mediated by the photon, denoted as γ . It is massless and has no electric charge itself, so it cannot couple to itself. The weak force is mediated by

the electrically charged $W^\pm$ bosons and the electrically neutral Z boson, which also carry a weak isospin. Left-handed up-type quarks and neutrinos have a weak isospin with its third component being $T_3 = +1/2$, while left-handed down-type quarks and charged leptons have a weak isospin with its third component being $T_3 = -1/2$. This means that the weak force couples to all left-handed fermions. Furthermore, self-couplings between the $W^\pm$ and the Z bosons are allowed. However, since the $W^\pm$ and Z bosons have non-zero masses, the spatial extent of the weak interaction is limited. Finally, the strong force is mediated by the gluons. As mentioned earlier, the quarks carry one of three colour charges, whereas the gluon has one of eight colour states, which leads to allowed gluon self-coupling. A pictorial summary of the allowed interactions of elementary particles is also given in Figure 2.1, indicated by the yellow-shaded areas.

2.2. Electroweak theory

Electromagnetic interactions and weak interactions are unified in the electroweak theory, which is subject to the $SU(2)_L \otimes U(1)_Y$ symmetry group. The generators of the $SU(2)_L$ are the weak isospin operators

$$T_i = \frac{\sigma_i}{2} \quad (2.1)$$

with σ_i being the Pauli matrices ($i = 1, 2, 3$), whereas the weak hypercharge Y is the generator of the $U(1)_Y$. The electric charge Q can be related to the third component of the weak isospin T_3 and the weak hypercharge by

$$Q = T_3 + \frac{Y}{2}. \quad (2.2)$$

The electroweak theory introduces flavour-preserving neutral currents, mediated by the Z boson and the photon, and flavour-changing charged currents, mediated by the $W^\pm$ boson. However, the flavour-changing currents are maximally parity violating, meaning that the $W^\pm$ boson only couples to left-handed fermions or right-handed antifermions. Since neutrinos are massless according to the SM (so they are always either left-handed or right-handed), the parity violation of the weak force (hence the subscript L , meaning that it couples only to left-handed fermions) leads to the existence of only left-handed neutrinos or right-handed antineutrinos. Also the fermion fields Ψ , defined as

$$\Psi_{L/R} = \frac{1}{2} (1 \mp \gamma^5) \Psi, \quad (2.3)$$

are split into handedness, with $\gamma^5 = \gamma^0 \gamma^1 \gamma^2 \gamma^3$ being the product of the Dirac matrices. The left-handed component now transforms as a doublet under an $SU(2)$ transformation, whereas the right-handed component transforms as a singlet [9]:

$$\begin{pmatrix} \nu^i \\ \ell^i \end{pmatrix}_L, \begin{pmatrix} u^i \\ d^i \end{pmatrix}_L, \ell_R^i, u_R^i, d_R^i, \quad (2.4)$$

with ν and ℓ being corresponding neutrino and charged lepton, u and d being an up-type quark and down-type quark respectively and i denoting the generation I, II and III.

The Cabibbo-Kobayashi-Maskawa matrix

Since the weak eigenstates of the quarks do not correspond to the mass eigenstates, there is a mixing matrix that mixes these states. In 1963, a predecessor of the **C**abibbo-**K**obayashi-**M**askawa (CKM) matrix was introduced by Nicola Cabibbo, the so-called Cabibbo matrix [10]. This matrix mixes the first two generations of quarks, but the discovery of indirect CP violation in 1964 [11] could not be explained in a four-quark model at that time. Instead, Kobayashi and Maskawa generalised the Cabibbo matrix into the unitary CKM matrix V_{CKM} [12] (predicting a third generation of quarks), which has the form

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \quad (2.5)$$

with the quark with the prime being in the weak eigenstates and quarks without in the mass eigenstates. One possibility of parametrising the CKM matrix is to use the “Standard parameters” [13] with three Euler angles θ_{12}, θ_{23} and θ_{13} and the CP-violating phase δ_{13} :

$$V_{\text{CKM}} = \begin{bmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{bmatrix} \quad (2.6)$$

with c_{ij} and s_{ij} denoting the cos and the sin of the angle θ_{ij} respectively. Transitions between quarks from quark flavour i to j have a probability proportional to $|V_{ij}|^2$. Recent combination of measurements of the magnitudes of the entries of the CKM matrix[14] can be seen in the following

$$|V_{\text{CKM}}| = \begin{bmatrix} 0.97435 \pm 0.00016 & 0.22500 \pm 0.00067 & 0.00369 \pm 0.00011 \\ 0.22486 \pm 0.00067 & 0.97349 \pm 0.00016 & 0.04182^{+0.00085}_{-0.00074} \\ 0.00857^{+0.00020}_{-0.00018} & 0.04110^{+0.00083}_{-0.00072} & 0.999118^{+0.000031}_{-0.000036} \end{bmatrix} \quad (2.7)$$

It can be seen that the CKM matrix has large diagonal elements and (especially for the top quark) small off-diagonal elements, which means that quarks prefer to decay into the quark of their own generation if kinematically possible. The standard parameterisation in Equation 2.6 also nicely shows the four free parameters needed to describe quark transitions.

Lagrange density of the electroweak theory

The Lagrange density, hereafter referred to as the Lagrangian $\mathcal{L}$, encodes the physics of particles in a quantum field theory. In general, for free massive fermions, the kinematics can be expressed in terms of the Dirac equation¹

$$(i\gamma^\mu \partial_\mu - m)\Psi = 0, \quad (2.8)$$

and a Lagrangian

$$\mathcal{L} = \bar{\Psi} (i\gamma^\mu \partial_\mu - m) \Psi \quad (2.9)$$

with $\partial_\mu = \frac{\partial}{\partial x^\mu}$ being the partial derivative. The mass m of a fermion can be written as a constant before the term $\bar{\Psi}\Psi$, where Ψ is the Dirac spinor. However, the requirement of local gauge invariance under both the $\text{SU}(2)_L$ and the $\text{U}(1)_Y$ symmetry group replaces the partial derivative ∂_μ by the covariant derivative for the left-handed field Ψ_L

¹In this formula and in the following, the Einstein summation notation is used, which simplifies the notation of sums: Whenever an index is used twice in a sum, it implies a summation over that index, i.e. $\gamma^\mu \partial_\mu \equiv \sum_{\mu=0}^4 \gamma^\mu \partial_\mu$.

$$D_\mu = \partial_\mu - ig \frac{\sigma_i}{2} A_\mu^i - ig' \frac{Y}{2} B_\mu \quad (2.10)$$

with g and g' being the coupling constants of the gauge fields A_μ^i and B_μ respectively. The covariant derivative for the right-handed field Ψ_R is similar but misses the term encoding interactions with the A_μ^i fields, i.e.

$$D_\mu = \partial_\mu - ig' \frac{Y}{2} B_\mu. \quad (2.11)$$

g and g' again are free parameters of the SM and have to be determined experimentally. A_μ^i with $i = 1, 2, 3$ corresponds to the gauge fields of the $SU(2)_L$ symmetry group and B_μ corresponds to the gauge field of the $U(1)_Y$ symmetry group. But even with the covariant derivative, the Lagrangian from Equation 2.9 is still not local gauge invariant, unless the masses of the fermions are set to zero, i.e. the Lagrangian equals

$$\mathcal{L} = -g \bar{\Psi} \gamma^\mu \frac{\sigma_i}{2} A_\mu^i \Psi - g' \bar{\Psi} \gamma^\mu \frac{Y}{2} B_\mu \Psi. \quad (2.12)$$

To implement the kinematics of the gauge bosons, terms including the field strength tensors corresponding to the gauge bosons must be added to the Lagrangian. With the field tensors $A_{\mu\nu}^i$ and $B_{\mu\nu}$ of $SU(2)$ and $U(1)$ respectively, defined as

$$A_{\mu\nu}^i = \partial_\mu A_\nu^i - \partial_\nu A_\mu^i + g \epsilon^{ijk} A_\mu^j A_\nu^k, \quad (2.13)$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \quad (2.14)$$

with ϵ^{ijk} being the totally antisymmetric Levi-Civita tensor, the gauge part of the Lagrangian can be written as

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4} A_{\mu\nu}^i A^{i,\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}. \quad (2.15)$$

Equation 2.12, together with Equation 2.15, forms the Lagrangian of the electroweak theory of the SM before spontaneous symmetry breaking. The latter is necessary

because the electroweak Lagrangian contains no mass terms, neither for the fermions nor for the bosons. Any mass terms in the Lagrangian at this point would again break the local gauge invariance. However, since it is known that fermions (except neutrinos in the SM) and the $W^\pm$ and Z bosons have mass, the so-called Higgs mechanism introduces mass terms into the SM without breaking local gauge invariance, which will be discussed in the next section.

Higgs Mechanism

The Brout-Englert-Higgs mechanism (or just Higgs mechanism) was developed in 1964 by three independent groups [15–17] and was implemented into the electroweak theory by Glashow, Weinberg and Salam [2, 9, 18]. This section briefly discusses how the Higgs mechanism is implemented and how masses for bosons and fermions emerge[1].

To obtain masses for the gauge bosons, a complex scalar field ϕ (i.e. the Higgs field) is introduced in the spinor representation of the $SU(2)$, together with the $U(1)$ gauge symmetry

$$\phi \rightarrow e^{i\alpha^a T^a} e^{i\frac{\beta}{2}} \phi \quad (2.16)$$

and a charge of $+\frac{1}{2}$ under the $U(1)$ symmetry. The Lagrangian of the Higgs field is described by

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \phi)^\dagger (D^\mu \phi) - \left(\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \right) \quad (2.17)$$

where $\phi^\dagger$ denotes the conjugate transpose of ϕ . The former part of Equation 2.17 encodes the dynamics between the Higgs field and the gauge bosons, whereas the latter part is known as the Higgs potential. The vacuum expectation value (VEV) of the Higgs field then takes the form

$$\langle \phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad v^2 = -\frac{\mu^2}{\lambda} \quad (2.18)$$

and is invariant under the gauge transformation from Equation 2.16 with $\alpha^1 = \alpha^2 = 0$, $\alpha^3 = \beta$. The parameter λ must be positive, because otherwise the Higgs potential would have no stable minima. Depending on μ^2 now, the potential either

has a single stable minimum at the origin for $\mu^2 > 0$, or an infinite number of minima for $\mu^2 < 0$, as can be seen in Figure 2.2 with its famous “Mexican-hat potential”. The choice of a minimum for the latter case spontaneously breaks the symmetry.

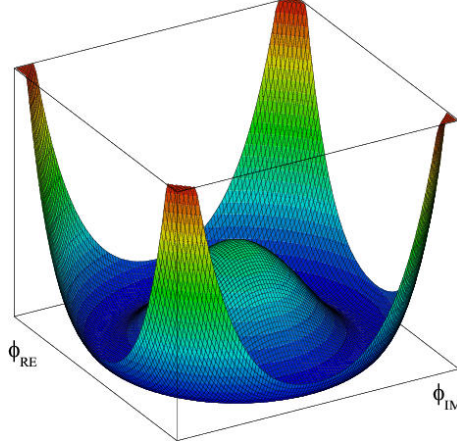


Figure 2.2: Higgs potential with its real and imaginary part for $\lambda > 0$ and $\mu^2 < 0$ [19].

The gauge boson mass terms can now be calculated by evaluating $D_\mu \phi$ at the VEV, with D_μ being the covariant derivative:

$$\mathcal{L}_{\text{Higgs}} = \frac{1}{2} \frac{v^2}{4} [g^2 (A_\mu^1)^2 + g^2 (A_\mu^2)^2 + (-gA_\mu^3 + g'B_\mu)^2]. \quad (2.19)$$

This results in one gauge boson remaining massless and the other gauge bosons acquiring mass from the Higgs mechanism. The physical photon and the Z boson are now a linear combination of the A_μ^3 and the B_μ field as follows:

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos(\theta_W) & -\sin(\theta_W) \\ \sin(\theta_W) & \cos(\theta_W) \end{pmatrix} \begin{pmatrix} A_\mu^3 \\ B_\mu \end{pmatrix} \quad (2.20)$$

In this equation, θ_W is the weak mixing angle or Weinberg angle. The physical $W^\pm$ boson is a linear combination of the A_μ^1 and A_μ^2 fields as follows:

$$W_\mu^\pm = \frac{A_\mu^1 \mp iA_\mu^2}{\sqrt{2}} \quad (2.21)$$

The Weinberg angle θ_W not only is related to g and g' via

$$\cos(\theta_W) = \frac{g}{\sqrt{g^2 + g'^2}}, \quad \sin(\theta_W) = \frac{g'}{\sqrt{g^2 + g'^2}}, \quad (2.22)$$

but also relates the actual mass of the $W^\pm$ boson m_W and of the Z boson m_Z by

$$m_W = m_Z \cos(\theta_W). \quad (2.23)$$

Also the bare electron charge e can be expressed as a function of the coupling constants of the gauge fields A_μ^i and B_μ by

$$e = \frac{gg'}{\sqrt{g^2 + g'^2}}. \quad (2.24)$$

The potential energy terms of an explicit renormalisable Lagrangian with the VEV for ϕ as in Equation 2.18 takes the form

$$\mathcal{L} = -\mu^2 h^2 - \lambda v h^3 - \frac{1}{4} \lambda h^4 \quad (2.25)$$

$$= -\frac{1}{2} m_h^2 h^2 - \sqrt{\frac{\lambda}{2}} m_h h^3 - \frac{1}{4} \lambda h^4, \quad (2.26)$$

where the Higgs boson is an excitation of the real-valued field h describing fluctuations around the minimum. The mass of the Higgs boson m_h equals to

$$m_h = \sqrt{2} \mu = \sqrt{2 \lambda} v \quad (2.27)$$

and one can choose m_h and the VEV v to be the two additional free parameters of the Higgs sector in the SM, which have to be determined experimentally.

Lastly, fermion masses can now be introduced into the SM by again using the mechanism of spontaneous symmetry breaking. Exemplarily for the electron, the gauge

invariant part of the Lagrangian, where the VEV is already inserted, can be seen in the following

$$\Delta\mathcal{L}_e = -\frac{1}{\sqrt{2}}\lambda_e v \bar{e}_L e_R + h.c. + \dots, \quad (2.28)$$

where $h.c.$ denotes the hermitian conjugate of the preceding term. It follows that the mass of the electron equals

$$m_e = \frac{1}{\sqrt{2}}\lambda_e v, \quad (2.29)$$

with λ_e being a new free parameter which has to be determined experimentally. It can also be seen that the coupling of the fermions to the Higgs field, also called the Yukawa coupling, is proportional to the mass of the fermions. Equation 2.29 can be developed analogously for each charged lepton and every quark, resulting in nine free parameters for the charged lepton- and quark masses.

2.3. Quantum chromodynamics

The theory of **quantum chromodynamics** (QCD) describes the strong force between quarks and gluons. Since the strong interaction is not affected by the Higgs mechanism, the gluons remain massless. The underlying symmetry group of the QCD is the $SU(3)_C$, with the index C denoting the newly introduced colour quantum number. The generator of the $SU(3)_C$ are

$$T^a = \frac{1}{2}\lambda^a \quad (2.30)$$

with λ^a , $a = 1, \dots, 8$, being the Gell-Mann matrices. The same amount of gauge fields G_μ^a are introduced as there are Gell-Mann matrices, each kind of gluon corresponding to one gauge field, which results in eight different kinds of gluons. Analogously to the electroweak theory, a covariant derivative D_μ is defined to maintain local gauge invariance:

$$D_\mu = \partial_\mu - ig_S T^a G_\mu^a \quad (2.31)$$

with g_S being the bare coupling constant of the strong interaction. Together with the gluon field tensor

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a + g_S f^{abc} G_\mu^b G_\nu^c \quad (2.32)$$

the Lagrangian can be expressed as

$$\mathcal{L}_{QCD} = -\frac{1}{4} G_{\mu\nu}^a G^{a,\mu\nu} + \sum_q \bar{\Psi}_q i\gamma^\mu D_\mu \Psi_q, \quad (2.33)$$

where the sum over q runs over all quark flavours. In Equation 2.32, f^{abc} denotes the structure constants of the $SU(3)_C$, fulfilling the relation $[\lambda_a, \lambda_b] = 2if_{abc}\lambda_c$. Since the Gell-Mann matrices in general do not commute, the QCD has a non-abelian group structure, which leads to self-couplings between gluons. This results in the coupling constant g_S actually not being constant, but instead vary with the energy scale. Moreover, g_S is again a free parameter of the theory and must be determined experimentally.

Running of the strong coupling constant

The strong coupling constant is usually expressed as

$$\alpha_S(Q^2) = \frac{g_S^2(Q^2)}{4\pi} \quad (2.34)$$

and depends of the energy scale Q of the interaction. At first order, the strong coupling constant can be expressed as

$$\alpha_S(Q^2) = \frac{\alpha_S(\mu^2)}{1 + \frac{\alpha_S(\mu^2)}{4\pi} \left(\frac{11}{3}N - \frac{2}{3}n_f \right) \log \left(\frac{Q^2}{\mu^2} \right)} \quad (2.35)$$

with a reference energy scale μ (usually the Z -boson mass since it can be measured precisely), the number of colour charges $N = 3$ and the number of quark flavours n_f for which holds $m(q_f) < Q$ [1]. From this dependence of the energy scale, two extreme cases for the strong coupling constant can be considered.

On the one hand, for small energy scales, the strong coupling constant becomes large ($\mathcal{O}(1)$), which is known as confinement [20]. As a result, quarks cannot be isolated and cannot be observed themselves, but only as colourless states known as hadrons. The exception is the top quark, which decays before it can form bound states, so the decay products of bare top quarks can be observed. Otherwise, either two quarks form a bound state called a meson, where the quarks have a colour and an anti-colour, or three quarks form a bound state called a baryon, where the quarks have either three different colours or three different anti-colours. However, more exotic baryon structures are also allowed as long as the state is colourless, such as a pentaquark consisting of four quarks and one antiquark, as discovered by the LHCb collaboration in 2015 [21].

On the other hand, at large energy scales, the strong coupling constant becomes small, which is known as asymptotic freedom [22]; quarks and gluons can be considered as free particles at high energy scales. Since higher order effects (i.e. loop contributions to processes) contribute in orders of α_S , sufficiently high collision energies allow perturbation theory to be applied to QCD.

2.4. Experimental validation

As previously stated in the introduction to this chapter, certain phenomena of the SM could only be explained by the existence of new particles. The prediction of the bottom quark and top quark as well as the Higgs boson, many years before their eventual discovery, has been a tremendous achievement for the SM. Over the past few decades, not only have the particles' existence been confirmed, but their characteristics have also been extensively measured. This section provides an overview of several remarkable measurements that verify specific aspects of the SM. Figure 2.3 demonstrates the predictive power of the SM for production cross sections (the probability/rate of a certain process to happen) of various processes. Note that the measurements from the ATLAS collaboration (denoted as coloured areas) almost always match with their respective SM prediction (denoted as grey areas) over 14 orders of magnitude.

As already discussed in the previous section, the strong coupling constant α_S varies with the energy scale. Figure 2.4 shows the world average of α_S and its dependence on the energy scale. Since α_S is one of the free parameters of the SM, it must be determined by experiments at a specific energy scale. This is demonstrated in Figure 2.4 (a), where different experiments using varying techniques produce a compatible result for the strong coupling constant. Predictions and experimental tests can be conducted at arbitrary energy scales, as illustrated in Figure

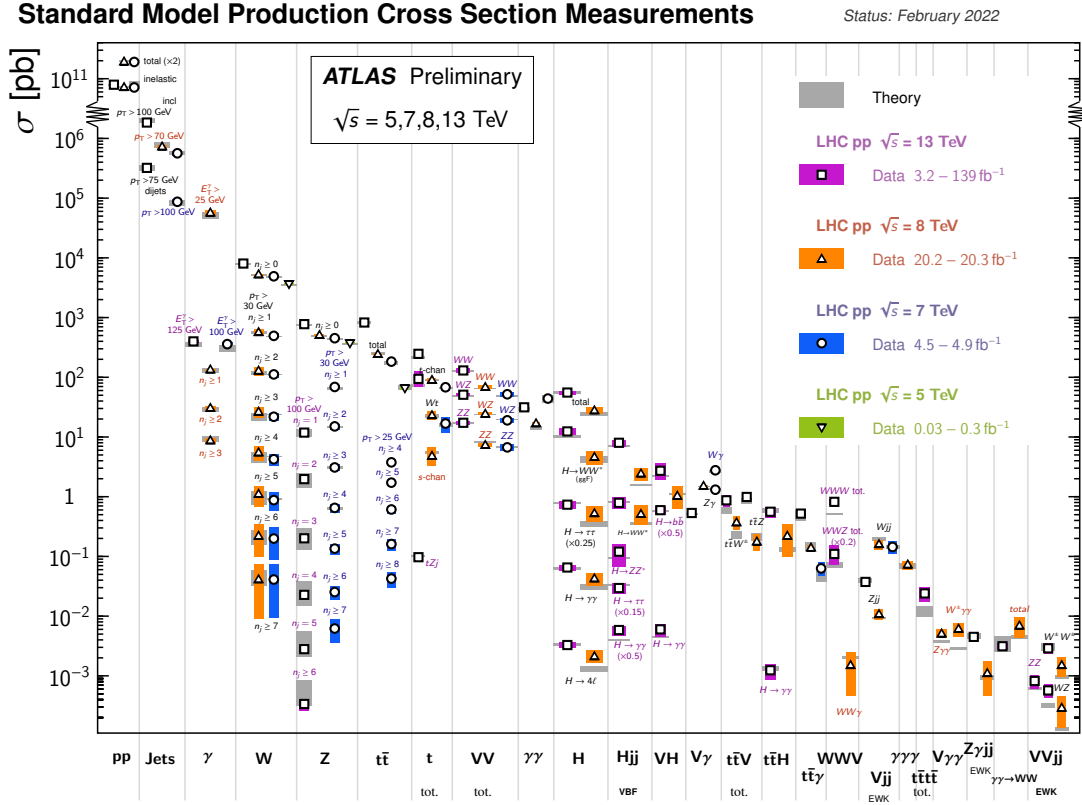


Figure 2.3: SM prediction of various production cross sections and their measurements [23].

2.4 (b). Again, a very good agreement between prediction and experiment can be observed.

As a last example, the direct measurement of the top quark and W boson masses is compared with the result of a global fit to the electroweak sector without these two parameters, shown in Figure 2.5. The global fit of the electroweak sector without the top quark and W-boson masses gives a prediction of just these, shown as grey ellipses. It is then possible to determine if the global fit of the electroweak sector is consistent with the mass measurements by independently measuring the top-quark mass and the W-boson mass. As illustrated, the SM effectively characterises the electroweak sector.

2.5. Limitations

Despite the numerous coherent precision measurements of the SM and its achievements over the past decades, there are fundamental problems that cannot be ex-

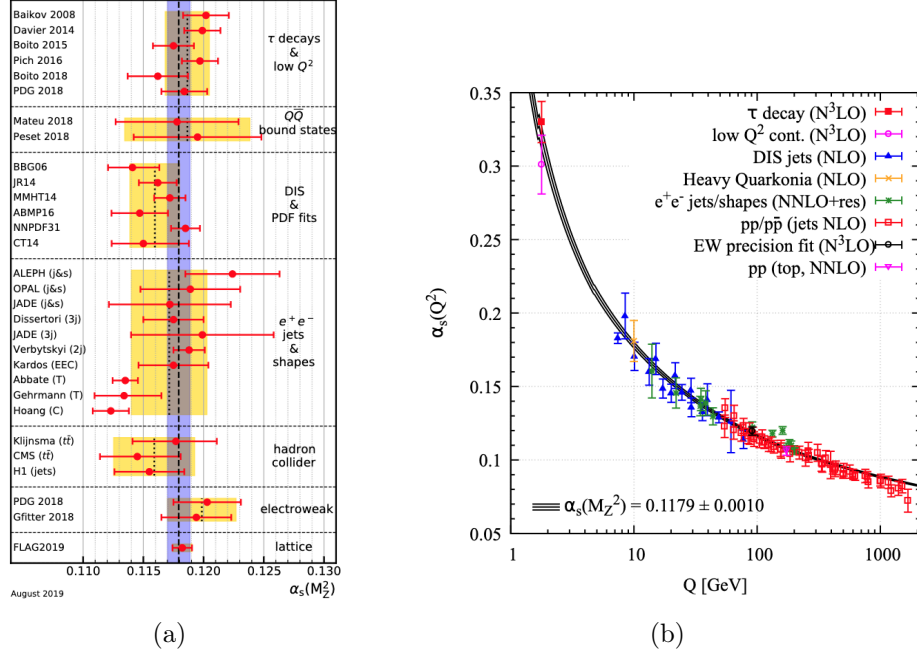


Figure 2.4: World average of the measurements of $\alpha_S(m_Z^2)$ (a) and the running of α_S , depending on the energy scale (b) [24].

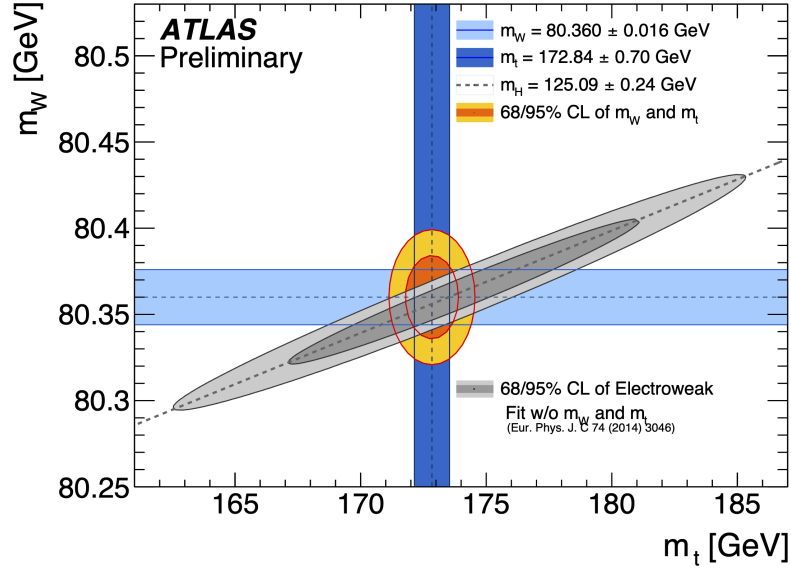


Figure 2.5: Measurement of the top-quark mass and the W-boson mass, compared to the prediction of a global fit of the electroweak sector [25].

plained by the SM at present. This section outlines several contradictions and unexplained phenomena associated with the Standard Model.

As discussed in the electroweak section of the SM theory, neutrinos are assumed to be massless in the SM. However, data from e.g. the Super-Kamiokande experiment [26] are incompatible with the assumption of massless neutrinos, because their data can only be explained by neutrino oscillation, meaning that the flavour of neutrinos can vary over time. This neutrino oscillation would require the neutrinos to have a mass greater than zero, in contradiction with the SM. However, at present it is not possible to determine the masses of the individual neutrinos, but only to provide confidence intervals for the sum of all three masses. Recent measurements show that this sum is approximately $\sum_\nu m_\nu = 0.1 \text{ eV}$ [27, 28].

The SM's next incongruity lies in its combination of only three of the four fundamental particle interactions: Electromagnetism, the weak force and the strong force. However, it omits gravity, which is described by Einstein's general theory of relativity [29]. Although gravitation has only a negligible effect on the physics of modern accelerators, at the Planck scale of approximately 10^{19} GeV , gravitation is expected to dominate over effects from the SM. At such high energies, a new Unified theory is expected to unify all four forces.

Yet another discrepancy arises from cosmology: When observing the rotational curves of galaxies, there is a mismatch between visible matter and the resulting rotation of the galaxies [30]. The accepted conclusion is that there must be a non-visible distribution of gravitationally interacting matter, named Dark Matter. The Standard Model offers no explanation of what this dark matter could be.

The final issue to be raised is primarily one of aesthetics rather than a genuine contradiction to the Standard Model. It concerns the quantity of unconstrained parameters, as outlined in Table 2.1.

Description	Number of free parameters
Quark masses	6
Charged lepton masses	3
CKM mixing angles + CP-violating phase	4
Coupling constants g , g' and g_s	3
QCD vacuum angle ¹	1
m_H and VEV	2
Total	19
Neutrino masses	3
PMNS ² mixing angles + CP-violating phase	4
New Total	26

Table 2.1.: Summary of free parameter of the Standard Model, which have to be determined experimentally.

It should be noted that this table additionally incorporates free parameters pertaining to the masses and mixing of neutrinos, which have been verified experimentally. In total, when also describing neutrino masses and mixing, 26 free parameters are necessary and must be determined in experiments. This huge amount of free parameters seems *unnatural* and the SM seems to be a “low-energy approximation” of a more fundamental theory. This thesis is part of the search to find out, where the low-energy assumption breaks down and where one can look in more detail for new physics.

¹This parameter corresponds to the potential CP violation in the QCD, but has up to now always been measured compatible with zero. [31, 32]

²The **P**ontecorvo–**M**aki–**N**akagawa–**S**akata (PMNS) matrix [33, 34] describes the mixing of the neutrinos, similar to the CKM matrix describing the mixing of the quarks.

3. Flavour-changing neutral currents

This chapter discusses **Flavour-changing neutral current** (FCNC) interactions and their theoretical implementation. Furthermore, the two processes contributing will be introduced.

3.1. FCNCs in the SM

As already discussed, in the SM the flavour of a quark may only be changed by a charged current, namely the $W^\pm$ boson, as can be seen in the Feynman diagram in Figure 3.1.

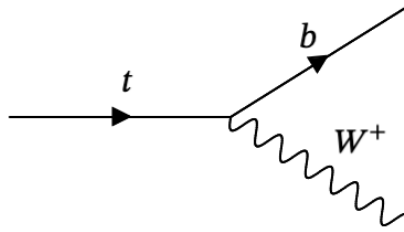


Figure 3.1: SM example for a flavour-changing charged current, where a top quark decays into a b quark and a W^+ boson.

There is no vertex for an FCNC included in the SM at leading order, and also higher orders are highly suppressed in the SM, as discovered by Glashow, Iliopoulos and Maiani in 1970 with their so-called GIM-mechanism [35]. The GIM-mechanism states that in loop diagrams of an FCNC process (e.g. Figure 3.2), potential contributions to the amplitude cancel each other due to the unitarity of the CKM matrix.

In fact, the loop diagrams of the SM FCNC contribution would cancel completely if the masses of the intermediate quarks were the same. Only due to the different masses of the b , s and d quark, the SM FCNC branching ratio prediction $\mathcal{B}(t \rightarrow qH)$ ($q = u, c$) is larger than zero, namely in the order of $\mathcal{O}(10^{-12}) - \mathcal{O}(10^{-16})$ (see Figure 3.3). Measuring such small branching ratios is far beyond the capabilities of today's accelerators and detectors.

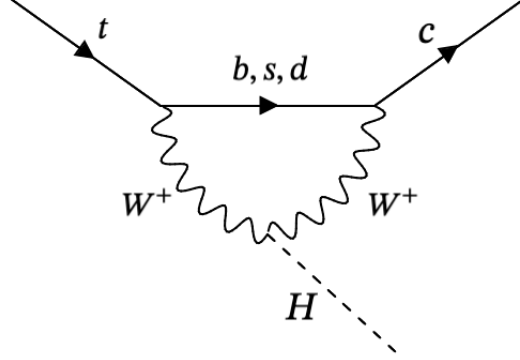


Figure 3.2: SM-allowed loop diagram of an FCNC process $t \rightarrow cH$, which is highly suppressed due to the GIM-mechanism.

3.2. FCNCs in Beyond-Standard-Model theories

Due to the acknowledged inadequacies of the SM, numerous extensions have been developed, called **Beyond-the-Standard-Model** (BSM) theories. These theories typically introduce new particles or newly allowed vertices, which consequently increase certain observables, such as the branching ratio $\mathcal{B}(t \rightarrow u/cH)$. Figure 3.3 shows a summary of various BSM theories and their impact on the prediction of the branching ratio $\mathcal{B}(t \rightarrow u/cH)$ displayed as coloured regions. Also visible in the plot is the SM prediction for FCNC processes presented as a solid black line, along with the present exclusion limits on the branching ratios. As can be seen, the first BSM theories can be probed by their FCNC contribution, especially in the $t \rightarrow qH$ and $t \rightarrow qg$ sectors.

Since there are numerous new theories predicting an FCNC contribution through various means, it would be unfeasible to parametrise each BSM theory separately and conduct a dedicated search. Instead, in order to establish a generally working approach, an **effective field theory** (EFT) is introduced, which parametrises any possible BSM contribution by effective operators. This principle is demonstrated in Figure 3.4, whereby Figure 3.4 (a) shows a hypothetical BSM theory with new particles X_1, X_2, X_3 and newly allowed vertices, eventually resulting in a $t \rightarrow u/cH$ FCNC interaction. However, detailed knowledge of the whole process producing the FCNC interaction is not necessary for the search, and the BSM part of the Feynman diagram can be replaced by an effective operator (see Figure 3.4 (b)). For aesthetic reasons, this operator is shown in the following as a small black dot (as can be seen in Figure 3.4 (c)).

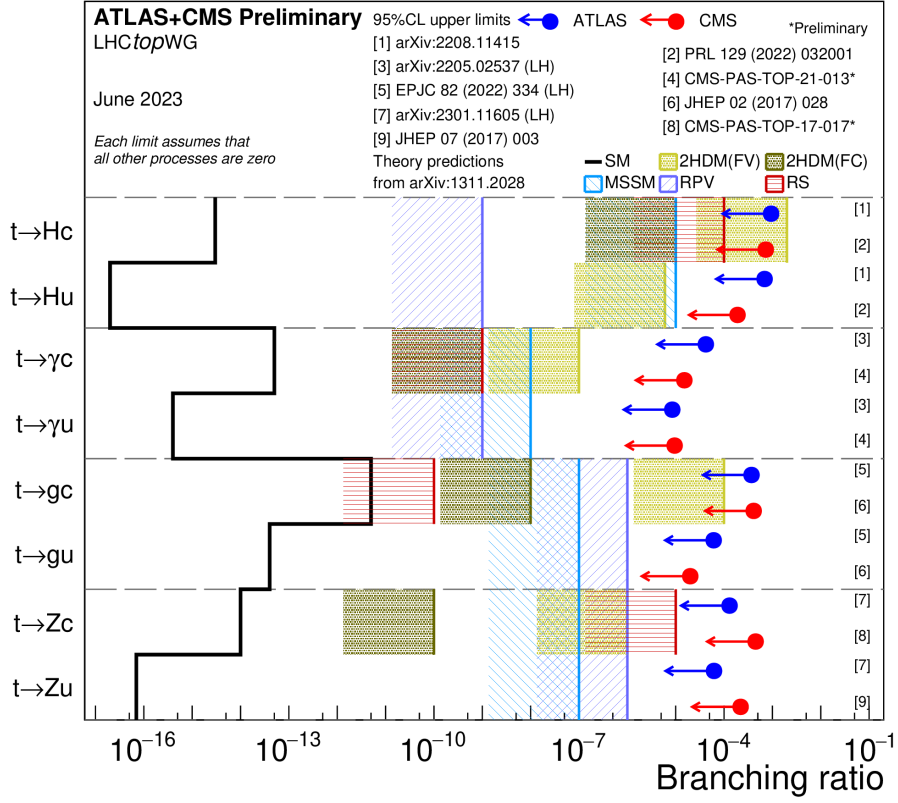


Figure 3.3: Summary of the current 95% confidence level observed limits on the branching ratios of the top quark decays via flavour changing neutral currents (FCNC) to a quark and a neutral boson $t \rightarrow Xq$ ($X = g, Z, \gamma$ or H ; $q = u$ or c) by the ATLAS and CMS Collaborations compared to several new physics models. [36]

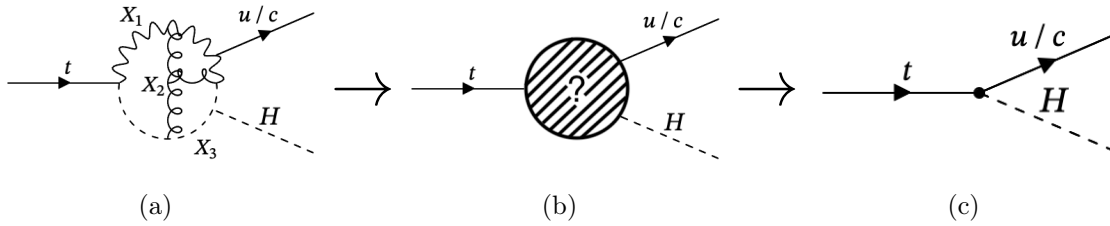


Figure 3.4: (a) A Feynman diagram of a hypothetical BSM theory with new particles X_1, X_2, X_3 and new vertices between each other and known elementary particles. Can be rewritten to (b) where there is no knowledge about the underlying theory and the overall vertex is approximated by a effective single one. This can then be reduced to (c), where the unknown vertex is just denoted by the small dot.

The Lagrangian of the SM is then extended by the sum of operators O_i , which are scaled by dimensionless Wilson coefficients C_i , as can be seen the following equation

$$\mathcal{L}_{\text{EFT}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{C_i}{\Lambda^2} O_i. \quad (3.1)$$

These operators O_i are encoding the kinematics of new physics in a general way, where each operator corresponds to a newly allowed vertex, e.g. a $t \rightarrow qH$ FCNC vertex. The Wilson coefficients scale the strength of the vertex (i.e. the probability of the vertex) and are constrained in analyses. Λ is the energy scale at which new physics is expected. Technically, the FCNC couplings are derived from an EFT for FCNC processes [37, 38] and implemented in the **TopFCNC** FEYNRULES model [39, 40]. This allows the simulation of FCNC events at the LHC at **next-to leading order** (NLO). Since there are potential differences not just between up- and charm quark couplings, but also between left- and righthanded couplings, Table 3.1 summarises the Wilson coefficients of interest. All other Wilson coefficients in Equation 3.1 corresponding to vertices not related to tqH FCNCs will be set to zero.

Block parameter	Wilson coefficient	Comment
RCtphi	$C_{u\phi}^{13}$	left-handed tuH coupling
RCuphi	$C_{u\phi}^{31}$	right-handed tuH coupling
RCtcphi	$C_{u\phi}^{23}$	left-handed tcH coupling
RCctphi	$C_{u\phi}^{32}$	right-handed tcH coupling

Table 3.1.: Summary of Wilson coefficients corresponding to operators causing tqH FCNC vertices. The block parameter denotes the internal naming of the corresponding Wilson coefficient, as used in configuration files.

However, dedicated studies (see Appendix C) indicate a negligible difference in the kinematic distributions between left- and right-handed couplings for utH and ctH respectively. To improve statistical accuracy, both couplings are used simultaneously with proper normalisation, and no distinction is made regarding handedness. As a notable distinction exists between the utH and ctH couplings, they are treated independently.

3.3. FCNC tqH processes

The FCNC vertex introduced in the previous section may occur in different processes, which are referred to as signal processes. The first process is called production-FCNC process (or prod-FCNC process), where an up- or charm quark receives enough energy to subsequently convert into a top quark and a Higgs boson, which can be seen in Figure 3.5 (a). The second process is called decay-FCNC process, where either the top- or the top antiquark from $t\bar{t}$ production decays via the FCNC vertex, as can be seen in Figure 3.5 (b). It is essential to consider both processes to obtain accurate results as both processes may occur when the corresponding Wilson coefficient has a non-zero value.

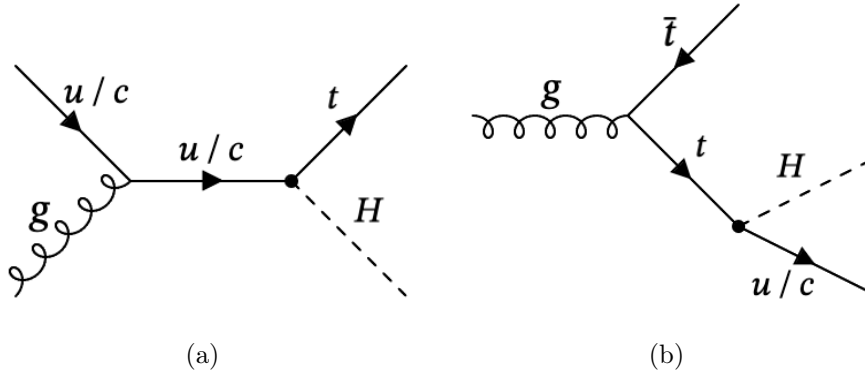


Figure 3.5: (a) The prod-FCNC process where a boosted up- or charm quark is converted into a top quark by radiating a Higgs boson and (b) the decay-FCNC process where either the top- or the top antiquark from $t\bar{t}$ production decays using the FCNC vertex.

It has to be ensured that both signal processes are normalised to the same value of the corresponding Wilson coefficient. In terms of the prod-FCNC process, the cross section σ (a measure for the probability of a process, see Section 4.1.2) is predicted by the simulation and is related to the Wilson coefficient by

$$\mathcal{M}_{tqH} \propto C_{u\phi} \quad \text{and} \quad \sigma \propto |\mathcal{M}_{tqH}|^2, \quad (3.2)$$

with $\mathcal{M}_{tqH}$ being the transition amplitude of the respective process. It follows that

$$\sigma_{tqH} = \alpha \cdot C_{u\phi}^2 \quad (3.3)$$

with α being a constant factor which can be determined by the simulation. The cross section of the decay-FCNC process, however, is determined by the $t\bar{t}$ production cross section, multiplied with the branching ratios for the top- and top antiquark to subsequently decay as enforced. First, the top-quark partial decay width for the FCNC process $\Gamma(t \rightarrow qH)$ must be calculated [41] using

$$\Gamma(t \rightarrow qH) = \frac{|C_{u\phi}|^2}{\Lambda^4} \frac{\sqrt{2}G_F m_t^7}{8\pi} \left(1 - \frac{m_H^2}{m_t^2}\right), \quad (3.4)$$

where m_t and m_H corresponds to the mass of the top quark and of the Higgs boson respectively, G_F denotes the Fermi constant and $\Lambda = 1 \text{ TeV}$ denotes the energy scale at which new physics is expected. Equation 3.4 assumes the light quark masses to be zero and only corresponds to the LO calculation. The branching ratio $\mathcal{B}(t \rightarrow qH)$ is then defined by

$$\mathcal{B}(t \rightarrow qH) = \frac{\Gamma_{t \rightarrow qH}}{\Gamma_{\text{total}}} = \frac{\Gamma_{t \rightarrow qH}}{\Gamma_{t \rightarrow qH} + \Gamma_{t \rightarrow bW}}, \quad (3.5)$$

with $\Gamma_{t \rightarrow bW}$ being the partial decay width for the known process $t \rightarrow bW$. Using equations 3.4 and 3.5, the arbitrary cross section given by the simulations for the decay-FCNC process can then be normalised to a meaningful cross section, which corresponds to a Wilson coefficient of $C_{u\phi} = 1$.

4. The Large Hadron Collider and the ATLAS experiment

This chapter discusses the Large Hadron Collider complex and subsequently the ATLAS experiment in their states during the 2015-2018 data collection period, also known as Run 2. Both the Large Hadron Collider complex and the ATLAS experiment, located near Geneva in Switzerland, were used to collide protons and subsequently record the data analysed in this paper.

4.1. The Large Hadron Collider complex

Since the **L**arge **H**adron **C**ollider (LHC) is not designed to accelerate particles at rest, the protons must first be accelerated to an energy of 450 GeV in various pre-accelerators before the LHC can accelerate them further to a centre-of-mass energy of 13 TeV.

4.1.1. The pre-accelerators

Figure 4.1 shows a schematic overview of the LHC including the pre-accelerators. The LHC and its pre-accelerators do not only provide the particle beams for the ATLAS experiment, but also for various other experiments. Those are however not further discussed in this thesis.

Linac 2 The **L**inear **a**ccelerator 2 (Linac 2) [43] uses a bottle of hydrogen gas as source. By passing the hydrogen gas through an electric field, the sole valence electron of the hydrogen atom is stripped away, resulting in bare protons. These protons are then accelerated by cylindrical conductors charged by high-frequency cavities. By the protons' arrival at Linac 2's end, their energy has increased to 50 MeV.

PSB From the Linac 2, the protons are subsequently accelerated in the **P**roton **S**ynchrotron **B**ooster (PSB) [43]. The PSB became operational in 1972 and underwent two upgrades to meet the increasing demands of the LHC. Since its second upgrade in 1999, the PSB is able to accelerate the protons to a maximum energy of 1.4 GeV in its four rings, each with a radius of 25 m.

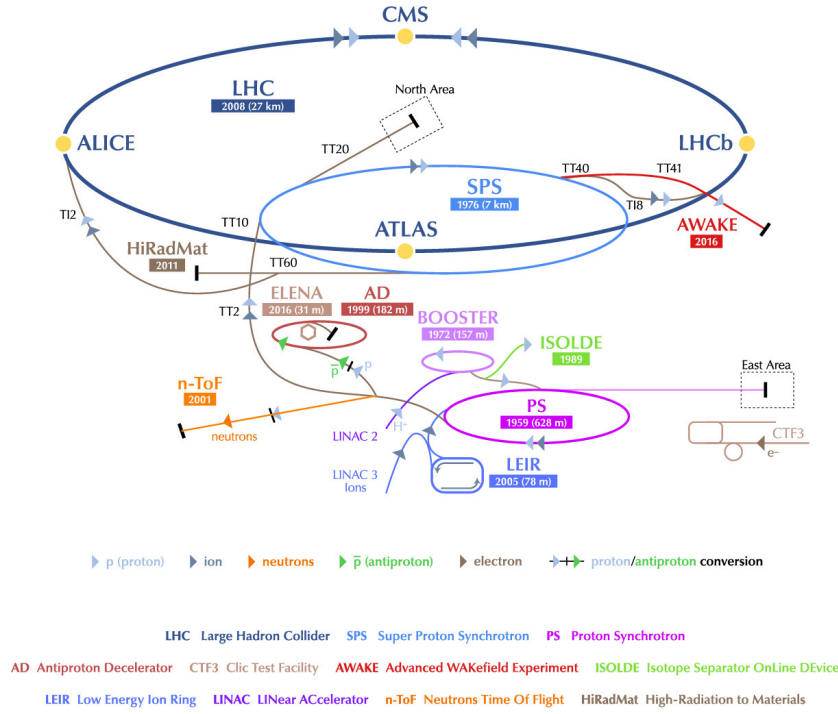


Figure 4.1: Schematic overview of the LHC complex including the pre-accelerators [42].

PS After the PSB, the protons are further accelerated in the **Proton Synchrotron** (PS) [43]. The PS is a synchrotron ring with a radius of 72 m where protons are accelerated to an energy of 25 GeV. The PS also prepares the bunch train structure used in the LHC, which will be discussed later.

SPS The final pre-accelerator of the LHC is the **Super Proton Synchrotron** (SPS) [43]. Here, protons are accelerated to an energy of 450 GeV in a synchrotron accelerator with a radius of about 1.1 km, before being injected into the LHC via two specially built transfer lines.

4.1.2. The Large Hadron Collider

The LHC [44] is the largest circular particle accelerator to date, with a circumference of approximately 27 km, colliding protons at a centre-of-mass energy of 13 TeV. The four major experiments ALICE[45], ATLAS[46], CMS[47] and LHCb[48] are located at four of its eight interaction points (IP). Protons (or heavy ions) are accelerated in opposite directions in two separate vacuum beam pipes and can be brought to collision in these IPs. The ALICE experiment specialises in heavy ion collisions and the LHCb experiment specialises in b hadron physics. ATLAS and CMS are both general-purpose detectors, built to cover a large solid angle around the IP to

precisely measure the SM and to search for new physics beyond the SM. The other four IPs are used for beam cleaning, radio frequency acceleration and the beam dumping system.

Instead of a continuous beam of protons, the protons are arranged in bunches. The energy of the protons in a bunch is not exactly the nominal energy for all protons, but some protons have slightly less energy and others slightly more. Because of the rising and falling flanks of the electric field of the RF cavities, the bunches can be accelerated so that the protons with too little energy receive a greater energy gain and vice versa for the protons with too much energy. This leads to the existence of stable positions for the bunches inside the LHC. Given the circumference of the LHC, the speed of the protons and the frequency of the RF cavities of $f_{RF} = 400$ MHz, a maximum number of positions where the protons of a bunch can oscillate stably around the nominal position can be calculated to be 35640. These positions are called *buckets*. However, since e.g. the beam-dump kicker system has a significant rise time, not every bucket is filled with a bunch. Instead, the bunches are arranged in *trains*, leaving some buckets empty in a row. Depending on the filling scheme of the given physics run of the LHC, up to a maximum of 2808 bunches can be filled into the LHC with a spacing of 25 ns to the next bunch within the same train. Each bunch contains about $n_p \approx 10^{11}$ protons.

To keep the protons on the nominal trajectory of the LHC, different types of magnets are used. To bend the protons on the circular path of the LHC, 1232 superconducting main dipole magnets are used, cooled to 1.9 K with superfluid helium. Quadrupole magnets are used throughout the LHC ring to prevent the protons in the bunches from spreading out, for example due to the repulsive electrical force of same-charged particles, which would eventually lead to beam loss. Quadrupole magnets consist of four magnets arranged so that their equally charged poles face each other. When the magnetic field inside a quadrupole magnet is expanded according to the multipole expansion, the dipole terms cancel out and the quadrupole terms become the most important terms. This results in a charged particle passing through the quadrupole experiencing a focusing force on one axis (depending on the charge of the quadrupole and the particle) and a defocusing force on the other axis. By using alternating quadrupole magnets, an overall focusing effect of the bunches can be achieved in all directions transversal to the moving direction. Higher pole magnets are also used to correct for higher order deviations of the beam from its nominal path.

An important characteristic of a particle collider is the instantaneous luminosity $\mathcal{L}$. Together with the cross section of a given process $\sigma_{pp \rightarrow X}$ (a measure of the probability of the process $pp \rightarrow X$), the rate of the process can be calculated by

$$\frac{dN_{pp \rightarrow X}}{dt} = \mathcal{L} \cdot \sigma_{pp \rightarrow X}. \quad (4.1)$$

Thus, a high instantaneous luminosity is desired, especially when searching for rare processes with small cross sections. With the assumption of head-on-colliding beams and Gaussians beam profiles, the instantaneous luminosity can be calculated by

$$\mathcal{L} = \frac{n_p^2 f N_b}{2\pi \sigma_x \sigma_y}, \quad (4.2)$$

with f being the revolution frequency in the LHC, N_b the number of bunches of the fill and σ_x and σ_y being the transverse widths of the beam profile. In the data taking run in 2017, the peak instantaneous luminosity exceeded $2 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ for the first time[49]. Figure 4.2 (a) shows the instantaneous luminosity integrated over time (called the integrated luminosity) per year, as recorded by the ATLAS experiment. Due to the number of protons per bunch and the highly focused beam profiles, multiple inelastic scattering processes can occur per bunch crossing, which is called in-time pile-up. Due to the high bunch crossing frequency of 40 MHz, remnants of the previous bunch crossing can also cause an out-of-time pile-up. Since the readout bandwidth of the detector apparatus is limited, the high peak luminosity since the 2017 run required levelling (e.g. β^* levelling¹) to keep the pile-up within acceptable limits[49]. Figure 4.2 (b) shows the average number of interactions per bunch crossing, weighted by luminosity.

4.2. The ATLAS Detector

The ATLAS detector[46] (short for “**A** Toroidal **L**HC **A**pparatu**S**”) is one of the four major experiments at the LHC and one of the two general-purpose detectors, along with CMS. The ATLAS detector consists of several sub-detectors, each specialised on a specific task to identify and measure particles produced at the IP around which the ATLAS detector is built. Figure 4.3 shows a schematic overview of the ATLAS detector, with its dimensions and the main subsystems labelled, whereas Figure 4.4 shows the working principle of particle identification of the ATLAS detector. In the following subsection, the sub-detectors of the ATLAS detector are discussed.

¹ β^* denotes the value of the β function (and thus the transverse size of the beam) at the interaction point

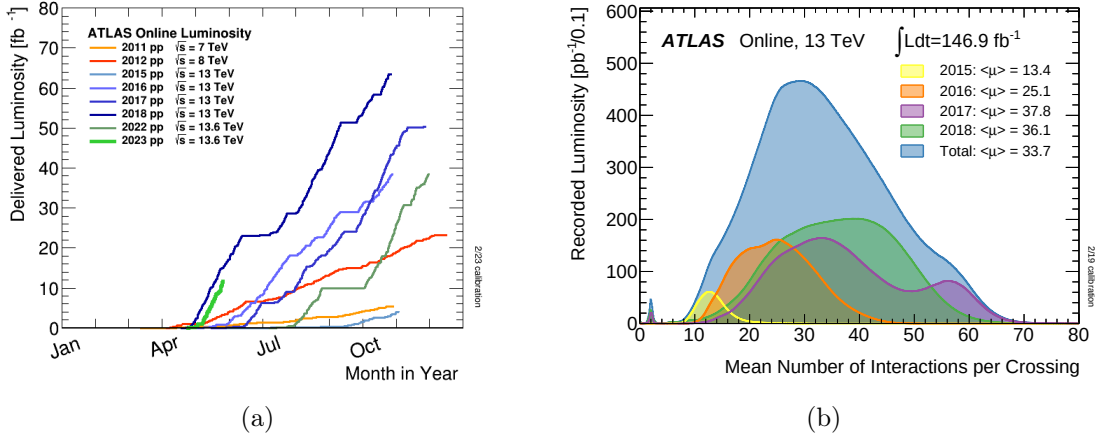


Figure 4.2: The delivered luminosity for pp collisions in the years 2011 - 2023 (a) [50] and the distribution of the mean number of interactions per bunch crossing and weighted by luminosity, split by the data-taking year 2015 - 2018 (b) [51].

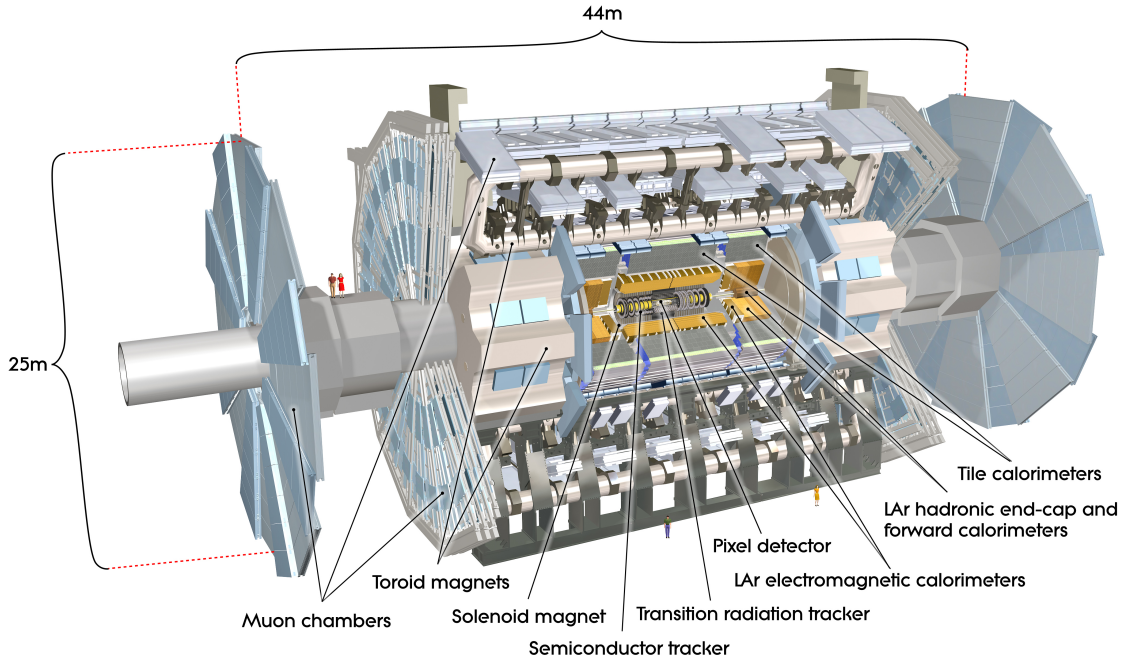


Figure 4.3: Schematic overview of the ATLAS detector including its dimensions [52].

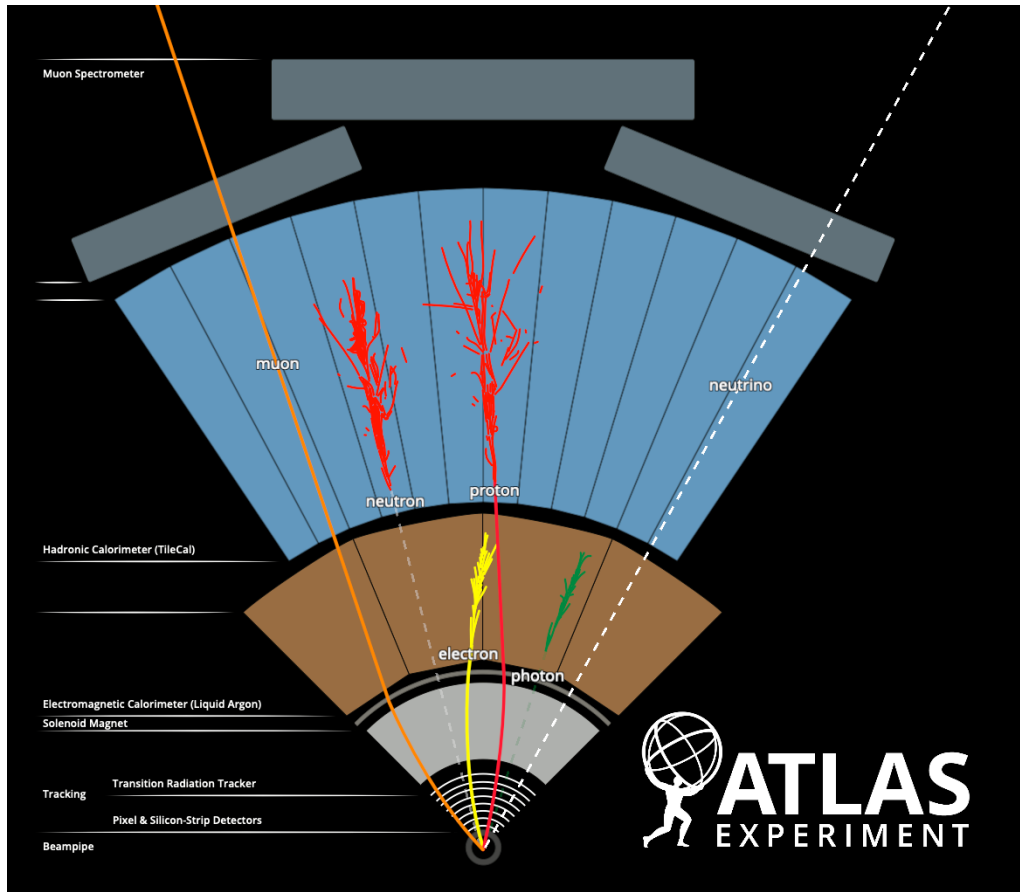


Figure 4.4: Cross-section of the ATLAS detector showing the different behaviour of particles in the sub-detectors. Dashed lines are invisible for the detector. Electrically charged particles are deflected by a magnetic field and tracked by the detector components shown in grey. The electromagnetic calorimeter (brown) then absorbs and measures the energy of, for example, electrons and photons. Hadrons, such as neutrons and protons, are absorbed by the hadronic calorimeter, shown in blue. As muons generally leave the calorimeter systems, their trajectory is again tracked in the muon spectrometer. Neutrinos leave the detector untracked. By combining the information from the sub-detectors, particles can be distinguished [53].

4.2.1. Coordinate system

The ATLAS detector uses a right-handed coordinate system centred at the IP. The x-axis points to the centre of the LHC ring, the y-axis points upwards and the z-axis points tangentially and anticlockwise along the beam pipe. Since the ATLAS detector is rotationally symmetric around the IP, an azimuthal angle ϕ and a polar angle θ are commonly used. The angle ϕ is measured in the transverse x - y -plane,

while the angle θ is measured with respect to the z -axis. Using θ , one can define the pseudorapidity η as

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right), \quad (4.3)$$

the rapidity y as

$$y = \frac{1}{2} \ln\left(\frac{E + p_z}{E - p_z}\right) \quad (4.4)$$

and the angular distance ΔR between two objects as

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \quad (4.5)$$

Thereby p_z is defined as the momentum in the z -axis. In the relativistic limit ($m \ll |\vec{p}|$), the pseudorapidity converges to the rapidity. The rapidity is preferred over the polar angle θ since differences in rapidity are Lorentz invariant under boosts along the z -axis.

4.2.2. Inner detector

Of all the detector subsystems, the Inner Detector (ID) is closest to the beam pipe and is designed to precisely track electrically charged particles in the range up to $|\eta| < 2.5$ [54, 55]. To measure the charge and momentum of the particles passing through, a superconducting solenoid magnet surrounds the ID and provides a 2 T magnetic field parallel to the z -axis, bending the trajectories of the particles according to their charge and their transverse momentum. The ID itself consists of three subsystems: The pixel detector, the semiconductor tracker and the transition radiation tracker. These will be briefly discussed in the following paragraphs. Figure 4.5 shows a cross-section of the ID with its subsystems in the barrel region.

Originally the pixel detector consisted of only three cylindrical layers in the barrel area (called B-layers, layer 1 and layer 2 with increasing distance from the beam pipe) and three discs in each end cap. However, in order to cope with the increasing pile-up in Run 2, among other things, the additional **I**nteractable **B**-layer (IBL) [57]

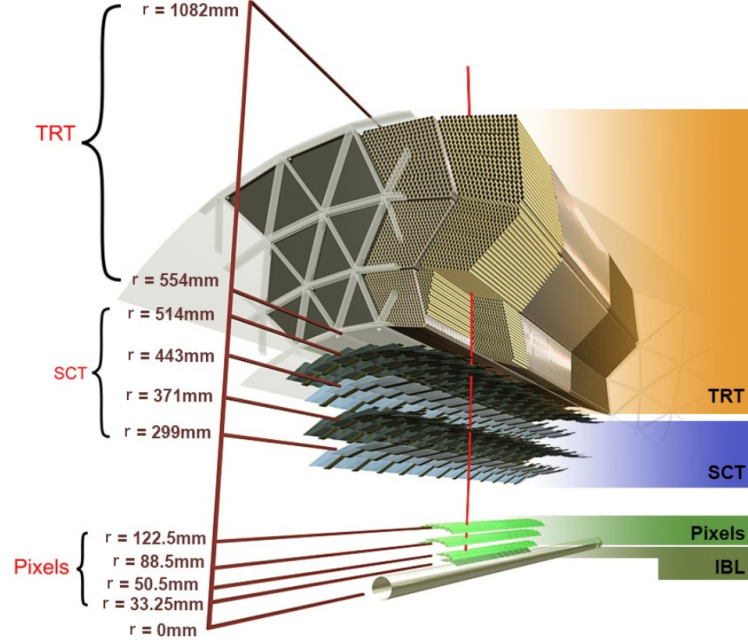


Figure 4.5: Cross section of the ID with its subsystems in the barrel region.[56]

was installed in the barrel region during Long Shutdown 1. The actual pixels in the silicon semiconductor sensors of the B-Layer, Layer 1 and Layer 2 have a size of $50\text{ }\mu\text{m} \times 400\text{ }\mu\text{m}$, whereas the pixels in the IBL sensors have a size of $50\text{ }\mu\text{m} \times 250\text{ }\mu\text{m}$, allowing for more precise tracking. Precise tracking is required to successfully reconstruct primary/secondary vertices and to distinguish from pile-up vertices. The spatial resolution of the original pixel detector is $10\text{ }\mu\text{m}$ in the $R\text{-}\phi$ -plane and $115\text{ }\mu\text{m}$ in the z -direction. The implementation of the IBL improves the spatial resolution in the z direction to $60\text{ }\mu\text{m}$.

The pixel detector is surrounded by the **semiconductor tracker** (SCT), which consists of four layers of silicon microstrip detectors in the barrel region and nine layers in the end-cap region, covering a pseudorapidity range of $|\eta| < 2.5$. Each sensor has a size of $63.6\text{ mm} \times 64\text{ mm}$ and is glued back-to-back to another sensor on a heat transfer plate at an angle of 40 mrad . Further alignment of parallel and perpendicular strips gives a spatial resolution of $16\text{ }\mu\text{m}$ in the $R\text{-}\phi$ -plane and $580\text{ }\mu\text{m}$ in the z -direction.

The SCT itself is surrounded by the **transition radiation tracker** (TRT). The TRT consists of 50 000 straw tubes in the barrel region and 320 000 straw tubes in the end cap region, each with a diameter of 4 mm. Each straw tube provides a drift time measurement when a charged particle passes through and ionises the gas mixture inside the tube, providing a spatial resolution of $130\text{ }\mu\text{m}$. A special feature of the TRT

is that it is interleaved with a specific transition radiation material. The amount of transition radiation produced depends on the mass of the passing charged particle, so the TRT can help to distinguish electrons/positrons from heavier particles, e.g. charged pions.

Combining all information from the previous discussed sub-detectors results in a transverse momentum resolution of charged particles of

$$\sigma_{p_T}/p_T = 0.05\% \cdot p_T[\text{GeV}] \oplus 1\%, \quad (4.6)$$

with $\oplus$ denoting an addition in quadrature.

4.2.3. Calorimeters

The calorimeters are designed to identify particles and measure their energy. The ATLAS calorimeter system[58] is designed so that the particles ideally lose all their energy in the calorimeters, with only muons and neutrinos leaving the detector. Since electrons and photons behave differently from hadrons in terms of energy loss, the ATLAS calorimeter system is divided into an electromagnetic calorimeter (ECAL) and a hadronic calorimeter (HCAL).

Electrons and photons lose energy mainly by bremsstrahlung and pair production, which influenced the design of the ECAL. It uses **L**iquid **A**rgon (LAr) as the active material and layers of lead as the absorbing material, arranged in an accordion geometry covering the full range of ϕ . The thickness of the lead layers is a function of the rapidity in order to optimise the energy resolution of the calorimeter. The ECAL is divided into a barrel section ($|\eta| < 1.475$) and an end cap on each side ($1.375 < |\eta| < 3.2$). The barrel part consists of two identical half-barrels separated by a 6 mm gap at $z = 0$. Each end cap of the ECAL is divided into an outer wheel of $1.375 < |\eta| < 2.5$ and an inner wheel of $2.5 < |\eta| < 3.2$. The granularity of the ECAL can be divided into a region with $|\eta| < 2.5$ with higher granularity and enhanced particle identification (e.g. γ/π^0 , e/π separation) and a region with lower granularity (the inner wheel of the calorimeter), which is still sufficient for jet reconstruction and measurement of the missing transverse momentum of the event. The thickness of the ECAL can be given in units of X_0 , where X_0 is the mean distance after which the energy of an electron is reduced to $1/e$ of its initial value. In the barrel region the ECAL has a thickness of about 22-23 X_0 , whereas the thickness of the end cap regions is between 24-38 X_0 . The relative energy resolution of the ECAL equals

$$\sigma(E)/E = \frac{10\%}{\sqrt{E[\text{GeV}]} } \oplus 0.7\%. \quad (4.7)$$

Hadrons lose energy mainly through inelastic hadronic interactions, and the thickness of the HCAL is optimised to absorb and measure the hadronic shower of particles, thus protecting the muon spectrometer from extensive radiation. Similar to X_0 for the ECAL, the characteristic interaction length λ can be defined for the HCAL. The tile calorimeter consists of steel absorber plates and scintillating tiles as active material and covers the barrel region ($|\eta| < 1.0$) and the extended barrel region ($0.8 < |\eta| < 1.7$) with a thickness of 7.2λ at $\eta = 0$. The end caps, however, use copper as the absorber and LAr as the active material. Each end cap is built from two discs, providing coverage for $1.5 < |\eta| < 3.2$. The relative energy resolution of the HCAL amounts to

$$\sigma(E)/E = \frac{50\%}{\sqrt{E[\text{GeV}]} } \oplus 3\%. \quad (4.8)$$

In order to measure particles with very high η , the range of the calorimeter system is extended with the so-called “forward calorimeter”. It also uses LAr as the active material, but the absorber layers differ in depth. The first absorber layer is made of copper to contain electromagnetic showers, while the later two layers are made of tungsten, optimised for hadronic showers. With the help of the forward calorimeter, particles with $3.1 < |\eta| < 4.9$ can be detected and measured with a relative energy resolution of

$$\sigma(E)/E = \frac{100\%}{\sqrt{E[\text{GeV}]} } \oplus 10\%, \quad (4.9)$$

which is crucial for a precise determination of the missing transverse momentum.

4.2.4. Muon spectrometer

The **muon spectrometer** (MS) [59] is the outermost subsystem of the ATLAS detector. It is designed to measure the p_T of electrically charged particles (in general muons because they are low-ionising particles) leaving the calorimeter system up to a pseudorapidity of $|\eta| < 2.7$. This is achieved by bending the trajectory of

the charged particles using strong magnetic fields and then measuring the trajectory. In the barrel region of $|\eta| \leq 1$, the magnetic field is provided by a system of eight superconducting air-core toroid magnets with a typical bending power of 3 Tm. At $1.4 \leq |\eta| \leq 2.7$, the magnetic field is provided by smaller superconducting air-core toroid end-cap magnets with a bending power of up to 6 Tm. In the pseudorapidity region of $1.0 \leq |\eta| \leq 1.4$ (called the transition region) there are no dedicated magnets, but a combination of the barrel and end-cap magnetic fields provide the magnetic field. The magnet arrangement, which can be seen in Figure 4.6, provides a magnetic field that is mostly orthogonal to the muon trajectories.

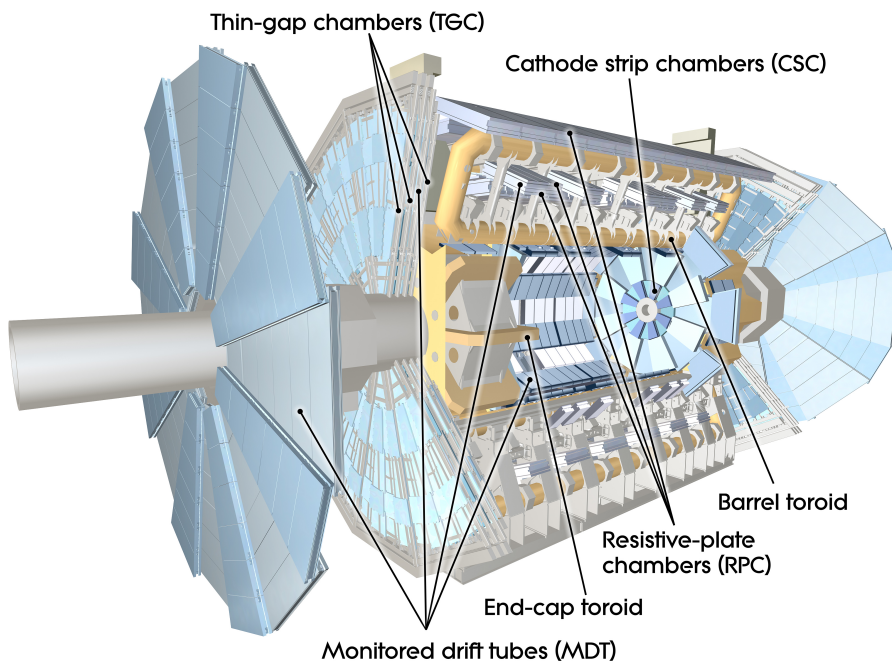


Figure 4.6: Cut-away view of the ATLAS muon spectrometer [60].

To measure the tracks, **M**onitored **D**rift **T**ubes (MDTs) are used in most of the pseudorapidity range. In the barrel region, the MDTs are arranged in three layers parallel to the z -axis, while in the transition region and in the end caps the MDTs are arranged in three layers perpendicular to the z -axis, thus allowing a track coordinate measurement in the principal bending direction of the magnetic field. These drift tubes are filled with an argon CO_2 mixture and have a maximum drift time of 700 ns with a spatial resolution of $50 \mu\text{m}$. As the maximum drift time of the MDTs is too long for high rate operation relevant at large pseudorapidities and close to the interaction point, **C**athode **S**trip **C**hambers (CSC) with a drift time of 40 ns and higher granularity are used.

A trigger system is installed in the pseudorapidity range $|\eta| \leq 2.4$. This trigger system not only allows triggering, but also provides a “second coordinate” measurement of the tracks in the non-bending direction, thus complementing the tracking information provided by the precision tracking chambers. In the barrel region, the trigger system uses **Resistive Plate Chambers** (RPCs), while in the end cap regions, **Thin Gap Chambers** (TGCs) are used. The RPCs consist of two parallel resistive plates, the volume between which is filled with a gas mixture. TGCs have a similar structure to multi-wire proportional chambers, but the density of the wires varies with η to match the required granularity.

4.2.5. Trigger system

Due to the high collision rate of 40 MHz and an average number of proton-proton interactions per bunch crossing of about 30, it is not feasible to record all events produced by the LHC. Moreover, the majority of the events consist of elastic pp scattering, rather than inelastic pp scattering. This is not of interest for analyses related to the ATLAS experiment, but rather for forward spectrometers such as ALFA[61]. To reduce the rate of events to be written to disk, a real-time trigger system [62, 63] is implemented to decide if an event was “interesting”, consisting of two stages.

The first stage of the trigger is a hardware-based trigger called “Level 1” (L1) which combines only partial information from the detector (i.e. from the calorimeters and the MS) to take no longer than 2.5 μ s to decide whether or not to keep an event. The L1 reduces the rate of potentially interesting events down to about 100 kHz. Events that are accepted by the L1 trigger are further processed by a software-based **high level trigger** (HLT). The HLT runs advanced reconstruction algorithms on regions of interest previously identified by the L1. Information from the full detector is only used in some cases to reconstruct physical objects. The HLT is designed to make a decision to keep or discard an event within a few hundred milliseconds, and is able to further reduce the number of events that are finally written to disk to about 1.2 kHz. This still gives a throughput to permanent storage of around 1.2 GB/s.

4.2.6. Luminosity measurement

Since knowledge of the delivered and recorded luminosity is essential for any analysis, ATLAS uses two independent subdetectors to measure the instantaneous luminosity. The first subdetector is the **Luminosity Cherenkov Integrating Detector 2** (LUCID2)[64], which is located at $z = \pm 17$ m at about 12 cm from the beamline,

corresponding to a pseudorapidity of about $\eta = 5.6$. It measures the Cherenkov radiation produced in thin quartz windows, which can be related to the instantaneous luminosity. The other subdetector is the **Beam Conditions Monitor** (BCM)[65], which consists of a total of 8 modules arranged symmetrically around the IP at $z = \pm 184$ cm and $r = 55$ mm to the beamline, corresponding to a pseudorapidity of $\eta = 4.2$. Each module contains two diamond sensors and fast electronics with a rise time of 2 ns. This allows a bunch-by-bunch luminosity measurement by counting in-time and out-of-time collisions. The absolute luminosity scale is then determined using van der Meer beam separation scans, taken at specific running periods in each year. The luminosity recorded by ATLAS under stable conditions and with all subsystems operating, including the absolute and relative uncertainties, is listed in Table 4.1[66].

Data sample	2015	2016	2017	2018	Combination
Integrated luminosity [fb^{-1}]	3.24	33.40	44.63	58.79	140.07
Total uncertainty [fb^{-1}]	0.04	0.30	0.50	0.64	1.17
Total uncertainty [%]	1.13	0.89	1.13	1.10	0.83

Table 4.1.: Integrated luminosity including its uncertainty per year of data taking and its combination.

Due to the partially correlated systematic uncertainties between the data-taking years, the integrated luminosity of Run 2 sums up to $\mathcal{L}_{int} = (140.1 \pm 1.2) \text{fb}^{-1}$.

5. Simulation of proton-proton collisions

Monte Carlo (MC) generators are employed to simulate the SM- and the BSM predictions. These MC predictions are then compared to the recorded data from the ATLAS experiment, as described in Section 4, to assess, whether the recorded data aligns with the SM or a BSM theory. This chapter explains how the proton-proton collisions are simulated. The process of simulating MC events involves different parts of the event generation (also shown in Figure 5.1), which are summarised in the following and subsequently explained in more detail.

- **Parton Distribution Functions:** At high energies, the colliding objects involved in the hard interaction are not the protons as a whole, but the constituents of the protons (called partons). **Parton Distribution Functions** (PDFs) provide the probability to find the partons in the proton as a function of the momentum fraction of the parton and the energy scale of the collision.
- **Hard Scattering Process:** The cross-section of the hard-scattering process can be calculated perturbatively in orders of the strong and weak coupling constant. The accuracy of the results can be improved by including higher orders, but the order of the perturbative expansion is computationally limited due to the increasing amount and complexity of diagrams to be considered.
- **Parton Shower:** Since the hard scattering process generally produces unstable high-energy particles, these particles decay further or radiate other particles. However, calculating the additional radiations and decays in the matrix element of the hard scattering process is only feasible up to a certain complexity. Therefore, this aspect is approximated using a parton shower algorithm. The radiation and decay of the particles using the parton shower continues until the hadronisation scale of $\simeq 1\text{ GeV}$ is reached. At this energy, α_S is no longer sufficiently small, resulting in colour-charged particles to form colourless hadrons. This transition is approximated by the parton shower generators through phenomenological simulations.
- **Detector Simulation:** Finally, the response of particles interacting with the detector sensors as well as the dead detector material (e.g. support structures, cable ducts, etc...) must be simulated in order to compare the measured data with the simulation.

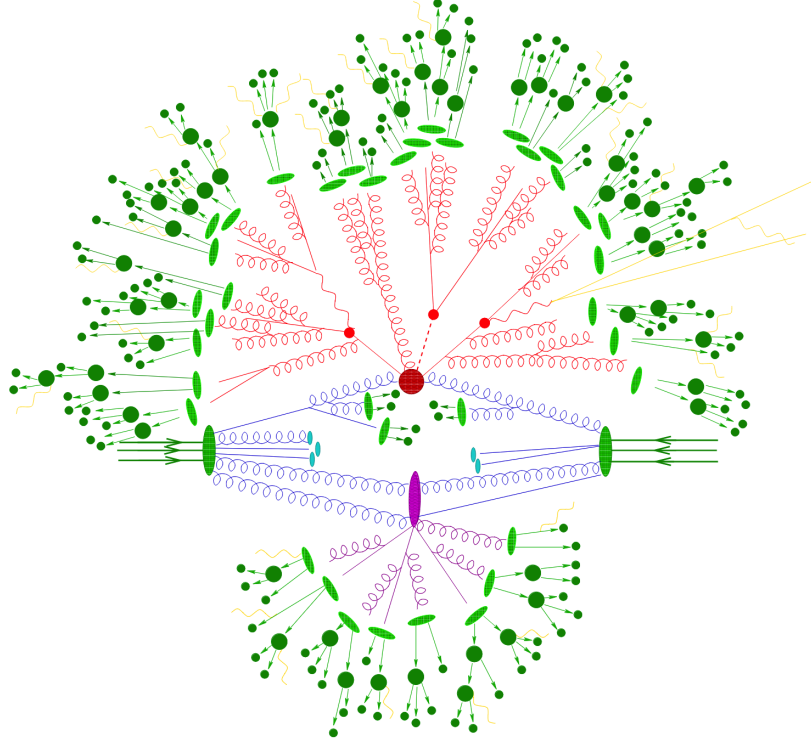


Figure 5.1: Representation of a proton-proton collision resulting in a $t\bar{t}H$ event as hard scattering process, denoted as large red blob. The parton shower is denoted with red lines, where the resulting particles at some point hadronise (light green blobs). The hadrons may decay further, until they decay into stable hadrons. Also the underlying event, evoked through the remnants of the colliding protons, and its resulting particles are shown in purple. Electromagnetic radiations are shown in yellow [67].

Because the different aspects of simulating MC events take place at different energies, the cross section of an arbitrary process $pp \rightarrow X$ can be calculated using the factorization theorem [68], where aspects of different energies can be separated/factorised:

$$\sigma(pp \rightarrow X) = \sum_{i,j} \int \int f_i(x_i, \mu_F^2) f_j(x_j, \mu_F^2) \cdot \hat{\sigma}_{ij \rightarrow X}(x_i x_j s, \mu_R^2, \mu_F^2) dx_i dx_j \quad (5.1)$$

The sum runs over the initiating particles i and j of the proton, namely quarks and gluons. $f_i(x_i, \mu_F^2)$ denotes the previously mentioned PDF of the proton for parton i which carries the momentum fraction x_i . μ_F is called the factorization scale, which is an somewhat arbitrary energy scale, separating the hard scattering process from

soft interactions. The PDF is dependent of the factorization scale, which can be seen in Figure 5.2 for two choices of factorization scale. However, since the PDF is an inherent characteristic of the proton, not depending on the experiment, the PDF can be measured at other experiments and their results can be used at the LHC. Furthermore, PDFs measured at an energy scale Q_1^2 can be transformed into PDFs at a different energy scale Q_2^2 using the **Dokshitzer–Gribov–Lipatov–Altarelli–Parisi** (DGLAP) evolution equations [69–71], therefore measurements of PDFs can be well combined to obtain precise predictions.

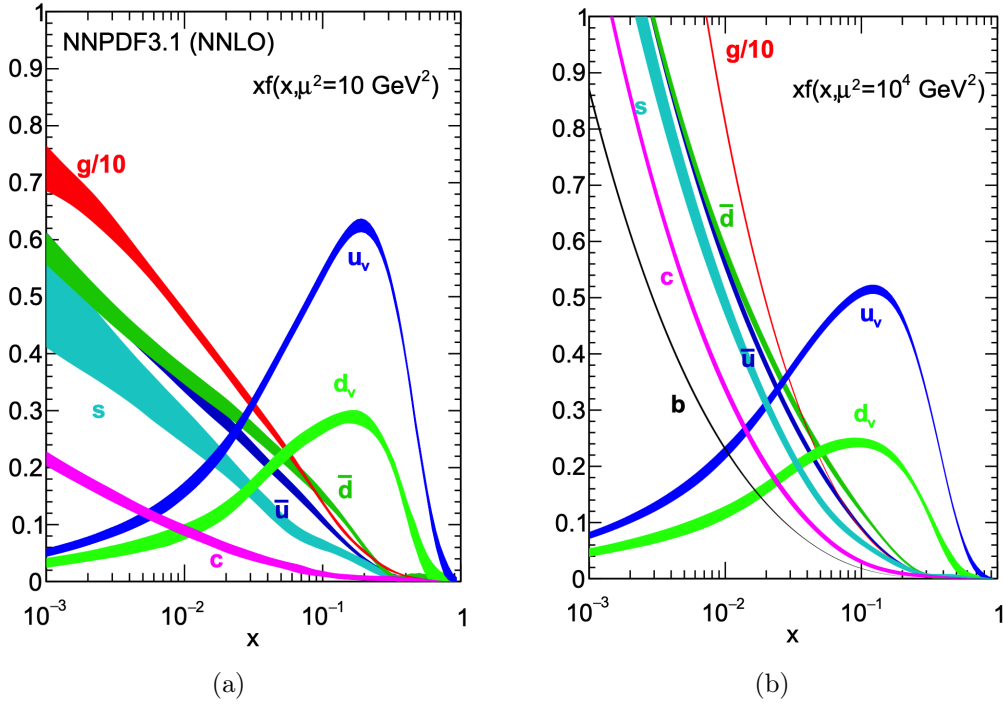


Figure 5.2: Results of a global fit by the NNPDF collaboration for the PDF of the proton at momentum transfer of (a) $Q^2 = 10 \text{ GeV}^2$ and (b) $Q^2 = 10^4 \text{ GeV}^2$ [72].

The other ingredient for the cross section calculation for the process $pp \rightarrow X$ in Equation 5.1 is the partonic cross section $\hat{\sigma}_{ij \rightarrow X}$, which can be calculated as follows:

$$\hat{\sigma}_{ij \rightarrow X} = \frac{1}{2x_i x_j s} \sum_{k=0}^{k_{max}} \int d\Phi_{X+k} \left| \sum_{l=0}^{l_{max}} \mathcal{M}_{X+k}^l(\Phi_{X+k}, \mu_F, \mu_R) \right|^2, \quad (5.2)$$

where s is the centre-of-mass energy squared. k represents the number of additional real emissions and l represents the number of virtual loops, whereas their

sum determines the order of the perturbative calculation. If $k = l = 0$, i.e. there are no additional real emissions or virtual loops, Equation 5.2 becomes maximally simple, called **leading order** (LO), but also potentially not very precise. Adding further real emissions and virtual loops enhances accuracy (**next-to leading order**, NLO, $k + l \leq 1$; **next-to next-to leading order**, NNLO, $k + l \leq 2$; etc.). However, this significantly increases computational demand, and hence NNLO is typically the highest order computed. $\mathcal{M}_{X+k}^l$ represents the **m**atrix **e**lement (ME) for the $ij \rightarrow X + k$ process, inclusive of l virtual loops, and is integrated within the permissible phase space Φ_{X+k} for the said process. An integration of the ME over the entire phase space Φ_{X+k} for $k > 0$ would include configurations where the additional emission could be collinear¹ or soft². Since this would lead to singularities in the calculations, the phase space is restricted at these regimes and a parton shower approximates such radiations in a later step. Furthermore, a renormalisation with a renormalisation scale μ_R is introduced such that $\alpha_S(\mu_R)$ is sufficiently small and higher orders of the perturbation theory lead to corrections in the order of $\alpha_S(\mu_R)^n$ only. In practice, the renormalisation and factorisation scale μ_R and μ_F may need to be selected somewhat arbitrarily. Nevertheless, an optimal choice reduces the impact of higher-order contributions that are not accounted for in fixed-order ME calculations[73]. In fact, if $k_{max} = l_{max} = \infty$, there would be no residual effect of the arbitrary choice of factorisation and renormalisation scale, but since this is not computationally possible, an effect of the choice remains.

Once the hard-scattering process is calculated, a **p**arton **s**hower (PS) is applied to simulate additional emissions by a $1 \rightarrow 2$ splitting of both the incoming (**i**nitial **s**tate **r**adiation, ISR) and outgoing (**f**inal **s**tate **r**adiations, FSR) partons, see Figure 5.3. Besides gluons, also quarks and photons can be radiated. In case of the FSR, the procedure of $1 \rightarrow 2$ splitting is repeated until around the hadronization scale of $Q_0 = 1 \text{ GeV}$, whereas for the ISR there is only a specific chance of a splitting to happen. However, when combining a ME generator with a PS, the risk of “double counting” of emissions exists, e.g. an additional radiation from the ME at NLO could correspond to the hardest emission from the PS of the LO ME. Every generator employs its own strategy to avoid such a double counting.

Once the partons are showered until approximately the hadronisation scale, colour-charged particles cannot be treated perturbatively any more and instead undergo confinement, leading to the formation of colourless hadrons. This process is modelled phenomenologically as it cannot be solved analytically nor perturbatively. Two commonly used models in MC generators are the Lund string model[74] and the cluster model[75]. In the Lund string model, a “string” connects colour-charged particles and represents the strong force. The potential of the string increases linearly with

¹The particle moves into the same direction as another particle in the process.

²The momentum of the particle tends to 0.

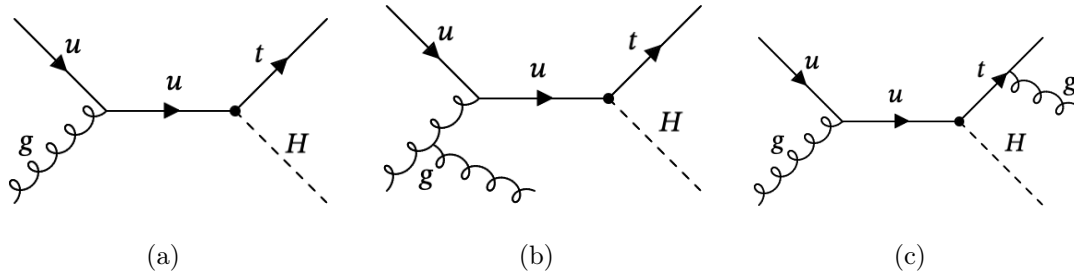


Figure 5.3: The LO Feynman diagram of the prod-FCNC process (a), with an ISR of a gluon (b) and an FSR of a gluon (c).

distance, until there is enough energy to create a new quark-antiquark pair from the string. In the cluster model, jet evolution is computed perturbatively in QCD until it consists of finite-mass, colourless clusters of quarks and gluons, called pre-confinement[76]. These clusters subsequently decay into smaller clusters or pairs of hadrons until only hadrons remain.

As a final stage of the MC production chain, simulating the interaction with the detector is necessary. This can be achieved either with a highly precise simulation of the interaction between the individual detector components and the particles or a quicker, though less precise, simulation. The former, known as *FullSim*, is integrated into the GEANT4[77] framework for simulating the passage and interaction of particles through matter across an energy spectrum spanning from 250 eV up to the TeV range. The drawback, however, is that a single event takes several minutes to be simulated in *FullSim*. Since simulating the calorimeter response is the most time-consuming aspect, the expedited alternative, called *FastSim*, or *AFII*[78], uses a parameterised calorimeter response to reduce the CPU time to a few seconds per event. This enables marginal samples of rare phase spaces to have sufficient statistics.

6. Object definitions

The raw data collected by the ATLAS experiment cannot directly be used for analyses as it only represents electrical signals from sensors throughout the detector. When an electron passes through the detector, for example, it produces an electric current above a certain threshold in the silicon semiconductors, an avalanche of ionised gas in certain drift tubes of the TRT and finally a cluster of electric signals in the ECAL. To conduct an analysis, objects must be inferred from the electrical signals, which is described in this chapter. The objects of concern in this analysis include electrons, muons, jets, and the missing transverse momentum, each of which will be defined with their respective quality criteria.

6.1. Tracks and vertices

The process of reconstructing the tracks and vertices of charged particles commences with the formation of *clusters* from energy deposits in the pixel and strip detectors [79]. In the $r - \phi$ plane, a traversing charged particle typically leaves an energy deposit in two adjacent pixels of the pixel sensors and in one to three adjacent pixels in the longitudinal direction, depending on η . A **connected component analysis (CCA)**[80] is first applied to form clusters from neighbouring active pixels. In the case of the pixel detector, each cluster corresponds to a space-point. However, for the SCT, a space-point can only be formed if both sides of a strip layer have a cluster. Due to the high occupancy of the detector, particularly near the IP, adjacent separate charged particles may produce clusters so close together that the CCA produces only a single cluster, see Figure 6.1. To identify and distinguish merged clusters from single-particle clusters, a neural network is employed combining information from the measured charge in the pixel sensors, the relative position of pixels in clusters, and the angle of particles provided by track candidates.

Once the space-points are given, track seeds are formed out of three space-points. This is done by ordering the track seeds by their purity (i.e. the proportion of seeds that result in good quality tracks), with SCT-only track seeds having the highest purity, followed by pixel-only track seeds and mixed track seeds. By utilizing a combinatorial Kalman filter [81], track candidates are constructed by integrating supplementary spatial information from the remaining layers of the pixel and SCT into the initial trajectory. A single track seed may produce several track candidates if a single layer has numerous space-point extensions. Once the track candidates are provided, a *track score* is computed for each candidate, indicating the probability

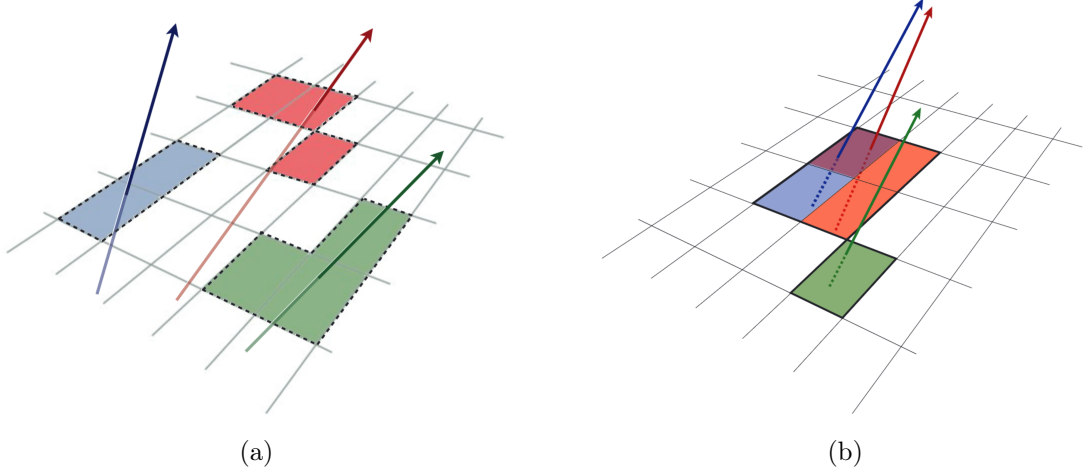


Figure 6.1: Clusters from the CCA for three charged particles traversing through a pixel sensor in case of (a) sufficient distance to have three distinct clusters and (b) insufficient distance to have partially merged clusters [79].

that it accurately depicts the path of a charged primary particle. The track score takes into account the number of clusters assigned to the track candidate, occurrences of holes (i.e. when the track candidate passes through the detector's active material without generating a signal), the χ^2 value of the track fit, and the logarithm of the track momentum. Then, beginning with the track candidate having the highest track score, an ambiguity solver compares merged clusters to track candidates, and, if necessary, eliminates clusters from track candidates. This is done when a cluster is used more than twice in separate track candidates or when a track candidate has more than two merged clusters. The track score is then recalculated for the track candidate, and the candidate is returned to the ordered list of remaining candidates. Furthermore, candidates for tracking are rejected by the ambiguity solver if they fail to meet any of the quality criteria outlined below:

- $p_T > 400 \text{ MeV}$
- $|\eta| < 2.5$
- At least 7 pixel and SCT clusters
- At most one shared pixel cluster or two shared SCT clusters on the same layer
- At most two holes in the pixel detector and the SCT together
- At most one hole in the pixel detector
- $|d_0^{\text{BL}}| < 2.0 \text{ mm}$
- $|z_0^{\text{BL}} \sin(\theta)| < 3.0 \text{ mm}$

with d_0^{BL} being the transverse impact parameter with respect to the measured beam-line position and z_0^{BL} being the longitudinal impact parameter denoting the difference along the beam line between the point where d_0^{BL} is measured and the primary vertex. The primary vertex is the location where the proton-proton collision took place, while θ represents the polar angle of the track. At this point, the primary vertex is extrapolated assuming a perfect helix of the track and a perfectly homogeneous magnetic field inside the detector. Successful candidates are extrapolated and combined with measurements from the TRT. As a final step, the track candidates that fulfil all requirements are fitted using a high-resolution fit that includes all available information about the track. The final track is determined by the fit, with parameters including the azimuthal angle ϕ , polar angle θ , ratio of charge to momentum $\frac{q}{p}$, and impact parameters d_0 and z_0 , representing the distance from the nominal IP in the transverse and longitudinal directions.

Since multiple collisions of protons occur in a single collision of a proton-proton bunch, it is crucial to distinguish these collisions by differentiating their primary vertices. To achieve this, an iterative procedure is used to determine the primary and secondary vertices once all the tracks have been calculated [82]. First, a seed position for a vertex is chosen based on the beam spot in the transverse plane and the z -coordinates of tracks at their respective points of closest approach to the reconstructed centre of the beam spot. Using this seed and all tracks, the optimal vertex position is computed through an iterative fit, excluding unsuitable tracks and eventually eliminating them from the procedure. Once the fit has determined the final vertex position, the previous steps are repeated with any tracks that do not match any of the previously determined vertex positions. After calculating all vertices, only those which satisfy certain quality criteria are kept. These criteria require that at least two tracks with $p_{\text{T}} > 500 \text{ MeV}$ are associated with a vertex. In the final step, vertices located within the estimated overlap of the proton bunches are considered primary vertices, while those outside of that region are counted as secondary vertices. The latter are subsequently used to identify B hadrons, as detailed in Section 6.5.

6.2. Electrons

Since electrons are charged particles that interact primarily electromagnetically, it is expected that they will leave tracks in the ID and deposit energy mainly in the ECAL. Therefore, electrons are reconstructed using ID tracks and corresponding energy deposits primarily from the ECAL [83]. To achieve this, calorimeter cells that are topologically connected and have a higher response than a certain noise threshold are first formed into what is known as “topo-clusters” [84]. These

topo-clusters are then aligned with tracks from the ID. As electrons generally emit bremsstrahlung when accelerated, bigger topo-clusters are paired with smaller “satellite” topo-clusters in their proximity to form so-called superclusters. The resulting superclusters are then refitted to tracks from the ID. As e.g. structural material of the ATLAS experiment in front of the calorimeters may cause undetected energy loss of the electrons, first a simulation-based correction procedure, then an adjustment of the absolute energy scale using $Z \rightarrow e^+e^-$ events is applied [85]. The universality of the obtained energy scale is then verified through the comparison with $J/\Psi \rightarrow e^+e^-$ events.

As there are various sources of misidentifying electrons (see e.g. Section 6.6), certain quality criteria are imposed on the resulting objects to reduce the contribution of incorrectly identified electrons. First, the previously obtained objects must have a pseudorapidity of $|\eta| < 2.47$, excluding the transition region from the electromagnetic barrel calorimeter to the electromagnetic endcap calorimeter of $1.37 < |\eta| < 1.52$. Additionally, the transverse momentum must meet the requirement of $p_T > 10$ GeV. Afterwards, a likelihood value is calculated using information about the general track quality (e.g. number of layers with a hit in the pixel detector and the SCT), transition radiation obtained from the TRT and the shape of the topo-clusters in the ECAL. On this likelihood value, a cut is applied on the internally called *TightLH* working point (WP) [86], which reduces the amount of incorrectly identified electrons by also sacrificing a certain amount of real electrons. Finally, to guarantee that the electrons originate from a primary vertex, a cut on impact parameters is implemented. Specifically, $|d_0/\sigma_{d_0}| < 5$ and $|\Delta z_0 \sin(\theta)| < 0.5$ mm must be met.

A further complication is posed by the misidentification of the charge of electrons. To reduce the amount of charge-misidentified electrons, the electron charge **ID** selector (ECIDS) tool [87] is used. The ECIDS tool utilises a boosted decision tree (BDT) to create a single score that incorporates various information about the electrons, such as their p_T and η as well as track and charge distribution data. In the BDT score, electrons with incorrectly identified charges are shifted to a BDT score of -1 , while correctly identified electrons are shifted to a BDT score of $+1$. A cut is then applied at the calibrated WP of a BDT score of -0.337671 to eliminate electron candidates that have a lower BDT score and are thus likely to have misidentified charge.

6.3. Muons

At the LHC, muons produced in the interactions are minimally ionising particles, usually passing through the calorimeter systems without losing all their energy,

creating tracks in the MS and then leaving the ATLAS detector. Consequently, while some information from the calorimeter system is used, muons are primarily reconstructed using tracks in the ID and MS [88, 89]. In the ID, muon tracks are reconstructed similar to the ones of electrons, see Section 6.2. In the MS, a Hough transform [90] is used to search for hits, which are expected to be aligned on a parabolic trajectory due to the magnetic field in the MS. Muon track candidates are then fitted from these hits in different layers (called segments), starting with a seed in the middle layer where more hits per segment are available. Hits from the inner and outer layer are then included to construct more complex track candidates. Because hits could potentially be used for more than one track candidate, an overlap removal algorithm consequently selects the best assignment of hits to an individual track or allows the hits to be shared between two tracks. The final MS tracks are then obtained by a global χ^2 fit to all hits in the MS and then satisfying certain selection criteria. Once the MS tracks are determined, four different muon types are reconstructed, depending on the information available from the other subsystems of the ATLAS detector:

Combined (CB) muons: After identifying tracks independently in the ID and the MS, CB muons are identified by extrapolating an MS track inwards into the ID and fitting it to an existing ID track. Subsequently, hits from the MS can potentially be removed during a fit to all hits of the ID and MS, to improve the fit quality. The complementary approach of extrapolating ID tracks and attempting to match them to tracks in the MS is also employed.

Segment-tagged (ST) muons: When muons cross only one layer of the MS due to low p_T or specific η regimes in the detector architecture, ID tracks may be classified as ST muons. This occurs when an extrapolated ID track matches a segment of hits in a single layer of the MS.

Calorimeter-tagged (CT) muons: When an ID track matches the calorimeter signature of a minimally ionising particle, it can be identified as a CT muon. CT muons have the lowest purity of all muon types, but they can recover acceptance in particular ATLAS detector regimes where the MS has a lot of dead material, e.g. due to cabling.

Extrapolated (ME) muons: A muon track may be identified as a muon even without any ID track if the extrapolated trajectory of the muon is compatible with originating from the IP. In this scenario, two or three layers of the MS must be traversed by the muon in the barrel and forward region, respectively. ME muons also restore acceptance of the detector, such as for $2.5 < |\eta| < 2.7$ where there is no coverage by the ID, but by the MS.

As there are four distinct definitions of muons, tracks may be used multiple times to reconstruct muons. Hence, a subsequent overlap removal process prioritises specific

types of muons over others. When two muon types have the same ID track, CB muons are given the highest priority, followed by ST muons and then CT muons. Conflicts with ME muons are resolved by analysing the MS tracks and selecting the track with the better fit quality and larger number of hits.

On the resulting objects, analogous to electrons, additional quality measures are applied. First, muon candidates must have a pseudorapidity of $|\eta| < 2.5$ and the transverse momentum must satisfy $p_T > 10 \text{ GeV}$. Additionally, $|d_0/\sigma_{d_0}| < 3$ and $|\Delta z_0 \sin(\theta)| < 0.5 \text{ mm}$ are required to ensure that the muons are originating from the IP. Beyond this, additional quality criteria based on the number of hits in the ID/MS and isolation aspects of the measurement are imposed, called *medium* WP [89, 91]. As with electrons, the energy scale of muons is calibrated and verified using $Z \rightarrow \mu^+ \mu^-$ and $J/\Psi \rightarrow \mu^+ \mu^-$ events [92].

6.4. Jets

Due to the confinement of QCD, colour-charged particles involved in the interaction eventually hadronise and produce a cone of hadrons and other particles, which can be detected in the ID and calorimeter systems. By analysing the characteristics of the jet, one can infer the particle that created it.

For the jets to be reconstructed, topological clusters in the calorimeters are used again, as in Section 6.2. Then, the particle-flow (PFlow) algorithm [93] combines information about the tracks in the ID and the topo-clusters to reconstruct hadronic jets and soft activity¹. This is done by first ordering tracks in the ID by their p_T and then matching the tracks to preferably single topo-clusters. However, it is common for particles to deposit their energy not only in a single topo-cluster but in multiple adjacent ones. Because of this, the PFlow algorithm calculates the probability for this to have happened based on the expected and observed energy distribution in the calorimeter and subsequently potentially adds topo-clusters to a certain ID track. Once the set of topo-clusters corresponding to a certain track is determined, a subtraction algorithm is performed to determine the energy which corresponds to the track. For this, the expected energy of the track is compared to the energy in the topo-clusters, which were matched to the track. If the expected energy exceeds the total energy of all matched topo-clusters of a track, then the topo-clusters are simply removed from the list of activity which has to be resolved in the calorimeter. Otherwise, a cell-by-cell subtraction is performed to subtract energy of cells from the expected energy, starting from the centre of the topo-cluster inside-out until the expected energy is reached. If the remaining energy in the

¹Activity in the calorimeters below the threshold used in the jet reconstruction, corresponding to hadronic recoil.

system is then consistent with the expected shower fluctuations, the remnants of the topo-cluster are removed. The output of the PFlow algorithm contains, in addition to a list of tracks and a list of remaining unmodified topo-clusters, also the set of new topo-clusters resulting from the energy subtraction algorithm. The ID tracks and the new set of energy-subtracted topo-clusters are then passed into a clustering algorithm to finally reconstruct the jets. This clustering algorithm uses the distances

$$d_{i,j} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \cdot \frac{\Delta_{i,j}^2}{R^2}, \quad (6.1)$$

$$d_{iB} = k_{ti}^{2p} \quad (6.2)$$

in an iterative procedure, where $d_{i,j}$ is a measure for the distance between entity i and j and d_{iB} a measure for the distance between entity i and the beam. k_{ti} denotes the transverse momentum of entity i , $\Delta_{i,j}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ (with y the rapidity and ϕ the azimuth angle) is a measure for the angular distance of the entities i and j , and R is called the radius parameter. In general, p can be chosen freely which imparts differing properties to the resulting algorithm. In this case, $p = -1$ is chosen, hence the name *anti- k_t* algorithm [94]. The choice of $p = -1$ results in an algorithm which is infrared and collinear safe, thus additional soft or collinear emissions do not alter the outcome of the algorithm. The clustering itself works by first calculating the distances $d_{i,j}$ and d_{iB} . If $d_{i,j} < d_{iB}$, entity i and j are recombined. If d_{iB} is smaller, then i is called a *jet* and removed from the list of remaining entities. The distances are recalculated and the procedure is repeated until no entities are left. The exemplary outcome of such a clustering is shown in Figure 6.2 with a radius parameter of $R = 1.0$, whereas in this analysis a radius parameter of $R = 0.4$ is used to cluster jets.

Once the calorimeter responses that compose the jets are determined, the **jet energy scale** (JES) must be calibrated to take into account detector effects and systematic errors during the simulation [95]. In a first step, excess energy from the jets is removed, which was caused by additional proton-proton collisions from in-time and out-of-time pile-up. These corrections are dependent of the in-time- and out-of-time pile up and the number of reconstructed primary vertices in the given event. Subsequently, the *absolute* JES is calibrated using truth-information from dijet MC events to correct the energy and the direction of the jet. Afterwards, the jet p_T resolution and its related uncertainties are improved by a so-called *global sequential calibration* by removing the dependence of the reconstructed jet response on observables, which are constructed using information from the tracking, the calorimeter and the MS. In a last step, an *in situ* calibration is applied on data to correct for

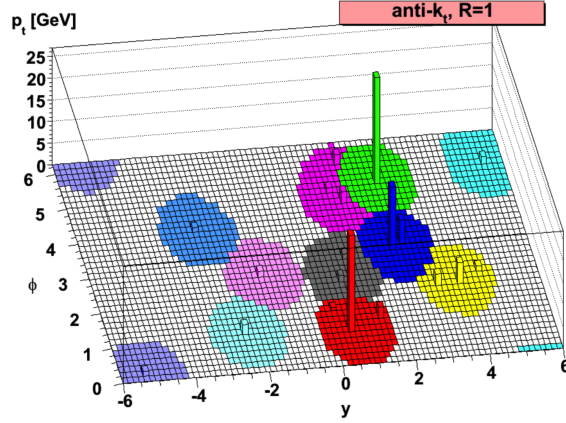


Figure 6.2: Clustering of jets in the rapidity- ϕ plane with the anti- k_t algorithm and a radius parameter of $R = 1.0$ [94].

remaining differences between data and MC simulations by comparing data to well-measured reference objects like e.g. Z bosons. Apart from the last correction, all corrections are applied to both data and MC simulation.

To reduce the impact of pile-up jets, a **jet-vertex-tagger** (JVT) [96] is used. The JVT is a discriminant combining information about the tracks forming the jet, yielding signal jet efficiencies of e.g. 80%, 90% and 95% with pile-up jet fake rates of 0.4%, 1.0% and 3% respectively. Differences in the JVT discriminant between data and MC simulation are compensated using scale factors derived from a tag-and-probe method in $Z \rightarrow \mu^+\mu^- + \text{jets}$ events. In this analysis, for a jet with $p_T < 60$ GeV and $|\eta| < 2.4$, a JVT output score above 0.5 is required, whereas jets in general must have a $p_T \geq 20$ GeV.

6.5. Flavour tagging

Hadrons, which are formed from b quarks, have a comparatively long mean lifetime of the order of 1.5 ps. As they move with nearly the speed of light, they can travel through the detector for a few millimetres before decaying, thus forming an additional secondary vertex at a distance from the primary vertex. This characteristic of b jets contrasts with jets resulting from lighter quarks, which promptly generate a shower of particles, with all tracks ideally pointing to the primary vertex. Figure 6.3 illustrates this behaviour. Using this information, it is possible to infer, with a certain probability and false-positive rate, which particle initiated a jet. The technique of flavour-tagging jets is fundamental for various analyses, including this one.

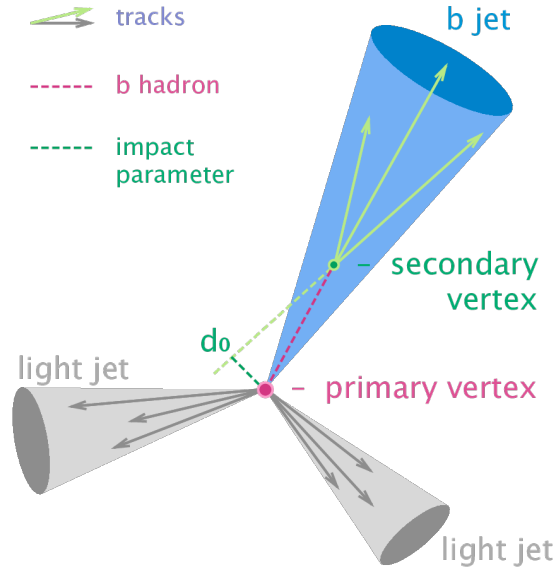


Figure 6.3: Schematic view showing main differences between light- and b jets, in particular the displaced secondary vertex [97].

For the identification of b jets, a powerful discriminant is obtained by combining the output of multiple low-level algorithms into a single deep neural network. The low-level reconstruction algorithms are:

IP2D/IP3D The impact-parameter-based algorithm (IP2D/IP3D) [98] solely utilises the trajectories of charged particles reconstructed in the ID and associated with jets. In particular, IP2D uses the transverse impact parameter significance as an input variable for its discriminant, while IP3D uses both the transverse and longitudinal impact parameter significance, taking into account correlations between the two input variables. The algorithms' output is based on a log-likelihood ratio (LLR).

SV1 The secondary vertex (SV1) [99] attempts to reconstruct a secondary vertex using an iterative χ^2 fit procedure. The method involves removing the track-to-vertex association with the highest χ^2 and conducting the fit repeatedly until an acceptable secondary vertex χ^2 is found. Applying several refinements and improved selection criteria of the track- and vertex candidate selection resulted in an overall improvement of the algorithm [99].

JetFitter The JetFitter algorithm [100] aims to fully reconstruct the b hadron decay chain. By using a modified Kalman filter [81], a line is searched on which the primary, the b -hadron and the subsequent c -hadron vertex lie. Even with just one track pointing to each the b - and c -hadron decays, this reconstruction is possible. Variables like the track multiplicity at the displaced vertices, the

invariant mass of tracks associated with the vertices, their energy fraction and their decay length are used as inputs.

The outputs of the aforementioned low-level algorithms are then combined into a high-level b -tagging algorithm called DL1[98, 101]. The DL1 is based on a deep feed-forward NN having a three-dimensional output corresponding to the probability for a jet to be a b -jet, a c -jet or a light-jet. The probabilities are referred to as p_b , p_c and p_{light} respectively. From these output values, the final DL1 b -tagging score can be calculated as follows:

$$D_{\text{DL1}} = \ln \left(\frac{p_b}{f_c \cdot p_c + (1 - f_c) \cdot p_{\text{light}}} \right), \quad (6.3)$$

with f_c being the effective c -jet fraction in the background training sample. The output of the DL1 score on the baseline $t\bar{t}$ sample for true b -jets, c -jets and light jets can be seen in Figure 6.4.

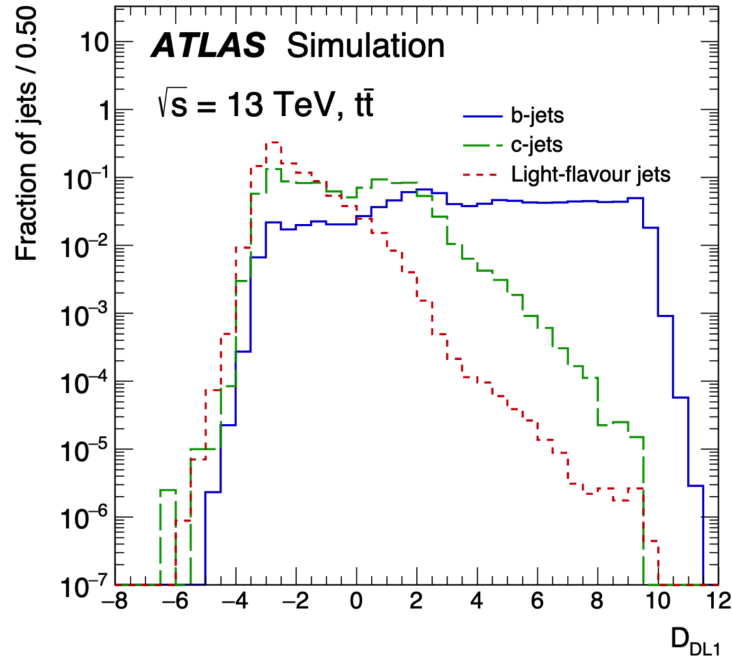


Figure 6.4: DL1 output discriminant for true b -jets, c -jets and light jets in the baseline $t\bar{t}$ sample [101].

Based on the DL1 discriminant, the performance of the b -jet identification can be evaluated at certain conventionally defined WPs ϵ_b , which refer to the fraction of true b -jets that are accepted as b -jets by the requirement on the output score. Table

ϵ_b	Selection	Rejection		
		c -jet	τ -jet	light jet
60%	>2.27	27	220	1300
70%	>2.02	9.4	43	390
77%	>1.45	4.9	14	130
85%	>0.46	2.6	3.9	29

Table 6.1.: Selection threshold of the DL1 score for true b -jets and the respective rejections of c -jets, τ -jets and light jets for selected WP [101].

6.1 displays this information, with “rejection” being defined as the inverse of the selection efficiency for each jet type.

As an additional input to the DL1 algorithm, a recurrent neural network (RNN) [102] is trained, where the variable length of the b hadron decay chain can be optimally analysed, yielding the DL1r algorithm. The DL1r output variable is then finally used in this analysis. The choice of a WP for an analysis depends on the respective phase space; a strict WP of $\epsilon_b = 60\%$ significantly reduces the amount of light-quark jets and c -jets falsely identified as b -jets by also sacrificing 40% of the actual b -jets and thus reducing the available statistics. A loose WP increases the amount of accepted b -jets, and thus usually increases the available statistics, by allowing more (also falsely identified) b -jets into the selected events. In this analysis a WP of

$$\epsilon_b = 70\%$$

is chosen. The choice of the specific WP is motivated by a study, documented in Appendix A.

6.6. Non-prompt lepton sources

In general, analyses involving multiple charged leptons in the final state have less statistics available. This is due to factors such as the branching ratio of bosons (W boson, Z boson, Higgs boson), which often appear in an intermediate state of the decay chain and prefer to decay into hadrons rather than charged leptons. As a result, analyses involving multiple charged leptons in the final state often exhibit notable contributions from backgrounds that mimic the detector signature of a charged lepton. These incorrectly identified leptons are referred to as fake

leptons and originate from various sources. The sources that impact this analysis are explained below:

HF-decay leptons: In the decay chain and fragmentation of a b quark in the detector, a shower of particles is produced, building a jet. In this jet of particles, electrons and muons can be produced, which are then deflected by the magnetic field thus potentially leave the reconstructed jet cone (see Figure 6.5). If this happens, the reconstruction algorithms can potentially misidentify the charged lepton from the jet shower as a prompt lepton¹. Another source is that a jet is not properly recognised as such (e.g. because its transverse momentum is below the threshold) and the charged lepton remains detected. In the case of a b quark initiating the jet of particles where the fake lepton originates from, the fake leptons are called **heavy-flavour (HF)-decay leptons**. HF-decay leptons build the largest source of fake leptons in this analysis.

Chargeflip e : The electron’s trajectory curvature in the ID is the key characteristic in identifying its charge. However, the bending radius of the trajectory is proportional to the velocity of the particle, meaning that a high energy electron will only have a slightly bent trajectory and can therefore potentially be identified with the wrong charge. Electrons with incorrectly identified charge are referred to as “Chargeflip e ” and their rate increases with higher transverse momentum. This problem is negligible for muons due to the additional tracking information provided by the MS.

Photon conversion: A photon can convert into an electron either through interactions near the IP (inner conversion) or interactions with detector material (external conversion). If this electron is misidentified as a prompt electron, it is classified as a “photon conversion electron”.

c - and LF-decay leptons: Similar to the HF-decay leptons but less probable, charged leptons may also potentially leave the cone of a c -jet or a light-flavoured jet (i.e. u -, d - or strange-jet) and be mistakenly identified as prompt leptons. These incorrectly identified charged leptons are referred to as c -decay or LF-decay leptons.

The identification of non-prompt leptons is necessary because the rate of appearance is often not well modelled by MC simulations and must be verified (and corrected if necessary) in a data-driven way. Reliably identifying non-prompt leptons is a major challenge in multi-lepton final state analyses. This analysis utilises the **PromptLeptonImprovedVeto (PLiV)** [103] tool, which combines information about

¹Prompt leptons are leptons originating from a “prompt” lepton source, i.e. coming from a W - and Z boson decay from the hard-scattering process or final state radiation / prompt quarkonium decays.

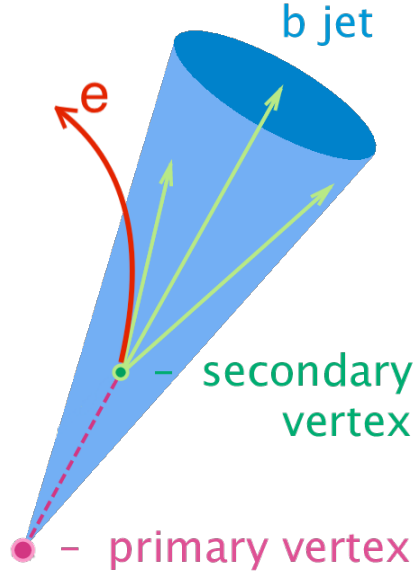


Figure 6.5: Schematic drawing showing an electron produced in the heavy-flavour decay of a b hadron leaving the jet cone.

the lepton isolation and lifetime about nearby jets in two separate BDTs, one for the barrel region and one for the endcap region. The isolation observables describe the energy depositions of particles in the vicinity of the lepton, e.g. the transverse energy/momentum of topological cell clusters in a cone with specific radius parameter around the lepton. Lifetime information of jets are usually used by b -tagging algorithms to search for a secondary displaced vertex; these information are also used by the PLiV algorithm to match lepton tracks to such vertices. After combining all these information in a BDT, multiple optimised WPs (as a function of the lepton p_T) are defined, whereby this analysis uses a cut on the so-called *PLImprovedTight* WP [104]. This WP is a compromise between high fake-lepton rejection and still retaining a high fraction of prompt leptons.

To classify charged leptons, retained by the PLiV tool, the *IFFTruthClassifier*[105] is used. This tool uses MC truth information (e.g. particles truth type, truth origin, etc.) to classify the leptons into categories describing their origin. These categories are:

Prompt electron: An electron originating from a prompt source (e.g. W/Z -boson decay). If the mother particle originated from a FSR photon or bremsstrahlung and then converted into an electron, it is also classified as prompt. The same is true for prompt quarkonium decays.

Prompt muon: Similar to the electron prompt classification.

Chargeflip electron: If the charge of the truth object is different compared to the reconstructed electron, its classified as charge flipped.

Prompt photon conversions: If the electron originates from a prompt photon, which is then converted, the electron is classified as a prompt photon conversion. No distinction is made between external and internal conversion.

Electrons from muons: If the truth-type of an electron corresponds to a non-isolated electron or photon and its truth-origin is a muon, the electron is classified as a reconstructed electron coming from a muon. However, due to the overlap removal procedure introduced in Section 6.8, this category is negligible.

τ decays: Similar to the “Electrons from muons” category, if an electron or muon is originating from an hadronic τ , it is categorised as such.

b -, c - or light-flavour decay: If an electron or muon originates from a b -, c , or light-flavoured meson or baryon, it is categorised as a b -, c - or light-flavour decay respectively. Next to the prompt leptons, these three categories are expected to be the most important ones.

KnownUnknown: In principle the lepton could be categorised as something from above, but certain MC truth information are missing to do so. However, the contribution from this category is negligible.

Unknown: If a lepton cannot be categorised into any category above, it is entering in the “Unknown” category. However, it is expected that this category is empty.

These categories will be used in a later step to define the non-prompt background processes in order to correct their rates in a data-driven way.

6.7. Missing transverse momentum

Because the beams in the LHC collide nearly head-on and due to momentum conservation, ideally the transverse momenta of all observed objects and tracks sum up to zero. Neutrinos, however, leave the ATLAS detector in general without being detected because they solely interact via the weak force. This leads to a momentum imbalance of the event, where the component in the transverse plane of the momentum imbalance can be detected by summing over all transverse momenta of all observed objects and tracks. The vector $\vec{E}_{x,y}^{\text{miss}}$ is thus defined as the negative of the vectorial sum of all reconstructed objects and soft tracks [106, 107]

$$\vec{E}_{x,y}^{\text{miss}} = - \sum_{i \in \{\text{hard objects}\}} \vec{p}_{x,y}^i - \sum_{j \in \{\text{soft tracks}\}} \vec{p}_{x,y}^j \quad (6.4)$$

where i is summed over all reconstructed objects and j is summed over all tracks, which are associated to the respective primary vertex but are not reconstructed as part of any calibrated object. The magnitude of $\vec{E}_{x,y}^{\text{miss}}$, calculated by

$$E_T^{\text{miss}} = \sqrt{(E_x^{\text{miss}})^2 + (E_y^{\text{miss}})^2}, \quad (6.5)$$

is called missing transverse momentum. Besides not having the longitudinal component of the E_T^{miss} , an additional drawback is that if multiple neutrinos are produced in an event, only the sum of all neutrino transverse momenta can be measured, not the individual components.

In the summation over the hard objects, it is important to only use mutually exclusive detector signals to avoid any double counting in the calculation. When there is an overlap between hard objects, priority is given to electrons, followed by photons, hadronically decaying τ -leptons and then to jets. Since muons are mostly determined by MS signals, the overlap of objects with muons is negligible.

6.8. Overlap removal

Because it is potentially possible to have used certain detector signals from the calorimeters or the ID multiple times to reconstruct different objects, a procedure called *overlap removal* avoids a double counting of signals. This is done by removing objects in the following order, often using the angular distance $\Delta R = \sqrt{(\Delta y)^2 + (\Delta \phi)^2}$ as measure:

- If an electron shares a track in the ID with a muon, the electron is removed.
- If a jet is within $\Delta R = 0.2$ of an electron, the jet is removed.
- If an electron is within $\Delta R = 0.4$ of a remaining jet, the electron is removed.
- If a jet is within $\Delta R = 0.2$ of a muon and the jet has two or less associated tracks, the jet is removed.
- If a muon is within $\Delta R = 0.4$ of a remaining jet, the muon is removed.

Removing a jet, which is within $\Delta R = 0.2$ of a lepton avoids double counting of energy deposits or tracks, whereas removing leptons within $\Delta R = 0.4$ of a jet reduces the rate of non-prompt leptons.

7. Data and simulated samples

This section provides a qualitative description of the recorded dataset and the signal- and background processes used, including the ME and PS generator used for the simulation. A summary of the samples including their `dataset identifier` (DSID)¹ and cross sections can be seen in Appendix B. Furthermore, the splitting and merging of the background processes into the non-prompt templates is explained.

The used ME generators are either `Powheg Box` [108–111] or `MadGraph_aMC@NLO` [112], both of which are capable of calculating QCD corrections at NLO. ME generators require an input PDF, which can either be `NNPDF3.0NLO` or `NNPDF2.3LO` [113]. Both ME generators can be interfaced with `MadSpin` [114, 115] to decay particles while preserving all spin correlations. `Pythia 8` [116] and `Herwig 7` [117, 118] are generators for PS simulation, which are executed subsequently to the ME generators. However, `Pythia 8` uses a tune of internal hyperparameters, tailored to the PDF set used. This set of tuned parameters is called “**ATLAS 2014**” (A14) [119]. `Sherpa` [120] is both a ME generator and a PS simulation.

7.1. Recorded data and trigger

As already mentioned in Section 4.2.6, the integrated luminosity of the recorded data with stable run conditions and all subsystems working properly sums up to $\mathcal{L}_{\text{int}} = (140.1 \pm 1.2) \text{fb}^{-1}$ for Run 2. This analysis then uses a single-lepton trigger which requires at least one charged lepton with a $p_T > 26 \text{ GeV}$. The possibility of using a di-lepton trigger was considered, which would allow the inclusion of events with at least two leptons where the leading- p_T lepton may have a p_T less than 26 GeV. However, because of the high contamination of HF-decay leptons in the low- p_T regime, it was decided to only use single-lepton triggers.

7.2. Signal FCNC samples

The prod-FCNC process describes the conversion of an up- or charm quark into a top quark by emitting a Higgs boson, which can be seen as Feynman diagram in Figure 7.1.

¹The DSID is an ATLAS intern numbering scheme for simulated samples.

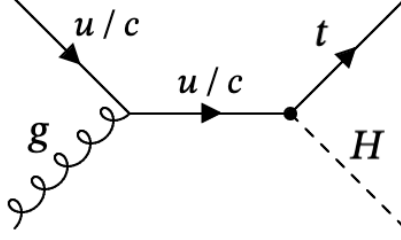


Figure 7.1: Feynman diagram of the prod-FCNC process, where a high-momentum up- or charm quark is converted into a top quark by emitting a Higgs boson. The black dot represents the FCNC vertex.

The ME of this process is calculated at NLO using `MadGraph aMC@NLO` in the 5 flavour scheme¹. The subsequent decay of the top quark to a b quark and a W boson where the W boson decays leptonically is performed in `MadSpin` to maintain spin correlations. The decay of the Higgs boson, as well as the PS, is simulated in `Pythia 8`. However, the three relevant decay modes of the Higgs boson, yielding at least one charged lepton ($H \rightarrow WW^*$, $H \rightarrow ZZ^*$ and $H \rightarrow \tau\bar{\tau}$) are split into three separate samples. One Wilson coefficient for the left/right-handed uH/ctH coupling is set to $C = 1$ with all other Wilson coefficients set to zero, resulting in four different processes with each process having three different samples (decay modes of the Higgs). Each sample consists of 10 million MC events passing an at-least-di-lepton filter, which are then interfaced with the *fullSim* detector simulation. To normalise the cross section, the cross section prediction of the ME generator with $C = 1$ is used. The left- and right-handed samples are subsequently combined and their cross sections are averaged. This method is employed as no noticeable distinctions in outcome are detected based on handedness. Appendix C also demonstrates no significant disparities in kinematic distributions between the left- and right-handed uH production samples.

The decay-FCNC process is similar to the $t\bar{t}$ production, but either the top quark or the top antiquark then decays into an up- or charm quark and a Higgs boson via the FCNC vertex. The Feynman diagram of the decay-FCNC process can be seen in Figure 7.2. `Pythia` with the A14 set of tuned parameters and the NNPDF2.3NLO PDF set was used for the ME generation and simulation of events at NLO, interfaced with `Pythia 8` as PS. In contrast to the prod-FCNC process, only left-handed FCNC interactions are simulated since it is not expected to see any differences in the kinematic distributions in $t\bar{t}$ events, based on the handedness. However, due to

¹In the 5 flavour scheme, in contrast to the 4 flavour scheme, the b^- and the b antiquark distribution is present in the proton PDF.

technical reasons, the decay-FCNC process is divided into two samples for both the utH and the ctH processes each.

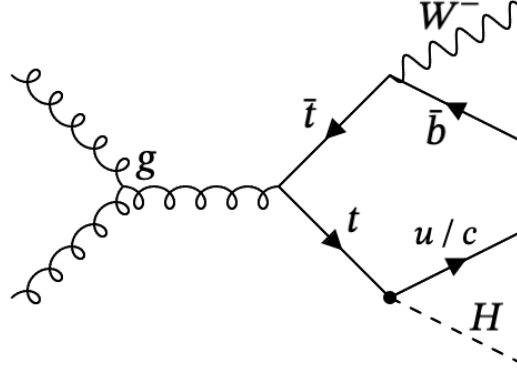


Figure 7.2: Feynman diagram of the decay-FCNC process, where the top quark from the $t\bar{t}$ production decays using the FCNC vertex, represented as black dot.

In the one sample, the top quark decays via the FCNC vertex into an up/charm quark, whereas in the other sample the top antiquark decays via the FCNC vertex. The remaining top antiquark (in case of the FCNC-decaying top quark) or top quark (in case of the FCNC-decaying top antiquark) decays according to the SM. Both samples, split by the top-quark decay, are then later combined to create the utH/ctH decay-FCNC sample. The Higgs boson decays inclusively in all cases. In total, each sample is generated with 37 million MC events that have passed an at-least-di-lepton filter. Subsequently, they are interfaced with the *fullSim* detector simulation. The normalisation of the decay-FCNC sample is done using Equation 3.4 and 3.5 together with the most-recent cross-section calculation for $t\bar{t}$ production at NNLO+NNLL with its central value of $\sigma_{t\bar{t}} = 833.9$ pb [121–123].

7.3. Background samples

Top quark-antiquark pair: The $t\bar{t}$ background is generated using *Powheg* and *Pythia* 8 as ME and PS generator respectively. $t\bar{t}$ enters the considered selections almost exclusively via non-prompt leptons: Either by charge-flipped electrons or HF-decay leptons. To optimise and speed up the workflow of the analysis, depending on the considered selection, $t\bar{t}$ samples with varying lepton filters are used: The $2\ell SS$ regions use a $\geq 1\ell$ filter on their $t\bar{t}$ samples to take into account HF contributions and charge-flipped electrons as well, whereas the 3ℓ regions use a $\geq 2\ell$ filter on their $t\bar{t}$ samples to have sufficient statistics and still take into account events with HF leptons. Taking into account events

with one prompt lepton and two non-prompt leptons in the 3ℓ regions is not necessary because their contribution is negligible.

single-top quark: The single-top quark production processes include the t -channel, the s -channel and the tW -channel and are produced using **Powheg** and **Pythia 8**. The single-top production processes have separate samples for the production of a top quark and of the top antiquark. These processes also contribute mostly via non-prompt leptons.

V + jets: W +jets and Z +jets, composing V +jets, are generated using **Sherpa 2.2.11** at NLO in perturbative QCD (pQCD) for 0-2 jets and at LO for more than two jets. V +jets only includes leptonic decays of the W - and Z boson but still mainly contributes via non-prompt leptons. The samples are split by the flavour of the respective vector boson's leptonic decay products and further grouped based on the flavour of the jets.

$t\bar{t}X$: The associated production of a W boson, Z boson or a Higgs boson together with a top quark-antiquark pair are denoted $t\bar{t}W$, $t\bar{t}Z$ or $t\bar{t}H$ respectively and together form the $t\bar{t}X$ background. These background processes contribute mainly by prompt leptons and together with their real b -jets in the final state, they are very similar to the signal processes in terms of their event signature, thus forming one of the major prompt background processes. $t\bar{t}W$ is generated using **Sherpa 2.2.10**, however, the cross section prediction for the $t\bar{t}W$ process is computed using **MadGraph aMC@NLO FxFx** [124]. $t\bar{t}Z$ is generated using **MadGraph aMC@NLO** at NLO with **Pythia 8** as PS and $t\bar{t}H$ is generated using **Powheg** and **Pythia 8**.

Diboson and Triboson: The Diboson and Triboson background denotes the production of two and three massive vector bosons respectively. Both backgrounds are generated and showered using **Sherpa 2.2.12**, split by their amount of leptons. The Diboson and Triboson processes are the second largest source of prompt background processes in this analysis. However, for the Diboson process to contribute significantly to the 2ℓ SS or 3ℓ regions, the vector bosons must decay leptonically which means that no b quarks are present on ME level. b -jets thus must be either falsely identified or b quarks must be produced in the PS. The rate of b -jets produced in the PS, however, is in general not well modelled, which often leads to an underestimation of MC prediction compared to data. This is why the Diboson sample is further split, not only by the number of leptons generated at ME level, but also by the truth flavour of the heaviest jet in an event. The flavours are then combined to a b/c sample and a $\tau/light$ sample per number of charged leptons. Doing so, the not-well-estimated rate of the additional radiations of b quarks in the PS can be accounted for in a later step of the analysis.

tWZ and tZq: The associated production of a top quark and a Z boson together with either a W boson or another quark is denoted tWZ and tZq respectively. Both samples are produced inclusively using MadGraph aMC@NLO with Pythia 8 as PS.

VH production: The associated production of a Higgs boson together with either a W - or a Z boson are denoted as VH . This background process is generated using Powheg at NLO in pQCD and Pythia 8 as PS.

Rare top-quark processes: Further minor processes are considered, which involve at least one top quark, namely $t\bar{t}t$, $t\bar{t}t$, $tWHq$ and $t\bar{t}XX$, with XX being WW , ZZ , HH , WZ and WH . All these backgrounds are simulated at NLO in pQCD (except $t\bar{t}XX$, which is generated at LO only and its cross section is normalised to the NLO prediction) using MadGraph aMC@NLO with Pythia 8 as PS.

7.4. Building the non-prompt lepton templates

To take into account and also to compensate the non-optimal rate prediction of non-prompt leptons, the previously introduced backgrounds are split, depending on the true origin of the leptons in a given event. Event by event for all background processes, the truth origin of the charged leptons is examined using the IFF **truth classifier** (see Section 6.6). If all charged leptons are either a prompt electron or a prompt muon, this event is collected in the respective prompt process. Otherwise, if exactly one electron or one muon originated from a b decay or one electron is charge-flipped, then this event is collected in the **HF-decay e**, **HF-decay μ** or the **charge-flipped e** template respectively. In the rare case of a different source of non-prompt leptons, the event is collected in one of the following templates:

Photon conversion e: This template collects events from all processes where one electron originates from a photon conversion.

Electron from τ - or μ decay: If exactly one electron originates from a τ - or μ decay, the event is collected in this template.

Electron or muon from a LF- or c decay: If exactly one electron or muon was produced in a light-flavour- or c -hadron decay, the event is collected in this template.

Known Unknown This template collects any other source or combination of non-prompt leptons, including rare events with two non-prompt leptons or events with insufficient MC information to be successfully classified. However, the contribution of this template is very small.

An schematic of the splitting of the processes can also be seen in Figure 7.3.

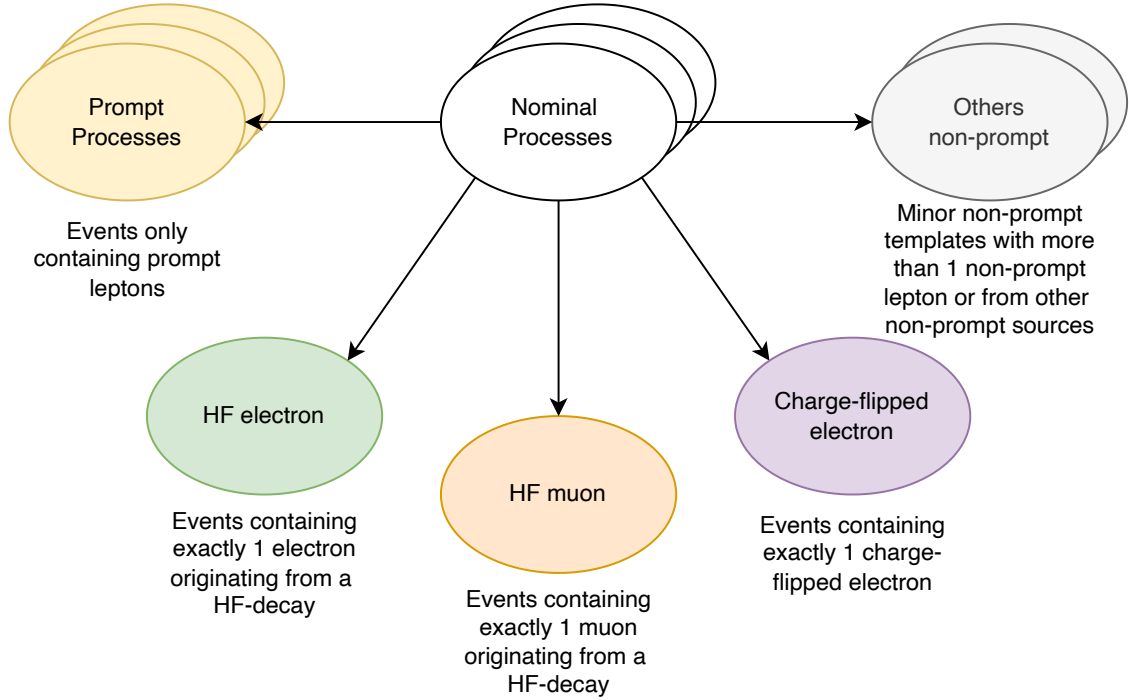


Figure 7.3: Schematic of how the splitting of background processes into different templates works. Events of the nominal background processes are sorted either into the respective prompt process, into the HF-decay electron/muon or charge-flipped-electron process or into other small non-prompt processes, depending on the truth-origin composition of the leptons in the event.

To summarise the process of MC sample splitting: Each background process has its corresponding prompt template kept separately, while events containing non-prompt leptons from different samples are combined into the categories listed above. This method causes no loss or double-counting of MC events, it is simply rearranging the MC events into distinct categories. Signal FCNC processes do not undergo the splitting. The benefit of amalgamating the non-prompt events from all background processes into the non-prompt templates is that these templates can be effectively managed in a binned profile likelihood fit. This means that their inaccurately predicted rate can be adjusted with specific normalisation factors. In Appendix D, figures can be seen showing the composition of the non-prompt templates by background process.

8. Event selection and region definition

A preliminary selection of events is carried out by imposing specific criteria, such as requiring a certain number of charged leptons and a minimum transverse momentum of objects, to eliminate as much background as possible and retain a maximum amount of signal. A basic event selection not only enhances the signal-to-background event ratio, thus improving the analysis sensitivity, but also significantly reduces processing time as a result of fewer MC events to process. Subsequently, specific sections of the phase space, referred to as “regions”, are defined by further selection criteria. The principal goal of a region may be to be enriched in signal events. Such a region is called a “**Signal Region**” (SR) and primarily improves the sensitivity of an analysis to the signal processes. However, backgrounds in the SR are occasionally not well understood. It is then necessary to correct the shape/rate of a background in an observable. To perform such a correction, “**Control Regions**” (CRs) are introduced, which are enriched in the specific background which is to be *controlled*. CRs must be orthogonal to the SRs but still have a similar phase space for the corrections to be applicable in the SRs. In this analysis, CRs are defined for non-prompt lepton backgrounds and for the $t\bar{t}W$ and $t\bar{t}Z$ backgrounds. In general, regions are defined so that they are *orthogonal* to each other, i.e. have no overlap of events. This is to avoid statistical correlations between the regions, which would be complicated to properly describe in the later statistical analysis. This chapter first describes the preliminary event selection and subsequently defines and motivates the regions containing three charged leptons (3ℓ) and two same-sign charged leptons ($2\ell\text{SS}$). Particular focus is given to the 3ℓ regions. Event yields including pie charts for all regions can be seen in Appendix E. The minimal criteria, objects must fulfil regardless of the region, are listed below. However, the criteria might be even more stringent for objects in certain regions.

Charged leptons: Charged leptons must fulfil $p_T \geq 10 \text{ GeV}$ and $|\eta| \leq 2.5$.

Jets: Jets must fulfil $p_T \geq 20 \text{ GeV}$ and $|\eta| \leq 2.4$.

b jets: b jets must pass the 70% WP of the DL1r b -tagging algorithm.

The optimization of the binning of regions is another important aspect, as the number and width of bins have an influence on the statistical uncertainty and thus influence the sensitivity of the entire analysis. *Few* bins have the advantage of low statistical uncertainty in the bins, however, *too few* bins can no longer accurately

reflect the shape of distributions and sensitivity is lost. A compromise must thus be made, which is done individually for each region.

8.1. Tri-lepton regions

The 3ℓ regions are defined by three charged leptons (electron or muon) with an electric charge adding up to ± 1 . HF-, c - and light-decay leptons, which form the major contribution of non-prompt leptons, tend to have lower transverse momentum, since they are not created in a decay of a weak boson but instead are decay products from b , c or light hadrons themselves. This behaviour can be seen in Figure 8.1, which shows the distribution of transverse momenta of the three leptons in a preliminary SR. In general, leptons are sorted according to their transverse momentum, with the lepton having the highest p_T in an event being called the *leading lepton* ℓ_0 , the lepton with the second-highest p_T being called the *sub-leading lepton* ℓ_1 , etc.

It can be seen that for the leading lepton there is only a small difference between the the FCNC processes and the non-prompt processes. However, for the second-leading but especially for the third-leading lepton, the non-prompt processes tend to lower transverse momenta, compared to the FCNC processes. To avoid an unnecessary high contamination of the non-prompt backgrounds in the SRs and because of the given p_T requirement of the triggers used, it is required for all 3ℓ regions that

$$p_T(\ell_0) \geq 28 \text{ GeV} \quad \text{and} \quad p_T(\ell_1) \geq 20 \text{ GeV}.$$

For the SRs, the minimum p_T of the third-leading lepton is required to be $p_T(\ell_2) \geq 16 \text{ GeV}$, since this removes a significant part of the non-prompt backgrounds and yet retains most of the signal. Furthermore, the cut on the p_T of the third-leading lepton allows for orthogonal CRs in a later step. The remaining cuts defining the SRs are made to be less restrictive to retain as much signal as possible. In total, there are two distinct SRs in the 3ℓ regions, which are derived from the two FCNC processes. One region is optimised for the prod-FCNC process (called SR3 ℓ Prod) and one is optimised for the decay-FCNC process (called SR3 ℓ Dec). Because on ME level, both FCNC processes possess a b quark (see Figure 8.2), $N_{b\text{-jets}} = 1$ is required at a 70% WP of the DL1r b -tagging algorithm for both SRs. It can also be seen that the decay-FCNC process possesses an additional u or c quark in its final state, compared to the prod-FCNC process. However, due to radiations of quarks in the PS

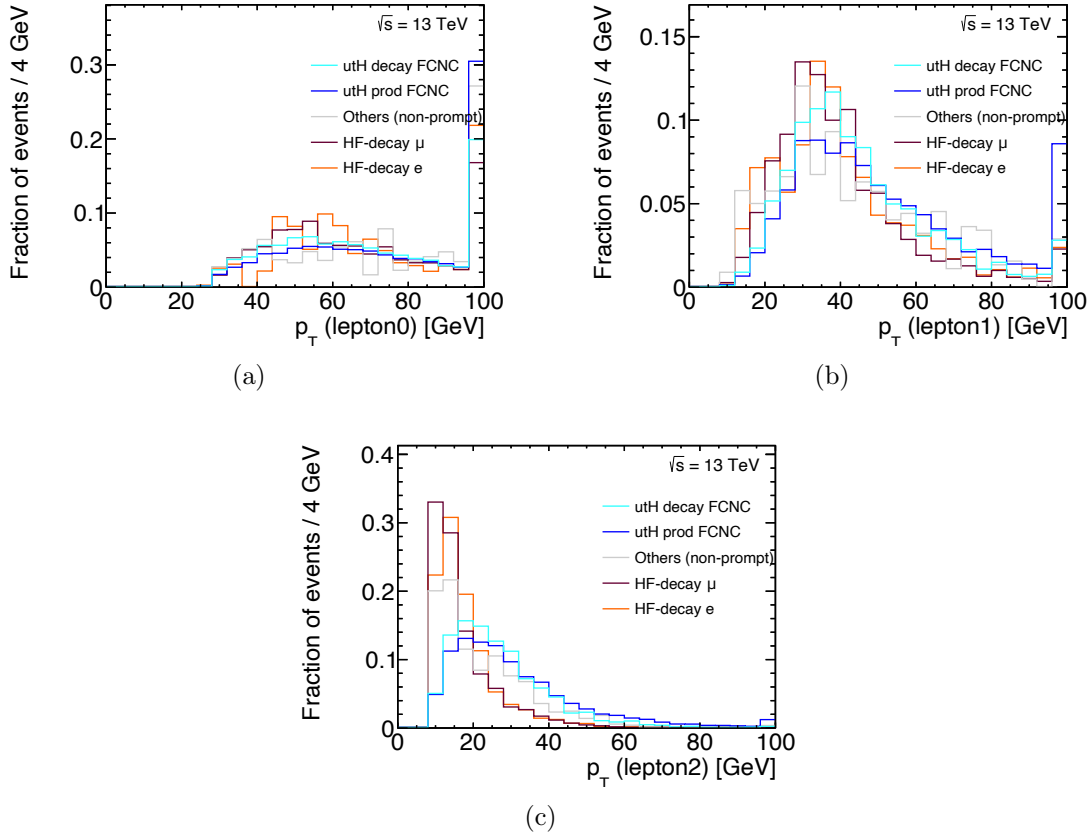


Figure 8.1: The transverse momentum of the charged leptons in a preliminary region, showing an enrichment of HF leptons in the low- p_T regime. The cuts for the preliminary region are one b -tagged jet and three leptons with $p_T > 10$ GeV each.

resulting in additional jets, the SR3 ℓ Prod requires $N_{l\text{-jets}} \leq 1$, counting c -jets also as l -jets. The SR3 ℓ Dec on the other hand requires $N_{l\text{-jets}} \geq 2$.

The shown decay of the Higgs boson in Figure 8.2 is just exemplary but likely to get a final state with three charged leptons for the FCNC process, as the $H \rightarrow WW^*$ decay channel constitutes around 70% of all decay channel of the FCNC processes. The exact breakdown of the decay channel of the Higgs boson, contributing to the 3 ℓ SRs, can be seen in Figure 8.3 for the utH process. Differences in the composition of the utH FCNC process depending on the region can be seen, i.e. requiring zero or one non- b -tagged jets (SR3 ℓ Prod) results in a significant increase in events from the $H \rightarrow WW^*$ channel at the expense of the $H \rightarrow ZZ^*$ channel, compared to requiring 2 or more non- b -tagged jets (SR3 ℓ Dec). Very similar behaviour can be observed in the case of the ctH FCNC process, which is not shown separately. Small differences can be observed when comparing the FCNC process composition of the 3 ℓ SRs

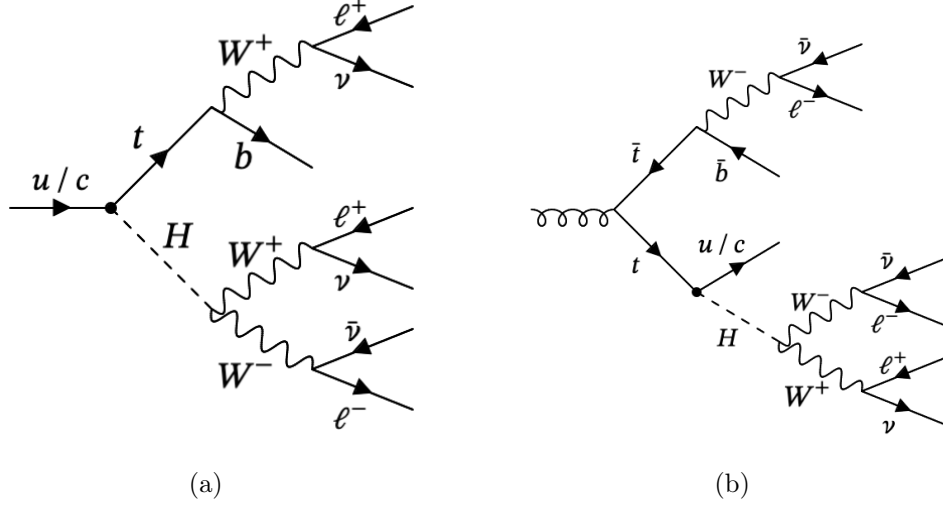


Figure 8.2: The prod-FCNC process (a) and the decay-FCNC process (b) with the $H \rightarrow WW^*$ decay channel resulting into three charged leptons.

with the 2ℓ SS SRs, however, the $H \rightarrow WW^*$ channel is always the most important channel with shares of around 75% to 80% in the 2ℓ SS SRs.

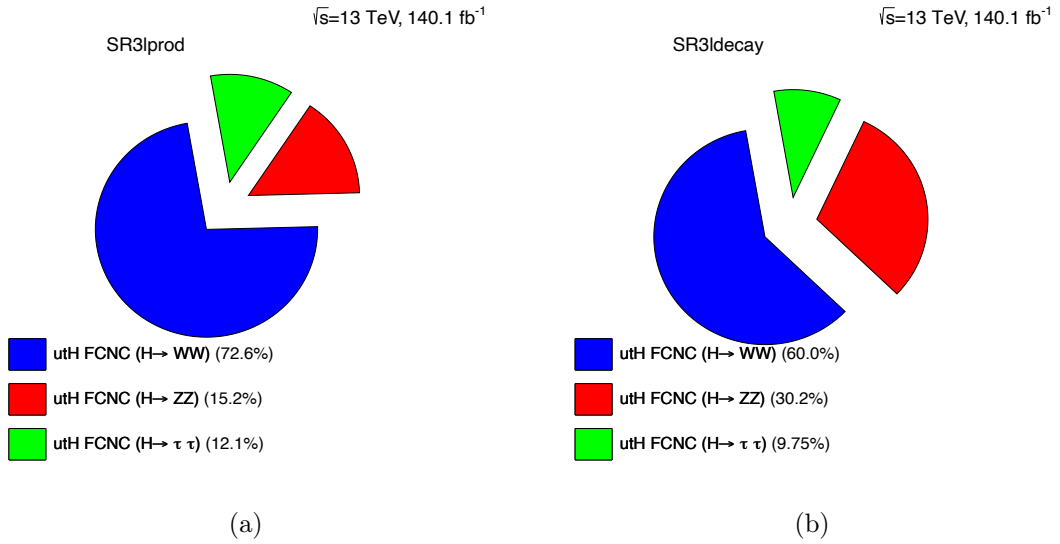


Figure 8.3: Composition of the utH FCNC process by the $H \rightarrow WW^*$, $H \rightarrow ZZ^*$ and $H \rightarrow \tau\tau$ decay channel in the SR3lProd (a) and in the SR3lDec (b).

Since it is expected that the rates of non-prompt processes are not well predicted by MC generators, two CRs are introduced, one for the HF-decay e process (CR3HF e)

and one for the HF-decay μ process (CR3 ℓ HF μ). They are orthogonal to the SRs by requiring the third-leading lepton to fulfil $10 \text{ GeV} \leq p_T(\ell_2) \leq 16 \text{ GeV}$. The third-leading lepton is then required to be an electron or a muon for the CR3 ℓ HF e and the CR3 ℓ HF μ region respectively. Because the HF-decay leptons are more likely to be the third-leading lepton of a given event since HF-decay leptons tend to have a small transverse momentum, applying this requirement ensures an enrichment of HF-decay events of the respective process in the HF CRs. The transverse momentum of the third-leading lepton in the HF-decay CRs can be seen in Figure 8.4 for all relevant processes, normalised to the SM expectation.

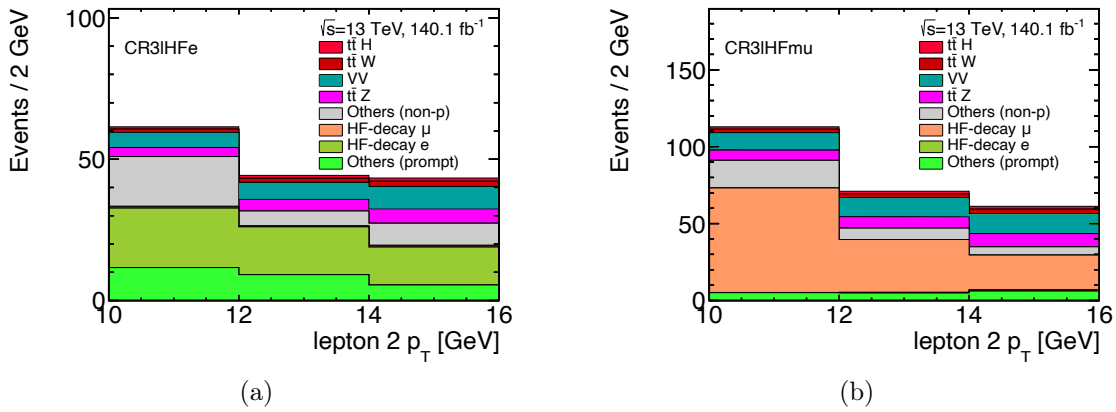


Figure 8.4: The transverse momentum p_T of the third-leading lepton in the CR3 ℓ HF e region (a) and in the CR3 ℓ HF μ region (b).

Lastly, two more CRs are defined to correct potential mismodelling for the $t\bar{t}W$ and the $t\bar{t}Z$ process. The CRs are called CR3 $\ell t\bar{t}W$ and CR3 $\ell t\bar{t}Z$ respectively. When comparing the probable decay chain of the FCNC processes in Figure 8.2 with the probable decay chain of the $t\bar{t}W$ background in Figure 8.5, it can be seen that the $t\bar{t}W$ background has one additional b quark at tree level. A similar observation can be made for the $t\bar{t}Z$ background. If the additional b jet is not properly recognised as a b jet (i.e. misidentified as l jet) or entirely lost due to other detector effects, a very similar final state to the one of the FCNC signals is achieved. Thus the CRs for the $t\bar{t}W$ and the $t\bar{t}Z$ backgrounds require the same objects as the SR, but without a specific number of light-flavoured jets and instead an additional b -tagged jet. Furthermore, the CR3 $\ell t\bar{t}Z$ region requires the invariant mass of a pair of charged leptons (called $ZRecoMass$, see Section 9.2) to be within a 10 GeV window around the Z -boson mass m_Z , whereas the CR3 $\ell t\bar{t}W$ region excludes this mass window. Using the window around the $ZRecoMass$, not only makes the two $t\bar{t}W/t\bar{t}Z$ CRs orthogonal to each other, the respective CRs are also enriched with the corresponding processes. The variable to compare data to the MC prediction is chosen to be the transverse momentum of the leading lepton, as the distribution of the transverse momentum is

relatively well modelled. A schematic summary of the six regions in the 3ℓ channel is given in Figure 8.6.

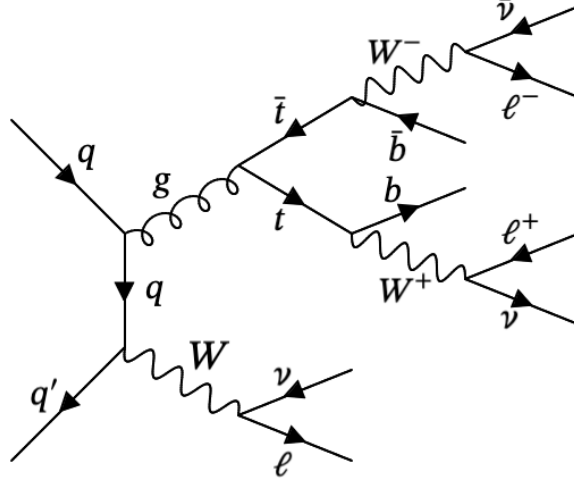


Figure 8.5: An exemplary production of the $t\bar{t}W$ background process together with the leptonic decay of all W bosons to get a similar final state as the FCNC processes.

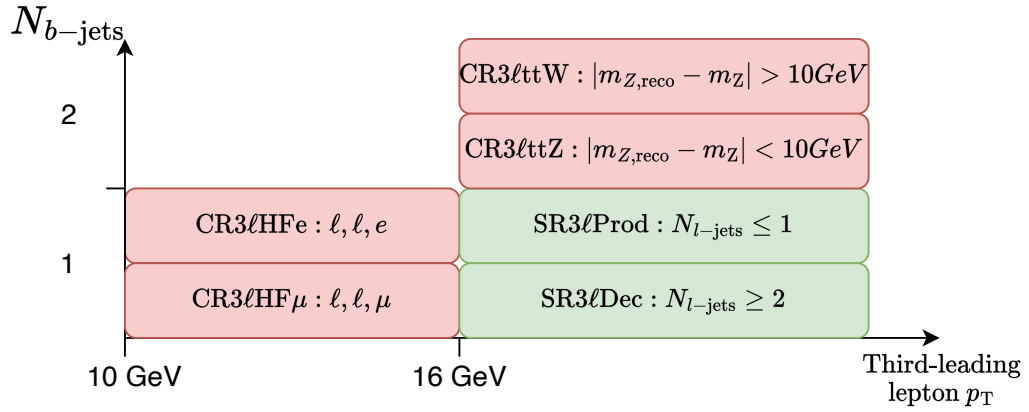


Figure 8.6: Schematic overview of the six 3ℓ regions with their most-defining cuts.

8.2. Di-lepton same sign regions

When searching for a rare process with two oppositely charged leptons, any region would be flooded by background events from the Z +jets process and $t\bar{t}$. To avoid this, only events with two leptons of the same charge are selected. For the signal-FCNC process, two same-charged leptons only occur for specific decays of the W bosons from the top quark and the Higgs boson decays (see Figure 8.7), but many backgrounds are significantly reduced by this requirement.

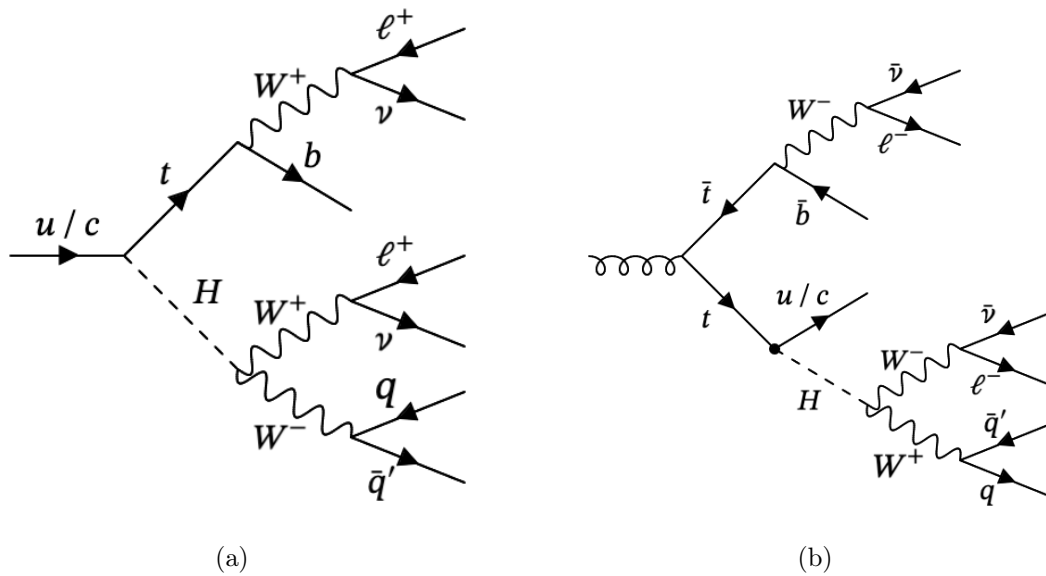


Figure 8.7: The prod-FCNC process (a) and the decay-FCNC process (b) with the $H \rightarrow WW^*$ decay channel resulting into two same-charged leptons.

However, a significant background arises from electrons which are reconstructed with an inverted charge, called “charge-flip electrons”. Since the rate of this background is hard to predict using MC generators, a dedicated CR in the 2ℓ SS channel is implemented for charge-flipped electrons, called CR2 ℓ q -flip. Complementing CRs for determining the rate of the HF-decay e and the HF-decay μ process (CR2 ℓ HF e and CR2 ℓ HF μ respectively), another CR is implemented for the $t\bar{t}W$ and the $t\bar{t}Z$ background. However, in contrast to the 3ℓ regions, there is only one combined region for the $t\bar{t}V$ backgrounds. In certain regions, if both selected charged leptons are electrons, the invariant mass $m(e_0, e_1)$ of both electrons is required to be further than 10 GeV away from the Z -boson mass, which ensures further suppression of the charge-flip background. A summary of all cuts defining the 2ℓ SS regions is given in Table (8.1).

Region	N_{jets}	$N_{b\text{-tags}}$	fl.	ℓ_0 ECIDS	fl.	ℓ_1 p_T [GeV]	m_Z -cut
SR2 ℓ Dec	≥ 4	1	e/μ	-	e/μ	≥ 12	$ m(e_0, e_1) - m_Z \geq 10 \text{ GeV}$
SR2 ℓ Prod	≤ 3	1	e/μ	≥ 0.65	e/μ	≥ 16	$ m(e_0, e_1) - m_Z \geq 10 \text{ GeV}$
CR2 ℓq -flip	≤ 3	1	e	< 0.65	e	-	$ m(e_0, e_1) - m_Z \geq 10 \text{ GeV}$
CR2 ℓ HFe	≤ 3	≥ 1	μ	-	e	< 16	-
CR2 ℓ HF μ	≤ 3	≥ 1	μ	-	μ	< 16	-
CR2 $\ell t\bar{t}V$	≥ 4	2	μ	-	e/μ	≥ 18	-

Table 8.1.: Summary of cuts, defining the 2 ℓ SS regions. $m_Z = 91.2 \text{ GeV}$ is the mass of the Z boson, which is used for the definition of certain regions.

By splitting the SRs into one SR enriched in the prod-FCNC process and one SR enriched in the decay-FCNC process, done both in the 2 ℓ SS and the 3 ℓ channel, the different kinematics between both FCNC processes can be better distinguished and exploited. In the SRs tailored to be enriched in the decay-FCNC process, the share between the prod-FCNC process and the decay-FCNC process in the expected event yield is approximately 1:6. On the other hand, the SRs tailored to be enriched in the prod-FCNC process, the share in the expected event yield is approximately 1:3, still in favour for the decay-FCNC process as this is the overall dominant FCNC process.

9. Reconstruction

This chapter describes the different approaches of the reconstruction methods used. Reconstructing events or specific event aspects serves as a vital step in distinguishing between signal and background processes. The reason behind this is that reconstruction yields variables that distinguish signal from background processes more effectively in comparison to bare kinematic distributions like the p_T of particles. In principle, a highly complex neural network could theoretically execute the reconstruction process on its own. However, this would necessitate substantial computing power, not only for training and evaluating the neural network but also for generating a significant number of MC events for the training process. To minimise the necessary computational resources and to expedite the overall workflow, the reconstruction is performed in advance, and then the reconstructed variables are fed into a neural network.

To better visualise and objectively quantify the separation power of variables, a receiver operating characteristic (ROC) curve is often used. A ROC curve plots the background selection efficiency against the signal selection efficiency for a varying cut from start to end. Each data point of the ROC curve is determined by the selection efficiency for a given cut. Variables with a high separation power will yield a large area A between the curve and the diagonal, whereas variables with small to no separating power will result in a curve close to the diagonal. The area A is therefore a measure for the separating power between signal and background for a given variable.

A significant challenge in reconstructing the given FCNC events in the signal regions is the presence of two or three neutrinos. Since neutrinos in general pass through the detector without interacting, only the sum of the p_T of all neutrinos can be measured as the negative of the missing transverse momentum E_T^{miss} . Consequently, there are too many degrees of freedom for a full analytical reconstruction. However, several of the presented reconstruction algorithms try to compensate for the lack of information by making certain assumptions or by working around the missing information. The following section outlines these reconstruction algorithms.

9.1. Neutrino weighting method

The neutrino weighting method tries to overcome the lack of information about the four-momenta of the neutrinos by imposing certain values for the η of the neutrinos in the event and evaluating the agreement with the measured E_T^{miss} . This is

done by first determining the η distributions of the involved neutrinos in the FCNC processes using truth MC information. This distribution is then split into ten bins with variable width, where each bin contains 10% of the events. As there is another combinatorial ambiguity regarding which charged lepton originated from the top quark and which originated from the Higgs boson, all possible permutations are determined. Subsequently, looping over all bins of the η distribution and using their mean value as assumption for an η , all possible *neutrino solutions*¹ are calculated for every permutation of leptons for every η bin. Once all possible neutrino solutions for the neutrinos from the top quark and from the Higgs boson are calculated for a given event, a loop over all neutrino solutions is performed and the agreement with the observed E_T^{miss} is evaluated. The resulting E_T^{miss} using the neutrino solutions as assumption is denoted as $E_{T\ i,x}^{\text{miss}}$ and $E_{T\ i,y}^{\text{miss}}$ for the x - and y component respectively, with i denoting the i -th permutation. Using the observed missing transverse momentum for the given event $E_{T\ \text{obs}}^{\text{miss}}$ and its variance $\sigma(E_T^{\text{miss}})$ taken from simulation, the neutrino weight ω can be calculated following Equation 9.1.

$$\omega = \frac{1}{N} \sum_{i=0}^N \exp \left[-\frac{(E_{T\ i,x}^{\text{miss}} - E_{T\ x,\text{obs}}^{\text{miss}})^2}{\sigma(E_{T\ x}^{\text{miss}})^2} - \frac{(E_{T\ i,y}^{\text{miss}} - E_{T\ y,\text{obs}}^{\text{miss}})^2}{\sigma(E_{T\ y}^{\text{miss}})^2} \right] \quad (9.1)$$

For the FCNC processes, one of the permutations is expected to be close to the actual decay of the event, resulting in a small exponent and thus a large value for the sum. The mean value over all permutations would subsequently be increased. For background processes on the other hand, no permutation is expected to yield a good match for the missing transverse momenta, thus a smaller neutrino weight is expected compared to signal processes. The resulting neutrino weight for the SR3 ℓ Prod region, comparing the utH FCNC processes to all relevant background processes, can be seen in Figure 9.1 (a). For a more comprehensive visualization of the separation power, the ROC curve for the neutrino weight in the SR3 ℓ Prod region can be seen in Figure 9.1 (b). Both the neutrino weight distribution and the ROC curve in the SR3 ℓ Dec are similar to the one presented in Figure 9.1.

Compared to the $t\bar{t}X$ backgrounds, the FCNC process tends to higher values in the neutrino weight. However, other important backgrounds like the non-prompt processes or VV are not well separated. Also the area A between the determined curve and the diagonal in the ROC plot is small, compared to other reconstructed variables. Nevertheless, the neutrino weight is still an important variable as it

¹With the mass constraints of the top-quark and Higgs-boson, as well as the W-boson mass constraints in the decay chain and the given η of the neutrino, always two four-momentum vector solutions for the neutrino emerge. These four-momentum vectors are called “neutrino solutions”.

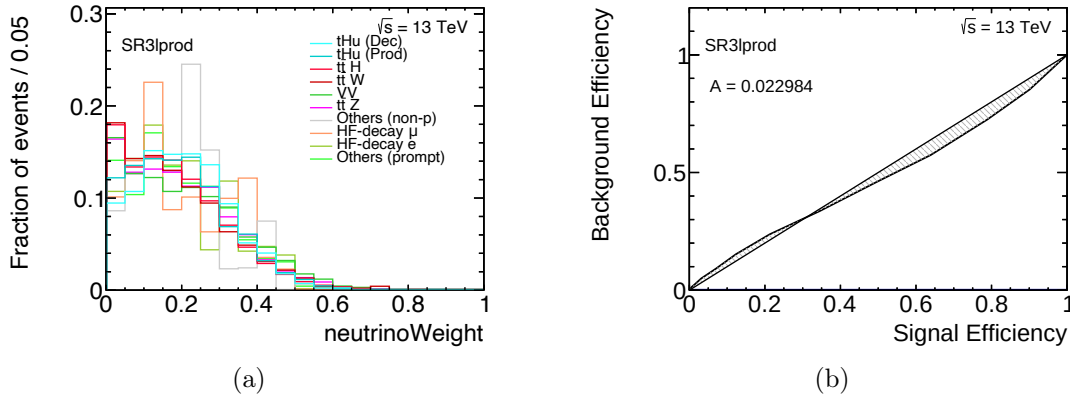


Figure 9.1: The distribution of the neutrino weight for the utH FCNC process and all background processes (a) and the corresponding ROC curve (b) in the SR3 ℓ Prod region.

introduces certain separating power against the $t\bar{t}X$ backgrounds, which is often not the case for other variables.

9.2. Reconstructed Z-boson mass

Some of the major backgrounds in the signal regions (like $t\bar{t}Z$, VV or tWZ) have a Z boson on tree level, whereas the FCNC processes mainly contribute to the 3ℓ final state via the Higgs boson decaying into two W bosons. In conclusion, a variable which is related to the Z -boson mass should yield a well-discriminating variable. The $ZRecoMass$ is calculated by looping over all permutations of leptons of the same generation with opposite charge. The invariant mass of the permutation of leptons being closest to the reference mass of $m_Z = 91$ GeV in a given event is defined as the $ZRecoMass$. It is expected that for the backgrounds containing a Z boson at tree level the distribution of the $ZRecoMass$ has a more pronounced peak at m_Z compared to the FCNC processes, thus yielding discriminating power. The distribution of the $ZRecoMass$ in the SR3 ℓ Prod region can be seen in Figure 9.2 (a) and the corresponding ROC curve can be seen in Figure 9.2 (b).

As expected, especially the VV , the $t\bar{t}Z$ and the Others(prompt) background processes have a pronounced peak at m_Z , which also results in a large area under the ROC curve. Notable is, that also processes which do not have a Z boson at tree level have a peak at m_Z in their $ZRecoMass$ distribution. This can be explained by the fact, that having multiple potential permutations of leptons and choosing the one with the invariant mass being closest to m_Z introduces a certain bias towards

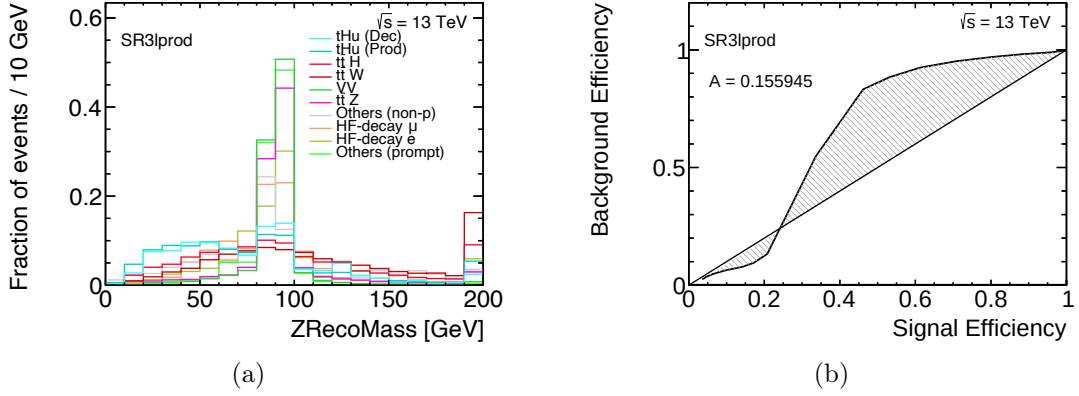


Figure 9.2: The distribution of the $ZRecoMass$ for the utH FCNC process and all background processes (a) and the corresponding ROC curve (b) in the $SR3lProd$ region.

the actual Z -boson mass. However, as long as a variable yields a shape difference between signal and background, it is very likely to be useful in the neural network at a later step.

9.3. Neutrino-independent combinatorics estimator

Another reconstruction method is employed in the 3ℓ regions, which is not dependent on any neutrino reconstruction or comparisons to E_T^{miss} but instead only compares the angular distance (ΔR) of different combinations of charged leptons and jets. This so-called “**neutrino-independent combinatorics estimator**” (NICE) reconstruction is based on the idea, that the top quark and the Higgs boson of the FCNC processes fly in opposite directions in their centre-of-mass system, which holds especially for the prod-FCNC process. Thus it is expected to observe two oppositely charged leptons with a small angular distance originating from the Higgs-boson decay and one charged lepton and one b -tagged jet with a small angular distance originating from the top-quark decay. The *NICE reco* checks if certain conditions concerning the angular distances are fulfilled and if so, creates a set of variables related to the matched decay products.

First, it is determined which two of the three charged leptons have the smallest ΔR between them. At this point, they are not required to have opposite charges, which is in general expected from Higgs-boson decay products. If the b -tagged jet and the remaining third charged lepton then have a smaller ΔR than any other combination of b -tagged jet and charged lepton, a boolean variable “NiceReco” is set to *true*,

otherwise it is *false*. If *NiceReco* is *true*, then the invariant masses and the ΔR of the two matched charged leptons (called “NiceHiggsChildren”) and the b -tagged jet and the charged lepton (called “NiceTopChildren”) are calculated. If in addition the two charged leptons, which are matched to the *NiceHiggsChildren* also have opposite charge, then a boolean called “ReallyNiceReco” is true and the respective variables are calculated analogously. If already the first check fails (i.e. the ΔR between the b -tagged jet and the remaining charged lepton is *not* smaller than the ΔR between the b -tagged jet and any other charged lepton), then the *NiceReco*-related variables are set to default values. This way, even the information of how often this reconstruction failed can be analysed and used in a later step.

Figure 9.3 shows the reconstruction success rate of the “niceReco” (a) and of the “reallyNiceReco” (b) in the SR3 ℓ Dec region for all background processes and the signal processes, whereas the value “0” indicates an unsuccessful reconstruction and “1” indicates a successful reconstruction.

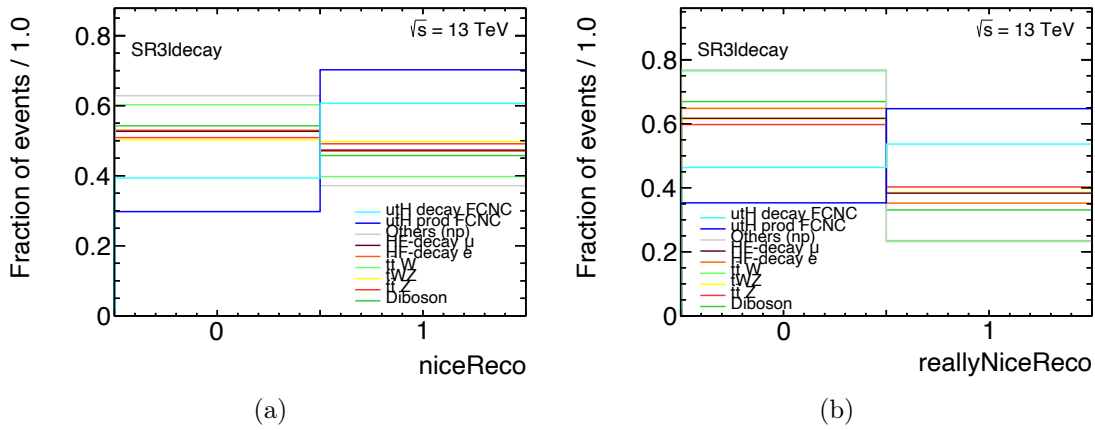


Figure 9.3: Success rate of the “niceReco” (a) and the “reallyNiceReco” (b) for all background processes and the signal processes in the SR3 ℓ Dec region.

It can be seen that the “(really)NiceReco” separates well between signal and background processes. Especially the $t\bar{t}W$ and Others (np) process have a low success rate in both reconstructions, whereas the FCNC processes successfully pass the reconstruction with 50% to 70% probability. Due to the more stringent requirement of the “reallyNiceReco” that the sum of charges of the leptons paired to the Higgs boson must be zero, the rate of successful reconstruction decreases for both the background and the signal processes. The prod-FCNC process has a higher success rate in both reconstruction methods compared to the decay-FCNC process. The rate of successful reconstructions in the SR3 ℓ Prod compared to the SR3 ℓ Dec regions are very similar for all processes. Figure 9.4 shows the invariant mass of the *HiggsChildren* for the *ReallyNiceReco* case in the SR3 ℓ Dec region.

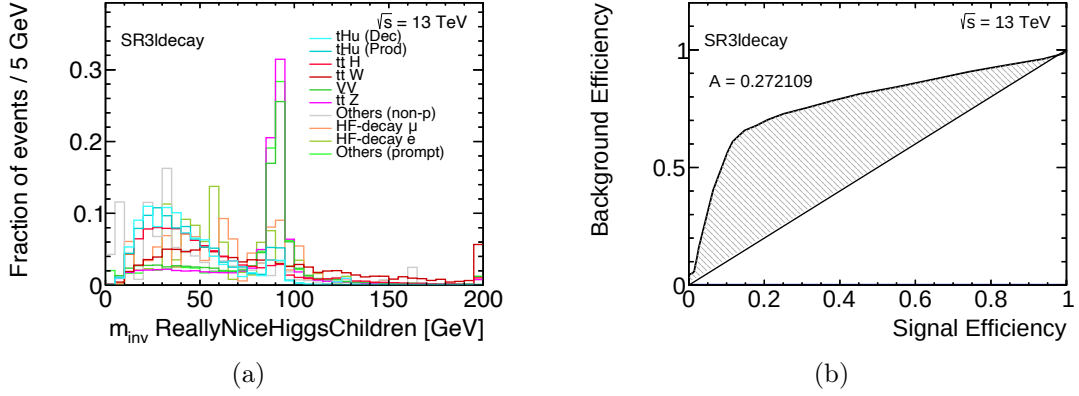


Figure 9.4: The distribution of the invariant mass of the *HiggsChildren* in the case of the *ReallyNiceReco* for the utH FCNC process and all background processes (a) and the corresponding ROC curve (b) in the SR3 ℓ Dec region.

It can be seen, that the FCNC processes tend to smaller values in the distribution of the invariant mass of the *HiggsChildren*, compared to most background processes. The systematic discrepancy between the reconstructed invariant mass and the actual Higgs-boson mass can be explained by not considering any neutrinos in the reconstruction. Since the neutrinos carry a significant fraction of the energy in the W -boson decay, the invariant mass of the reconstructed *HiggsChildren* tends to lower values. Also it can be observed that backgrounds involving a Z boson on tree level again have a pronounced peak at m_Z , similar to the distribution of the *ZRecoMass*. This happens because the charged leptons from the Z -boson decay naturally have a small angular distance and are thus usually misidentified as the Higgs-boson decay products. Since in this decay chain no neutrinos are involved, the invariant mass peak energy corresponds to the actual mass of the Z -boson. The distributions for the *NiceReco* variables and also the corresponding distributions in the SR3 ℓ Prod region show similar behaviour as shown here.

Figure 9.5 shows exemplarily the distribution of the invariant mass of the *TopChildren* in case of the *NiceReco* in the SR3 ℓ Dec region. In this case, however, no significant differentiating power can be observed between the FCNC processes and the background processes. This is likely due to the fact that most backgrounds also involve top-quark decays. Therefore, the reconstruction works equally well for the FCNC processes as it does for most background processes. This results in distributions that do not exhibit significant shape differences. Also in other regions or for the ctH FCNC process, the top-quark related reconstructed invariant masses do not provide a high separation power.

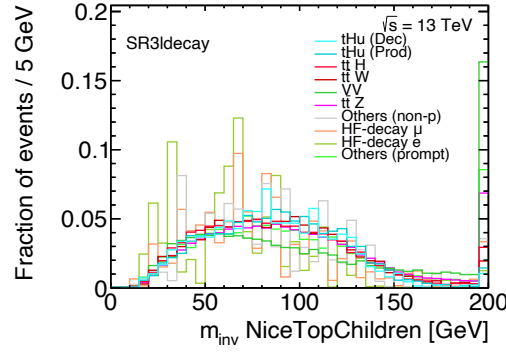


Figure 9.5: The distribution of the invariant mass of the *TopChildren* in the case of the *NiceReco* for the utH FCNC process and all background processes in the SR3lDec region.

9.4. Same-charged and opposite-charged leptons

In the 3ℓ regions, besides ordering the charged leptons according to their transverse momenta, the charged leptons may also be classified by being either one of the two same-charged leptons or being the single opposite-charged lepton. The two same-charged leptons are then again ordered by the transverse momentum (i.e. the leading and the sub-leading same-charged lepton). This additional ordering scheme of charged leptons allows for more possibilities of calculations of invariant masses or angular distances between objects (e.g. the invariant mass between the sub-leading lepton and the opposite-charged lepton, etc...). Certain invariant masses and angular distances proved to be well-discriminating variables in a later step, as can be seen e.g. in Table 10.1, summarising the most important input variables for the neural network.

9.5. Recursive jigsaw reconstruction

The **recursive jigsaw reconstruction** (RJR) [125] is a technique to reconstruct any given event following certain *jigsaw* rules, implemented into the **RestFrames** package [126]. These jigsaws may contain information about intermediate particles' masses, which and how many particles originate from which particle or other kinematic behaviour. The RJR then reconstructs particles in their inertial frames taking into account kinematic and combinatoric ambiguities. The RJR is implemented by imposing a decay tree, tailored to the signal process, as can be seen in Figure 9.6 for the decay-FCNC process.

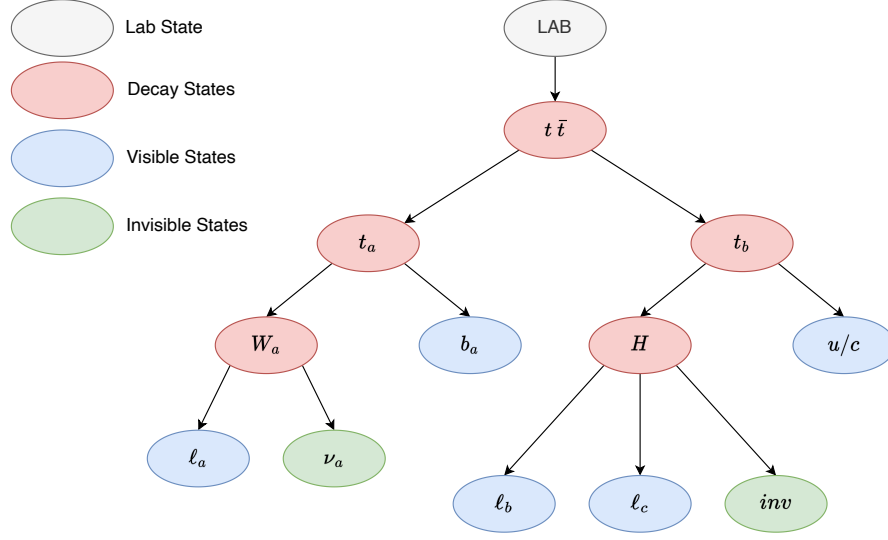


Figure 9.6: A decay tree with its jigsaws, tailored for the decay-FCNC process. The intermediate particles between the Higgs boson and the charged leptons are not specified to be inclusive in the Higgs decay modes. The “inv” object may consist of multiple invisible particles.

The jigsaws e.g. define that two charged leptons together with invisible particles originate from the Higgs boson. Such jigsaws are defined for every vertex in the decay tree and the algorithm then applies the jigsaws recursively through every stage of the decay tree, starting at the lab frame. The output of the algorithm are the four-vectors of every reconstructed intermediate particle, which can subsequently be used to obtain invariant masses, angular distances between certain objects and other variables of interest. However, in the decay tree used, it is chosen not to use the Higgs-boson mass as constraint at the respective vertex because this variable is supposed to be used as a discriminating variable in a later step. Since many background processes do not have a Higgs boson in their final state, it is expected that the RJR will discern the disparity between signal- and background processes. The resulting distribution of the reconstructed Higgs-boson mass using the RJR algorithm, together with the corresponding ROC curve can be seen in Figure 9.7.

Using the RJR algorithm, a distinct difference between the signal and most of the background processes can be observed. The FCNC processes, as well as $t\bar{t}H$ background process, tend to have smaller reconstructed Higgs boson masses still close to the actual Higgs boson mass, while many other backgrounds besides $t\bar{t}H$ have a pronounced peak at about 200 GeV. It is noteworthy that the reconstructed mass of the Higgs-boson is now closer to its actual value, owing to the RJR method incor-

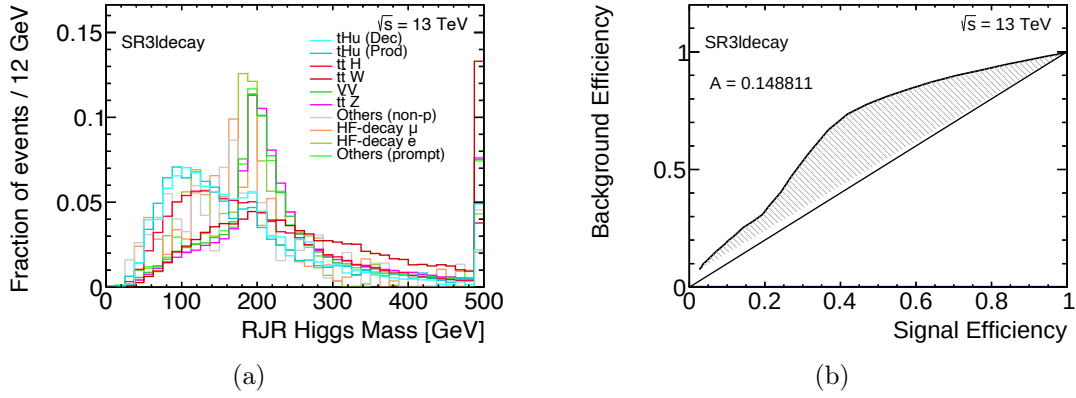


Figure 9.7: Distribution of the invariant mass of the Higgs boson decay produces as reconstructed by the RJR (a) and the corresponding ROC curve (b) in the SR3 ℓ Dec region.

porating the contribution of *invisible* particles, as indicated in the decay tree shown in Figure 9.6. This phenomenon also accounts for the shift of the Z mass peak towards 200 GeV for the backgrounds containing a Z boson, as the RJR is likely to utilize the charged leptons originating from the Z boson, but then adds further energy from the *invisible* particles.

10. Multivariate analysis

Once a number of discriminating variables have been selected, a naive way to proceed would be to use the variable that discriminates the best between signal and background (e.g. in terms of area under the ROC curve) to assess the agreement between MC and data in a statistical analysis. However, this would neglect the discriminative power of all other variables, which may be of some use as long as they are not 100% correlated. On the other hand, a statistical analysis using all constructed variables from Section 9 and moreover all possible combinations of kinematic/spatial relations between particles is not feasible. The computational power required for this statistical evaluation would be immense, as every bin of every distribution introduces more parameters to be optimised in the fit. Additionally, correctly propagating systematic uncertainties and correlations would be an enormous technical task. For these reasons, the discriminative power of the previously created variables is combined into a single variable using a **neural network** (NN).

10.1. Feed-forward neural networks

A feed-forward NN consists of a layer of input nodes, any number of hidden layers, where each hidden layer can have any number of hidden nodes, and an output layer with any number of output nodes. The flow of information is, as the name suggests, only in the forward direction. Each node in the input layer corresponds to one input variable. The number of hidden layers and nodes reflects the ability of the NN to learn complex correlations between input variables. However, using an excessively large numbers of hidden layers and nodes can be computationally challenging and lead to overtraining¹. Each node m_i in layer i is connected to every node n_j in the subsequent layer j by the weight $w_{m,n}^{i \rightarrow j}$. The input h_m^i of a node beyond the first layer is given by Equation 10.1

$$h_m^i = \sum_{l=1}^{l_{\max}} w_{l,m}^{i-1 \rightarrow i} x_l^{i-1} \quad (10.1)$$

with x_l^{i-1} being the output of the node l in layer $i-1$. The output of a node is given by the input of the respective node, passed through an activation function S , i.e.

¹Overtraining refers to the phenomenon where the NN is erroneously learning statistical fluctuations and, as a result, is unable to accurately classify unseen data.

$x_m^i = S(h_m^i)$. The choice of activation function can in principle be made arbitrary, however, different activation functions have different advantages and drawbacks. The activation function used in the NN used in this thesis is the sigmoid function, described by Equation 10.2

$$S(x) = \frac{2}{1 + e^{-x}} - 1. \quad (10.2)$$

The advantages of the sigmoid function are, that the output of the function is limited between -1 and 1 . This limit prevents nodes to become *saturated*, i.e. dominated by single huge values and unresponsive to small changes. Furthermore, the magnitude of negative input values of nodes is not ignored (as it is e.g. the case for the ReLU activation function) and the sigmoid function is differentiable which is beneficial for certain backpropagation models which use the gradient of the weight-parameter space. Every node in the output layer corresponds to one classification of the input data. In the case of the NN used, one output node is sufficient, where background events have the target value of -1 and signal events have the target value of $+1$. The *loss function* compares the target value and the resulting output value in the output node for all events. The *training* of an NN consists of the optimisation of the weights between the nodes to minimise the loss function, which is synonymous for events being evaluated according to their respective target value.

10.2. NeuroBayes

The NN package used in this analysis is called “NeuroBayes”[127–129], a single-hidden-layer feed-forward NN. It features a strong and robust preprocessing of the input variables, together with the highly efficient low-memory¹ **BFGS Bound** constrains² (L-BFGS-B) algorithm [130] for optimising the weights in the back-propagation. Main features of the NeuroBayes package can be seen in Figure 10.1.

The preprocessing of the variables is a vital step and the main reason why a single hidden layer is sufficient for good separation between signal and background for NeuroBayes. As a first step, with the signal processes being scaled to make up 50% of the input weights, the input variables are flattened. This is achieved by a rebinning of each variable, such that the sum of signal events plus background events in each

¹The BFGS algorithm only uses a reduced Hesse matrix in its search for an optimal next weight phase space point, therefore only requires a comparable small amount of memory.

²The BFGS algorithm has certain boundary constraints on the values the weights may take.

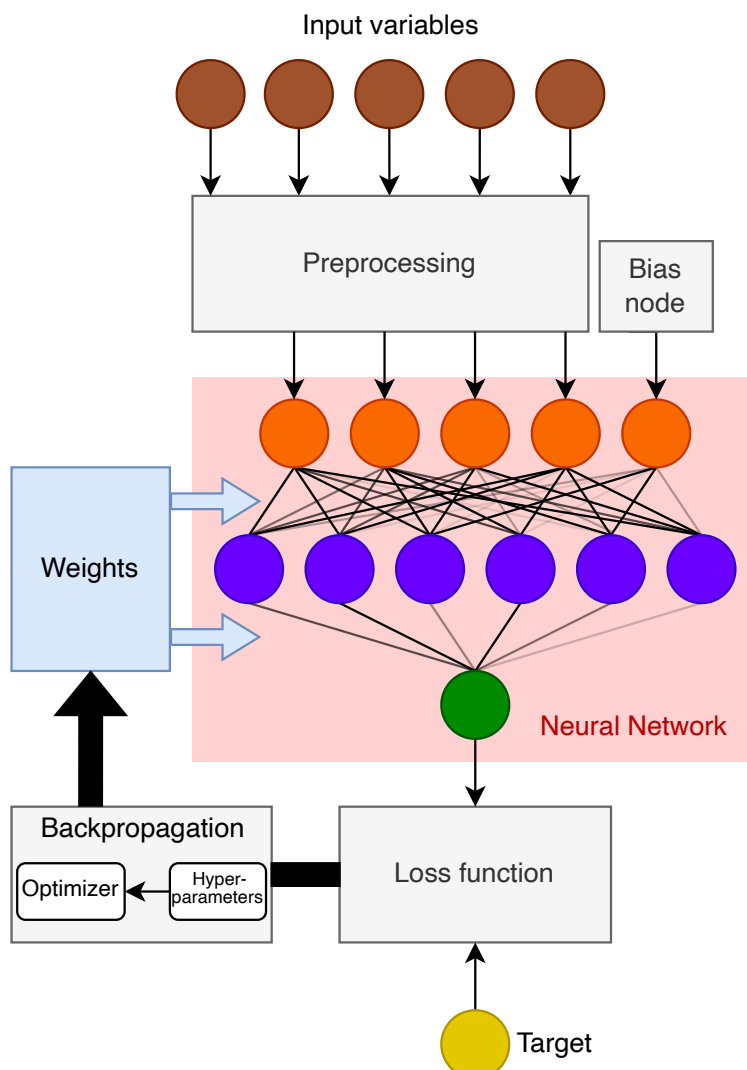


Figure 10.1: Diagram depicting the main characteristics of NeuroBayes, including its pre-processing of input variables, the single-hidden-layer structure including the bias node of the neural network and the backpropagation to optimise the weights between nodes.

bin is constant. This prevents extreme outliers to have a negative effect on the NN like e.g. saturating specific nodes. Afterwards, a spline fit is performed through the purity¹, which yields a continuous transformation of the original variable to the purity. As a next step, the input variables are converted into distributions

¹The purity is defined as $\frac{s}{s+b}$ with s and b being the number of signal- and backgrounds events respectively.

with a mean value of 0 and a standard deviation of 1, where the mean value of 0 prevents large eigenvalues in the Hesse matrix of the weights and the standard deviation of 1 leads to inputs for the next layer also distributed with mean 0 and standard deviation 1. Both features lead to optimal learning conditions, allowing for fast initial learning rates and preventing node saturation [131]. Subsequently, the transformed input variables are decorrelated, which is an enormous help for the NN since it does not have to learn the correlations between variables itself. The decorrelation of variables is performed by first calculating the covariance matrix of the input variables. After diagonalising and dividing the diagonal entries by the square root of the corresponding eigenvalue, the transformation is applied to the input variables. This results in the covariance matrix being the unit matrix. As a last step of the preprocessing, a ranking of the input variables by their correlation to the target is performed. This feature allows NeuroBayes to only use the most important variables for the training, rather than use all given variables (up to 99). The ranking of the variables is performed by first calculating the total correlation to the target for all variables. The variable with the smallest correlation to the target is removed and ranked lowest. This procedure is then repeated for the remaining $N - 1$ input variables, until the entire ranking is determined.

Besides the previously determined most important input variables, one additional *bias node* contributes to the input layer. This bias node is always at the value +1 which allows the NN to easily shift certain inputs. The number of hidden nodes in the hidden layer is a hyperparameter of the NN and can be adjusted as desired. The NNs trained for this analysis, depending on the specific SR and the signal (utH or ctH FCNC signal), consist of 15-25 hidden nodes. The magnitude of the weight between nodes is often depicted as varying opacity, as can be seen in Figure 10.1. The loss function used by NeuroBayes is the cross-entropy loss function with an additional weight-decay term, given in Equation 10.3.

$$E_D = \sum_i \log \left(\frac{1}{2} (1 + T_i \cdot o_i + \epsilon) \right) + \sum_j \frac{1}{2} w_j^2 \quad (10.3)$$

with T_i being the target value of event i and o_i being the output value of event i . ϵ is a small value added to the cross entropy in the first few training iterations to avoid numerical problems for untrained networks. After a few iterations, ϵ is set to zero. An advantage of the cross entropy loss function is that the NN learns very fast to avoid completely wrong classification of events. For the weight-decay term, the sum over j sums up all weights w_j in the network. This leads to the preference of smaller weights, again preventing overtraining from unlikely large weights for single nodes. The backpropagation of information from the loss function and information about the weight phase space in the optimiser allows for an adjustment of the weights to

minimise the loss function. During the training, a pruning of weights is performed. If a weight has a negligible impact on the loss function, the weight is permanently set to exactly zero, thus reducing the number of degrees of freedom of the network. This improves the signal-to-noise ratio and prevents overtraining. Several hyperparameters have an influence on the quality and speed of the training, e.g. preventing the optimizer getting stuck in a local minima. For every FCNC process in every SR, a dedicated optimisation of the hyperparameters was performed to ensure optimal results.

Furthermore, to entirely eliminate the possibility of overtraining, during the training process of the NN the training samples are split by their MC event number¹ into an even and an odd training set. A dedicated *even* and *odd* NN is then trained on the respective training set but applied on the other one in the analysis itself. This ensures that the NN output value for MC events in the analysis is always produced by a NN, which has *never seen* this event before, i.e. was not trained on this event. For data, the even/odd NN is chosen based on the run- and event number of the data event, yielding an approximate share of 50/50 of both NNs. However, a comparison between the even and the odd NN has shown that in terms of statistical uncertainties, the even and odd NN perform equally well for both their own and the unseen MC events.

10.3. Applying the neural network

The NN framework “NeuroBayes”, discussed above, is applied to separate the FCNC processes from the background processes in the SRs. In each of the four SRs, a dedicated NN is trained for the case of a *utH* or *ctH* FCNC signal. In the training, the *prod*-FCNC and the *decay*-FCNC process are used as signal simultaneously with their share of the expected FCNC event yields corresponding to the share in the NN. Also taking into account the even/odd splitting of the NN, in total $4 \cdot 2 \cdot 2 = 16$ NNs are trained. An ordered list of input variables for a training in the SR3 ℓ Prod and the SR3 ℓ Dec region for a *ctH* signal can be seen in Table 10.1 in descending order of their *added significance*²(AS). Within the list of input variables for the NN, $m(X, Y)$ denotes the invariant mass of particle X and Y and (Really)NiceTop/Higgs denotes the sum of the four vectors of the top-quark / Higgs-boson children in case of the (Really)NiceReco being fulfilled.

¹The MC event number is an integer, numbering the MC events. The MC event number has no correlation whatsoever to any kinematic distribution of the MC event.

²The added significance can only be used to compare the importance of input variables within the same training, since for its calculation the amount of MC events is used, which differs from region to region. Higher values correspond to a higher importance of the variable.

¹ $H_T(\text{jets})$ denotes the scalar sum of the transverse momenta of all jets.

Table 10.1.: Ordered list of input variables in a training for a ctH signal in the 3ℓ SRs. The variables are sorted according to their added significance in descending order. ω denotes the neutrino weight.

SR3 ℓ Prod		SR3 ℓ Dec	
Variable	AS	Variable	AS
$ZRecoMass$	38.90	$m(\text{ReallyNiceHiggs})$	43.96
$m(\text{ReallyNiceHiggs})$	19.82	$ZRecoMass$	23.07
ω	16.61	$H_T(\text{jets})$	16.29
$m(\text{ReallyNiceTop})$	10.88	$m(\text{ReallyNiceTop})$	11.79
$\Delta R(\text{ReallyNiceTop}, \ell_{OC})$	7.48	ω	10.94
$m(b\text{-jet}, \text{ReallyNiceTop})$	7.27	$\Delta R(\text{ReallyNiceTop}, \ell_{OC})$	7.79
$H_T(\text{jets})^1$	6.58	$m(b\text{-jet}, \text{ReallyNiceTop})$	6.48
$m(\ell_{SC,0}, \ell_{OC})$	3.75	$m(b\text{-jet}, \ell_1)$	2.78
$\Delta R(\text{ReallyNiceTopChildren})$	3.66	$\Delta R(b\text{-jet}, \ell_1)$	5.00
$m(\ell_{SC,1}, \ell_{OC})$	3.74	$m(\ell_{SC,1}, \ell_{OC})$	4.56
$m(b\text{-jet}, \ell_1)$	3.31	$m(b\text{-jet}, \ell_{SC,0})$	3.84
$\Delta R(b\text{-jet}, \ell_1)$	2.95	$m(\ell_{SC,1}, \ell_{OC})$	4.08
$\Delta R(\ell_{SC,1}, \ell_{OC})$	2.89	$p_T(t_{RJR})$	4.52

It can be observed that the reconstruction plays a vital role in the discrimination of signal and background, as six of the seven leading variables in both SRs are either related to the reconstruction of the Z -boson mass or to the *ReallyNiceReco*. Trainings for a different FCNC signal in the 3ℓ SRs use the same or very similar input variables. Trainings in the 2ℓ SS SRs use slightly different input variables, related to the reconstruction algorithms used in the 2ℓ SS regions.

The NN output distribution for the signal and the background can be seen in Figure 10.2 (a) exemplarily for the ctH signal in the SR3 ℓ Dec region, whereas Figure 10.2 (b) shows the respective purity per NN output. For the cross-entropy loss function omitting the weight-decay term, it is possible to show[128] that for an optimally trained NN, the purity is expected to be distributed around the diagonal, which is given after an optimisation of the hyperparameters. The NN output distribution per process in the SR3 ℓ Dec region can be seen in Figure 10.3 (a), whereas Figure 10.3 (b) shows the SM expectation of background processes exemplarily with the ctH process, normalised to a Wilson coefficient of $C = 1$. The same plots for the SR3 ℓ Prod region can be seen in Figure 10.4.

It can be seen that the performance of the NN in separating signal from background is depending on the background process. Certain backgrounds (like VV or $t\bar{t}Z$) can be separated very well, whereas backgrounds like $t\bar{t}H$ or the HF processes can

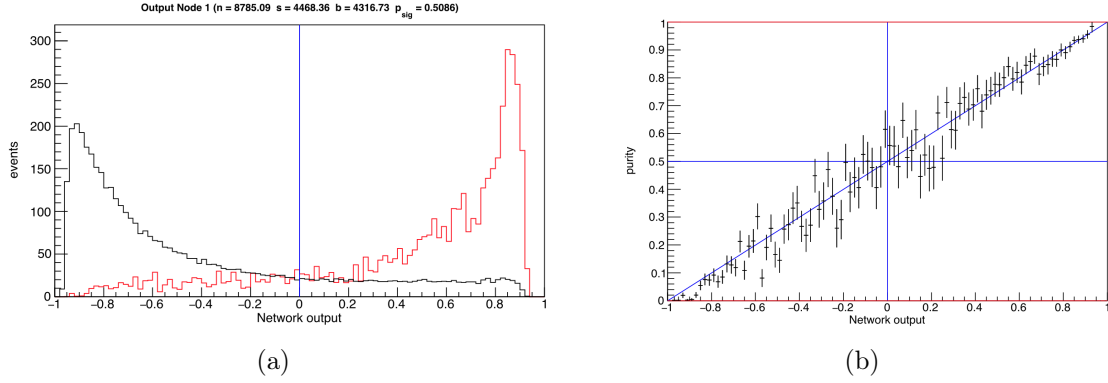


Figure 10.2: The NN output distribution (a) and the corresponding purity (b) for the ctH-based training in the SR3 ℓ Dec region.

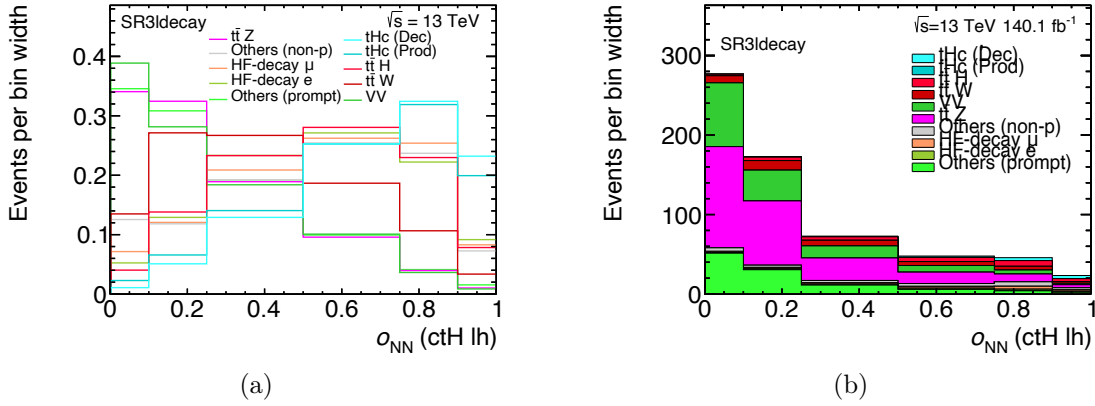


Figure 10.3: The NN output distribution per process (a) and the SM expectation of the background processes including a ctH signal normalised to $C = 1$ (b) for the ctH-based training in the SR3 ℓ Dec region.

not be separated equally well. One possible explanation for this behaviour is that certain processes exhibit greater kinematic similarity to FCNC processes compared to others. For example, variables related to the reconstructed Higgs boson are very important for the NN, thus background processes possessing a Higgs boson (like $t\bar{t}H$) are more difficult to separate. Another aspect the NN takes into account is the share of a specific background process at the total background yield. Given the limited complexity of the NN, *more effort* is invested into separating important background from the signal than unimportant ones (with a small expected total event yield). Also interesting to see is that in both SRs the prod-FCNC and the decay-FCNC process are not equally well separated from the background processes: In both regions, the decay-FCNC process is better separated than the prod-FCNC process.

11. Statistical analysis techniques

After constructing regions enriched in signal events and using NNs to discriminate between signal and background, the agreement between the MC prediction and the observed data needs to be evaluated. This is done using a binned profile likelihood fit simultaneously in all CRs and SRs to obtain the best-fit scenario. If there is no observed excess of data over MC, hypothesis tests are used to derive exclusion limits on the signal normalisation. This chapter introduces both the technique of binned profile likelihood fits and hypothesis tests. For a more comprehensive description of both the binned profile likelihood fit and the hypothesis testing, refer to [132].

11.1. Binned profile likelihood fits

The advantage of a binned profile likelihood fit is that the compatibility of a binned expected MC distribution (which is given in the regions considered here) can be evaluated with data. Additionally, the model is able to accommodate statistical and systematic uncertainties that originate from detector effects or the modelling of the MC samples. *Profiling* the correlations and constraints of systematic uncertainties found within the distributions of the CRs and SRs can lead to a significant improvement in the outcome, in contrast to basic cut-and-count methods.

The number of expected events N in a single bin i can be expressed as

$$N_i = \mu \cdot (s_i^{\text{prod}} + s_i^{\text{decay}}) + \sum_j \mu_j b_{j,i} \quad (11.1)$$

with s_i^{prod} and s_i^{decay} being the number of prod-FCNC and decay-FCNC events in bin i respectively. μ is the **parameter of interest (POI)**, which scales the amount of FCNC events. $\mu = 1$ corresponds to a Wilson coefficient of $C = 1$. The summation over the index j runs over all background processes j with events $b_{j,i}$ in bin i , where certain backgrounds have a fixed normalisation of $\mu_j = 1$ and other backgrounds (i.e. the non-prompt background processes and the $t\bar{t}W/t\bar{t}Z$ process) have a free floating normalisation. The probability of finding n_i data events in a bin i with a true expected number of events N_i can then be expressed in terms of the Poisson distribution $\mathcal{P}$, i.e.

$$\mathcal{P}(n_i; N_i) = \frac{\left(\mu \cdot (s_i^{\text{prod}} + s_i^{\text{decay}}) + \sum_j \mu_j b_{j,i}\right)^{n_i}}{n_i!} \cdot e^{-\mu \cdot (s_i^{\text{prod}} + s_i^{\text{decay}}) - \sum_j \mu_j b_{j,i}}. \quad (11.2)$$

A reasonable (preliminary) likelihood function L , which would subsequently be maximised in a fit by varying the POI and the normalization factors of the respective free-floating background processes, would therefore be the product of the Poisson probabilities of every bin in every CR and SR, which can be seen in the following equation

$$L(\mu) = \prod_i \mathcal{P}(n_i; N_i), \quad (11.3)$$

where the dependence of L of μ_i is in general not explicitly written any more. To also incorporate the effects of systematic uncertainties, the likelihood function from Equation 11.3 is extended by certain *prior terms*, where each systematic uncertainty has its own prior term with a corresponding **n**uisance **p**arameter (NP). Systematic uncertainties are implemented with a Gaussian prior around the nominal value. The $\pm 1\sigma$ deviation of a systematic uncertainty is in general determined by processing MC events with the alternative characteristic (e.g. with the $\pm 1\sigma$ jet energy scale calibration) through the entire analysis framework to obtain the $\pm 1\sigma$ kinematic/NN distribution in the respective regions. The corresponding constraint term for the NP θ_j of a systematic uncertainty is then given by

$$\rho(\theta_j) = \frac{1}{\sqrt{2\pi}\sigma_j} \cdot \exp\left(-\frac{\theta_j^2}{2\sigma_j^2}\right) \quad (11.4)$$

with its standard deviation $\sigma_j = 1$. Pulling a NP by a factor $\alpha \in \mathbb{R}$ means replacing the corresponding nominal sample by a certain procedure. At $\alpha = \pm 1$, the nominal sample is entirely replaced by the $\pm 1\sigma$ variation. However, since there are in general only two additional *points* per systematic uncertainty, interpolation (for $|\alpha| < 1$) and extrapolation (for $|\alpha| > 1$) is performed. The shape of a systematic uncertainty in bin i is then inter/extrapolated linearly according to

$$\psi_i(\alpha) = \psi_i^0 + I_{i,\text{lin}}, \text{ where } I_{i,\text{lin}} = \begin{cases} \alpha \cdot (I_i^+ - I_i^0) & \alpha \geq 0 \\ \alpha \cdot (I_i^0 - I_i^-) & \alpha < 0 \end{cases}, \quad (11.5)$$

where I_i^+ , I_i^0 and I_i^- denote the event yield of the up-variation, the nominal and the down-variation of the respective systematic uncertainty in bin i . The normalisation of a systematic uncertainty, on the other hand, is inter/extrapolated exponentially as in

$$\eta(\alpha) = I_{\text{exp}}, \text{ where } I_{\text{exp}} = \begin{cases} (I^+/I^0)^\alpha & \alpha \geq 0 \\ (I^-/I^0)^{-\alpha} & \alpha < 0 \end{cases}. \quad (11.6)$$

Using the exponential inter/extrapolation for the normalisation removes the possibility of unphysical negative values for the event yield for any given α . By adding the Gaussian constraint terms from Equation 11.4 into the preliminary likelihood function from Equation 11.3, a more sophisticated likelihood function is obtained which allows for systematic uncertainties to have an effect on the fit and to be pulled/constrained. The resulting likelihood function is given by

$$L(\mu, \vec{\theta}) = \prod_i \mathcal{P}(n_i; N_i) \cdot \prod_j \rho(\theta_j) \quad (11.7)$$

with the vector of all NPs $\vec{\theta}$. To lastly also incorporate statistical uncertainties originating from the MC samples being produced with limited statistics, the Poisson term $\mathcal{P}(n_i; N_i)$ in the likelihood is modified. For this, the absolute statistical uncertainty in bin i , denoted as δ_i , is used to derive the relative statistical uncertainty σ_i by $\sigma_i = \delta_i/N_i$. The *effective number of MC events* m_i is then defined as the number of events, which would result in the relative statistical uncertainty σ_i , i.e. $\sigma_i = \frac{1}{\sqrt{m_i}} \Rightarrow m_i = \frac{1}{\sigma_i^2}$. The Poisson term in the likelihood function is then replaced by the following

$$\mathcal{P}(n_i; N_i) \rightarrow \mathcal{P}(n_i; \gamma_i N_i) \cdot \mathcal{P}(m_i; \gamma_i m_i) \quad (11.8)$$

with γ_i being additional NPs of the fit[133]. The former term allows for certain variation of the number of MC events due to the limited MC statistics, whereas the latter term constraints the newly introduced γ factors depending on the available MC statistics.

Technically, the given likelihood function is not maximised, but instead the negative logarithm of the likelihood function is minimised. Especially the logarithm is computationally advantageous because the product of (up to hundreds of) small

values is turned into a sum, preventing problems with numerical precision. NPs can become *constrained* in the fit which means that their post-fit uncertainty is reduced, compared to their input prior uncertainty. The constraint is calculated for a given NP by varying the NP around the best-fit scenario. The plus and minus values of the NP, where $-\log(L)$ is 0.5 larger than the minimum define the constrained plus and minus uncertainties of the NP.

11.2. Hypothesis testing

Fitting the likelihood function from the previous section will yield a certain best-fit value for the signal normalisation μ , quite possibly larger than zero. But this does not necessarily imply a meaningful discovery of a BSM process, as an excess in data could also be caused by a statistical fluctuation and does not have to be significant. To evaluate the compatibility between the H_0 hypothesis of only SM processes to data or the H_1 hypothesis of the SM processes plus the FCNC process to data, certain statistical techniques are employed [134], which will be introduced in this section.

According to the Neyman-Pearson lemma [135], the ratio of likelihoods is the most powerful test statistic to distinguish between the two hypotheses. The test statistics $\lambda(\mu)$ is thus defined as

$$\lambda(\mu) = \frac{L(\mu, \hat{\vec{\theta}})}{L(\hat{\mu}, \hat{\vec{\theta}})}, \quad \lambda(\mu) \in (0, 1] \quad \forall \mu \geq 0, \quad (11.9)$$

with $\hat{\vec{\theta}}$ being the set of NPs which maximise the likelihood function for a given μ . $\hat{\mu}$ and $\hat{\vec{\theta}}$ are the POI and the set of NPs which overall maximise the likelihood function. However, since the given BSM contribution can only add events to the event yield but not decrease the expected SM event yield, negative values for $\hat{\mu}$ are unphysical. Therefore the test statistics is modified to

$$\tilde{\lambda}(\mu) = \begin{cases} \frac{L(\mu, \hat{\vec{\theta}})}{L(\hat{\mu}, \hat{\vec{\theta}})} & \hat{\mu} \geq 0 \\ \frac{L(\mu, \hat{\vec{\theta}})}{L(0, \hat{\vec{\theta}}(0))} & \hat{\mu} < 0 \end{cases} \quad (11.10)$$

to exclude such cases. To map the range of the test statistics onto the range $(\infty, 0)$ and to have larger values corresponding to increasing incompatibility, the test statistics is transformed to

$$\tilde{q}_\mu = -2 \ln(\tilde{\lambda}(\mu)). \quad (11.11)$$

Furthermore, in case of no significant excess, we will be interested in an *upper limit* of the POI. Best-fit values $\hat{\mu} > \mu$ should therefore not increase the level of incompatibility between MC and data, so that $\tilde{q}_\mu$ becomes

$$\tilde{q}_\mu = \begin{cases} -2 \ln \frac{L(\mu, \hat{\hat{\theta}})}{L(\hat{\mu}, \hat{\hat{\theta}})} & \mu \geq \hat{\mu} \geq 0 \\ -2 \ln \frac{L(\mu, \hat{\hat{\theta}})}{L(0, \hat{\hat{\theta}}(0))} & \hat{\mu} < 0 \\ 0 & \hat{\mu} > \mu \end{cases}. \quad (11.12)$$

Because the calculation of $\tilde{q}_\mu$ can become computationally very expensive, an asymptotic formula[132] is used in this analysis:

$$\tilde{q}_\mu = \begin{cases} \frac{(\mu - \hat{\mu})^2}{\sigma^2} & \mu \geq \hat{\mu} \geq 0 \\ \frac{\mu^2}{\sigma^2} - \frac{2\mu\hat{\mu}}{\sigma^2} & \hat{\mu} < 0 \\ 0 & \hat{\mu} > \mu \end{cases}. \quad (11.13)$$

where $\hat{\mu}$ follows a Gaussian distribution with standard deviation σ . The asymptotic formula introduces corrections of the order $\mathcal{O}(1/\sqrt{N})$ with N being the data sample size, therefore being well justified for large data sample sizes. From Equation 11.13, finally the so-called p -value can be calculated following

$$p_\mu = \int_{\tilde{q}_\mu^{\text{obs}}}^{\infty} f(\tilde{q}_\mu | \mu) d\tilde{q}_\mu \quad (11.14)$$

with $f(\tilde{q}_\mu | \mu)$ being the probability density function of the observable $\tilde{q}_\mu$ given a specific μ and $\tilde{q}_\mu^{\text{obs}}$ being the observed value of $\tilde{q}_\mu$ with data. The p -value corresponds to the probability to obtain by chance a result, which is as extreme or more extreme as observed, assuming H_0 . The H_1 hypothesis is rejected, if the p -value for the respective hypothesis is below a specific value. In searches, it is of particular interest

to find the maximum value of the POI μ , such that the p -value is still above the specific value (conventionally chosen to be 5 %). This value of the POI is then called the *upper exclusion limit* and is usually (after implementation of the CL_S method, see below) the main result of a search for new physics. It is common to convert small p -values p_0 into significances Z , which are defined by

$$Z = \Phi^{-1}(1 - p) \quad (11.15)$$

where Φ^{-1} is the inverse of the cumulative function of the Gaussian distribution, called the quantile. A p -value of 5% then corresponds to a significance of $Z = 1.64$, whereas the thresholds for an *evidence* and a *discovery* of new physics are at $Z = 3$ and $Z = 5$ respectively.

However, when searching for a signal which is much smaller than the backgrounds, obtaining limits with p -values has a drawback in case of downward-fluctuations. Even when not expecting sensitivity for a measurement, the signal might still become *unrealistically* strongly excluded. To avoid this drawback, the CL_S method [136, 137] is used in this analysis to define the confidence levels. The CL_S is defined as

$$CL_S \equiv \frac{p_\mu}{1 - p_0} \quad (11.16)$$

where p_μ is the p -value for the H_1 hypothesis and p_0 is the p -value for the H_0 hypothesis.

The entire workflow starting with the binned profile-likelihood fit and the subsequent statistical analysis is implemented in the ATLAS-internal `TRExFitter`[138, 139] framework.

12. Systematic uncertainties

In contrast to statistical uncertainties, systematic uncertainties cannot be reduced at will by repeating the measurement. They reflect e.g. limited knowledge of certain aspects of the detector and arise, for example, from the potentially imperfect calibrations of the energy of the objects measured in the detector. Additionally, the modelling of MC events has non-negligible systematic uncertainties, e.g. the choice of certain parameters in the simulation chain, which are not given by the theory but are assumed ad hoc. This chapter discusses all systematic uncertainties considered in this search and then covers the implementation of systematic uncertainties in the statistical analysis, which was introduced in Chapter 11.

12.1. Experimental systematic uncertainties

Systematic uncertainties associated with the measurement and calibration of the various objects are presented below.

Lepton-related uncertainties

The identification, reconstruction, and selection efficiencies of electrons and muons are calibrated to the well-understood Z and J/Ψ resonances utilising scale factors. The scale factor uncertainties are determined using a tag-and-probe method applied to the respective lepton from the above-mentioned resonances. While each scale factor for electrons has one combined variation, those for muons are divided into a statistical and a systematic variation.

Systematic uncertainties, accounting for the lepton scale and resolution, are introduced by the electron and muon momentum calibration and are taken into account. As the direction of curvature of muon tracks in the MS is dependent on the charge of the muons, a bias could be introduced due to imperfections in the detector-to-magnetic-field layout, which is dependent on the p_T . To account for such a bias, a specific sagitta bias uncertainty is implemented for muons. Additionally, the momentum scale variation is divided into an ID variation and an MS variation for muons.

Jet-related uncertainties

The calibration of the scale and the resolution of the energy measurement of jets (**J**et **E**nergy **S**cale/**R**esolution, JES/JER) introduces systematic uncertainties. The full set of JES-related systematic uncertainties[140] includes 125 individual terms, e.g. from *in-situ* calibrations, pile-up effects or the flavour-dependent response of the detector. However, not all of the 125 individual terms are relevant for many analyses as there are significant correlations between specific terms. To significantly reduce computational power, a reduced scheme of 30 JES uncertainties is commonly used, as it is the case in this analysis. In detail, the following sources of JES-related systematic uncertainties are taken into account:

- 15 NPs for the 15 most important systematic uncertainties from a set of *effective*¹ NPs
- 6 NPs corresponding to the η -intercalibration (correcting the energy scale of forward jets ($0.8 < |\eta| < 4.5$) to match those of central jets ($|\eta| < 0.8$))
- 4 NPs corresponding to pile-up corrections (removing excess energy from pile-up events)
- 2 NPs corresponding to the flavour of the initial particle (one for the detector response which differs especially between quark-initiated and gluon-initiated jets and one for the composition)
- 1 NP each for
 - punch-through, when the energy of a jet is not fully contained in the calorimeter system
 - BJES, difference in detector response for true b -jets between different MCs
 - singleParticle high- p_T , difference in detector response for individual hadrons with p_T up to 2.4 TeV

In addition to the above, if the detector simulation of MC samples is simulated in *fast-sim*, an additional *non-closure* systematic uncertainty related to the JES is used. For the JER, a jet resolution smearing method is employed when the resolution in MC is better than in data, resulting in a set of 13 variations called *FullJER*. For MC samples with the detector simulation performed in *fast-sim*, again an additional systematic uncertainty related to the JER is used. Finally, a single NP accounts for variations in the scale factors related to the JVT.

¹As many of the 125 individual terms are functions only of the p_T , an eigenvector decomposition is performed on the covariance matrix of these terms. The resulting orthogonal terms are the *effective* NPs.

Flavour-tagging-related uncertainties

Differences in the selection efficiency and mis-tagging rates of b -jets between data and MC can arise from statistical sources, detector calibration and modelling effects. The uncertainties due to these effects are propagated into variations of the b -tagging related scale factors. In total, 45 b -jet related variations and 20 c /light-jet related variations each are considered, resulting in a total of 85 flavour-tagging related variations. Due to the construction of the variations, the anticipated impact on an analysis is sorted in descending order per jet flavour related group, where the zeroth variation is expected to have the largest impact.

E_T^{miss} -related uncertainties

All systematic uncertainties mentioned earlier are propagated through the calculation of the E_T^{miss} . Additionally, when calculating the soft term within the E_T^{miss} (see Section 6.7), scale and resolution variations are taken into account. A total of three variations are considered, where the resolution variation is split into a parallel direction and a perpendicular direction relative to the soft term. The scale variation accounts for the uncertainty of the parallel scaling of the soft term.

Other experiment-related uncertainties

The uncertainty on the scale factor used to correct the pile-up in simulation to match those observed in data, is taken into account by a single NP. Furthermore, as already discussed in Section 4.2.6, the measured integrated luminosity has an uncertainty. This overall uncertainty of $\pm 0.83\%$ is taken into account.

12.2. Theoretical uncertainties

Systematic uncertainties, related to theoretical aspects of this analysis, are discussed below.

MC modelling-related uncertainties

The modelling of MC events is subject to inherent systematic uncertainties, first and foremost the choice of the MC generator itself. Additionally, certain parameters within the modelling chain are chosen ad hoc or as a result of a tune to measured

data. To estimate potential systematic effects arising from the chosen values, these values are varied and the resulting variations is taken into account as systematic uncertainty.

Choice of the MC ME generator: For $t\bar{t}W$, the choice of the MC ME generator (nominal **Sherpa** 2.2.10) is compared to **aMC@NLO+Pythia8** FxFx and the difference is used as a one-sided variation.

Choice of the MC PS generator: For several major backgrounds, the choice of the PS generator is compared to an alternative PS generator and their difference is used as a one-sided systematic variation. Namely, for the processes $t\bar{t}$, $t\bar{t}H$, $t\bar{t}Z$ and for both FCNC signals, the **Pythia8** shower generator is compared to **Herwig7.1.3**. For $t\bar{t}W$, a shower systematic uncertainty is implemented by producing a $t\bar{t}W$ sample using **Powheg Pythia8** and comparing it to a $t\bar{t}W$ **Powheg Herwig** sample. The difference between both samples is then applied to the nominal $t\bar{t}W$ **Sherpa** sample.

Choice of μ_R and μ_F : The choice of the renormalization scale and factorization scale μ_R and μ_F in the ME computation is each varied by a factor of 0.5 downwards and a factor of 2 upwards while keeping the other one constant.

Var3c and final state radiations: The tuned set of internal parameters of **Pythia8** (A14 tune) is varied and the resulting variations are used as systematic uncertainties. For this “Var3c” variation, among other parameters, the strong coupling constant α_S is set to its up- and down variation. An FSR uncertainty is taken into account by varying the renormalization scale μ_R of the PS generator by a factor of 0.5 downwards and a factor of 2 upwards. This is done for the $t\bar{t}$ and $t\bar{t}H$ backgrounds as well as for the decay-FCNC process.

hdamp: *hdamp* is a parameter used by **Powheg** controlling the radiation of the first emission, thus describing the recoil against the $t\bar{t}$ system. The nominal choice of $hdamp = 1.5 m_t$ is somewhat arbitrary and so a systematic uncertainty is applied on this parameter by producing an alternative sample with $hdamp = 3 m_t$. The difference between the two samples is taken as variation for the *hdamp* parameter. This uncertainty is only considered for the $t\bar{t}$ background.

p_T^{hard} : The p_T^{hard} parameter is a flag in **Powheg**, changing the definition of the vetoed region of the parton shower to avoid overlap between **Powheg** and **Pythia8**[141]. An alternative sample for $t\bar{t}$ with the p_T^{hard} parameter set to 1 is used to estimate the uncertainty, whereas $p_T^{\text{hard}} = 0$ is the default.

Parton distribution function: The used PDF set is provided with a set of 30 eigenvectors, which can be varied as an systematic uncertainty. All processes with the PDF4LHC15 PDF set (namely $t\bar{t}$, $t\bar{t}H$ and the decay-FCNC process) have 30 individual NP, corresponding to the 30 eigenvectors available for the

PDF variation. A variation is determined by comparing the reweighted MC sample to its respective nominal sample.

A-priori cross section uncertainty

Since the likelihood fit is designed to control the normalization of certain background processes, these processes have a free-floating normalisation factor, i.e. there is no penalty in the likelihood function for normalising these samples at will. Other background processes have an a-priori cross section uncertainty, where the given value corresponds to the 1σ uncertainty derived from a theoretical prediction or measurements of these processes. The processes with a free-floating normalisation factor are:

- HF-decay e
- HF-decay μ
- q -flip e
- $t\bar{t}W$
- $t\bar{t}Z$

The a-priori cross section uncertainties for the other background processes are taken from publications with similar phase spaces and are given below:

- tZq : 30% [142]
- tWZ : 30% [142]
- $t\bar{t}H$: 15% [143]
- Diboson: 50% [144] (decorrelated among number of charged leptons and flavour of additional jet)
- Minor non-prompt templates: Ad hoc 50%

12.3. Implementation of systematic uncertainties

Systematic uncertainties are propagated to the non-prompt templates whenever possible. This means that for systematic uncertainties, which are available for all processes, the respective variation is used in all processes to define the variation for the non-prompt template. E.g. the b -tagging-related uncertainty “ b_PC_01 ” for the HF-decay e process is composed of “ $b_PC_01(t\bar{t})$ ” + “ $b_PC_01(V+jets)$ ” + “ $b_PC_01(\dots)$ ”. On the other hand, systematic uncertainties which are only available for certain processes are propagated to the non-prompt templates by using the nominal template for all other backgrounds, which do not have the respective variation. E.g. the shower variation for $t\bar{t}$ for the HF-decay e template is composed of “shower($t\bar{t}$)” + nominal($V+jets$) + “nominal($\dots$)”. Table 12.1 presents an overview of all implemented systematic uncertainties, including the number of NPs they add to the likelihood fit. Because certain systematic uncertainties are decorrelated (have an independent NP in the likelihood fit) across specific regions or processes, the stated number of NP may be larger than described in the dedicated section.

Optimising a likelihood fit with that many free parameters (also considering all γ factors) is a huge computational task. To simplify the challenge of finding a global minimum of the likelihood function and also to stabilise the fit itself, the systematic uncertainties undergo the following steps before entering the actual fit.

Smoothing: The MC samples of systematic uncertainties which require separate samples typically have lower statistical accuracy compared to their nominal counterparts. Because of this, such uncertainties are prone to be more affected by statistical fluctuations. These statistical fluctuations, however, are not taken into account in the likelihood function. To mitigate their effect, the distribution of systematic uncertainties are smoothed. Following a specific temporary rebinning and back-interpolation procedure, the smoothing algorithm in general alters the distribution which is passed into the likelihood fit.

Symmetrisation: Symmetrisation improves the stability of the fit. However, a distinction is made between one-sided and two-sided symmetrisation. Certain systematic uncertainties are only available in *one direction*, e.g. shower systematic uncertainties, where only a single variation is available. The normalisation of such uncertainties is then bin-by-bin mirrored to the other side, which is called “one-sided symmetrisation”. However, the majority of systematic uncertainties has an up- and a down variation defined. For some of these uncertainties, a “two-sided symmetrisation” is implemented, where the mean of the up- and the down variation is calculated bin-by-bin. This bin-by-bin mean is then used as up-variation and its mirrored version as the down-variation.

Table 12.1.: Summary of all systematic uncertainties, implemented in the analysis, accompanied by their number of NP in the likelihood fit and a short description.

Category	Systematic	N_{NP}	Description
Experimental	Muon SF	10	Stat/Syst comp. for Trigger, TTVA, Isol., ID, ID_low p_{T}
	Electron SF	4	Trigger, Reco., Isol., ID
	JES	30	See Section 12.1
	JER	13	<i>FullJER</i> scheme
	Flavour-tagging	85	45 b -jet and 20 c /light-jet related
	$E_{\text{T}}^{\text{miss}}$	3	Soft term variation, 1 para., 2 perp.
	Other	3	Pile-up, luminosity and JVT
Total number of experimental-related NP: 148			
Theory	μ_R, μ_F	20	Decorrelated across processes
	FSR, Var3c	2	Decorrelated across processes
	MC generator	6	See Section 12.2
	PDF variation	30	Variations for PDF4LHC
	Free normalization	5	HF-decay e , HF-decay μ , q -flip e , $t\bar{t}W$ and $t\bar{t}Z$
	A-priori cross section	5	tZq , tWZ , $t\bar{t}H$, VV and minor non-prompt templates
Total number of theory-related NP: 68			
Total number of NP: 216			

For certain systematic uncertainties, however, a specific difference between the up- and the down variation is expected. In such cases, no symmetrisation is applied.

Pruning: Depending on the phase space of an analysis, specific systematic uncertainties can have a larger or a smaller impact on the resulting variation. In order to filter out insignificant systematic uncertainties, a pruning step is applied which reduces the required computing power. For this, each variation is divided into a shape component and a normalisation component. Each component can be not taken into account individually, if the respective component is not above a certain threshold. This component is then *dropped*. This pruning procedure is done separately for every systematic uncertainty in every region. In the statistical analysis at hand, the threshold is set to be 1%, meaning that the shape and the normalisation effect of a systematic uncertainty must be larger than 1% to not be dropped. Studies with a lower threshold of 0.001

have been performed to validate that the pruning does not induce significant changes in the result.

Two examples of important systematic uncertainties, also showing the smoothing and symmetrisation, can be seen in Figure 12.1. The black line shows the expected event yield of the nominal sample in the respective region, whereas the solid red and blue line shows the expected event yield of the up- and down variation respectively after symmetrisation and smoothing. The dotted red/blue crosses shows the actual value of the variation before being altered by the smoothing algorithm and the symmetrisation. The b -tagging related uncertainties are so-called weight systematics, which means that in the calculation of the systematic variation, a weight is exchanged in comparison to the nominal calculation. As a result, the systematic variations of weight systematic uncertainties have the same MC statistics available as the nominal sample.

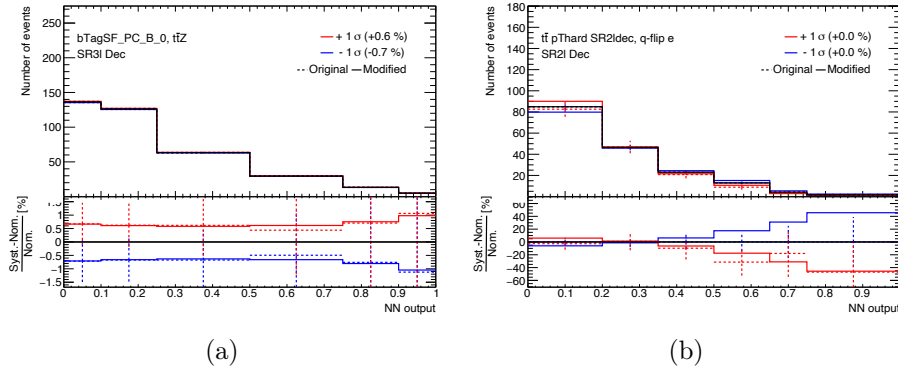


Figure 12.1: The principle component 0 of the b -jet tagging related uncertainties of the $t\bar{t}Z$ background in the SR3 ℓ Dec region (a) and the $t\bar{t} p_T^{\text{hard}}$ systematic of the q -flip e template in the SR2 ℓ Dec region (b).

This is the reason why the magnitude of the MC statistical uncertainty of the systematic b -tagging variation is only in the order of 0.5% (visible by the dashed “Pluses” as data points in the ratio plot in Figure 12.1). The smoothing algorithm only slightly changes the shape of the distribution into the solid red and blue line, which can be seen best in the fourth and sixth bin. In contrast, the p_T^{hard} variation of the q -flip e template is derived from a single alternative sample for the $t\bar{t}$ process. Alternative systematic variation samples in general have less MC statistics available, which results in larger MC statistical uncertainties. However, the smoothing algorithm nicely transforms the fluctuating systematic variation into a smooth distribution. The correct processing of all systematic variations including a reasonable smoothing by the smoothing algorithm is checked in order to prevent erroneous results and to prevent the likelihood minimisation algorithm from not finding the global minimum.

13. FCNC search results

This chapter presents the result of the search for FCNC interactions. However, because the paper of the corresponding analysis is still ATLAS internal and has not been published at the time of writing this dissertation, the fit setup does not use measured data in the most signal-sensitive high-NNout bins¹. Instead, the fit setup is as follows: First, a background-only fit is performed using data in all CRs and the low-NNout SRs to obtain a realistic estimation for the normalisation of the free-floating background processes and also to observe and understand pulls/constraints of the NPs corresponding to systematic uncertainties. Afterwards, a “Realistic Asimov Fit” is performed which uses the MC expectation, modified by the previously determined normalisation factors and post-fit NPs, as data. This way, the analysis is still blinded² but a realistic estimate for the sensitivity of the analysis is obtained. Unblinding the analysis to observe an FCNC signal or to set observed upper exclusion limits for the FCNC signal normalisation is thus beyond the scope of this dissertation. In Section 13.1 the background-only fit is discussed along with important observed discrepancies between data and MC expectation and the resulting pulls and constraints. Afterwards, Section 13.2 contains the results of the realistic Asimov fit, whereas Section 13.3 presents the sensitivity of the analysis together with a comparison to similar analyses. Since the kinematic distributions between the left-handed and right-handed production-FCNC process are equal within statistical uncertainties, the mean of both left- and right-handed components is used in the fit. Because the results are very similar between the utH and the ctH fit, only plots related to the utH fit are shown in the main body.

13.1. Background-only fit

In Appendix F, Figures F.1 and F.2 show all 2ℓ SS regions whereas Figures F.3 and F.4 show all 3ℓ regions, with the pre-fit plots in the upper row and the post-fit plots in the lower row respectively. The error bands in the pre-fit plots always contain all systematic uncertainties added in quadrature, however, neglecting any free-floating normalisation factors and correlations between the systematics. The

¹The most signal-sensitive high-NNout bins are defined as bins which have a signal-to-background ratio larger than 5% with the FCNC signals scaled up to the current upper exclusion limit of $C_{u\phi} = 1.53$ ($C_{u\phi} = 1.47$) for the utH (ctH) coupling respectively.

²The analysis avoids examining the most signal-sensitive bins to prevent any bias, as required by the ATLAS collaboration for analyses. Otherwise, optimising on statistical fluctuations in data could result in either a false observation or limits that are too stringent for the given data.

error bands in the post-fit plots also include the (now determined and constrained) normalisation factors and the correlations between the uncertainties. Exemplarily, the distributions of the CR3 ℓ HFe and the SR3 ℓ Prod region are shown in Figure 13.1. A reasonably good post-fit agreement between MC prediction and data can be observed in all regions.

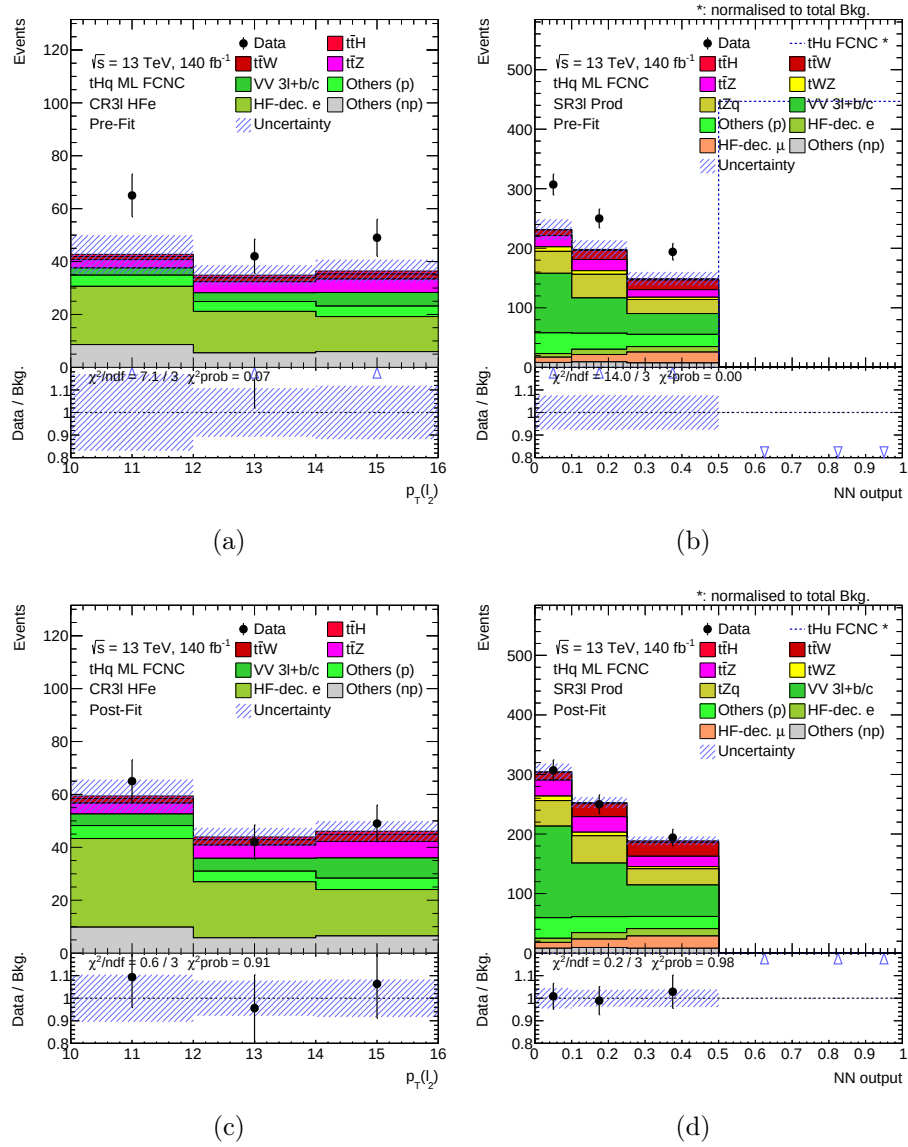


Figure 13.1: The CR3 ℓ HFe and the SR3 ℓ Prod region with their fitted distribution once pre-fit (upper row) and once post-fit (lower row) in the background-only fit.

It can be seen that the MC prediction, especially of the HF-decay e template, underestimates the data, which is compensated by the free floating normalisation factor

$\beta_{HF e}$, as can be seen from the agreement in the post-fit plots. Furthermore, the low NNout regime of the SR3 ℓ Prod yields an excellent χ^2 probability¹. The resulting normalisation factors can be seen in Figure 13.2.

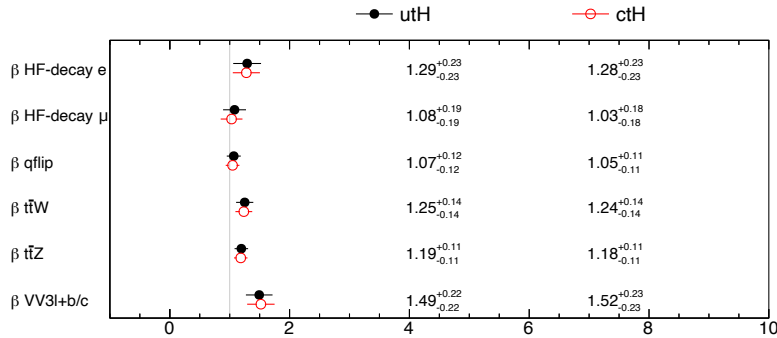


Figure 13.2: Resulting normalisation factors for the free-floating backgrounds in the background-only fit.

The normalisation of the HF-decay e template, the $VV3\ell + b/c$ process and both $t\bar{t}V$ processes is increased, compared to the prior MC prediction. The VV process requires both the W and Z bosons to decay leptonically in order to enter the 3ℓ regions, resulting in no b -jets at tree-level. As the region definitions for the SRs require a b -tagged jet, either this additional heavy-flavour jet must be created in the PS simulation or a light-flavour jet must be mis-tagged. However, PS simulations are prone for inaccurately predicting additional heavy quark radiation, which can explain why the normalisation factor for $VV3\ell + b/c$ is greater than 1. The correlation matrix for the NPs in the fit that are correlated stronger than 25% with any other NP other than themselves can be seen in Figure F.5 in the appendix. The correlation matrix reveals correlations between the normalisation factors of background processes and other systematic uncertainties. Additionally, some modelling-related systematic uncertainties of the $t\bar{t}$ background exhibit a high level of correlation to each other within the CR3 $\ell t\bar{t}W$ region. Exploiting correlations in addition to constraining uncertainties is the reason for reduced uncertainty bands in the post-fit plots and the resulting good sensitivity of analyses utilising profile-likelihood fits, compared to e.g. simple cut-and-count analyses.

An interesting observation can be made inspecting the γ -factors in Figure 13.3, which are related to the MC statistics. Each bin in every region has its own γ -factor, which describes the statistical uncertainty introduced by limited MC statistics in

¹The χ^2 probability is the p-value of the Mahalanobis distance as test statistic and is therefore a measure for the compatibility between MC and data, taking into account correlations and constraints. Higher values of the χ^2 probability $\in (0, 1]$ correspond to better agreement between data and MC.

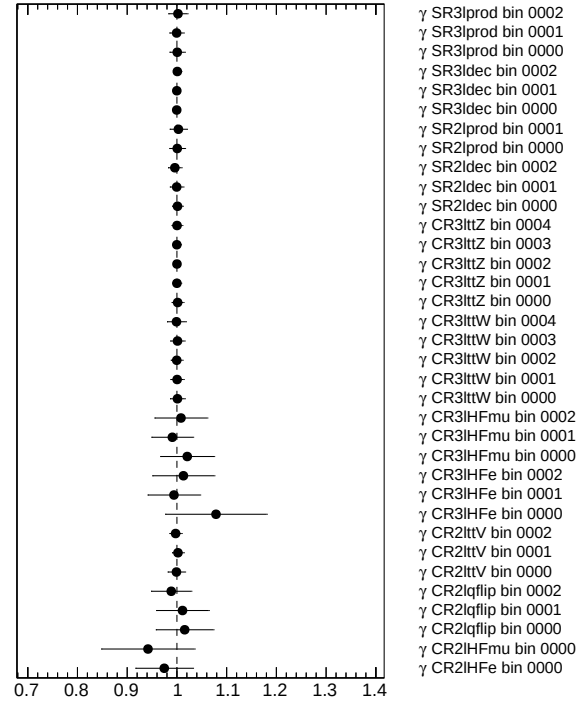


Figure 13.3: Pulls and constraints of the γ -factors in the background-only fit, exemplarily for the utH case.

that bin. While all CRs associated with prompt background processes and all SRs have comparably good MC statistics (as indicated by the strong constraint/small uncertainty of the respective NP), the γ factors in the CRs associated with the non-prompt templates are significantly less constrained. A study implementing separate γ -factors for each process in each bin and region discovered that the low MC statistics originate from the non-prompt templates. Unfortunately, it is not possible to selectively generate more non-prompt MC events, so it is not straightforward to reduce the uncertainty introduced by the low MC statistics of the non-prompt templates.

In Appendix F, Figures F.6 - F.10 show the resulting pulls and constraints of all NPs related to systematic uncertainties for both the utH fit and the ctH fit. As most NPs are neither significantly pulled nor constrained, only the interesting NPs are shown and discussed in the main body. In the background-only fit, the differences between the utH and ctH fit are small, as almost all NPs are equally pulled/constrained in the utH and ctH fit. Nevertheless, slight variations in the trained NNs can cause differences in the acceptance of events for systematic uncertainties, leading to the pruning of certain NPs for one fit but not for the other. This phenomenon can be

observed for specific systematic uncertainties, wherein an entry is absent for one fit setup but not for the other. The first interesting pulls can be observed in the group of b -tagging related systematic uncertainties, where the most important ones can be seen in Figure 13.4.

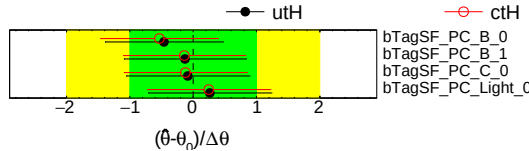


Figure 13.4: Most important pulls and constraints in the regime of b -tagging related NP in the background-only fit.

With the requirement for one or two b -tagged jets for the regions, the main components of the b -tagged NP are likely to have an impact on the fit. Such an effect is noticeable in minor pulls, especially for the b -jet-related NPs, but also for the light-flavour-jet-related NPs. Nevertheless, the other 81 b -tagging related NPs are neither constrained nor pulled at all.

The next group of NPs to be discussed are the modelling-related NPs. Figure 13.5 presents the complete set of modelling systematic uncertainties. Some larger pulls (up to 0.6σ) can be observed for the $t\bar{t}W$ FxFx bin-to-bin migration, the $t\bar{t}Z$ aMCHw7 and the $t\bar{t}Z$ Var3 related NPs. These pulls originate in the CR3 $l\bar{t}tW$ and CR3 $l\bar{t}tZ$ region respectively, as can be seen in Figure F.3 (c) and F.4 (a). Besides the overall excess of data over MC in these two regions, a shape difference is present. Especially the first bin in the CR3 $l\bar{t}tZ$ region exhibits this behaviour. The free-floating normalisation factors cannot compensate such a shape difference, as they can only overall scale the respective processes. To correct the shape difference, instead, the above mentioned systematic uncertainties are pulled. The remaining modelling-related systematic uncertainties are only slightly constrained and not significantly pulled. Some $t\bar{t}$ -related systematic uncertainties are slightly constrained, which can be attributed to the fact that the non-prompt background processes are strongly affected by systematic uncertainties altering the (behaviour of the) PS. As the HF-decay processes rely on leptons emerging from jet cones in the PS, such systematic uncertainties naturally have a large impact and are thus prone to be pulled or constrained in case of slight differences between data and MC.

Figure 1 is a plot showing the difference in the effective number of degrees of freedom, $\Delta\theta$, between the NP and EFT models, for various EFT models and NP models. The x-axis is labeled $(\theta - \theta_{\text{EFT}})/\Delta\theta$ and ranges from -2 to 2. The y-axis lists 20 different models. The plot shows two data series: uTH (black circles) and cTH (red circles). The background is green for values between -1 and 1, and yellow for values outside this range. The models are listed on the right side of the plot.

Model	uTH $\Delta\theta$	cTH $\Delta\theta$
JET BJES Response	0.0	0.0
JET Eft. NP Detector 1	0.0	0.0
JET Eft. NP Detector 2	0.0	0.0
JET Eft. NP Mixed 1	0.0	0.0
JET Eft. NP Mixed 2	0.0	0.0
JET Eft. NP Mixed 3	0.0	0.0
JET Eft. NP Modelling 1	0.0	0.0
JET Eft. NP Modelling 2	0.0	0.0
JET Eft. NP Modelling 3	0.0	0.0
JET Eft. NP Modelling 4	0.0	0.0
JET Eft. NP Statistical 1	0.0	0.0
JET Eft. NP Statistical 2	0.0	0.0
JET Eft. NP Statistical 3	0.0	0.0
JET Eft. NP Statistical 4	0.0	0.0
JET Eft. NP Statistical 5	0.0	0.0
JET Eft. NP Statistical 6	0.0	0.0
JET η -Intercalib. Modelling	0.0	0.0
JET η -Intercalib. NonClosure 2018data	0.0	0.0
JET η -Intercalib. NonClosure highEta	0.0	0.0
JET η -Intercalib. NonClosure negEta	0.0	0.0
JET η -Intercalib. NonClosure posEta	0.0	0.0
JET η -Intercalib. TotalStat	0.0	0.0
JET Flavor Comp.	-0.5	-0.5
JET Flavor Response	0.0	0.0
JET Pileup OffsetMu	0.0	0.0
JET Pileup OffsetNPV	0.0	0.0
JET Pileup PTerm	0.0	0.0
JET Pileup RhoTopology	-0.5	-0.5
JET PunchThrough MC16	0.0	0.0
JET PunchThrough AFII	0.0	0.0
JET RelativeNonClosure AFII	0.0	0.0

For the JER (see Figure F.8 (a)) and the JES systematic uncertainties (see Figure 13.6), some larger pulls are observed. In particular, for the JES-related NPs, the *JET Pileup ρ topology*, *Jet Flavour Composition/Response* and the *Jet η intercalibration modelling* are pulled up to 0.6σ . Splitting these pulls by region reveals that the origin of the pulls are again the CR3 $\ell t\bar{t}W$ and the CR3 $\ell t\bar{t}Z$ region. Additional smaller pulls can be observed in the theoretical cross section prediction of the tZq and tWZ processes, but also for the $VV4\ell + b/c$ process without the free-floating NF, as can be seen in Figure 13.7. The pulls of the former process also originate in the CR3 $\ell t\bar{t}Z$ and region, while the tWZ process is normalised downwards due to the SR3 ℓ Dec region. The slight increase in the $VV4\ell + b/c$ process normalisation matches with the observation of the $VV3\ell + b/c$ process normalisation.

Validation of the input variable modelling for the NN is another important aspect of an analysis, as without well-modelled input variables, the output of NNs cannot be employed for the analysis. It can generally be discussed whether to validate the input variables pre- or post-fit. However, in this analysis, certain background processes and templates are significantly underestimated by the MC prediction and are therefore corrected by normalisation factors in different CRs, so the input variables are validated post-fit. Otherwise, it would not be clear whether the discrepancy between the MC prediction and the data is due to the underestimation of certain backgrounds or to other issues in the MC modelling. A comprehensive set of validation plots including pre- and post-fit plots of various kinematic distributions (mainly NN input variables) can be seen in Appendix G. It is important to bear in mind that the kinematic distributions were not subjected to fitting and therefore, an equally good χ^2 probability as the fitted variables is not expected. One important variable, the number of jets in the CR3 $\ell t\bar{t}W$ and the CR3 $\ell t\bar{t}Z$ region exhibit significant differences between data and MC, as can be seen in Figure 13.8. Even post-fit, the agreement between data and MC is still not good, especially in the CR3 $\ell t\bar{t}Z$ region, as indicated by the χ^2 probability. The source of this shape discrepancy between the data and MC prediction is intriguing, especially because a similar surplus in data has been identified within the ATLAS differential $t\bar{t}W$ cross-section measurement, examining a similar phase space[145]. Nonetheless, identifying the origin of this mismatch is beyond the scope of this analysis. However, the remaining kinematic distributions are judged to be well modelled. Some other variables, e.g. the transverse momentum of the sub-leading lepton in the CR3 ℓHFe region (Figure G.7 (b) and (e)), show not so good agreement between the data and the MC prediction (as indicated by the χ^2 probability), but for statistical reasons it is to be expected that some variables will have bad agreement when looking at many distributions.

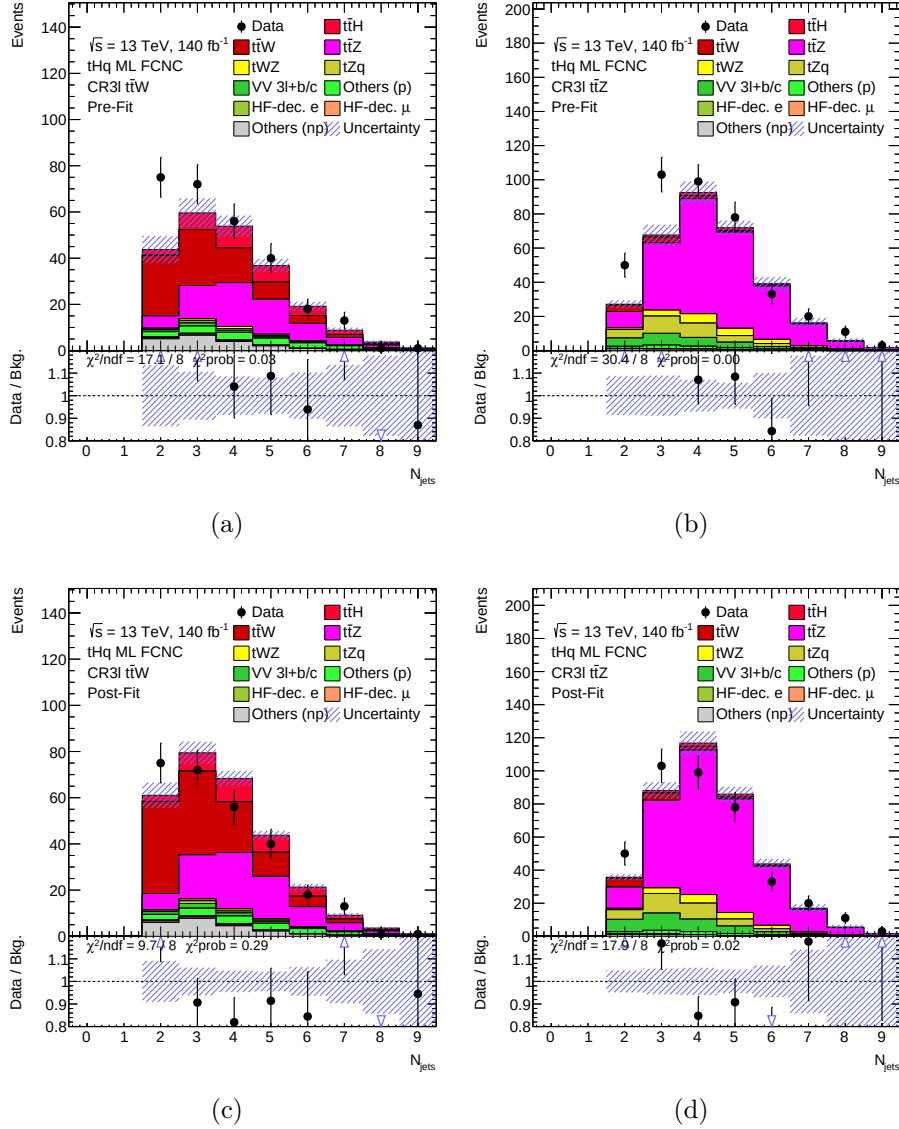


Figure 13.8: The distribution of the number of jets in the $CR3l t\bar{t}W$ region (left) and the $CR3l t\bar{t}Z$ region (right) once pre-fit (upper row) and once post-fit (lower row) in the background-only fit.

With a thorough understanding of all pulls and constraints, the absence of any extreme pulls or highly constrained NPs, obtaining realistic normalisation factors for the free-floating backgrounds and the validation of the NN input variables, the analysis's sensitivity is determined in the following section.

13.2. Realistic Asimov fit

Using the normalisation factors and pulls from the background-only fit, now a realistic Asimov fit is performed. For this, the normalisation factors and the pulls of NPs from the background-only fit are applied on the MC prediction and this is used as data. As a consequence, similar pulls will be observed in the realistic Asimov fit, compared to the background-only fit. This way, the full SRs can be used to determine the sensitivity of the analysis without the need for unblinding. In Appendix H, Figures H.1-H.2 show all 2ℓ SS region plots whereas Figures H.3-H.4 show all 3ℓ region plots. Exemplarily, the distributions of the CR3 ℓ HF μ and the SR3 ℓ Dec region are shown in Figure 13.9.

As expected, and similar to all other regions, the uncertainty bands are reduced by the fit in the post-fit plots by exploiting constraints and correlations between systematic uncertainties. Because the realistic Asimov dataset was created using the pulls and normalisation factors from the background-only fit, fitting the MC expectation to the Asimov dataset leads to similar pulls and normalization factors in the realistic Asimov fit as in the background-only fit. The full set of pulls and constraints in the realistic Asimov fit can be seen in Appendix H Figure H.6-H.10, which confirms the similarity between both fit setups. The resulting normalisation factors can be seen in Figure 13.10, which are compatible with the normalisation factors from the background-only fit.

Figure 13.11 shows a ranking of systematic uncertainties in the utH fit (a) and in the ctH fit (b) in a descending order in terms of their respective impact on the POI (using the scale on the upper end of the plot), including their respective pulls and constraints (using the scale at the bottom of the plot). For this ranking, every systematic uncertainty was set to its $\pm 1\sigma$ pre-fit and post-fit value and the impact on the POI is calculated. Thereby, the pre-fit impact can be seen as bordered bar and the post-fit impact can be seen as filled bar, with the up-variation as a dark-blue bar and the down-variation as a light-blue bar. At the same time, the pull and constraint of the respective NP can be seen as a black dot with its uncertainty band.

The first thing one notices is that nearly all top-ranked systematic uncertainties in the utH fit are also top-ranked in the ctH fit, with the only exception of the $b\text{Tag}SF_PC_B_0$ and the $t\bar{t}W\ Fx\ Fx\ (shape)$ systematic uncertainty. Because of the similarity between both fit setups (same background processes, same kinematic distributions in the CRs and similar NN output distributions in the SRs) this is a sign of stable fit setups, not being affected strongly by random statistical effects. However, in both fit setups nearly all systematic uncertainties are related to the estimation of the non-prompt templates. The normalisation factors in the ranking plot are of course directly related to the estimation of the non-prompt templates, but

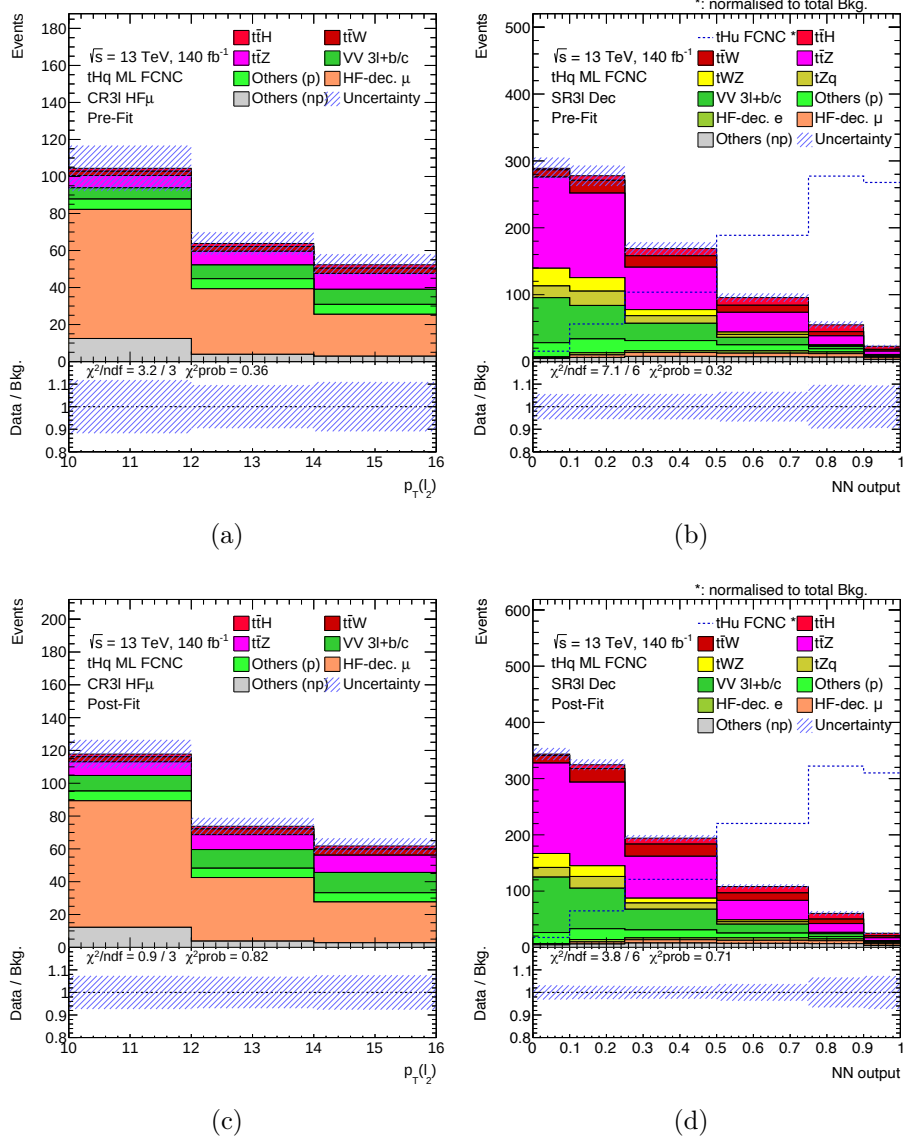


Figure 13.9: The CR3lHF μ and the SR3lDec region with their fitted distributions once pre-fit (upper row) and once post-fit (lower row) in the realistic Asimov fit.

also all $t\bar{t}$ modelling systematic uncertainties are affecting mainly the non-prompt templates since $t\bar{t}$ is the major component of the non-prompt templates. Beside the presence of the principle component of the b -jet tagging and a muon-related NP, also two NPs are related to the $t\bar{t}W$ background, likely entering the ranking of the most-important NPs due to the high contribution of the $t\bar{t}W$ process to in the high-NNout bins especially in the 2lSS SRs.

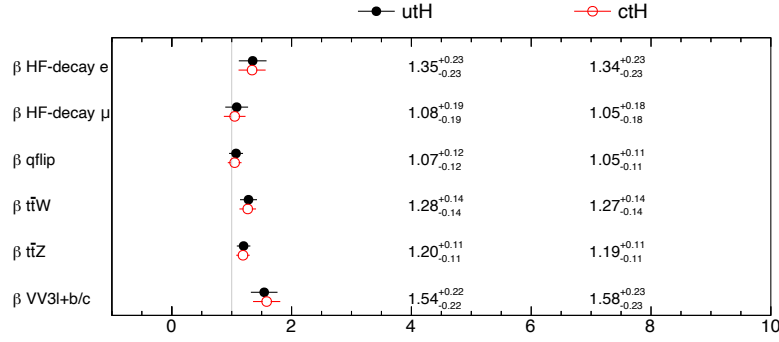


Figure 13.10: Resulting normalisation factors for the free-floating backgrounds in the realistic Asimov fit.

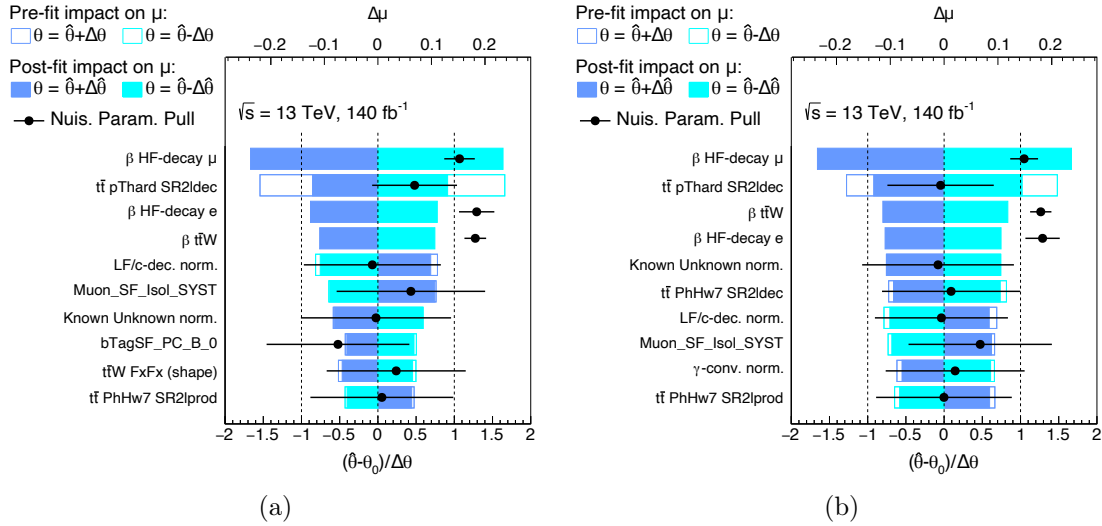


Figure 13.11: Ranking plots of the systematic uncertainties in the utH fit (a) and the ctH fit (b), showing the importance of systematic uncertainties in a descending order, including their respective pulls and constraints.

For highly ranked and thus important systematic uncertainties, a thorough investigation of the behaviour of the likelihood function and the impact on the POI itself can be performed. This is done by performing a so-called likelihood scan and impact scan, wherein the likelihood value and the impact on the POI in dependence of the varied NP value is evaluated. These two scans are performed exemplarily for the $t\bar{t} p_{T\text{hard}} SR2\ell Dec$ systematic uncertainty in the ctH fit to validate the proper response of the likelihood function and the POI on changes of the respective NP. The results of this test can be seen in Figure 13.12.

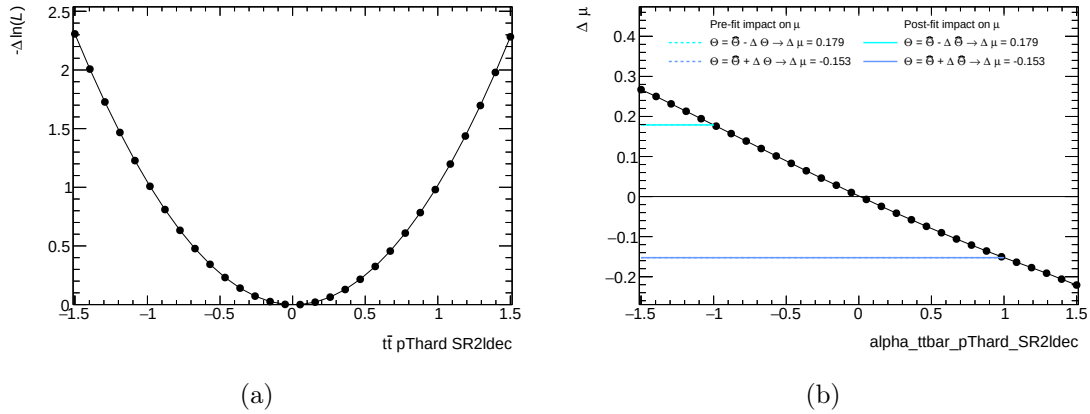


Figure 13.12: A likelihood scan (a) and an impact scan (b) of the “ $t\bar{t} p_{T\text{hard}} \text{SR2lDec}$ ” systematic uncertainty in the case of the ctH fit.

In the likelihood scan a smooth parabola-like curve can be observed, which is directly expected from the constraint term for systematic uncertainties in the likelihood function. This parabola can in principle also be deformed due to other effects in the likelihood fit, however, in this case the systematic uncertainty responds *normally*¹. Spikes or other non-smooth behaviour in the likelihood scan *could* have indicated instabilities in the fit. In the impact scan an approximate linear response of the POI on the NP related to the $t\bar{t} p_{T\text{hard}} \text{SR2lDec}$ systematic uncertainty can be observed. A linear response indicates that the systematic uncertainty is dominated by a shape effect rather than a normalisation effect, as the shape alters the yield linearly (see Equation 11.5) whereas the normalisation alters the yield exponentially (see Equation 11.6). A change in the yield is then directly related to a change in μ . This observation also fits with the red/blue plot shown in Figure 12.1 (b), which shows the very same systematic uncertainty with a strong shape effect and a pruned normalisation effect. With the validation of the properly working likelihood fit, now the evaluation of the best-fit value and the expected exclusion limits can be performed in the next section.

¹The likelihood function is a function with hundreds of potentially correlated parameters, where the *nominal* response for systematics can be altered due to correlation/constraint effects. Also the impact can be non-linear and distorted for e.g. large systematics which experience physical boundaries (like the non-negativity of event yields). In principle, an abnormal-looking systematic uncertainty must be thoroughly investigated for correct implementation and smooth likelihood-fit response, but an abnormal-looking systematic must not automatically be an indication of problems.

13.3. Realistic Asimov fit results and exclusion limits

This section summarises and discusses the outcome of the realistic Asimov fit, both for the utH and the ctH fit. First the best-fit value is examined in the case of no signal injection. The resulting values of the POI μ , i.e. the FCNC signal normalisation factor, can be seen in Table 13.1.

Signal	$\mu_{\text{best-fit}}$
utH	$-0.027^{+0.371}_{-0.371}$
ctH	$-0.034^{+0.453}_{-0.453}$

Table 13.1.: The best-fit values of the signal-FCNC normalisation factor μ for both the utH and the ctH fit case, including the full set of systematic uncertainties.

For the utH fit as for the ctH fit, the signal-FCNC normalisation factor μ is well compatible with 0, meaning that no excess of data over MC expectation is found, as it is expected for a realistic Asimov fit. As a next step, the exclusion limits on μ are calculated and documented in Table 13.2. Because it is of interest to see the worsening of the limits from a fit setup using only statistical uncertainties to a fit setup using the full set of systematic uncertainties, both cases are documented in the respective table.

Fit type	Signal	95% CL_S expected exclusion limits				
		-2σ	-1σ	μ	$+1\sigma$	$+2\sigma$
stat. only	utH	0.2625	0.3524	0.4891	0.6807	0.9157
	ctH	0.3342	0.4487	0.6227	0.8668	1.166
syst incl	utH	0.3924	0.5268	0.7311	1.001	1.338
	ctH	0.4717	0.6333	0.8789	1.211	1.610

Table 13.2.: The 95% CL_S exclusion limits of the normalisation factor μ on the utH/ctH signal-FCNC samples. The results of a fit using only statistical uncertainties and the results of a fit using the full systematic uncertainty set are shown.

The first fact to be observed is, that the utH-related limits are in both fit setups always lower/better, compared to the ctH-related limits. This is because the utH prod-FCNC process has a larger cross section compared to the ctH prod-FCNC process, which is explained by the proton PDF having real up-quarks but only sea-charm-quarks. A larger expected event yield can naturally be better excluded.

Furthermore, the limits which are extracted from the fit including the full set of systematic uncertainties are approximately 40-50% larger compared to the limits extracted from the statistical-only fit. A certain effect is always to be expected by the inclusion of the full set of systematic uncertainties, but this means that the analysis is still statistically limited.

Because the exclusion limits are dependent on the FCNC samples used (e.g. Wilson coefficient used for producing, branching ratio of W/Z boson forced to leptonic decays or not, etc...), it can become complicated to compare the results of this analysis with other analyses, searching for the same coupling in different channels. For this reason, the limits on the signal strength (including the systematic uncertainties) are transformed into limits on the Wilson coefficient $C_{u\phi}$. To also be able to compare to older analyses using other EFT approaches, the limits on the Wilson coefficients are transformed into limits on the branching ratio of the top quark decaying to an up-type quark and the Higgs boson $\mathcal{B}(t \rightarrow qH)$, following Equation 3.4 and 3.5. Both alternative limits are given in Table 13.3.

Limit setting	Signal	95% CL_S expected exclusion limits				
		-2σ	-1σ	mean	$+1\sigma$	$+2\sigma$
$C_{u\phi}$	utH	0.626	0.726	0.855	1.00	1.16
	ctH	0.687	0.796	0.937	1.10	1.27
$\mathcal{B}(t \rightarrow qH) [10^{-4}]$	utH	2.2	3.0	4.1	5.6	7.6
	ctH	2.7	3.6	5.0	6.8	9.1

Table 13.3.: The 95% CL_S exclusion limits of the Wilson coefficient $C_{u\phi}$ and the branching ratio $\mathcal{B}(t \rightarrow qH)$ on the utH/ctH signal-FCNC samples. All limits are given for the fit setup including the full set of systematic uncertainties.

The exclusion limits on the branching ratio are then compared to similar analyses. The limits obtained in this analysis are always compared to expected exclusion limits and not to observed exclusion limits, since the expected limits reflect the sensitivity of the analysis, whereas the observed limits are affected by fluctuations in data. The first analysis to compare with is the tHq FCNC multi-lepton analysis with the early-Run 2 ATLAS dataset of $\mathcal{L}_{\text{int}} = 36 \text{ fb}^{-1}$ [146]. The comparison with this analysis is of particular interest because this analysis also used $2\ell\text{SS}$ and 3ℓ regions with similar definitions, however, the early-run 2 tHq FCNC multi-lepton analysis did not include the prod-FCNC process. The comparison to the ctH channel (considering the fourfold integrated luminosity) thus shows the advancements made in the analysis technique, whereas the comparison to the utH channel shows the improvement by including the prod-FCNC process in addition to the advanced analysis techniques. The other analysis to which the presented results are compared to is the combination

of the various tHq FCNC analyses of the early-run 2 dataset[147]. The combination includes the $H \rightarrow b\bar{b}$, $H \rightarrow \tau\bar{\tau}$, $H \rightarrow$ multi-lepton and the $H \rightarrow \gamma\gamma$ decay channels, where the $H \rightarrow$ multi-lepton and the $H \rightarrow \gamma\gamma$ decay channel were the most sensitive analyses of the four. The comparison of the expected upper exclusion limits of the previously mentioned analyses to the limits obtained in this analysis is shown in Figure 13.13.

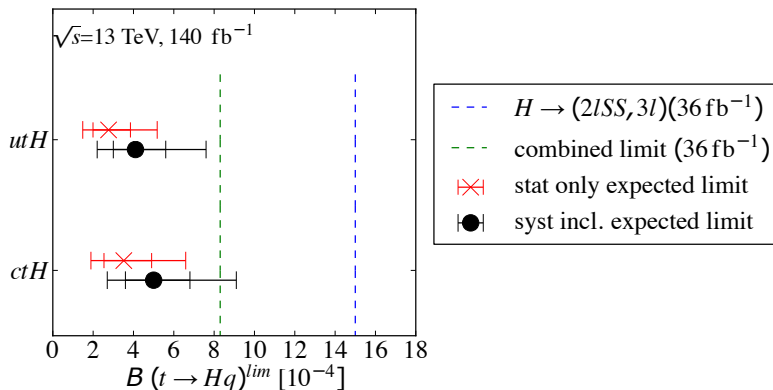


Figure 13.13: Comparison of the expected upper exclusion limits of the tHq FCNC multi-lepton analysis based on the $\mathcal{L} = 36 \text{ fb}^{-1}$ data set and the combination of the $\mathcal{L} = 36 \text{ fb}^{-1}$ tHq FCNC analyses to the expected upper exclusion limits obtained in this analysis.

Comparing with the early-run2 dataset multi-lepton analysis, it can be observed that the expected limits were improved much more compared to the naive expectation when using four times more data statistics. One reason for this improvement is the implementation of the prod-FCNC process in this analysis, which was not the case for the old multi-lepton analysis. This also explains the larger improvement in the utH channel compared to the ctH channel. However, improved analysis techniques, in particular the usage of advanced reconstruction methods and a sufficiently complex NN to discriminate signal from background, can be identified as the main reasons for the improvement. Even the expected exclusion limits of the early-run 2 FCNC combination analysis can be surpassed by nearly a factor of 2. Comparing with recent results with the full run 2 dataset (e.g. the $H \rightarrow \gamma\gamma$ analysis from CMS [148] or the $H \rightarrow \tau^+\tau^-$ analysis from ATLAS [149]), this analysis achieves comparable expected exclusion limits. Comparing to the multi-lepton tHq FCNC analysis utilising full run 2 by CMS [150], the expected upper exclusion limits in the analysis presented here are 20% to 30% better.

14. Conclusion and outlook

This dissertation presents a search for flavour-changing neutral currents in di-lepton same-sign and tri-lepton final states in the top-quark Higgs boson sector with the ATLAS detector at a centre-of-mass energy of $\sqrt{s} = 13 \text{ TeV}$ using the full Run 2 data set. Although flavour-changing neutral currents are strongly suppressed in the Standard Model of particle physics, they are significantly enhanced by several theories beyond the Standard Model. The FCNC interaction is considered as a production-FCNC process and a decay-FCNC process, which has also influenced the definition of the signal regions: Two signal regions are defined in both the di-lepton and tri-lepton regions, each tailored to be enriched in either the prod-FCNC process or the decay-FCNC process. Several control regions are utilised to estimate the rate of specific background processes, including $t\bar{t}W$, $t\bar{t}Z$ and processes involving non-prompt leptons. After reconstructing specific aspects of the event kinematics, various distributions are used as input to a single hidden layer feed-forward neural network to discriminate between signal and background events. A subsequent statistical analysis takes into account the effect of systematic uncertainties due to detector effects or theoretical assumptions. Expected upper 95% exclusion limits on the branching ratio of the top quark decaying into an up-type quark and a Higgs boson $\mathcal{B}(t \rightarrow qH)$ are derived from the statistical analysis. The results show that $\mathcal{B}(t \rightarrow uH) < 4.1 \cdot 10^{-4}$ and $\mathcal{B}(t \rightarrow cH) < 5.0 \cdot 10^{-4}$ for the uH -FCNC vertex and the cH -FCNC vertex respectively.

Since this search for flavour-changing neutral current interactions is primarily statistically limited, the most direct approach to improving sensitivity is to collect additional data. However, since the LHC has transitioned to Run 3 in July 2022, with a higher centre-of-mass energy of $\sqrt{s} = 13.6 \text{ TeV}$, it will not be possible to increase the data statistics of Run 2. But not only the available data statistics is limiting, also the MC statistics (especially of the non-prompt templates) has a non-negligible effect. The optimisation of the neural network structure is another notable factor in improving outcomes. Recent research by the working group in Wuppertal suggests that implementing more complex neural network architectures, such as deep neural networks with a higher number of hidden layers or modern graph neural networks, can lead to a considerable improvement in signal and background separation. The impact of this improved separation on the exclusion limit remains to be established through a detailed examination of the complete statistical analysis chain that factors in systematic uncertainties. Nonetheless, the preliminary results appear promising. Finally, this analysis will be part of a combination of different tHq FCNC analyses, which will then provide the final limit on the tHq FCNC interaction by the ATLAS collaboration.

Appendices

A. b -tagging WP study

For the b -tagging using the DL1r tagging algorithm, there are four different calibrated WPs available, namely the 85%, 77%, 70% and the 60% WP. The percentage relates to the fraction of true b -jets being identified as a b -jet by the algorithm, resulting in the 85% WP being a loose selection criteria and the 60% WP being the strictest WP. The choice of a WP influences the entire analysis since not only the region definition depends on the b -tagging WP, but also the b -tagging related uncertainties potentially strongly vary between the choice of the WP. To determine the optimal WP for this analysis, a study is performed using a preliminary region selection and the measure of $\frac{S}{\sqrt{B}}$ to determine the best WP, where S refers to the number of signal events and B to the number of background events. The preliminary region definition used for this study is listed below:

- exactly 3 leptons with $p_T > 12$ GeV
- $E_T^{\text{miss}} > 30$ GeV
- $n_{\text{jets}} \geq 1$
- exactly 1 b -jet with variable WP

The event yields split by process and the resulting $\frac{S}{\sqrt{B}}$, depending on the WP, can be seen in Table A.1.

As can be seen, the 85% and the 60% WP are not favoured by this study because either too less background or too much signal is removed by the choice of the WP, whereas the 77% and the 70% WP yield similar $\frac{S}{\sqrt{B}}$ values. Because for both signals, the 70% WP performs slightly better than the 77% WP, the 70% b -tagging WP is used throughout the analysis.

Table A.1.: Event yields and the $\frac{S}{\sqrt{B}}$ ratio for different b -tagging WPs for the left- and right-handed utH/ctH signal in a preliminary region. The highest $\frac{S}{\sqrt{B}}$ for each signal case is highlighted.

Process	WP = 85%	WP = 77%	WP = 70%	WP = 60%
VV	928.82	420.54	233.64	126.28
$t\bar{t}V$	445.88	536.97	578.61	594.63
$Z + \text{jets}$	46.14	11.58	3.20	2.44
$t\bar{t}$	4.76	5.70	5.69	5.37
HFdecay e	145.07	132.42	120.85	104.82
HFdecay μ	327.80	297.54	258.10	221.81
ChargeFlip e	31.37	21.36	9.79	8.74
Others	112.13	104.31	101.41	99.81
utH FCNC	60.17	56.17	52.08	45.55
$\frac{S}{\sqrt{B}}$	1.331	1.436	<u>1.438</u>	1.335
ctH FCNC	45.36	46.33	44.23	38.96
$\frac{S}{\sqrt{B}}$	1.003	1.186	<u>1.235</u>	1.150

B. Monte Carlo Samples

This appendix contains a quantitative summary of all used MC samples, including their DSID, the order in which the perturbative QCD was calculated, the used MC generator and PS and the used cross section (X-sec).

Table B.1.: Summary of important information about the prod-FCNC and the decay-FCNC MC samples. The stated cross section corresponds to a Wilson coefficient of $C = 1$.

Process	Description	DSID	pQCD	Generator	Shower	X-sec [pb]
tHq Decay	tHc ($\bar{c}$)	411229	NLO	Powheg	Pythia 8	0.1844
	tHc (c)	411230				0.1844
	tHu ($\bar{u}$)	411231				0.1810
	tHu (u)	411232				0.1810
tHq Prod ($H \rightarrow WW$)	tHu (lh)	411420	NLO	MadGraph aMC@NLO	Pythia 8	0.0092
	tHc (lh)	411421				0.0013
	tHu (rh)	411422				0.0092
	tHc (rh)	411423				0.0013
tHq Prod ($H \rightarrow ZZ$)	tHu (lh)	411424	NLO	MadGraph aMC@NLO	Pythia 8	0.0016
	tHc (lh)	411425				0.0002
	tHu (rh)	411426				0.0016
	tHc (rh)	411427				0.0002
tHq Prod ($H \rightarrow \tau^+ \tau^-$)	tHu (lh)	412098	NLO	MadGraph aMC@NLO	Pythia 8	0.0032
	tHc (lh)	412100				0.0006
	tHu (rh)	412102				0.0032
	tHc (rh)	412104				0.0006

Table B.2.: Summary of important information about the $t\bar{t}$ MC samples.

Process	Description	DSID	pQCD	Generator	Shower	X-sec [pb]
$t\bar{t}$	$\geq 1\ell$	410470	NLO	Powheg	Pythia 8	396.87
	2ℓ	410472				76.95

Table B.3.: Summary of important information about the single-top quark MC samples.

Process	Description	DSID	pQCD	Generator	Shower	X-sec [pb]
single- t	t -channel (t)	410658	NLO	Powheg	Pythia 8	36.99
	t -channel ($\bar{t}$)	410659				22.18
	tW^- ($\geq 1\ell$)	410646				37.94
	$\bar{t}W^+$ ($\geq 1\ell$)	410647				37.91
	tW^- (2ℓ)	410648				4.00
	$\bar{t}W^+$ (2ℓ)	410649				3.99
	s -channel (t)	410644				2.03
	s -channel ($\bar{t}$)	410645				1.27

Table B.4.: Summary of important information about the V +jets MC samples.

Process	Description	DSID	pQCD	Generator	Shower	X-sec [pb]
W +jets	$W \rightarrow e\nu$ (b -filt.)	700338	0,1,2j@NLO, $\geq 3j$ @LO	Sherpa 2.2.11		204.45
	$W \rightarrow e\nu$ (c -filt.)	700339				3194.77
	$W \rightarrow e\nu$ (l -filt.)	700340				18342.86
	$W \rightarrow \mu\nu$ (b -filt.)	700341				201.21
	$W \rightarrow \mu\nu$ (c -filt.)	700342				3207.66
	$W \rightarrow \mu\nu$ (l -filt.)	700343				18395.32
	$W \rightarrow \tau\nu$ (τ_{lep}, b -filt.)	700344				69.40
	$W \rightarrow \tau\nu$ (τ_{lep}, c -filt.)	700345				1102.46
	$W \rightarrow \tau\nu$ (τ_{lep}, l -filt.)	700346				6508.11
	$W \rightarrow \tau\nu$ (τ_{had}, b -filt.)	700347				129.03
	$W \rightarrow \tau\nu$ (τ_{had}, c -filt.)	700348				2024.96
	$W \rightarrow \tau\nu$ (τ_{had}, l -filt.)	700349				11970.09
Z +jets	$Z \rightarrow e^+e^-$ (b -filt.)	700320	0,1,2j@NLO, $\geq 3j$ @LO	Sherpa 2.2.11		55.54
	$Z \rightarrow e^+e^-$ (c -filt.)	700321				286.39
	$Z \rightarrow e^+e^-$ (l -filt.)	700322				1879.38
	$Z \rightarrow \mu^+\mu^-$ (b -filt.)	700323				54.15
	$Z \rightarrow \mu^+\mu^-$ (c -filt.)	700324				287.35
	$Z \rightarrow \mu^+\mu^-$ (b -filt.)	700325				1880.00
	$Z \rightarrow \tau^+\tau^-$ ($\tau_{\text{lep}}\tau_{\text{lep}}, b$ -filt.)	700326				6.68
	$Z \rightarrow \tau^+\tau^-$ ($\tau_{\text{lep}}\tau_{\text{lep}}, c$ -filt.)	700327				34.51
	$Z \rightarrow \tau^+\tau^-$ ($\tau_{\text{lep}}\tau_{\text{lep}}, l$ -filt.)	700328				234.13
	$Z \rightarrow \tau^+\tau^-$ ($\tau_{\text{lep}}\tau_{\text{had}}, b$ -filt.)	700329				24.80
	$Z \rightarrow \tau^+\tau^-$ ($\tau_{\text{lep}}\tau_{\text{had}}, c$ -filt.)	700330				127.00
	$Z \rightarrow \tau^+\tau^-$ ($\tau_{\text{lep}}\tau_{\text{had}}, l$ -filt.)	700331				861.44
	$Z \rightarrow \tau^+\tau^-$ ($\tau_{\text{had}}\tau_{\text{had}}, b$ -filt.)	700332				23.12
	$Z \rightarrow \tau^+\tau^-$ ($\tau_{\text{had}}\tau_{\text{had}}, c$ -filt.)	700333				117.18
	$Z \rightarrow \tau^+\tau^-$ ($\tau_{\text{had}}\tau_{\text{had}}, l$ -filt.)	700334				792.69

Table B.5.: Summary of important information about the $t\bar{t}X$ MC samples.

Process	Description	DSID	pQCD	Generator	Shower	X-sec [pb]
$t\bar{t}W$	$t\bar{t}W$ QCD+EW	700168	LO	Sherpa 2.2.10		0.6747
	$t\bar{t}W$ EW	700205	NLO			0.0477
$t\bar{t}Z$	$Z \rightarrow e^+e^-$	504330	NLO	MadGraph aMC@NLO	Pythia 8	0.0369
	$Z \rightarrow \mu^+\mu^-$	504334				0.0369
	$Z \rightarrow \tau^+\tau^-$	504342				0.0367
$t\bar{t}H$	1ℓ	346344	NLO	Powheg	Pythia 8	0.223
	2ℓ	346345				0.053

Table B.6.: Summary of important information about the Diboson and Triboson MC samples.

Process	Description	DSID	pQCD	Generator	Shower	X-sec [pb]
Diboson	$llll$	700600	NLO	Sherpa 2.2.12		1.2974
	$lll\nu$	700601				4.6610
	$ll\nu\nu$ (SS)	700603				0.0222
Triboson	$lllljj$	700587	NLO	Sherpa 2.2.12		$11.510 \cdot 10^{-3}$
	$lll\nu jj$	700588				$48.453 \cdot 10^{-3}$
	$ll\nu\nu jj$ (SS)	700590				$45.598 \cdot 10^{-3}$

Table B.7.: Summary of important information about the tZq and the tWZ MC samples.

Process	Description	DSID	pQCD	Generator	Shower	X-sec [pb]
tZq	tZq ($\geq 1\ell$)	410560	NLO	MadGraph aMC@NLO	Pythia 8	0.2404
tWZ	tWZ ($\geq 1\ell$, DR1)	412118	NLO	MadGraph aMC@NLO	Pythia 8	0.0161

Table B.8.: Summary of important information about the VH MC samples.

Process	Description	DSID	pQCD	Generator	Shower	X-sec [pb]
VH	WH (inc.)	342284	NLO	Powheg	Pythia 8	1.1021
	ZH (inc.)	342285	NLO			0.6007

Table B.9.: Summary of important information about the MC samples including rare top-quark processes.

Process	Description	DSID	pQCD	Generator	Shower	X-sec [pb]
$t\bar{t}t$	$t\bar{t}t$ (lept.)	516978	NLO	MadGraph aMC@NLO	Pythia 8	0.0016
$t\bar{t}t\bar{t}$	$t\bar{t}t\bar{t}$ (lept.)	412043	NLO	MadGraph aMC@NLO	Pythia 8	0.0120
$tWHq$	tHW	346678	NLO	MadGraph aMC@NLO	Pythia 8	0.01672
	tHq	346676				0.06014
$t\bar{t}XX$	$t\bar{t}WW$	410081	NLO	MadGraph aMC@NLO	Pythia 8	0.00810
	$t\bar{t}HH$	500460				0.00074
	$t\bar{t}WH$	500461				0.00114
	$t\bar{t}ZZ$	500462				0.00149
	$t\bar{t}WZ$	500463				0.00247

C. Comparison of the left- and right handed prod-FCNC processes

This appendix compares kinematic distributions between the lefthanded and the righthanded $u\bar{t}H$ process in the 3ℓ SRs. Figure C.1 - C.3 show six chosen kinematic distributions in the SR3 ℓ Dec region whereas Figure C.4 - C.6 show the same six kinematic distributions in the SR3 ℓ Prod region. The error bars in the plots denote the MC statistical uncertainty. It can be seen, that within MC statistics, no significant differences can be observed based on the handness of the prod-FCNC process. Similar behaviour can be observed for the $c\bar{t}H$ processes, as well as in the 2ℓ SS regions. As a result, both left- and righthanded prod-FCNC samples are combined to make use of the improved MC statistics and limits are derived on both the left- and the righthanded coupling at the same time.

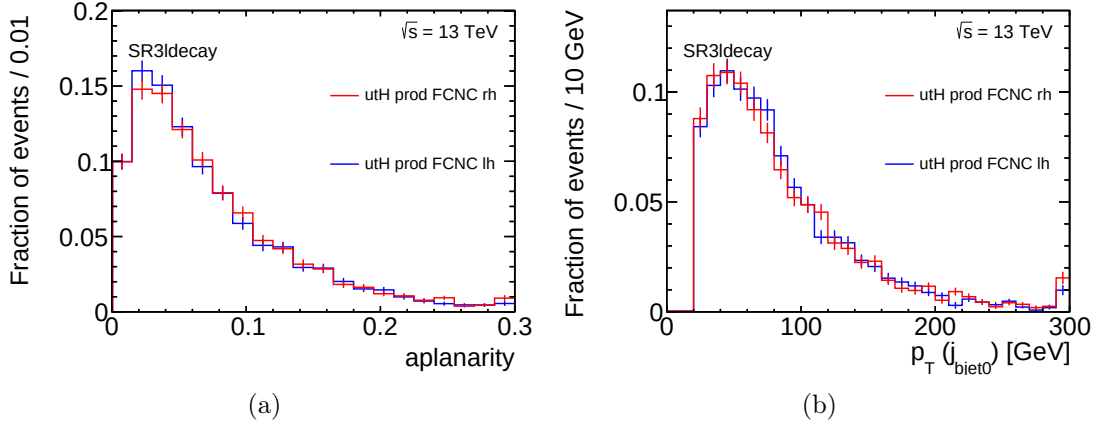


Figure C.1: Comparison between the left- and righthanded $u\bar{t}H$ prod-FCNC process of the aplanarity (a) and the p_T of the b -tagged jet (b) in the SR3 ℓ Dec region.

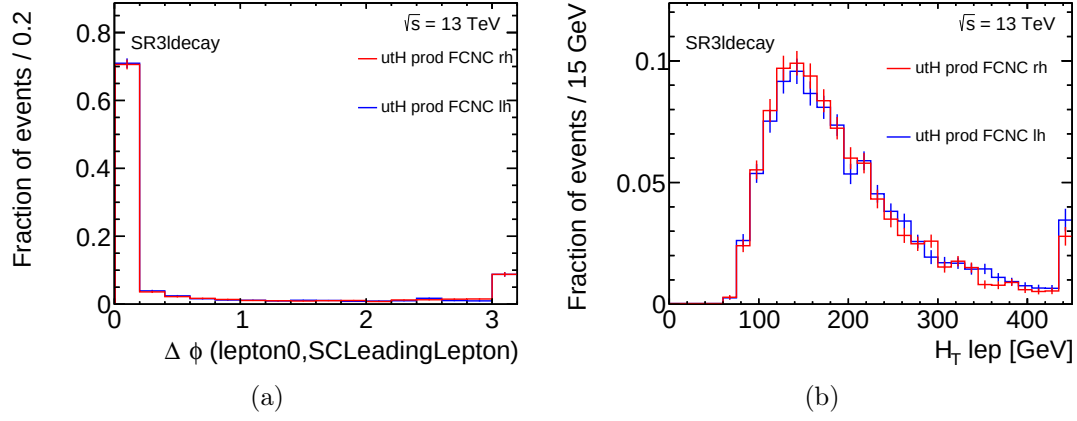


Figure C.2: Comparison between the left- and righthanded utH prod-FCNC process of the $\Delta\phi$ between the leading lepton and the leading SC lepton (a) and the p_T of the scalar sum of all lepton p_T 's (b) in the $SR3\ell Dec$ region.

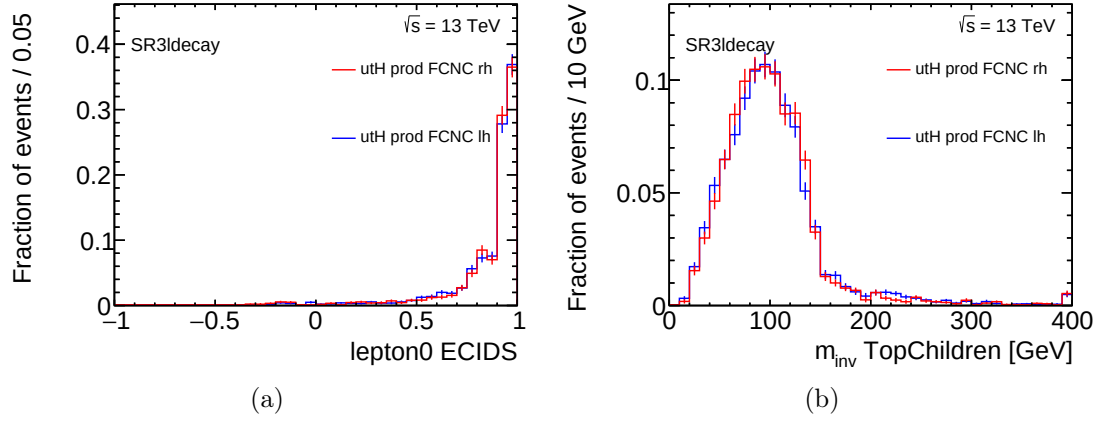


Figure C.3: Comparison between the left- and righthanded utH prod-FCNC process of the ECIDS score of the leading lepton (a) and the invariant mass of the top children, coming from the NICE reco (b) in the $SR3\ell Dec$ region.

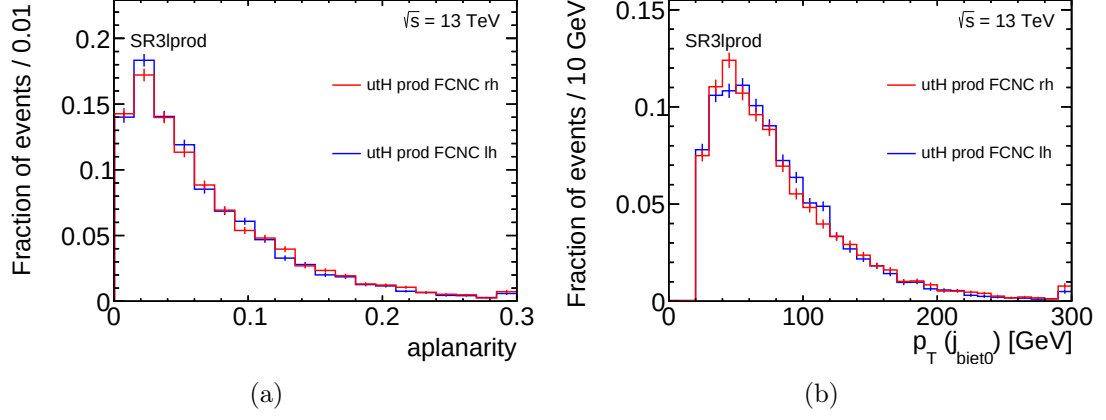


Figure C.4: Comparison between the left- and righthanded utH prod-FCNC process of the aplanarity (a) and the p_T of the b -tagged jet (b) in the SR3lProd region.

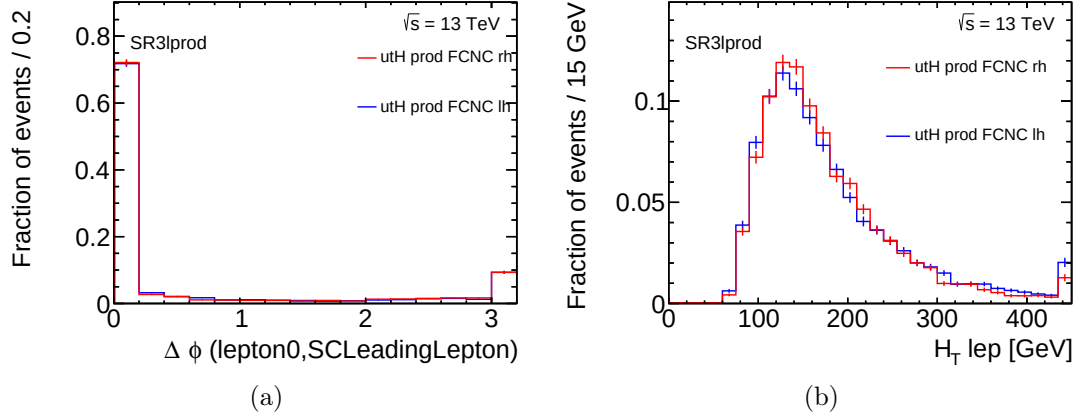


Figure C.5: Comparison between the left- and righthanded utH prod-FCNC process of the $\Delta\phi$ between the leading lepton and the leading SC lepton (a) and the p_T of the scalar sum of all lepton p_T 's (b) in the SR3lProd region.

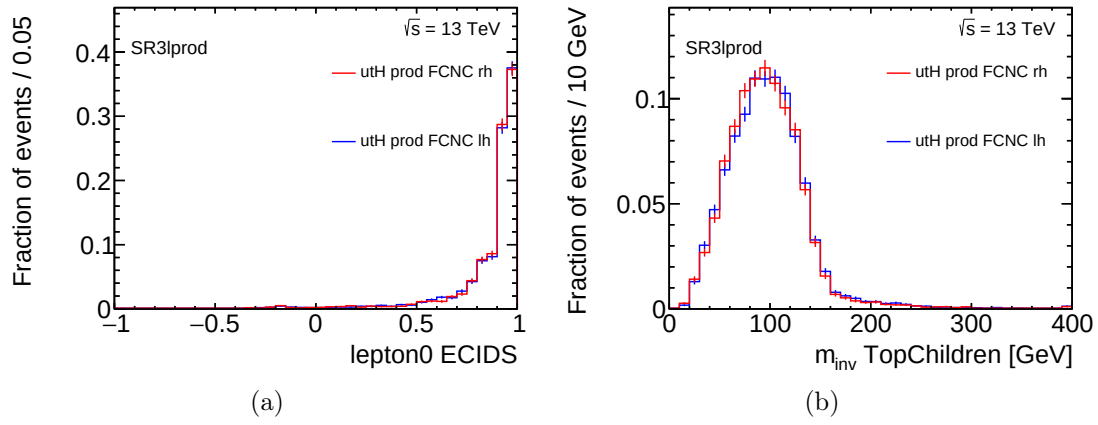


Figure C.6: Comparison between the left- and righthanded utH prod-FCNC process of the ECIDS score of the leading lepton (a) and the invariant mass of the top children, coming from the NICE reco (b) in the SR3 ℓ Prod region.

D. Composition of the non-prompt templates

This appendix shows the composition of the relevant non-prompt templates in selected regions. See Section 7.4 for the definition of the non-prompt templates.

D.1. Composition in the 3ℓ regions

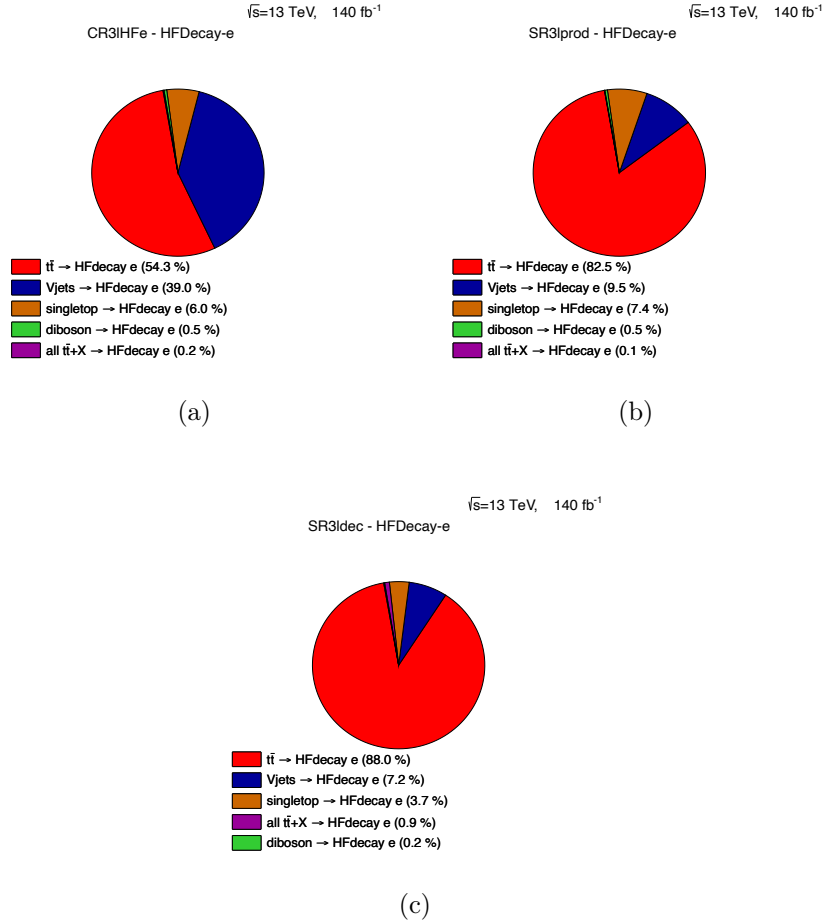


Figure D.1: The composition of the non-prompt HFe template in the CR3 ℓ HFe (a), SR3 ℓ Prod (b) and the SR3 ℓ Dec (c) region.

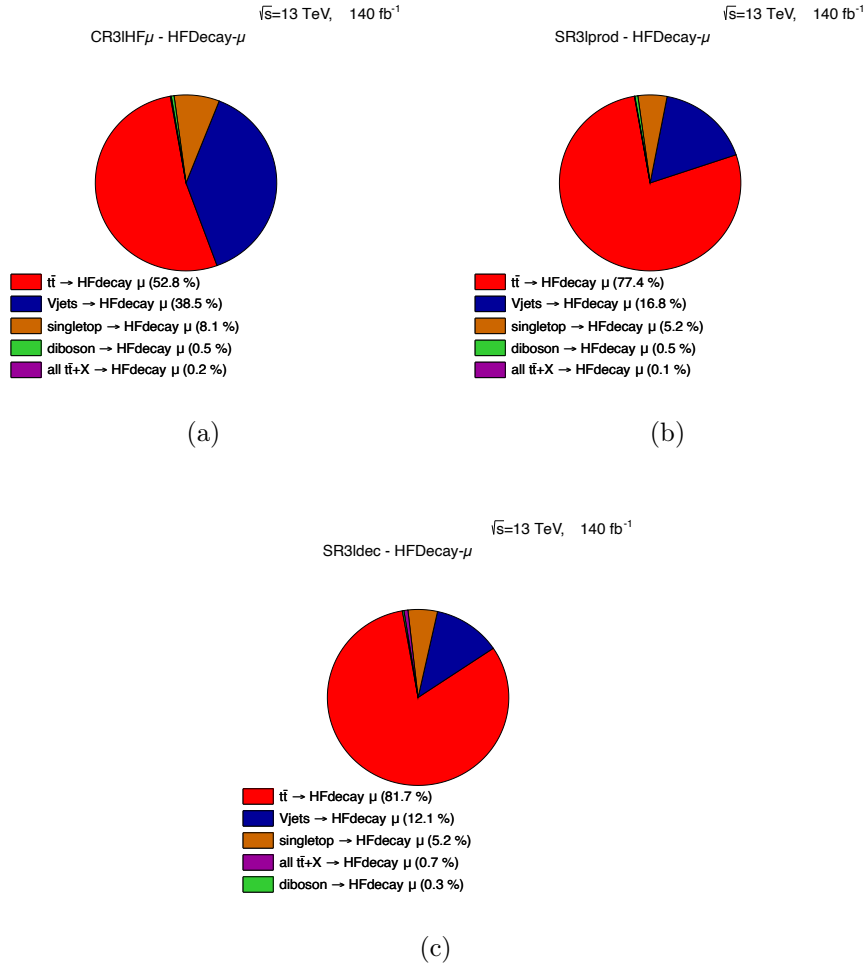


Figure D.2: The composition of the non-prompt HF μ template in the CR3lHF μ (a), SR3lProd (b) and the SR3lDec (c) region.

D.2. Composition in the 2ℓ SS regions

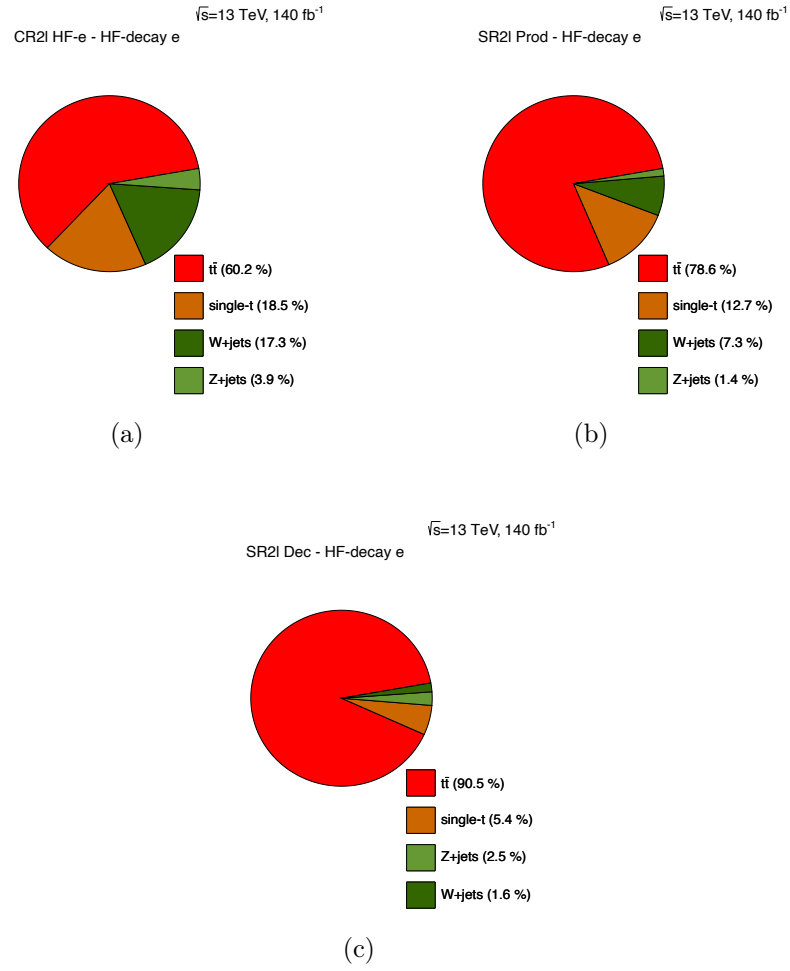


Figure D.3: The composition of the non-prompt HFe template in the CR2 ℓ HFe (a), SR2 ℓ Prod (b) and the SR2 ℓ Dec (c) region.

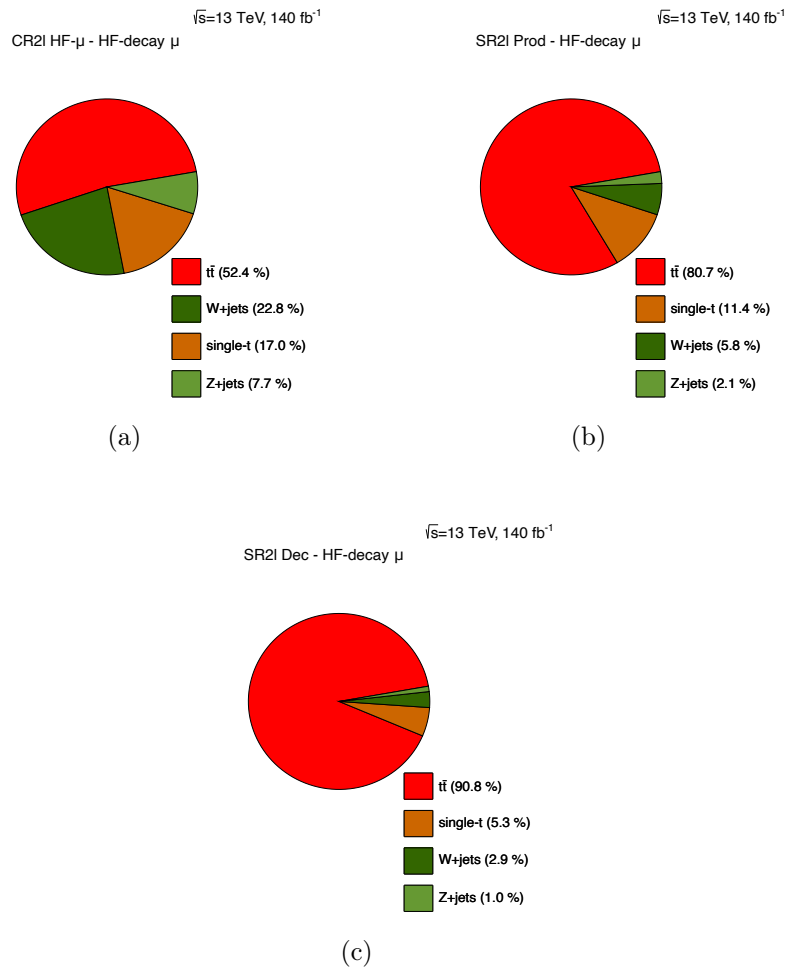


Figure D.4: The composition of the non-prompt HF μ template in the CR2lHF μ (a), SR2lProd (b) and the SR2lDec (c) region.

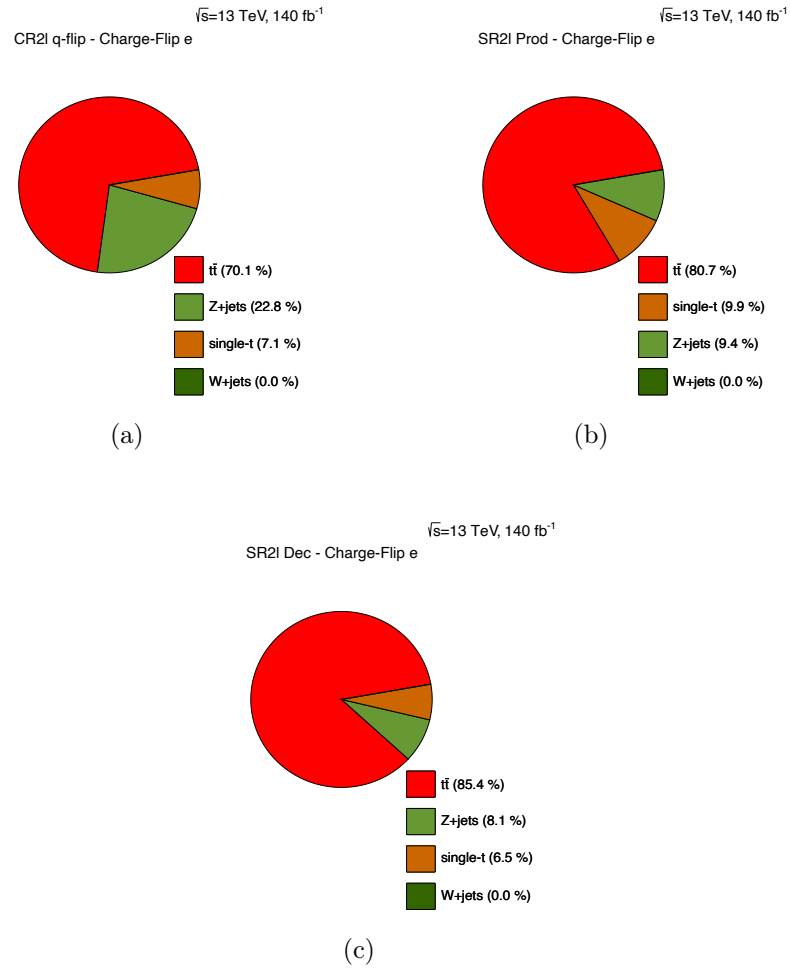


Figure D.5: The composition of the non-prompt electron charge flip template in the CR2l q -flip (a), SR2lProd (b) and the SR2lDec (c) region.

E. Pre-fit event yields

This appendix shows the pre-fit event yields (i.e. the predicted events per region and process before any corrections or adjustments from a fit). The event yields are split for the 3ℓ regions in Table E.1 and for the 2ℓ SS regions in Table E.2. Furthermore, the event yields are graphically represented via pie charts in Figure E.1 and E.2. The amount of data events in the SRs is not presented because the analysis is blinded.

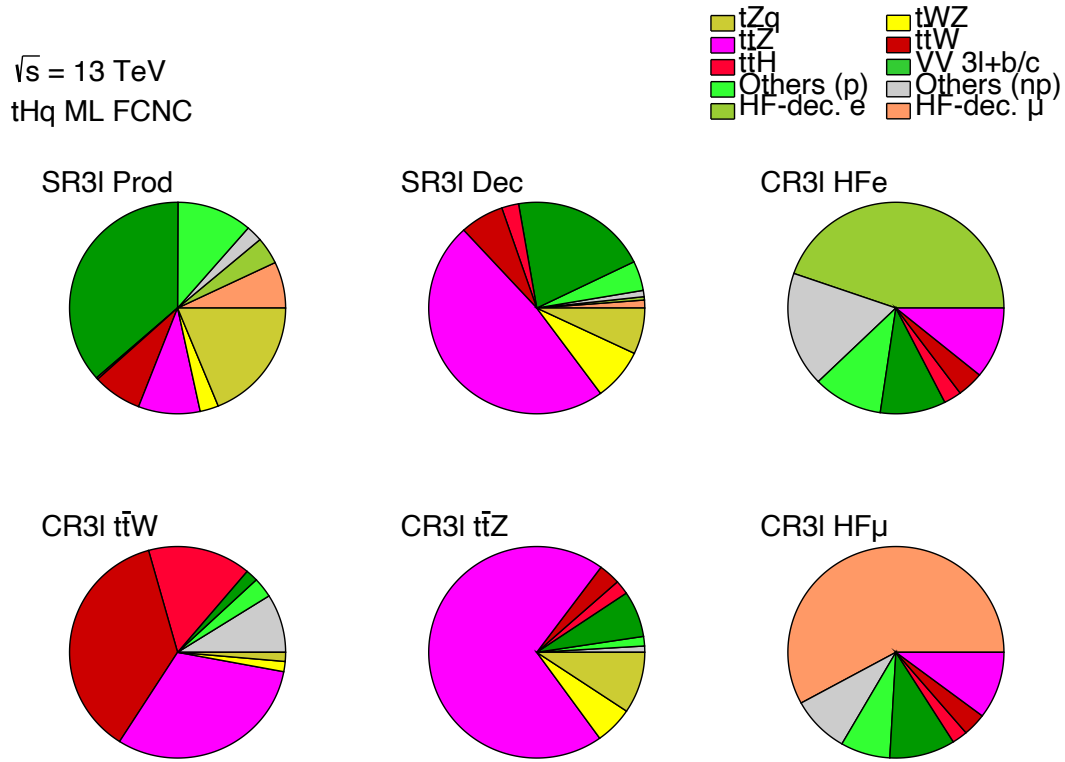


Figure E.1: Pie charts of the SM process composition in the 3ℓ regions.

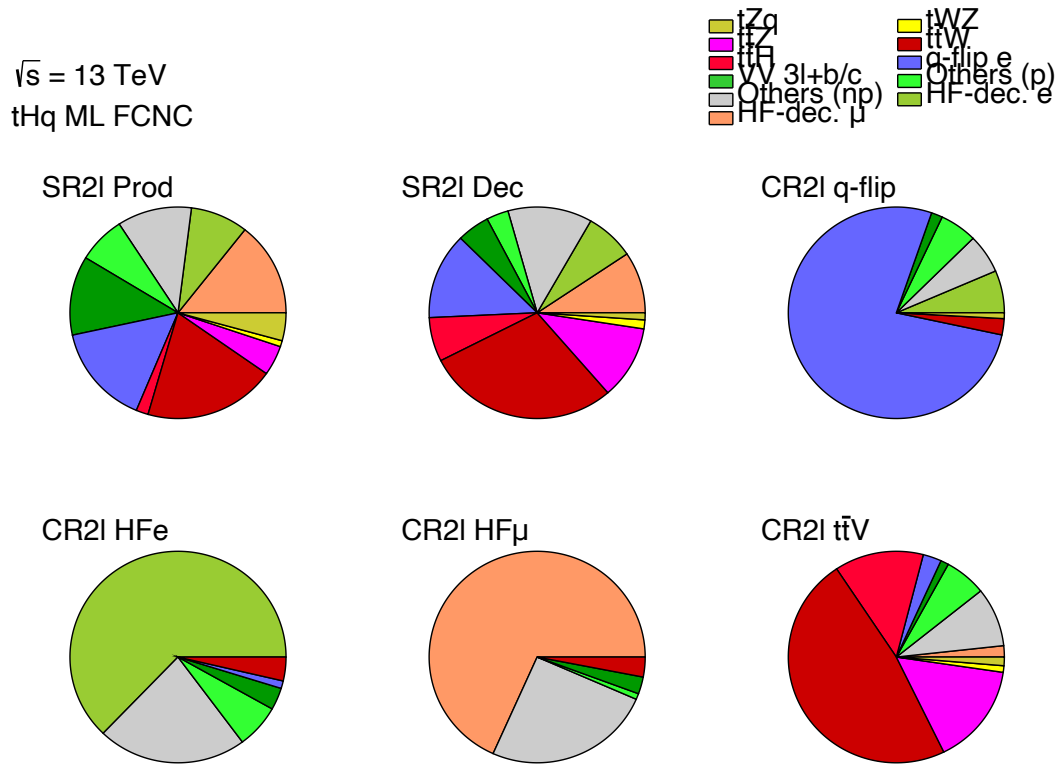


Figure E.2: Pie charts of the SM process composition in the $2\ell\text{SS}$ regions.

Table E.1.: Pre-fit event yields in the 3ℓ regions. All systematic uncertainties are taken into account and are presented by the number in round brackets. However, processes with less than 0.5 events in a specific regions are omitted in that region. The FCNC processes are normalised to a Wilson coefficient of $C = 1$.

Process	SR3 ℓ Prod	SR3 ℓ Dec	CR3 $\ell t\bar{t}W$	CR3 $\ell t\bar{t}Z$	CR3 ℓ HF e	CR3 ℓ HF μ
tHu	20 (1)	20 (1)	1 (1)	-	3 (1)	5 (1)
tHc	14 (1)	17 (1)	2 (1)	1 (1)	2 (1)	4 (1)
HF-decay e	37 (5)	14 (2)	-	-	50 (7)	-
HF-decay μ	66 (4)	25 (3)	-	-	-	126 (11)
q -flip e	-	-	-	-	-	-
Others (non-p)	40 (12)	35 (11)	18 (7)	3 (1)	20 (8)	19 (9)
$t\bar{t}H$	9 (2)	47 (7)	33 (5)	7 (1)	3 (1)	5 (1)
$t\bar{t}W$	61 (7)	67 (5)	77 (14)	10 (1)	5 (1)	8 (1)
$t\bar{t}Z$	61 (8)	374 (10)	65 (5)	220 (25)	13 (1)	23 (2)
$VV\ 2\ell$	-	-	-	-	-	-
$VV\ 3\ell+b$	102 (4)	93 (6)	4 (1)	21 (1)	5 (1)	11 (1)
$VV\ 3\ell+c$	110 (9)	66 (8)	-	1 (1)	6 (1)	10 (1)
$VV\ 3\ell+l/\tau$	23 (13)	13 (7)	-	-	1 (1)	2 (1)
$VV\ 4\ell$	57 (28)	34 (17)	2 (1)	4 (2)	7 (4)	13 (7)
tWZ	19 (6)	61 (18)	3 (1)	18 (5)	-	-
tZq	109 (12)	56 (7)	3 (1)	29 (3)	-	-
Rare t	-	11 (2)	-	-	-	-
VVV	6 (1)	8 (1)	-	-	-	-
$tWHq$	2 (1)	3 (1)	1 (1)	-	-	1 (1)
Rare t	-	11 (2)	-	-	-	-
VH	-	-	-	-	-	-
$t\bar{t}$	-	-	4 (1)	-	4 (1)	1 (1)
single- t	-	-	-	-	-	-
Z +jets	13 (3)	4 (1)	-	-	-	-
Total BG	716 (44)	909 (45)	211 (19)	314 (27)	113 (13)	217 (21)
Data	-	-	277	399	159	263

Table E.2.: Pre-fit event yields in the 2ℓ SS regions. All systematic uncertainties are taken into account and are presented by the number in round brackets. However, processes with less than 0.5 events in a specific regions are omitted in that region. The FCNC processes are normalised to a Wilson coefficient of $C = 1$.

Process	SR2 ℓ Dec	SR2 ℓ Prod	CR2 ℓ HF e	CR2 ℓ HF μ	CR2 ℓ q -flip	CR2 $\ell t\bar{t}V$
tHu	112 (7)	116 (5)	5 (4)	9 (8)	5 (4)	9 (8)
tHc	92 (5)	89 (4)	2 (1)	3 (1)	1 (1)	6 (1)
HF-decay e	130 (12)	102 (8)	63 (7)	-	10 (1)	-
HF-decay μ	237 (22)	196 (11)	-	133 (22)	-	7 (1)
q -flip e	173 (18)	148 (7)	1 (1)	-	121 (6)	11 (2)
Others (non-p)	242 (73)	134 (41)	22 (8)	49 (32)	9 (4)	35 (14)
$t\bar{t}H$	136 (21)	24 (5)	-	-	-	54 (9)
$t\bar{t}W$	454 (26)	217 (18)	4 (1)	6 (1)	4 (1)	189 (9)
$t\bar{t}Z$	189 (8)	52 (5)	-	-	-	65 (5)
$VV\ 2\ell$	9 (4)	9 (5)	-	-	-	-
$VV\ 3\ell+b$	48 (4)	68 (2)	2 (1)	2 (1)	1 (1)	5 (1)
$VV\ 3\ell+c$	36 (5)	70 (6)	2 (1)	3 (1)	1 (1)	-
$VV\ 3\ell+l/\tau$	11 (6)	28 (15)	1 (1)	1 (1)	1 (1)	-
$VV\ 4\ell$	11 (6)	26 (13)	-	-	-	-
tWZ	25 (8)	12 (4)	-	-	-	4 (1)
tZq	24 (3)	50 (5)	-	-	1 (1)	6 (1)
Rare t	26 (5)	2 (1)	-	-	-	15 (3)
VVV	-	-	-	-	-	-
$tWHq$	9 (1)	8 (1)	-	-	-	2 (1)
Rare t	26 (5)	2 (1)	-	-	-	15 (3)
VH	4 (1)	-	-	-	-	-
$t\bar{t}$	31 (3)	22 (2)	6 (1)	-	8 (1)	8 (1)
single- t	3 (1)	-	-	-	-	-
Z +jets	6 (1)	14 (2)	-	-	-	-
Total BG	1804 (115)	1183 (63)	101 (12)	195 (40)	157 (8)	401 (25)
Data	-	-	120	195	176	453

F. Plots related to the background-only fit

This appendix contains a full set of plots from the utH fit, related to Section 13.1.

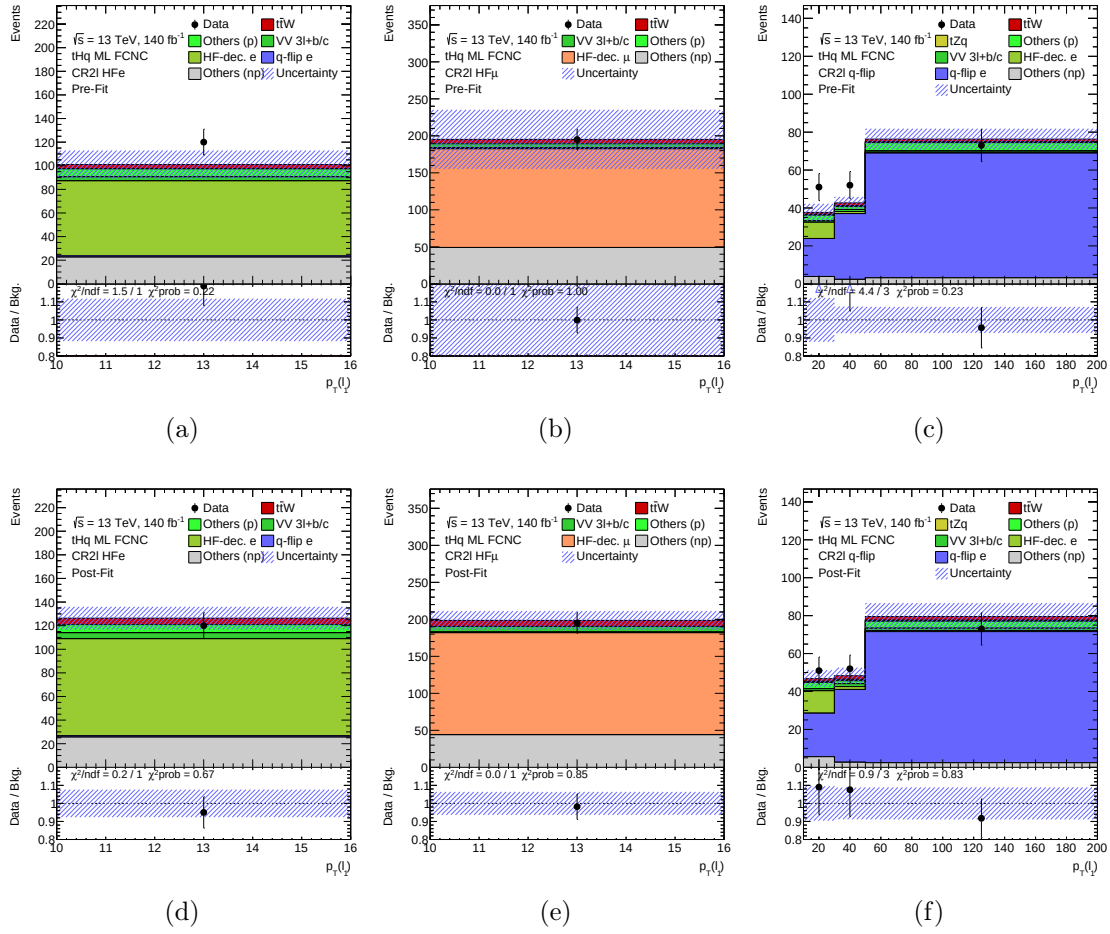


Figure F.1: The 2LSS regions CR2ℓHFe, CR2ℓHFμ and CR2ℓq-flip once pre-fit (upper row) and once post-fit (lower row) in the background-only fit.

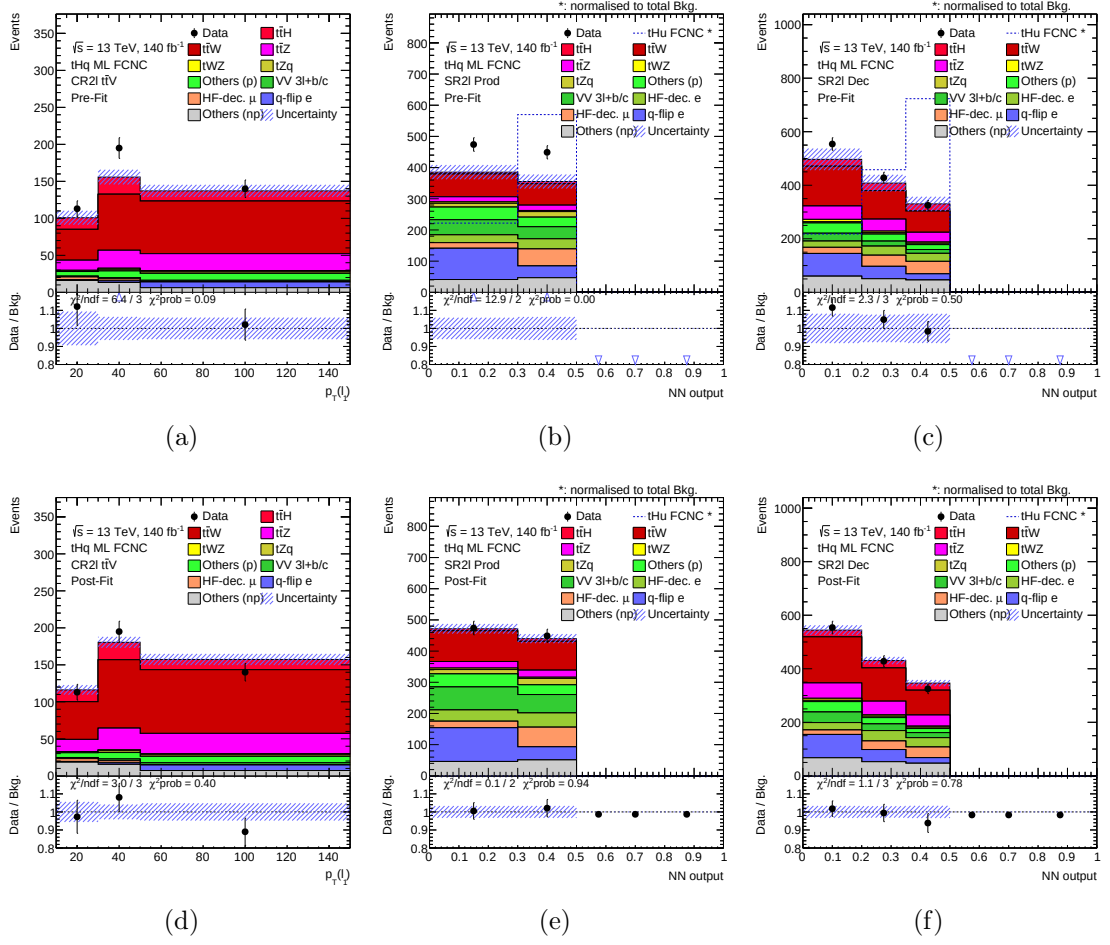


Figure F.2: The 2ℓSS regions CR2ℓtV̄, SR2ℓProd and SR2ℓDec once pre-fit (upper row) and once post-fit (lower row) in the background-only fit.

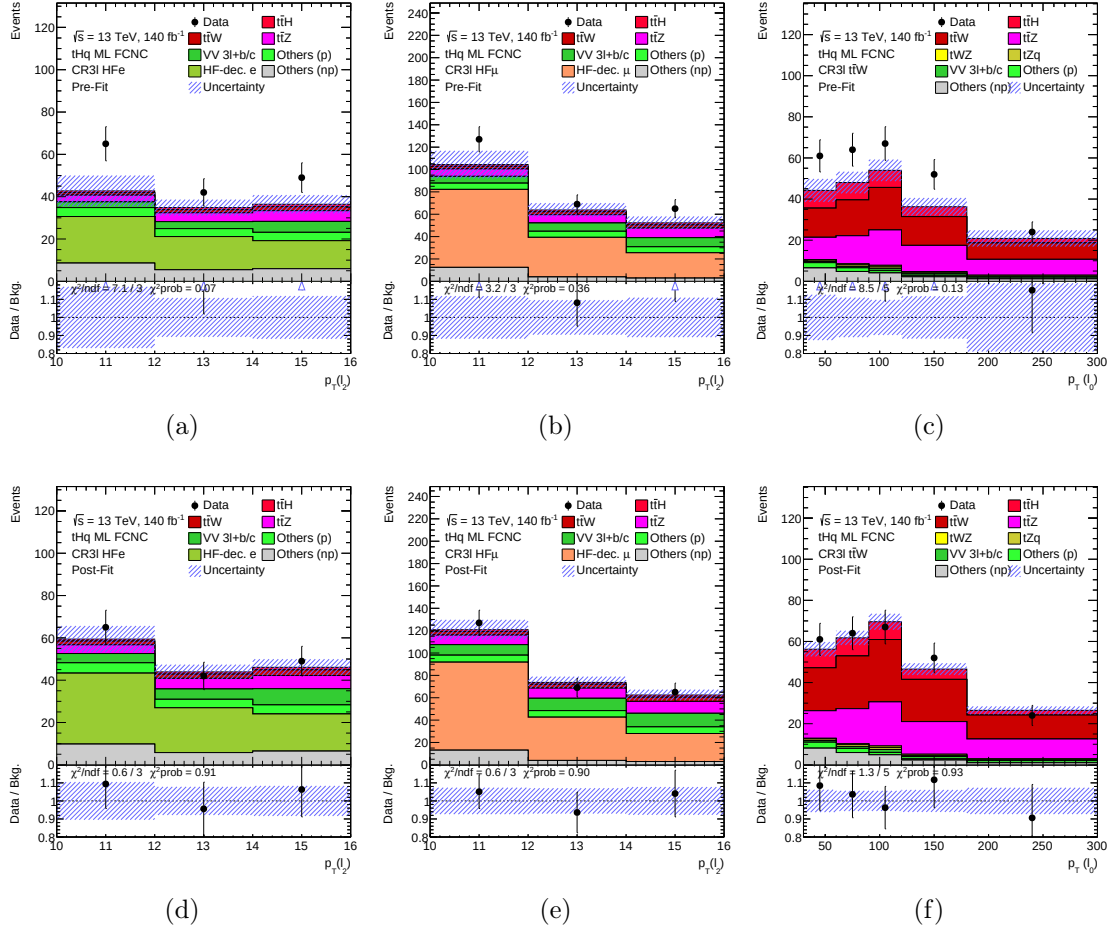


Figure F.3: The 3ℓ regions CR3 ℓ HFe, CR3 ℓ H μ and CR3 ℓ $t\bar{t}W$ once pre-fit (upper row) and once post-fit (lower row) in the background-only fit.

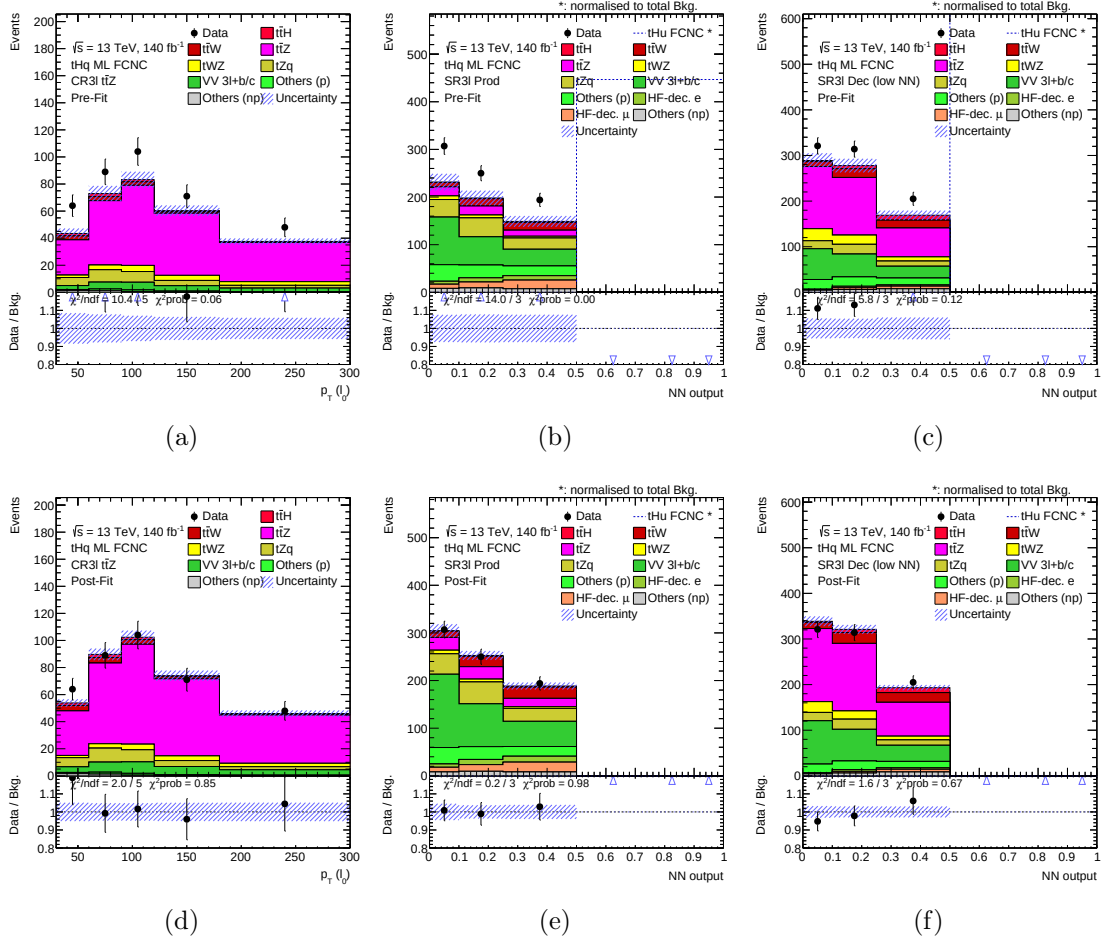


Figure F.4: The 3 ℓ regions CR3 ℓ t ℓ Z, SR3 ℓ Prod and SR3 ℓ Dec once pre-fit (upper row) and once post-fit (lower row) in the background-only fit.

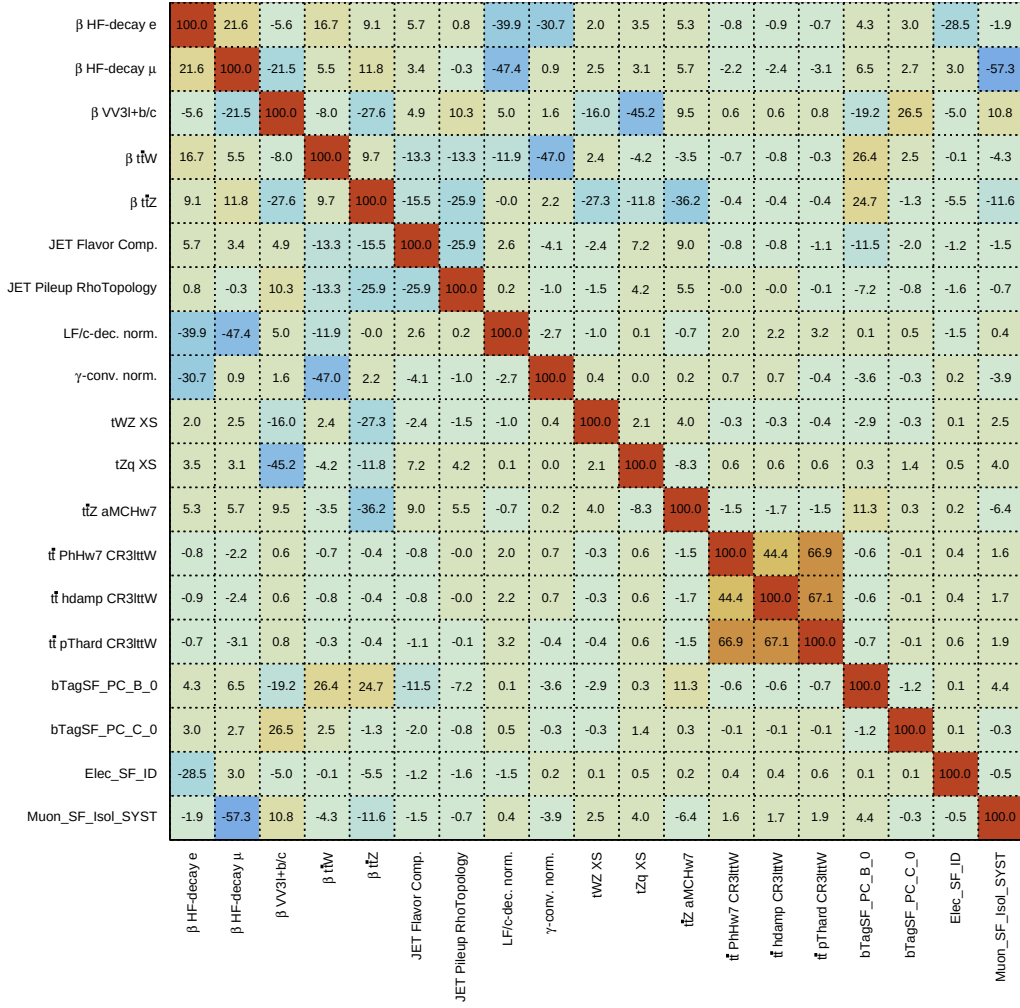


Figure F.5: Correlation matrix of NPs in the background-only fit with a linear correlation coefficient larger than 25% to any other NP but themselves, exemplarily for the utH case.

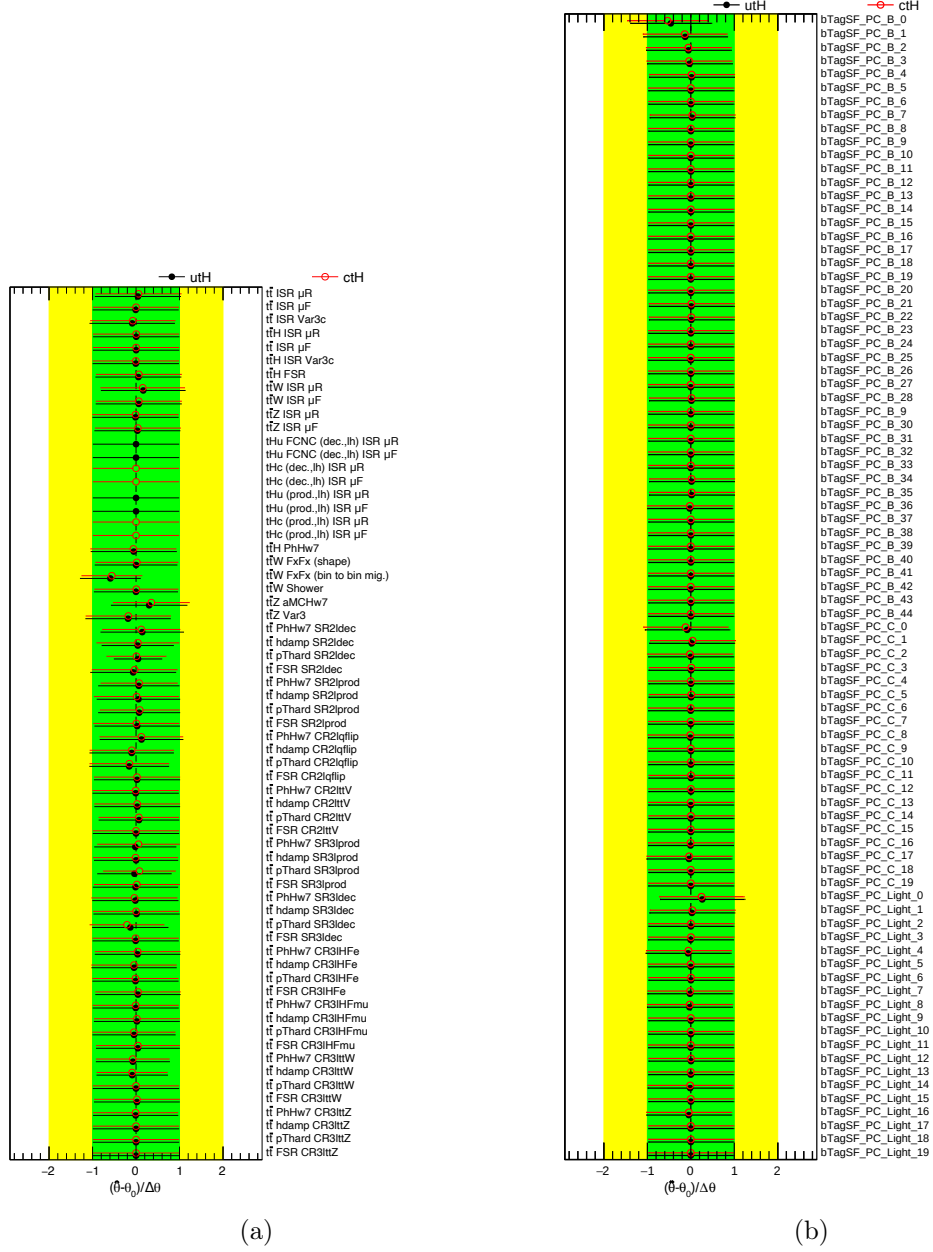


Figure F.6: Pulls and constraints of the NPs related to modelling systematic uncertainties (a) and b -tagging systematic uncertainties (b) in the background-only fit.

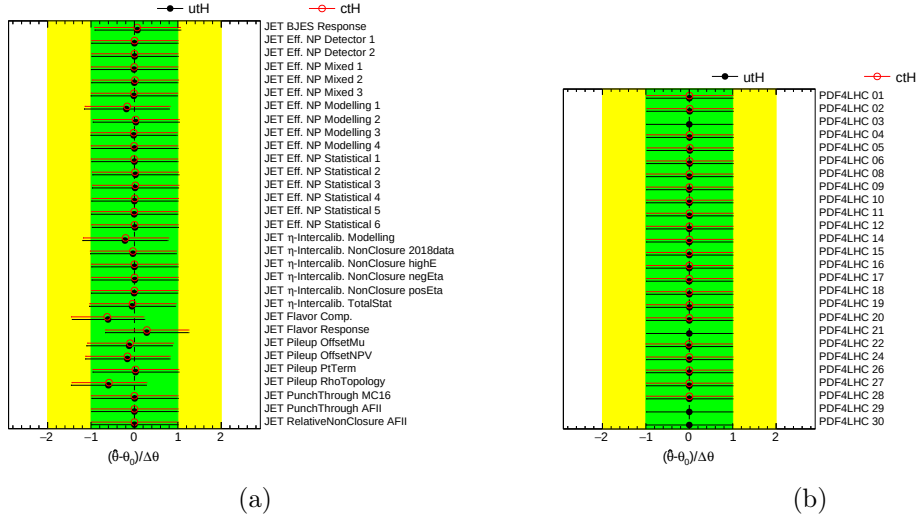


Figure F.7: Pulls and constraints of the NPs related to jet systematic uncertainties (a) and PDF systematic uncertainties (b) in the background-only fit.

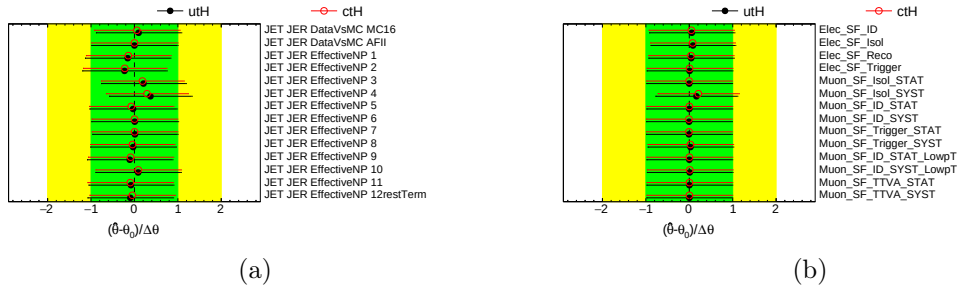


Figure F.8: Pulls and constraints of the NPs related to JER systematic uncertainties (a) and lepton systematic uncertainties (b) in the background-only fit.

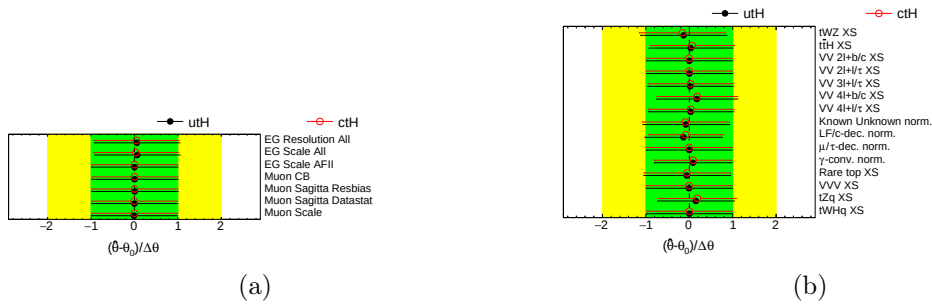


Figure F.9: Pulls and constraints of the NPs related to electron and muon systematic uncertainties (a) and theory cross-section systematic uncertainties (b) in the background-only fit.

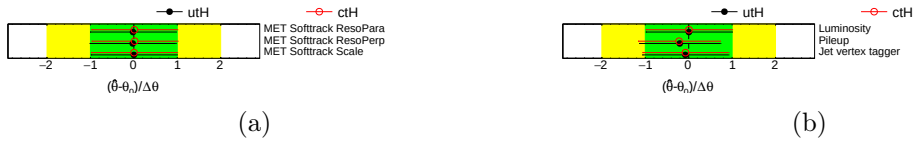


Figure F.10: Pulls and constraints of the NPs related to E_T^{miss} systematic uncertainties (a) and other systematic uncertainties (b) in the background-only fit.

G. Validation plots of kinematic distributions

This appendix contains pre- and post-fit plots of kinematic distribution for validation purposes.

G.1. Validation plots for the 2ℓ SS regions

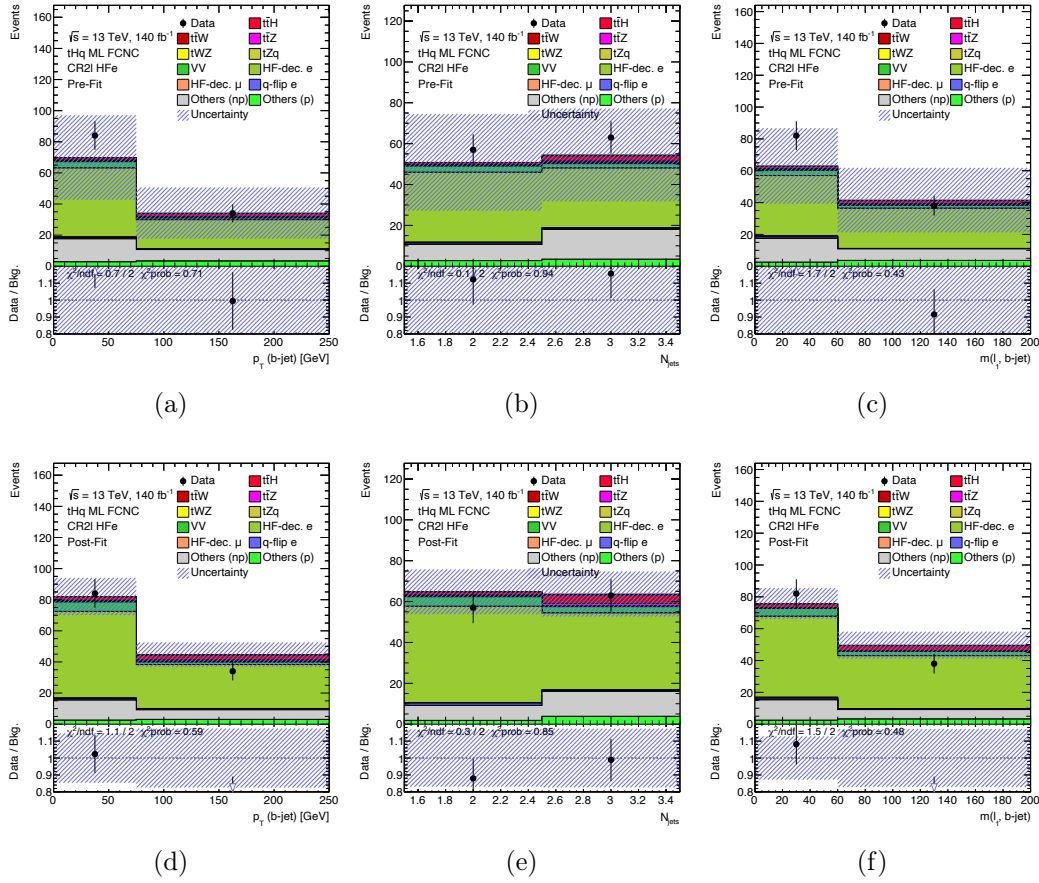


Figure G.1: Validation plots of p_T of the b -tagged jet (left), the number of jets (middle) and invariant mass of the second-leading lepton and the b -tagged jet (right) in the CR2 ℓ HFe region. Upper row contains pre-fit plots and lower row contains post-background-only-fit plots.

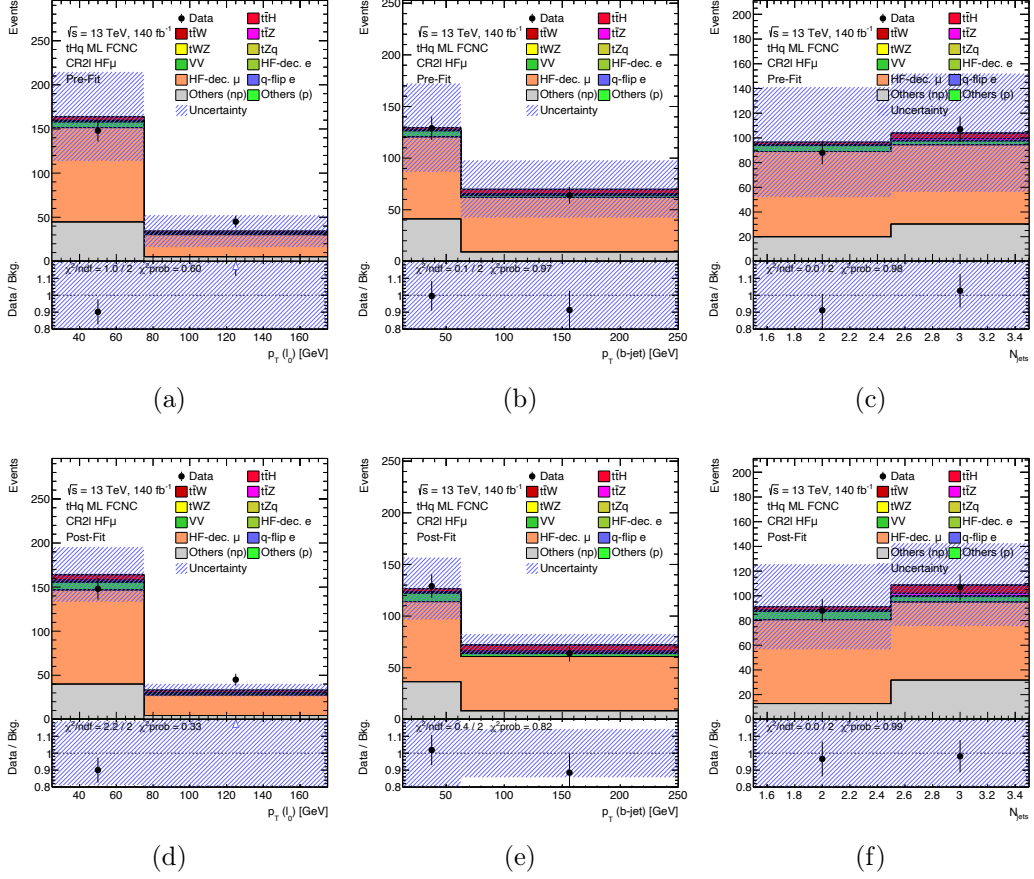


Figure G.2: Validation plots of the leading-lepton p_T (left), p_T of the b -tagged jet (middle) and the number of jets (right) in the CR2 ℓ HF μ region. Upper row contains pre-fit plots and lower row contains post-background-only-fit plots.

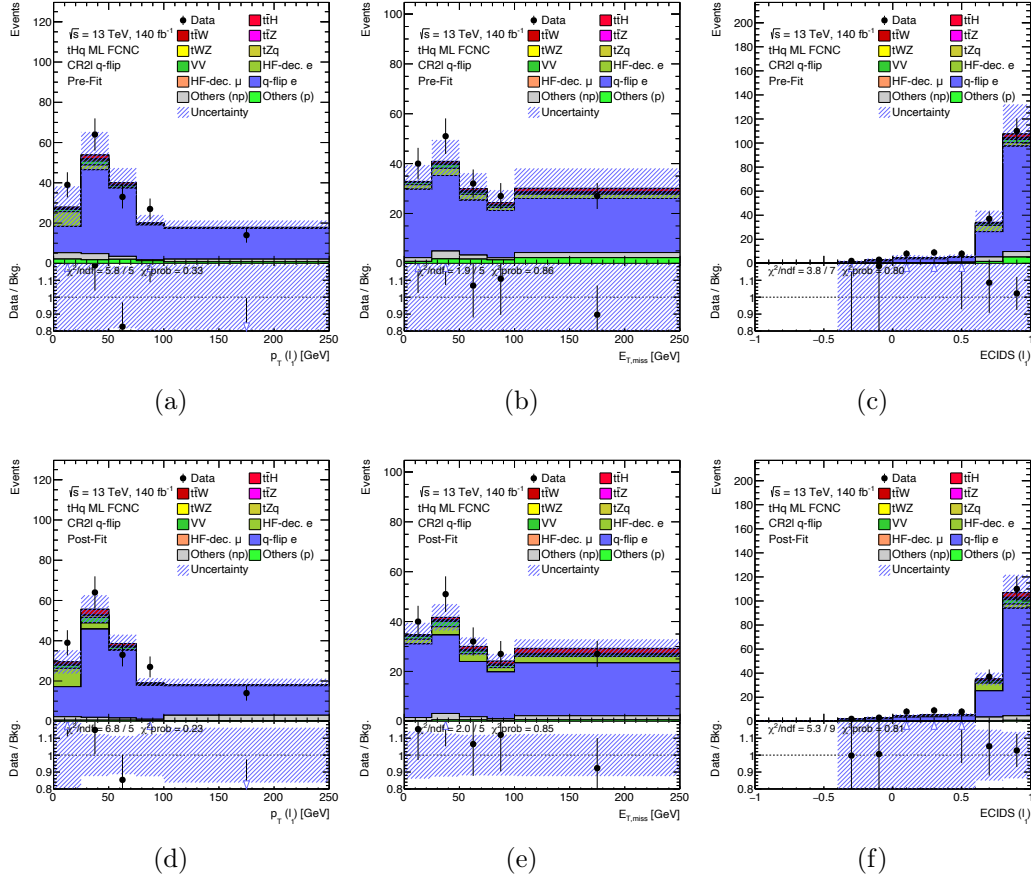


Figure G.3: Validation plots of the sub-leading lepton p_T (left), the E_T^{miss} (middle) and the ECIDS score of the sub-leading lepton (right) in the CR2 lq -flip region. Upper row contains pre-fit plots and lower row contains post-background-only-fit plots.

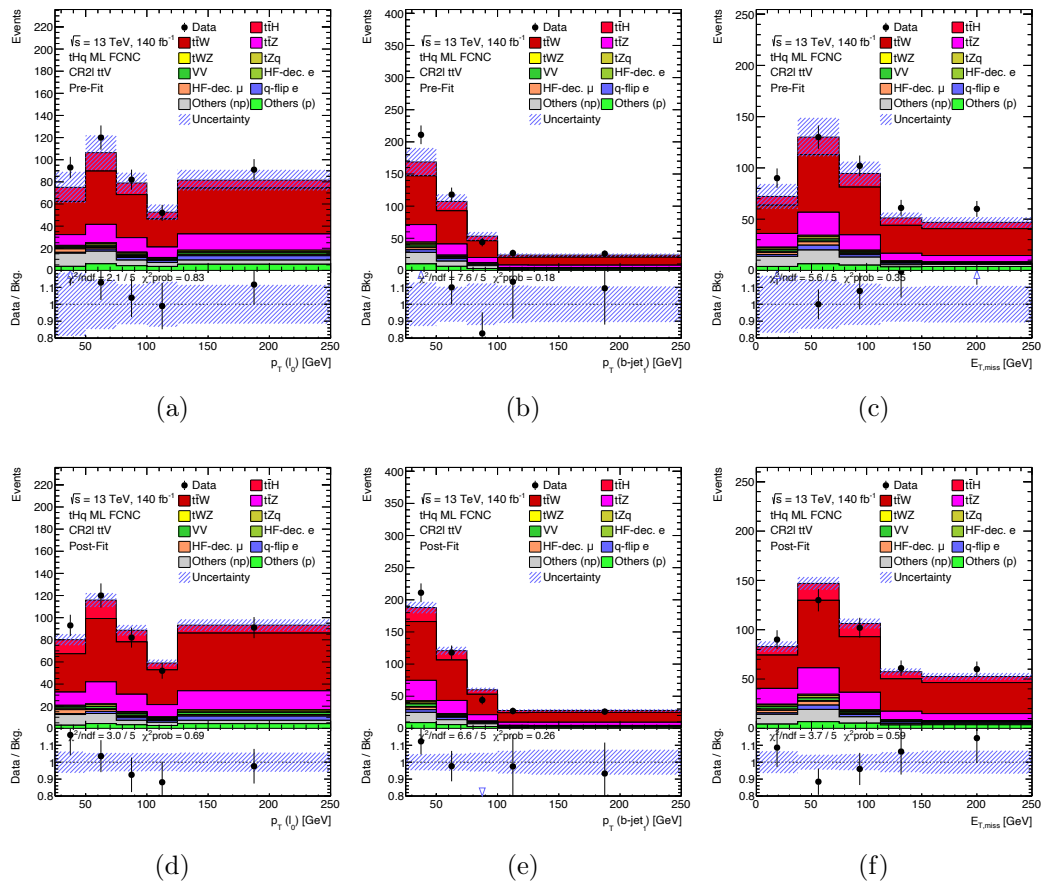


Figure G.4: Validation plots of the leading-lepton p_T (left), the p_T of the second-leading b -tagged jet (middle) and the E_T^{miss} (right) in the CR2 $\ell t\bar{t}V$ region. Upper row contains pre-fit plots and lower row contains post-background-only-fit plots.

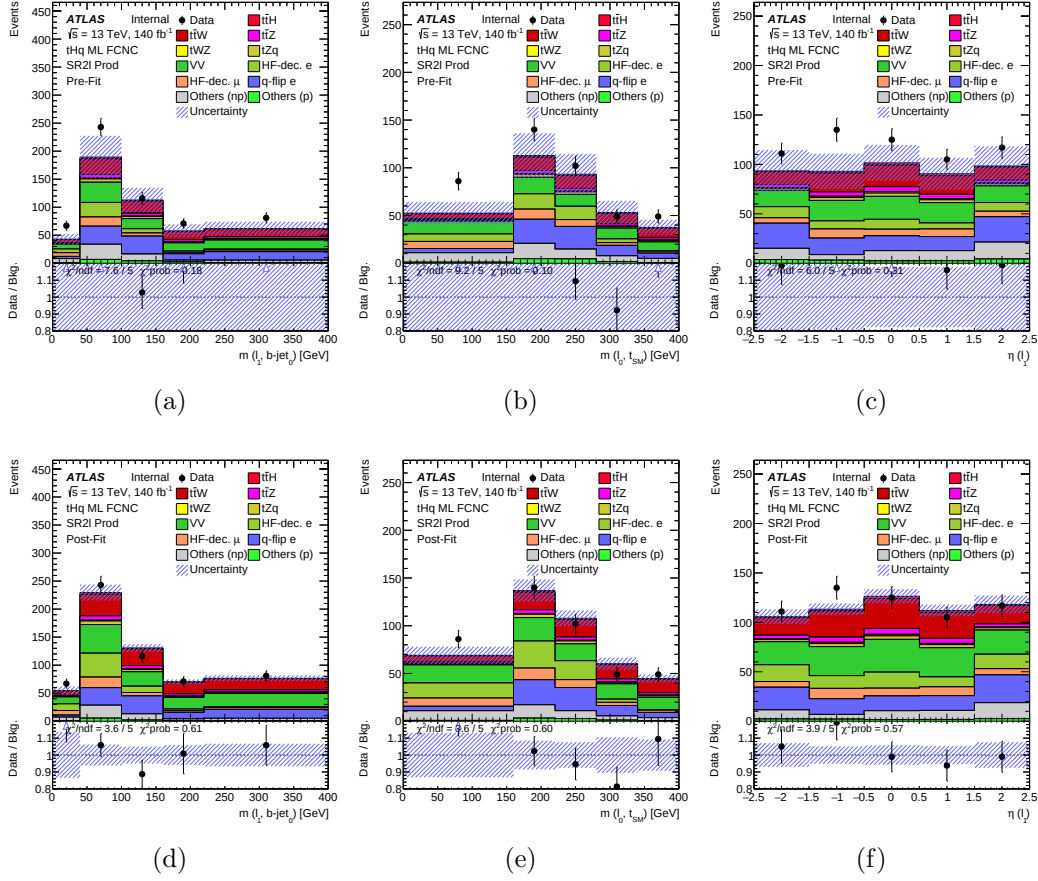


Figure G.5: Validation plots of the invariant mass of the second-leading lepton and the b -tagged jet (left), the invariant mass of the leading lepton and the reconstructed top-quark (middle) and the η of the sub-leading lepton (right) in the SR2 ℓ Prod region. Upper row contains pre-fit plots and lower row contains post-background-only-fit plots.

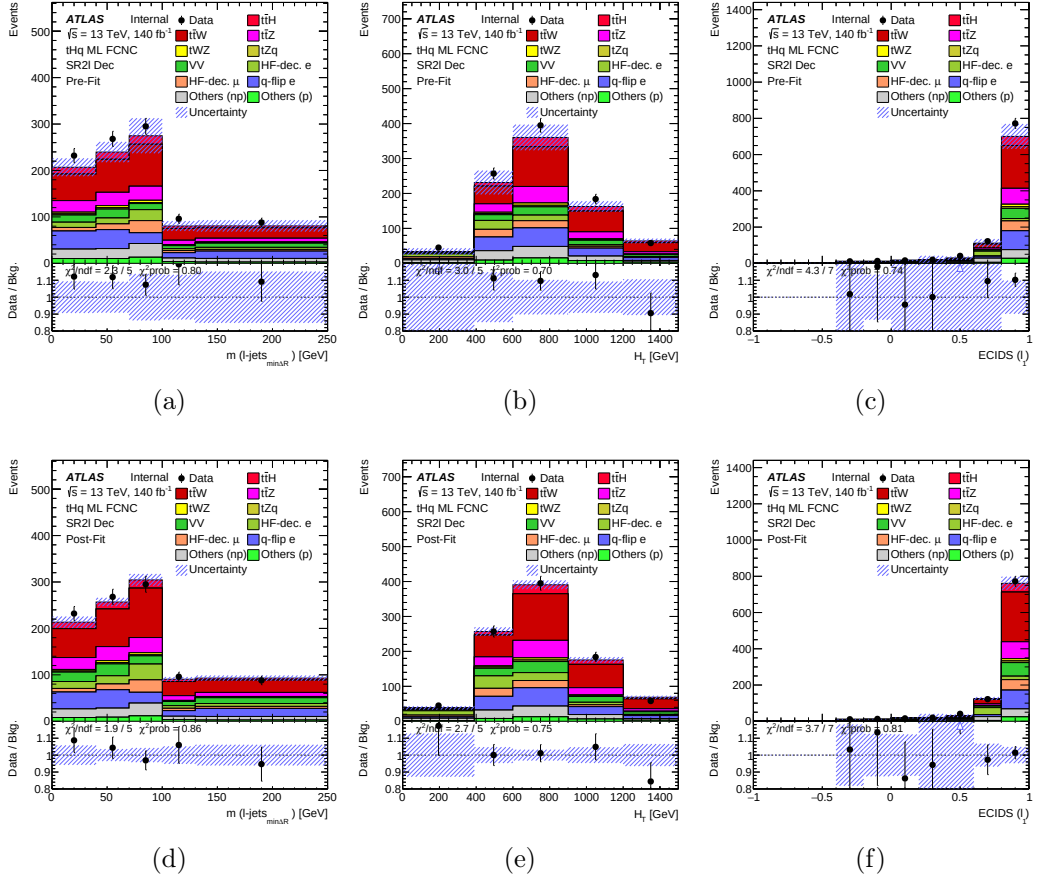


Figure G.6: Validation plots of the invariant mass of the jets with the smallest ΔR (left), the H_T (middle) and the ECIDS score of the sub-leading lepton (right) in the SR2 ℓ Dec region. Upper row contains pre-fit plots and lower row contains post-background-only-fit plots.

G.2. Validation plots for the 3ℓ regions

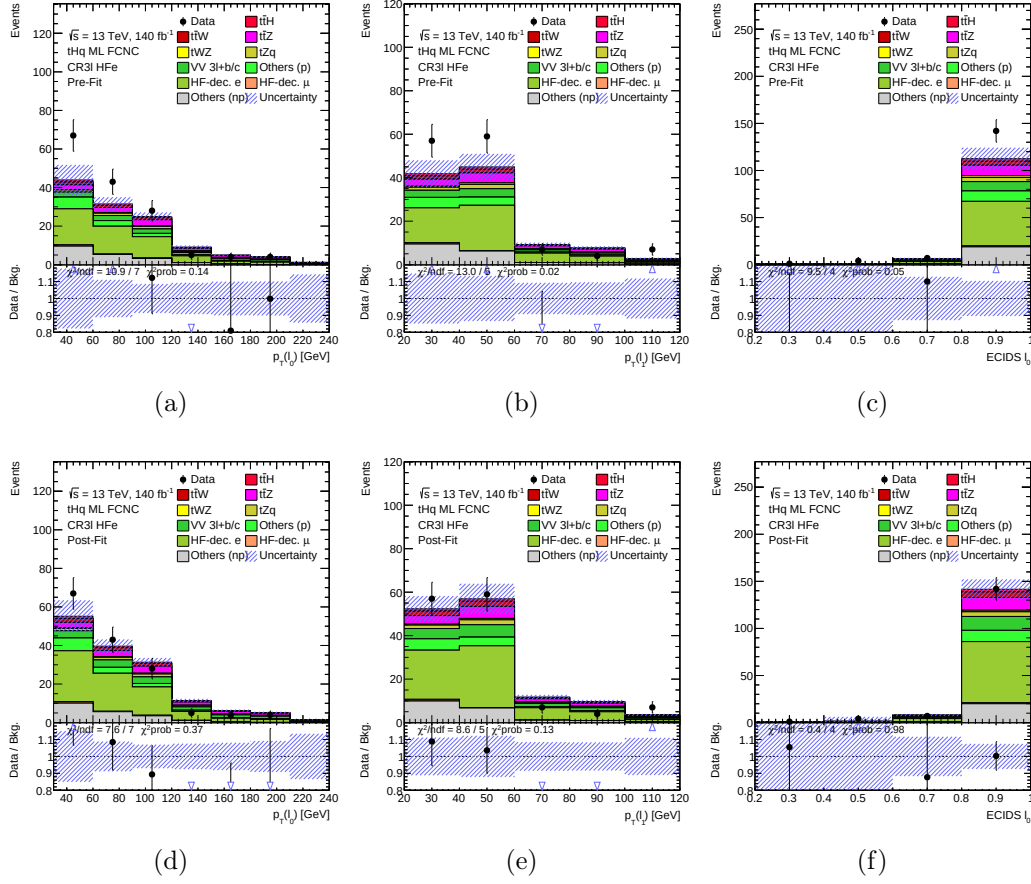


Figure G.7: Validation plots of the leading-lepton p_T (left), the second-leading lepton p_T (middle) and the ECIDS score of the leading lepton (right) in the CR3 ℓ HFe region. Upper row contains pre-fit plots and lower row contains post-background-only-fit plots.

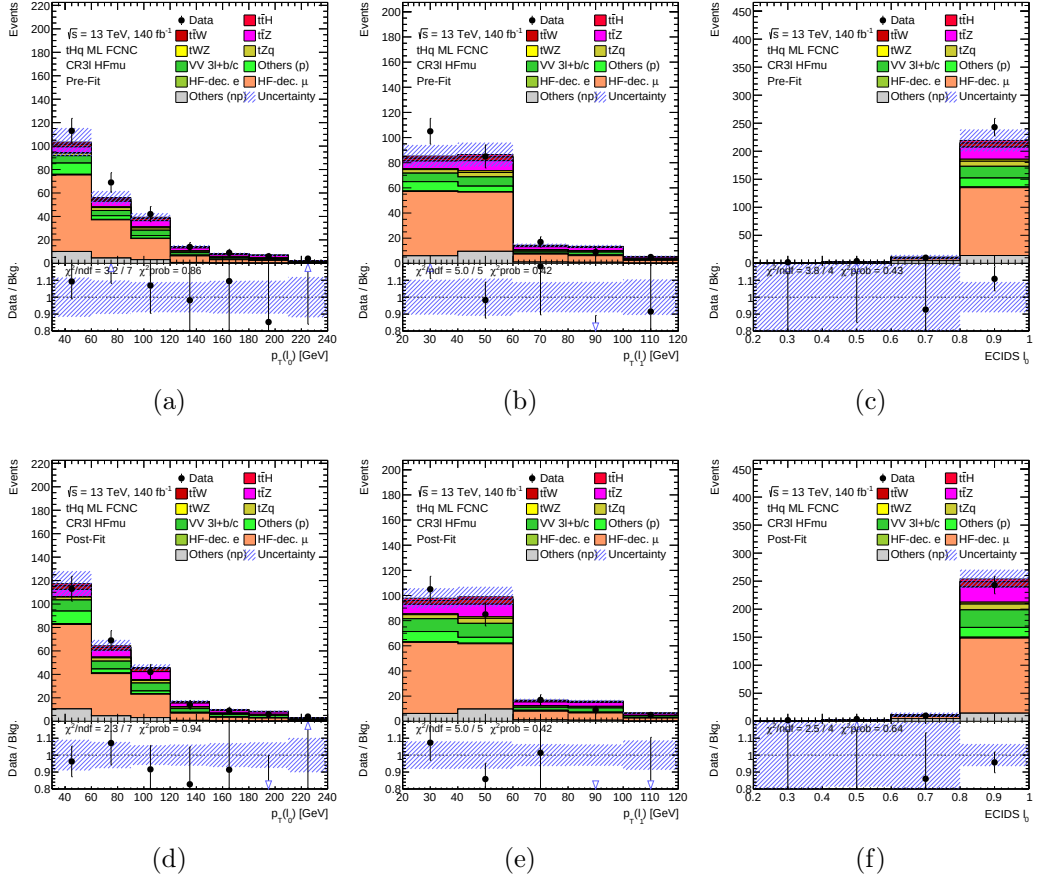


Figure G.8: Validation plots of the leading-lepton p_T (left), the second-leading lepton p_T (middle) and the ECIDS score of the leading lepton (right) in the CR3 ℓ HF μ region. Upper row contains pre-fit plots and lower row contains post-background-only-fit plots.

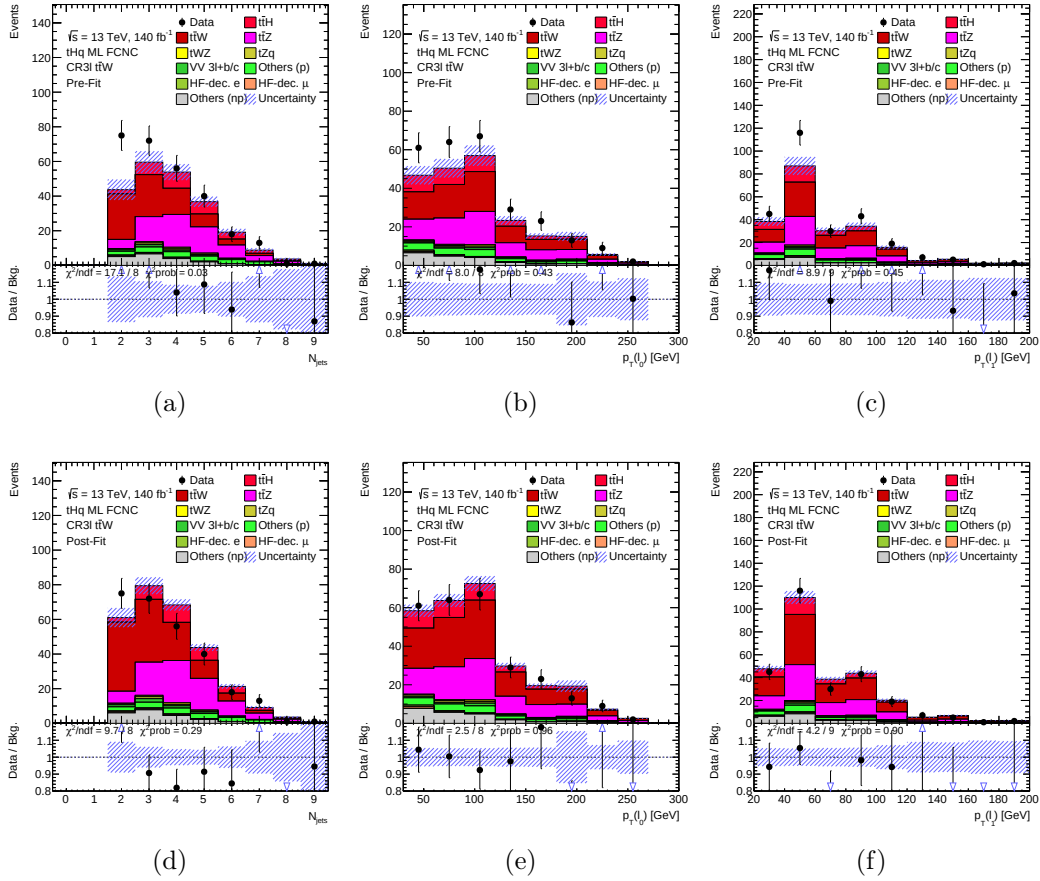


Figure G.9: Validation plots of the number of jets (left), the leading lepton p_T (middle) and the second-leading lepton p_T (right) in the $CR3l\bar{t}tW$ region. Upper row contains pre-fit plots and lower row contains post-background-only-fit plots.

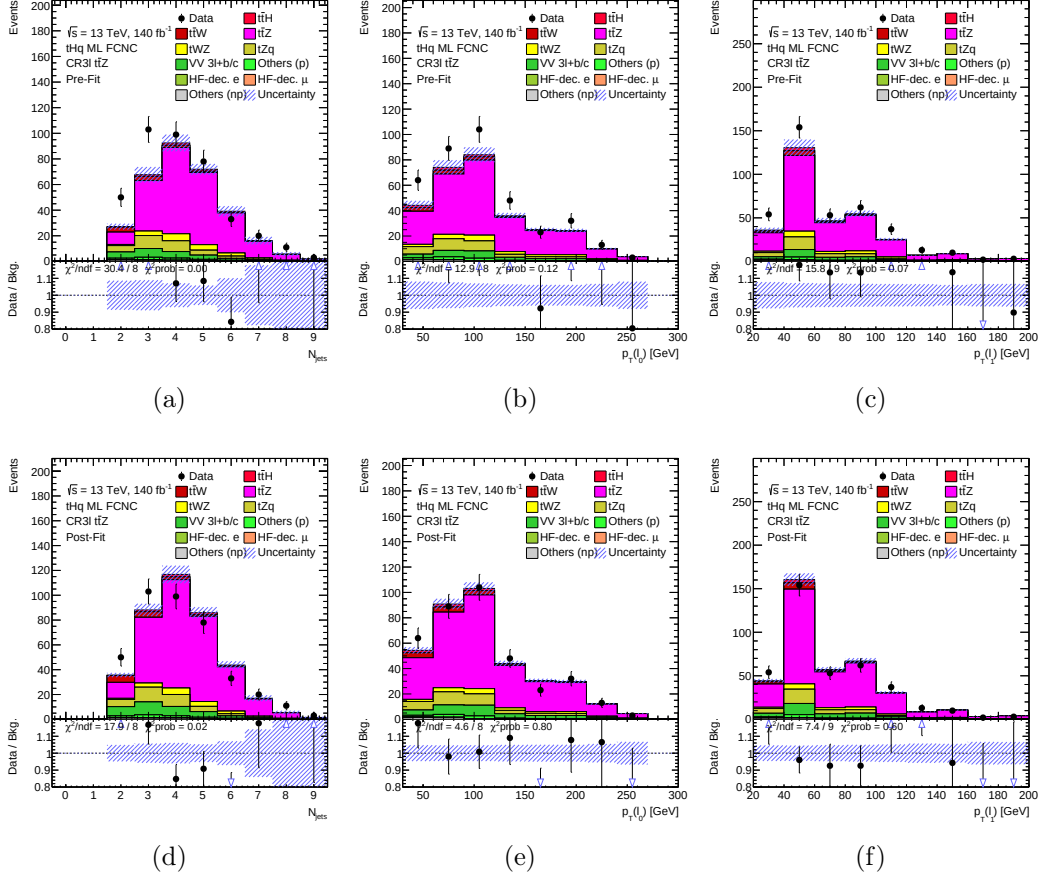


Figure G.10: Validation plots of the number of jets (left), the leading lepton p_T (middle) and the second-leading lepton p_T (right) in the $CR3l ttZ$ region. Upper row contains pre-fit plots and lower row contains post-background-only-fit plots.

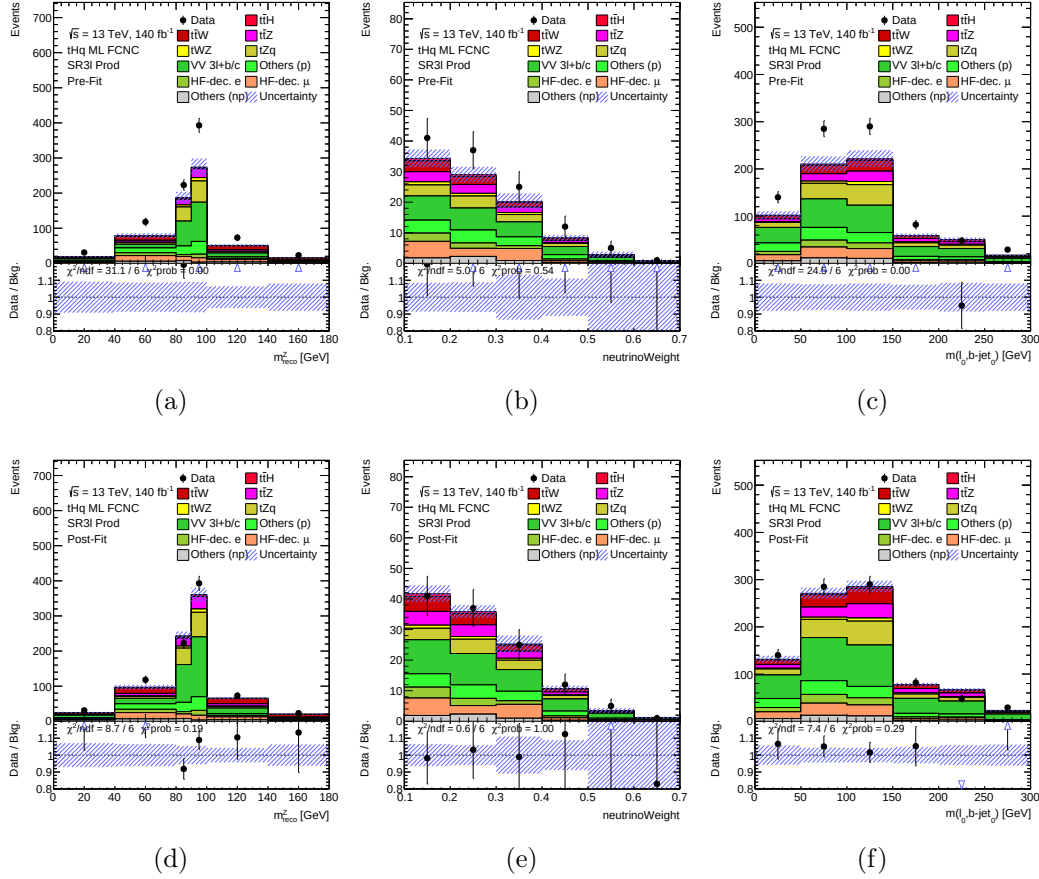


Figure G.11: Validation plots of the $ZRecoMass$ distribution (left), the $neutrinoWeight$ (middle) and the invariant mass of the second-leading lepton and the b -tagged jet (right) in the SR3lProd region. Upper row contains pre-fit plots and lower row contains post-background-only-fit plots.

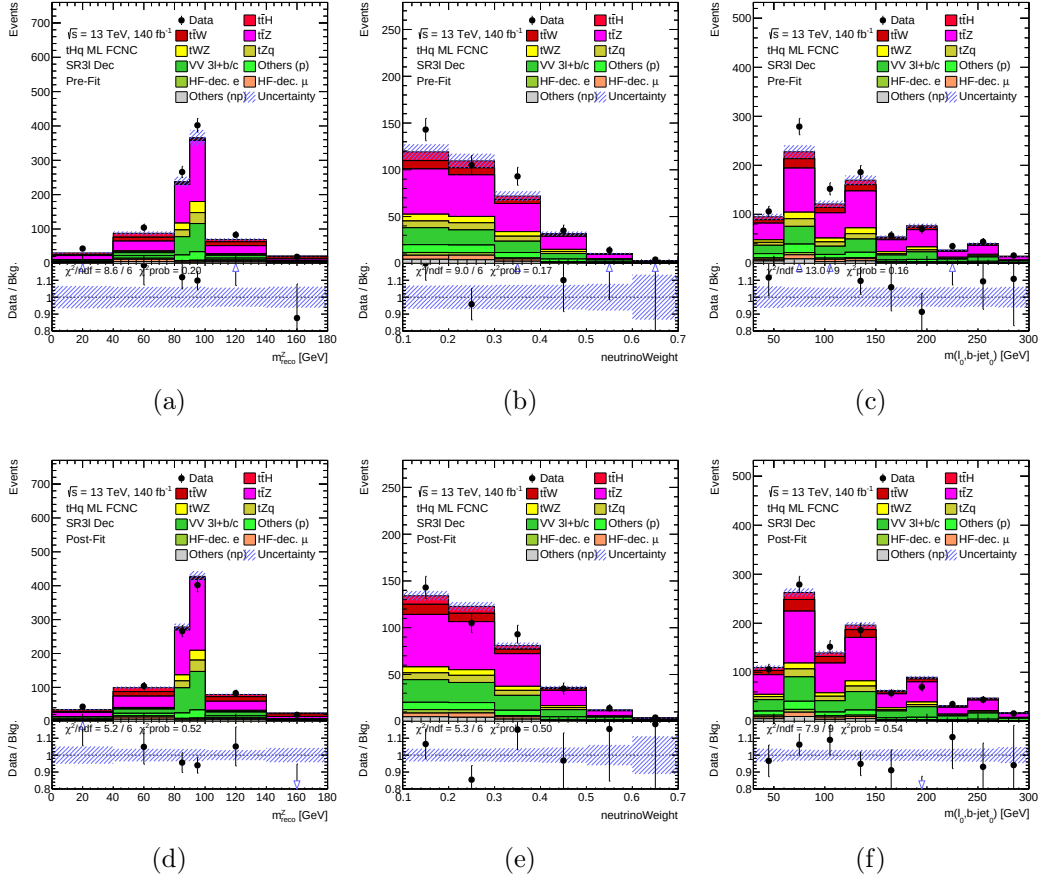


Figure G.12: Validation plots of the $ZRecoMass$ distribution (left), the $neutrinoWeight$ (middle) and the invariant mass of the second-leading lepton and the b -tagged jet (right) in the SR3 ℓ Dec region. Upper row contains pre-fit plots and lower row contains post-background-only-fit plots.

H. Plots related to the realistic Asimov Fit

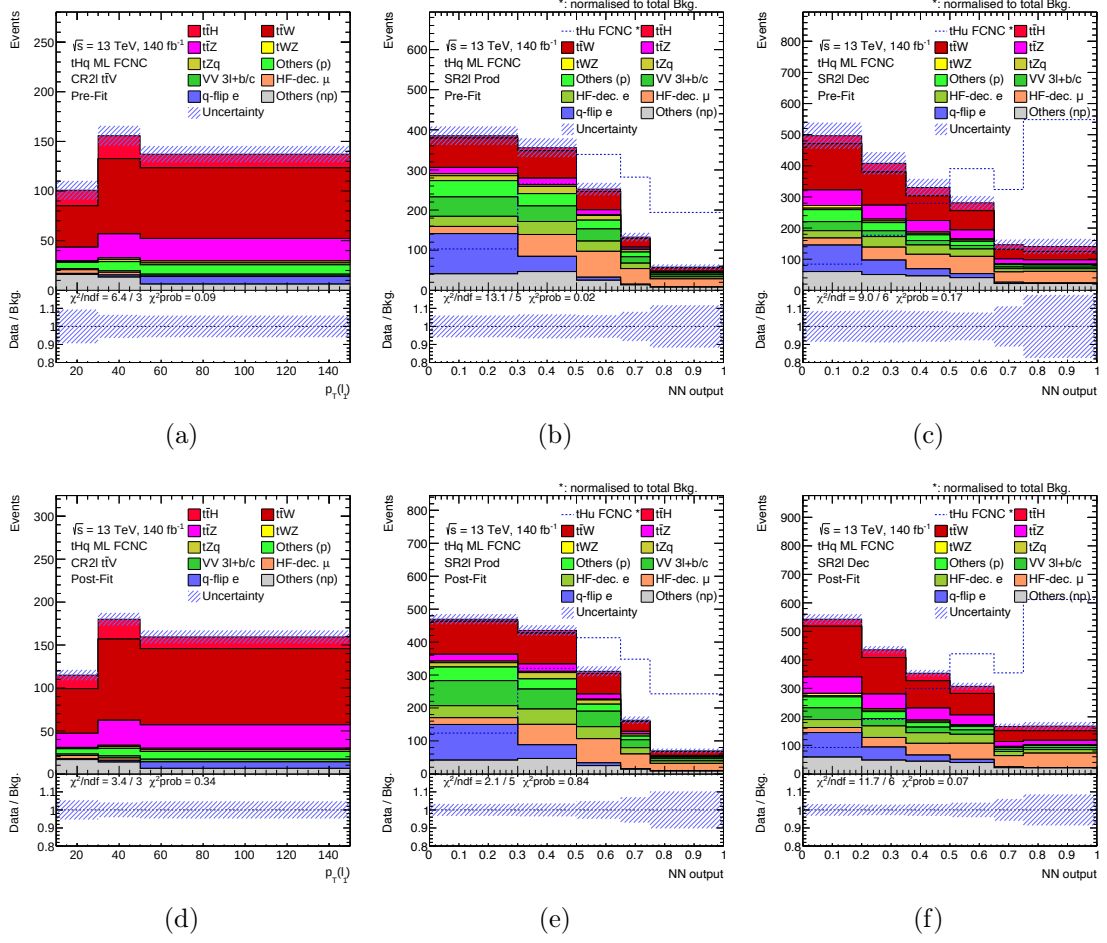


Figure H.2: The 2ℓSS regions CR2ℓ $t\bar{t}V$, SR2ℓProd and SR2ℓDec once pre-fit (upper row) and once post-fit (lower row) in the realistic Asimov fit.

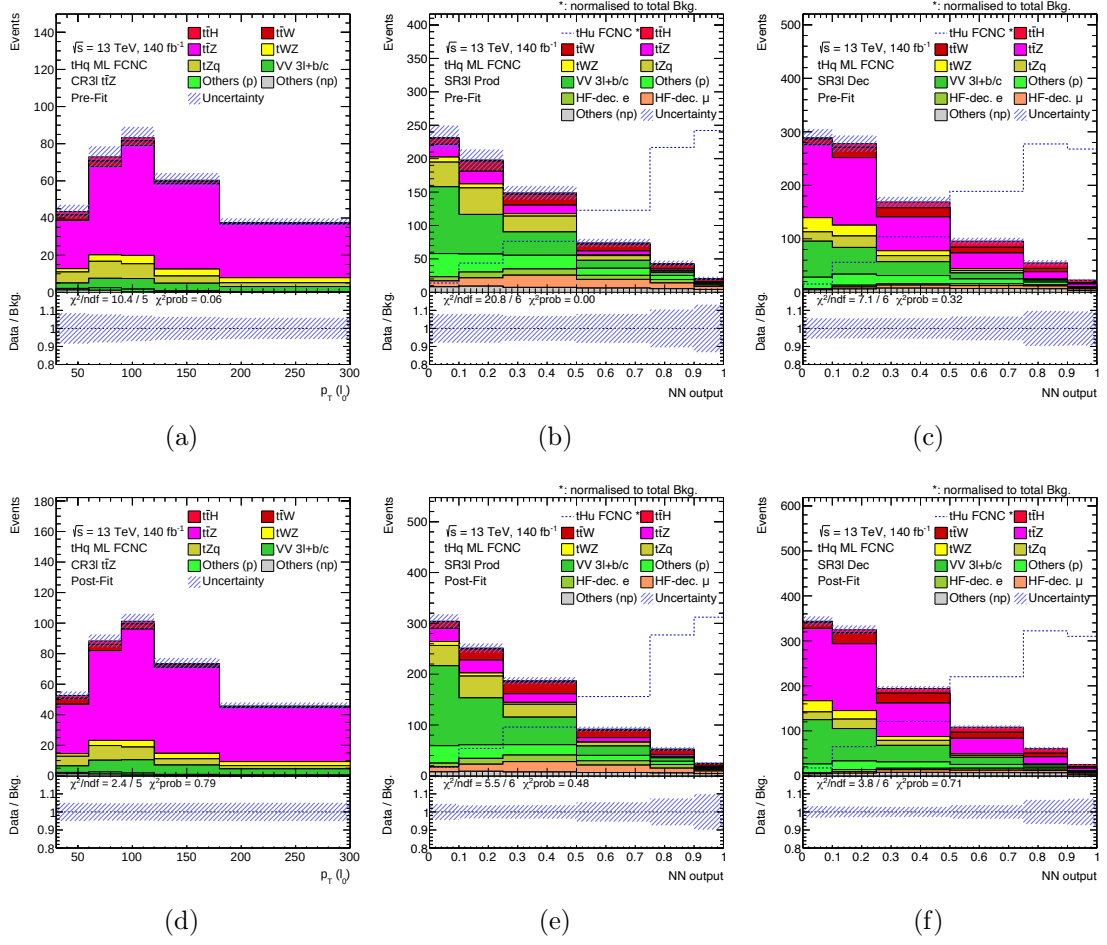


Figure H.4: The 3 ℓ regions CR3 ℓ t τ Z, SR3 ℓ Prod and SR3 ℓ Dec once pre-fit (upper row) and once post-fit (lower row) in the realistic Asimov fit.

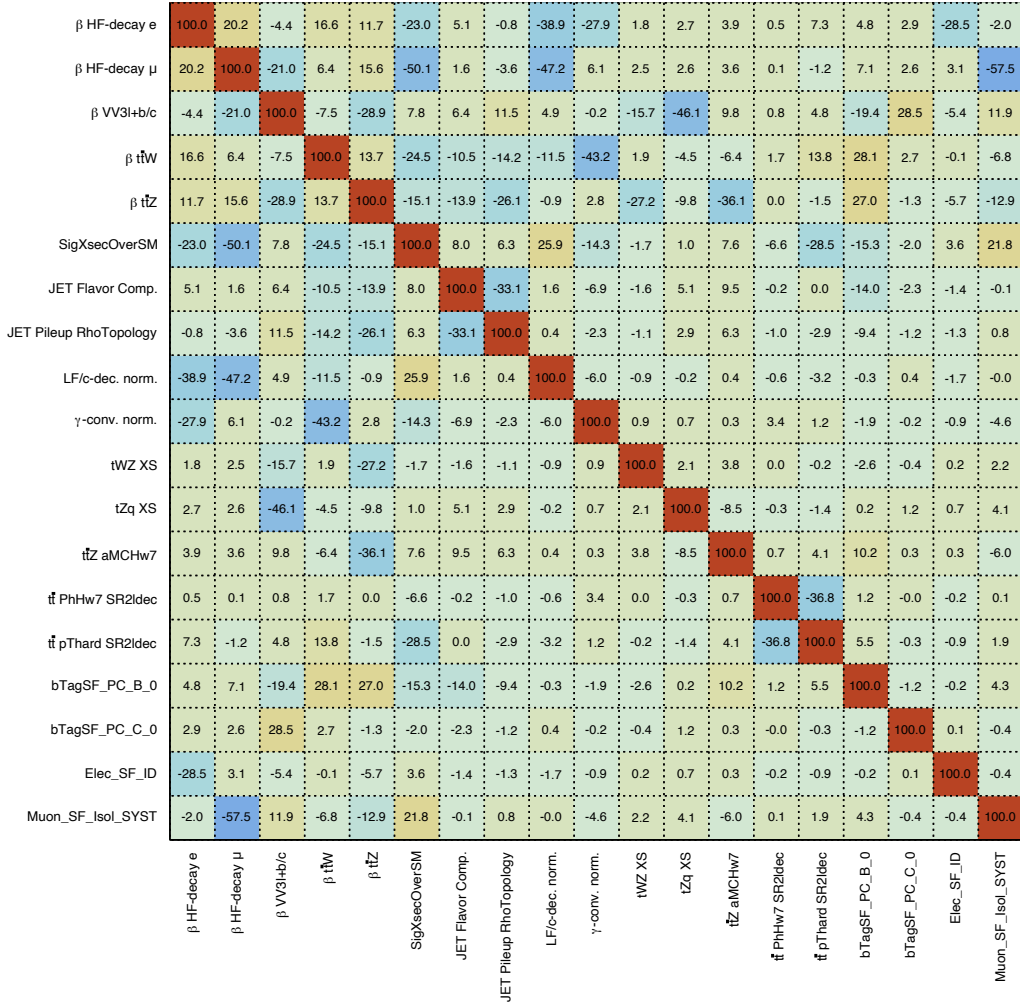


Figure H.5: Correlation matrix of NPs in the realistic Asimov fit with a linear correlation coefficient larger than 25% to any other NP but themselves, exemplarily for the utH case.

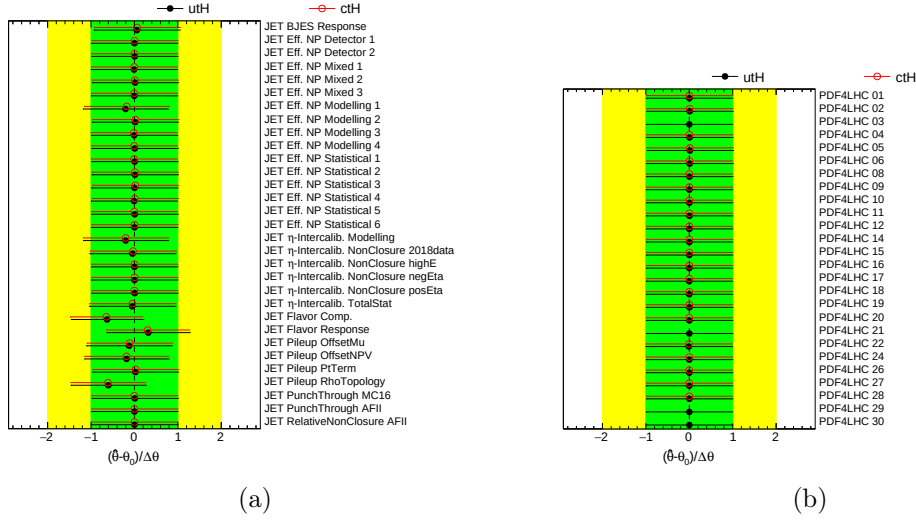


Figure H.7: Pulls and constraints of the NPs related to jet systematic uncertainties (a) and PDF systematic uncertainties (b) in the realistic Asimov fit.

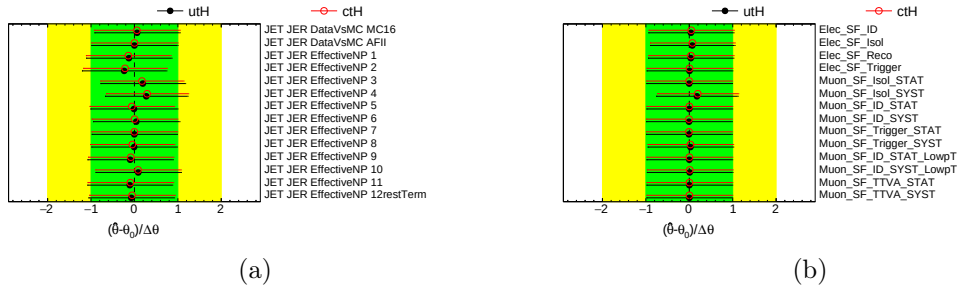


Figure H.8: Pulls and constraints of the NPs related to JER systematic uncertainties (a) and lepton systematic uncertainties (b) in the realistic Asimov fit.

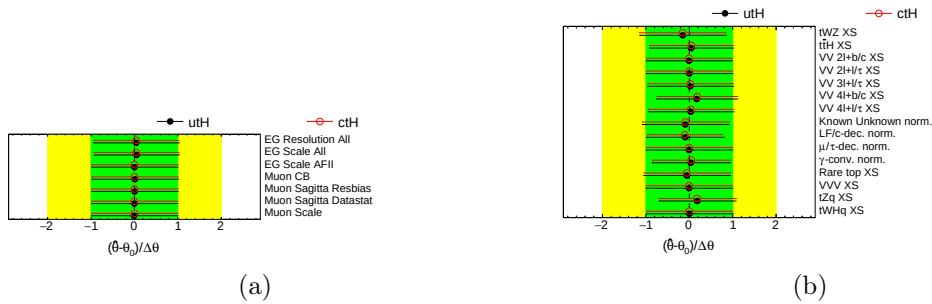


Figure H.9: Pulls and constraints of the NPs related to electron and muon systematic uncertainties (a) and theory cross-section systematic uncertainties (b) in the realistic Asimov fit.

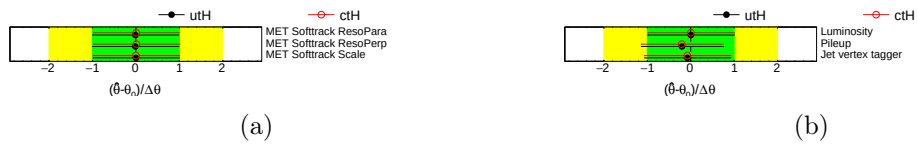


Figure H.10: Pulls and constraints of the NPs related to E_T^{miss} systematic uncertainties (a) and other systematic uncertainties (b) in the realistic Asimov fit.

Bibliography

- [1] Michael Edward Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory*. Reading, USA: Addison-Wesley (1995) 842 p. Westview Press, 1995.
- [2] Sheldon L. Glashow. “Partial-symmetries of weak interactions”. In: *Nuclear Physics* 22.4 (1961), pp. 579–588. ISSN: 0029-5582. DOI: [https://doi.org/10.1016/0029-5582\(61\)90469-2](https://doi.org/10.1016/0029-5582(61)90469-2). URL: <https://www.sciencedirect.com/science/article/pii/0029558261904692>.
- [3] CDF Collaboration. “Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab”. In: *Phys. Rev. Lett.* 74 (14 Apr. 1995), pp. 2626–2631. DOI: 10.1103/PhysRevLett.74.2626. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.74.2626>.
- [4] S. et al. Abachi. “Observation of the Top Quark”. In: *Phys. Rev. Lett.* 74 (14 Apr. 1995), pp. 2632–2637. DOI: 10.1103/PhysRevLett.74.2632. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.74.2632>.
- [5] ATLAS collaboration. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Physics Letters B* 716.1 (Sept. 2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. URL: <https://doi.org/10.1016%2Fj.physletb.2012.08.020>.
- [6] CMS collaboration. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: *Physics Letters B* 716.1 (Sept. 2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. URL: <https://doi.org/10.1016%2Fj.physletb.2012.08.021>.
- [7] R. P. Feynman. “Space-Time Approach to Quantum Electrodynamics”. In: *Phys. Rev.* 76 (6 Sept. 1949), pp. 769–789. DOI: 10.1103/PhysRev.76.769. URL: <https://link.aps.org/doi/10.1103/PhysRev.76.769>.
- [8] Cush MissMJ. *Standard Model of Elementary Particles*. 2023. URL: https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg (visited on 13/06/2023).
- [9] Steven Weinberg. “A Model of Leptons”. In: *Phys. Rev. Lett.* 19 (21 Nov. 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.19.1264>.
- [10] Nicola Cabibbo. “Unitary Symmetry and Leptonic Decays”. In: *Phys. Rev. Lett.* 10 (12 June 1963), pp. 531–533. DOI: 10.1103/PhysRevLett.10.531. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.10.531>.

- [11] J. H. Christenson et al. “Evidence for the 2π Decay of the K_2^0 Meson”. In: *Phys. Rev. Lett.* 13 (4 July 1964), pp. 138–140. DOI: 10.1103/PhysRevLett.13.138. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.138>.
- [12] Makoto Kobayashi and Toshihide Maskawa. “CP-Violation in the Renormalizable Theory of Weak Interaction”. In: *Progress of Theoretical Physics* 49.2 (Feb. 1973), pp. 652–657. ISSN: 0033-068X. DOI: 10.1143/PTP.49.652. eprint: <https://academic.oup.com/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf>. URL: <https://doi.org/10.1143/PTP.49.652>.
- [13] Ling-Lie Chau and Wai-Yee Keung. “Comments on the Parametrization of the Kobayashi-Maskawa Matrix”. In: *Phys. Rev. Lett.* 53 (19 Nov. 1984), pp. 1802–1805. DOI: 10.1103/PhysRevLett.53.1802. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.53.1802>.
- [14] R. L. Workman et al. “Review of Particle Physics”. In: *PTEP* 2022 (2022), p. 083C01. DOI: 10.1093/ptep/ptac097.
- [15] P.W. Higgs. “Broken symmetries, massless particles and gauge fields”. In: *Physics Letters* 12.2 (1964), pp. 132–133. ISSN: 0031-9163. DOI: [https://doi.org/10.1016/0031-9163\(64\)91136-9](https://doi.org/10.1016/0031-9163(64)91136-9). URL: <https://www.sciencedirect.com/science/article/pii/0031916364911369>.
- [16] F. Englert and R. Brout. “Broken Symmetry and the Mass of Gauge Vector Mesons”. In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.321>.
- [17] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble. “Global Conservation Laws and Massless Particles”. In: *Phys. Rev. Lett.* 13 (20 Nov. 1964), pp. 585–587. DOI: 10.1103/PhysRevLett.13.585. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.585>.
- [18] Abdus Salam. “Weak and electromagnetic interactions”. In: *Selected Papers of Abdus Salam*, pp. 244–254. DOI: 10.1142/9789812795915_0034. eprint: https://www.worldscientific.com/doi/pdf/10.1142/9789812795915_0034. URL: https://www.worldscientific.com/doi/abs/10.1142/9789812795915_0034.
- [19] Gonix. *The potential for the Higgs field*. 2007. URL: https://commons.wikimedia.org/wiki/File:Mecanismo_de_Higgs_PH.png (visited on 23/06/2023).
- [20] Kenneth G. Wilson. “Confinement of quarks”. In: *Phys. Rev. D* 10 (8 Oct. 1974), pp. 2445–2459. DOI: 10.1103/PhysRevD.10.2445. URL: <https://link.aps.org/doi/10.1103/PhysRevD.10.2445>.

- [21] LHCb collaboration. “Observation of J/Ψ resonances consistent with pentaquark states in $\Lambda_0 b \rightarrow J/\Psi K$ -p decays”. In: *Physical Review Letters* 115.7 (Aug. 2015). DOI: 10.1103/physrevlett.115.072001. URL: <https://doi.org/10.1103/physrevlett.115.072001>.
- [22] David J. Gross and Frank Wilczek. “Ultraviolet Behavior of Non-Abelian Gauge Theories”. In: *Phys. Rev. Lett.* 30 (26 June 1973), pp. 1343–1346. DOI: 10.1103/PhysRevLett.30.1343. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.30.1343>.
- [23] ATLAS Collaboration. *Summary of several Standard Model total and fiducial production cross-section measurements*. 2022. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2022-009/fig_03a.pdf (visited on 28/06/2023).
- [24] Particle Data Group. “Review of Particle Physics”. In: *Progress of Theoretical and Experimental Physics* 2020.8 (Aug. 2020), p. 083C01. ISSN: 2050-3911. DOI: 10.1093/ptep/ptaa104. eprint: <https://academic.oup.com/ptep/article-pdf/2020/8/083C01/34673722/ptaa104.pdf>. URL: <https://doi.org/10.1093/ptep/ptaa104>.
- [25] ATLAS Collaboration. *Measurement of the top-quark mass and the W-boson mass, compared to the prediction of a global electroweak fit*. 2023. URL: <https://cds.cern.ch/images/ATLAS-PHOTO-2023-022-3> (visited on 28/06/2023).
- [26] Y. et al. Fukuda. “Evidence for Oscillation of Atmospheric Neutrinos”. In: *Phys. Rev. Lett.* 81 (8 Aug. 1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.81.1562>.
- [27] Pablo F. de Salas et al. “Neutrino Mass Ordering from Oscillations and Beyond: 2018 Status and Future Prospects”. In: *Frontiers in Astronomy and Space Sciences* 5 (Oct. 2018). DOI: 10.3389/fspas.2018.00036. URL: <https://doi.org/10.3389/fspas.2018.00036>.
- [28] N. Aghanim et al. “Planck 2018 results”. In: *Astronomy and Astrophysics* 641 (Sept. 2020), A6. DOI: 10.1051/0004-6361/201833910. URL: <https://doi.org/10.1051/0004-6361/201833910>.
- [29] A. Einstein. “Die Grundlage der allgemeinen Relativitätstheorie”. In: *Annalen der Physik* 354.7 (1916), pp. 769–822. DOI: <https://doi.org/10.1002/andp.19163540702>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.19163540702>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.19163540702>.

- [30] K. G. Begeman, A. H. Broeils and R. H. Sanders. “Extended rotation curves of spiral galaxies: dark haloes and modified dynamics”. In: *Monthly Notices of the Royal Astronomical Society* 249.3 (Apr. 1991), pp. 523–537. ISSN: 0035-8711. DOI: 10.1093/mnras/249.3.523. eprint: <https://academic.oup.com/mnras/article-pdf/249/3/523/18160929/mnras249-0523.pdf>. URL: <https://doi.org/10.1093/mnras/249.3.523>.
- [31] Matthew D. Schwartz. *Quantum Field Theory and the Standard Model*. 2013. ISBN: 9781107034730.
- [32] C. A. Baker et al. “Improved Experimental Limit on the Electric Dipole Moment of the Neutron”. In: *Phys. Rev. Lett.* 97 (13 Sept. 2006), p. 131801. DOI: 10.1103/PhysRevLett.97.131801. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.97.131801>.
- [33] B. Pontecorvo. “Inverse beta processes and nonconservation of lepton charge”. In: *Zh. Eksp. Teor. Fiz.* 34 (1957), p. 247.
- [34] Ziro Maki, Masami Nakagawa and Shoichi Sakata. “Remarks on the Unified Model of Elementary Particles”. In: *Progress of Theoretical Physics* 28.5 (Nov. 1962), pp. 870–880. ISSN: 0033-068X. DOI: 10.1143/PTP.28.870. eprint: <https://academic.oup.com/ptp/article-pdf/28/5/870/5258750/28-5-870.pdf>. URL: <https://doi.org/10.1143/PTP.28.870>.
- [35] S. L. Glashow, J. Iliopoulos and L. Maiani. “Weak Interactions with Lepton-Hadron Symmetry”. In: *Phys. Rev. D* 2 (7 Oct. 1970), pp. 1285–1292. DOI: 10.1103/PhysRevD.2.1285. URL: <https://link.aps.org/doi/10.1103/PhysRevD.2.1285>.
- [36] ATLAS and CMS Collaboration. *Summary of the current 95% confidence level observed limits on the branching ratios of the top quark decays via flavor changing neutral currents (FCNC) to a quark and a neutral boson $t \rightarrow Xq$ ($X = g, Z, \gamma$ or H ; $q = u$ or c) by the ATLAS and CMS Collaborations compared to several new physics models*. 2023. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2023-013/fig_04.png (visited on 30/06/2023).
- [37] Celine Degrande et al. “Automatic computations at next-to-leading order in QCD for top-quark flavor-changing neutral processes”. In: *Phys. Rev. D* 91 (3 Feb. 2015), p. 034024. DOI: {10.1103/PhysRevD.91.034024}. URL: <https://link.aps.org/doi/10.1103/PhysRevD.91.034024>.
- [38] Gauthier Durieux, Fabio Maltoni and Cen Zhang. “Global approach to top-quark flavor-changing interactions”. In: *Phys. Rev. D* 91 (7 Apr. 2015), p. 074017. DOI: 10.1103/PhysRevD.91.074017. URL: <https://link.aps.org/doi/10.1103/PhysRevD.91.074017>.

- [39] Adam Alloul et al. “FeynRules 2.0 - A complete toolbox for tree-level phenomenology”. In: *Computer Physics Communications* 185.8 (2014), pp. 2250–2300. ISSN: 0010-4655. DOI: <https://doi.org/10.1016/j.cpc.2014.04.012>. URL: <http://www.sciencedirect.com/science/article/pii/S0010465514001350>.
- [40] Celine Degrande et al. “UFO – The Universal FeynRules Output”. In: *Computer Physics Communications* 183.6 (2012), pp. 1201–1214. ISSN: 0010-4655. DOI: <https://doi.org/10.1016/j.cpc.2012.01.022>. URL: <http://www.sciencedirect.com/science/article/pii/S0010465512000379>.
- [41] Cen Zhang and Fabio Maltoni. “Top-quark decay into Higgs boson and a light quark at next-to-leading order in QCD”. In: *Phys. Rev. D* 88 (5 Sept. 2013), p. 054005. DOI: 10.1103/PhysRevD.88.054005. URL: <https://link.aps.org/doi/10.1103/PhysRevD.88.054005>.
- [42] Cinzia De Melis. “The CERN accelerator complex. Complexe des accélérateurs du CERN”. In: (2016). General Photo. URL: <https://cds.cern.ch/record/2119882>.
- [43] Michael Benedikt et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: 10.5170/CERN-2004-003-V-3. URL: <https://cds.cern.ch/record/823808>.
- [44] Lyndon Evans and Philip Bryant. “LHC Machine”. In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001. URL: <https://dx.doi.org/10.1088/1748-0221/3/08/S08001>.
- [45] The ALICE Collaboration. “The ALICE experiment at the CERN LHC”. In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002. URL: <https://dx.doi.org/10.1088/1748-0221/3/08/S08002>.
- [46] The ATLAS Collaboration. “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003. URL: <https://dx.doi.org/10.1088/1748-0221/3/08/S08003>.
- [47] The CMS Collaboration. “The CMS experiment at the CERN LHC”. In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004. URL: <https://dx.doi.org/10.1088/1748-0221/3/08/S08004>.
- [48] The LHCb Collaboration. “The LHCb Detector at the LHC”. In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005. URL: <https://dx.doi.org/10.1088/1748-0221/3/08/S08005>.
- [49] B Salvachua. “Overview of Proton-Proton Physics during Run 2”. In: (2019), pp. 7–14. URL: <https://cds.cern.ch/record/2750272>.

- [50] ATLAS Collaboration. *LHC delivered luminosity per year*. 2023. URL: <https://atlas.web.cern.ch/Atlas/GROUPS/DATAPREPARATION/PublicPlots/2023/DataSummary/figs/intlumivsy.png> (visited on 30/05/2023).
- [51] ATLAS Collaboration. *Mean Number of Interactions per Bunch Crossing*. 2023. URL: https://atlas.web.cern.ch/Atlas/GROUPS/DATAPREPARATION/PublicPlots/2018/DataSummary/figs/mu_2015_2018.png (visited on 26/05/2023).
- [52] ATLAS Collaboration. *Schematic of the ATLAS Detector*. 2023. URL: <https://cds.cern.ch/record/1095924> (visited on 30/05/2023).
- [53] Sascha Mehlhase. “ATLAS detector slice (and particle visualisations)”. In: (2021). URL: <https://cds.cern.ch/record/2770815>.
- [54] ATLAS Collaboration. *ATLAS inner detector: Technical Design Report, 1*. Technical design report. ATLAS. Geneva: CERN, 1997. URL: <https://cds.cern.ch/record/331063>.
- [55] S Haywood et al. *ATLAS inner detector: Technical Design Report, 2*. Technical design report. ATLAS. Geneva: CERN, 1997. URL: <https://cds.cern.ch/record/331064>.
- [56] ATLAS collaboration. “Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2”. In: *The European Physical Journal C* 77.10 (Oct. 2017). DOI: 10.1140/epjc/s10052-017-5225-7. URL: <https://doi.org/10.1140/epjc/s10052-017-5225-7>.
- [57] ATLAS collaboration. “Production and integration of the ATLAS Insertable B-Layer”. In: *Journal of Instrumentation* 13.05 (May 2018), T05008–T05008. DOI: 10.1088/1748-0221/13/05/t05008. URL: <https://doi.org/10.1088/1748-0221/13/05/t05008>.
- [58] ATLAS collaboration. *ATLAS calorimeter performance: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1996. URL: <https://cds.cern.ch/record/331059>.
- [59] ATLAS collaboration. *ATLAS muon spectrometer: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1997. URL: <https://cds.cern.ch/record/331068>.
- [60] Joao Pequeno. “Computer generated image of the ATLAS Muons subsystem”. 2008. URL: <http://cds.cern.ch/record/1095929>.
- [61] S. Abdel Khalek et al. “The ALFA Roman Pot detectors of ATLAS”. In: *Journal of Instrumentation* 11.11 (Nov. 2016), P11013–P11013. DOI: 10.1088/1748-0221/11/11/p11013. URL: <https://doi.org/10.1088/1748-0221/11/11/p11013>.

- [62] ATLAS collaboration. “Performance of the ATLAS trigger system in 2015”. In: *The European Physical Journal C* 77.5 (May 2017). DOI: 10.1140/epjc/s10052-017-4852-3. URL: <https://doi.org/10.1140%2Fepjc%2Fs10052-017-4852-3>.
- [63] The ATLAS collaboration. “Operation of the ATLAS trigger system in Run 2”. In: *Journal of Instrumentation* 15.10 (Oct. 2020), P10004–P10004. DOI: 10.1088/1748-0221/15/10/p10004. URL: <https://doi.org/10.1088%2F1748-0221%2F15%2F10%2Fp10004>.
- [64] G. Avoni et al. “The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS”. In: *Journal of Instrumentation* 13.07 (July 2018), P07017. DOI: 10.1088/1748-0221/13/07/P07017. URL: <https://dx.doi.org/10.1088/1748-0221/13/07/P07017>.
- [65] V Cindro et al. “The ATLAS Beam Conditions Monitor”. In: *Journal of Instrumentation* 3.02 (Feb. 2008), P02004. DOI: 10.1088/1748-0221/3/02/P02004. URL: <https://dx.doi.org/10.1088/1748-0221/3/02/P02004>.
- [66] “Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC”. In: (Dec. 2022). arXiv: 2212.09379 [hep-ex].
- [67] Frank Siegert. “Monte-Carlo event generation for the LHC”. Doctoral thesis. Durham University, 2010.
- [68] John C. Collins, Davison E. Soper and George Sterman. *Factorization of Hard Processes in QCD*. 2004. arXiv: hep-ph/0409313 [hep-ph].
- [69] Yuri L. Dokshitzer. “Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics.” In: *Sov. Phys. JETP* 46 (1977), pp. 641–653.
- [70] V. N. Gribov and L. N. Lipatov. “Deep inelastic $e p$ scattering in perturbation theory”. In: *Sov. J. Nucl. Phys.* 15 (1972), pp. 438–450.
- [71] G. Altarelli and G. Parisi. “Asymptotic freedom in parton language”. In: *Nuclear Physics B* 126.2 (1977), pp. 298–318. ISSN: 0550-3213. DOI: [https://doi.org/10.1016/0550-3213\(77\)90384-4](https://doi.org/10.1016/0550-3213(77)90384-4). URL: <https://www.sciencedirect.com/science/article/pii/0550321377903844>.
- [72] Luigi Del Debbio. “Parton distributions in the LHC era”. In: *EPJ Web of Conferences* 175 (Jan. 2018), p. 01006. DOI: 10.1051/epjconf/201817501006.
- [73] John C. Collins. “Choosing the factorization / renormalization scale in perturbative qcd calculations”. In: *25th Rencontres de Moriond: High-energy Hadronic Interactions*. 1990, pp. 123–128.

- [74] B. Andersson et al. “Parton fragmentation and string dynamics”. In: *Physics Reports* 97.2 (1983), pp. 31–145. ISSN: 0370-1573. DOI: [https://doi.org/10.1016/0370-1573\(83\)90080-7](https://doi.org/10.1016/0370-1573(83)90080-7). URL: <https://www.sciencedirect.com/science/article/pii/0370157383900807>.
- [75] B.R. Webber. “A QCD model for jet fragmentation including soft gluon interference”. In: *Nuclear Physics B* 238.3 (1984), pp. 492–528. ISSN: 0550-3213. DOI: [https://doi.org/10.1016/0550-3213\(84\)90333-X](https://doi.org/10.1016/0550-3213(84)90333-X). URL: <https://www.sciencedirect.com/science/article/pii/055032138490333X>.
- [76] D. Amati and G. Veneziano. “Preconfinement as a property of perturbative QCD”. In: *Physics Letters B* 83.1 (1979), pp. 87–92. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(79\)90896-7](https://doi.org/10.1016/0370-2693(79)90896-7). URL: <https://www.sciencedirect.com/science/article/pii/0370269379908967>.
- [77] S. Agostinelli et al. “Geant4—a simulation toolkit”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. ISSN: 0168-9002. DOI: [https://doi.org/10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8). URL: <https://www.sciencedirect.com/science/article/pii/S0168900203013688>.
- [78] ATLAS collaboration. *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*. Tech. rep. Geneva: CERN, 2010. URL: <https://cds.cern.ch/record/1300517>.
- [79] ATLAS collaboration. “Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2”. In: *The European Physical Journal C* 77.10 (Oct. 2017). DOI: 10.1140/epjc/s10052-017-5225-7. URL: <https://doi.org/10.1140%2Fepjc%2Fs10052-017-5225-7>.
- [80] Azriel Rosenfeld and John L. Pfaltz. “Sequential Operations in Digital Picture Processing”. In: *J. ACM* 13.4 (Oct. 1966), pp. 471–494. ISSN: 0004-5411. DOI: 10.1145/321356.321357. URL: <https://doi.org/10.1145/321356.321357>.
- [81] R. Frühwirth. “Application of Kalman filtering to track and vertex fitting”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 262.2 (1987), pp. 444–450. ISSN: 0168-9002. DOI: [https://doi.org/10.1016/0168-9002\(87\)90887-4](https://doi.org/10.1016/0168-9002(87)90887-4). URL: <https://www.sciencedirect.com/science/article/pii/0168900287908874>.
- [82] ATLAS collaboration. “Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC”. In: *The European Physical Journal C* 77.5 (May 2017). DOI: 10.1140/epjc/s10052-017-4887-5. URL: <https://doi.org/10.1140%2Fepjc%2Fs10052-017-4887-5>.

- [83] ATLAS collaboration. “Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data”. In: *Journal of Instrumentation* 14.12 (Dec. 2019), P12006–P12006. DOI: 10.1088/1748-0221/14/12/p12006. URL: <https://doi.org/10.1088/1748-0221/14/12/p12006>.
- [84] ATLAS collaboration. “Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1”. In: *The European Physical Journal C* 77.7 (July 2017). DOI: 10.1140/epjc/s10052-017-5004-5. URL: <https://doi.org/10.1140/epjc/s10052-017-5004-5>.
- [85] ATLAS collaboration. “Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data”. In: *Journal of Instrumentation* 14.03 (Mar. 2019), P03017–P03017. DOI: 10.1088/1748-0221/14/03/p03017. URL: <https://doi.org/10.1088/1748-0221/14/03/p03017>.
- [86] ATLAS Collaboration. *Electron and Photon Selection and Identification for Run2 and Run3*. 2023. URL: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/EGammaIdentificationRun2> (visited on 12/07/2023).
- [87] ATLAS Collaboration. *Electron Charge ID Selector Tool*. 2023. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ElectronChargeFlipTaggerTool> (visited on 12/07/2023).
- [88] ATLAS collaboration. “Muon reconstruction performance of the ATLAS detector in proton proton collision data at $\sqrt{s} = 13$ TeV”. In: *The European Physical Journal C* 76.5 (May 2016). DOI: 10.1140/epjc/s10052-016-4120-y. URL: <https://doi.org/10.1140/epjc/s10052-016-4120-y>.
- [89] G. et al Aad. “Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV”. In: *The European Physical Journal C* 81.7 (July 2021), p. 578. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-021-09233-2. URL: <https://doi.org/10.1140/epjc/s10052-021-09233-2>.
- [90] J. Illingworth and J. Kittler. “A survey of the hough transform”. In: *Computer Vision, Graphics, and Image Processing* 44.1 (1988), pp. 87–116. ISSN: 0734-189X. DOI: [https://doi.org/10.1016/S0734-189X\(88\)80033-1](https://doi.org/10.1016/S0734-189X(88)80033-1). URL: <https://www.sciencedirect.com/science/article/pii/S0734189X88800331>.
- [91] ATLAS Collaboration. *MuonSelectionTool*. 2023. URL: <https://twiki.cern.ch/twiki/bin/viewauth/Atlas/MuonSelectionTool> (visited on 13/07/2023).

- [92] ATLAS collaboration. “Performance of the ATLAS muon triggers in Run 2”. In: *Journal of Instrumentation* 15.09 (Sept. 2020), P09015–P09015. DOI: 10.1088/1748-0221/15/09/p09015. URL: <https://doi.org/10.1088/1748-0221/15/09/p09015>.
- [93] ATLAS collaboration. “Jet reconstruction and performance using particle flow with the ATLAS Detector”. In: *The European Physical Journal C* 77.7 (July 2017). DOI: 10.1140/epjc/s10052-017-5031-2. URL: <https://doi.org/10.1140/epjc/s10052-017-5031-2>.
- [94] Matteo Cacciari, Gavin P Salam and Gregory Soyez. “The anti- k_t clustering algorithm”. In: *Journal of High Energy Physics* 2008.04 (Apr. 2008), pp. 063–063. DOI: 10.1088/1126-6708/2008/04/063. URL: <https://doi.org/10.1088/1126-6708/2008/04/063>.
- [95] ATLAS collaboration. “Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Eur. Phys. J. C* 81.8 (2021), p. 689. DOI: 10.1140/epjc/s10052-021-09402-3. arXiv: 2007.02645 [hep-ex].
- [96] ATLAS collaboration. “Performance of pile-up mitigation techniques for jets in proton-proton collisions at $\sqrt{s}$ TeV using the ATLAS detector”. In: *The European Physical Journal C* 76.11 (Oct. 2016). DOI: 10.1140/epjc/s10052-016-4395-z. URL: <https://doi.org/10.1140/epjc/s10052-016-4395-z>.
- [97] Nazar Bartosik. *Diagram showing the common principle of identification of jets initiated by b-hadron decays*. 2016. URL: <https://commons.wikimedia.org/w/index.php?curid=49738737> (visited on 17/07/2023).
- [98] *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*. Tech. rep. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2273281>.
- [99] *Secondary vertex finding for jet flavour identification with the ATLAS detector*. Tech. rep. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2270366>.
- [100] *Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC*. Tech. rep. Geneva: CERN, 2018. URL: <https://cds.cern.ch/record/2645405>.
- [101] ATLAS collaboration. “ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13\text{TeV}$ ”. In: *The European Physical Journal C* 79.11 (Nov. 2019). DOI: 10.1140/epjc/s10052-019-7450-8. URL: <https://doi.org/10.1140/epjc/s10052-019-7450-8>.

- [102] *Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment*. Tech. rep. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2255226>.
- [103] Fudong He et al. *Tagging non-prompt electrons and muons using lifetime and isolation information*. Tech. rep. Geneva: CERN, 2022. URL: <https://cds.cern.ch/record/2800472>.
- [104] Fudong He et al. *PLImprovedWPs*. 2022. URL: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/PLImprovedWPs> (visited on 20/07/2023).
- [105] Tom Neep et al. *IFFTruthClassifier*. 2020. URL: <https://gitlab.cern.ch/atlas-physics/sm/ew/wgg/IFFTruthClassifier> (visited on 20/07/2023).
- [106] ATLAS collaboration. “Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{13}$ TeV”. In: *The European Physical Journal C* 78.11 (Nov. 2018). DOI: 10.1140/epjc/s10052-018-6288-9. URL: <https://doi.org/10.1140/epjc/s10052-018-6288-9>.
- [107] E_T^{miss} performance in the ATLAS detector using 2015-2016 LHC p-p collisions. Tech. rep. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2018-023>. Geneva: CERN, 2018. URL: <https://cds.cern.ch/record/2625233>.
- [108] Stefano Frixione, Giovanni Ridolfi and Paolo Nason. “A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction”. In: *Journal of High Energy Physics* 2007.09 (Sept. 2007), pp. 126–126. DOI: 10.1088/1126-6708/2007/09/126. URL: <https://doi.org/10.1088/1126-6708/2007/09/126>.
- [109] Simone Alioli et al. “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”. In: *Journal of High Energy Physics* 2010.6 (June 2010). DOI: 10.1007/jhep06(2010)043. URL: [https://doi.org/10.1007/jhep06\(2010\)043](https://doi.org/10.1007/jhep06(2010)043).
- [110] Stefano Frixione, Paolo Nason and Carlo Oleari. “Matching NLO QCD computations with parton shower simulations: the POWHEG method”. In: *Journal of High Energy Physics* 2007.11 (Nov. 2007), pp. 070–070. DOI: 10.1088/1126-6708/2007/11/070. URL: <https://doi.org/10.1088/1126-6708/2007/11/070>.
- [111] Paolo Nason. “A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms”. In: *Journal of High Energy Physics* 2004.11 (Nov. 2004), pp. 040–040. DOI: 10.1088/1126-6708/2004/11/040. URL: <https://doi.org/10.1088/1126-6708/2004/11/040>.

- [112] J. Alwall et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In: *Journal of High Energy Physics* 2014.7 (July 2014). DOI: 10.1007/jhep07(2014)079. URL: <https://doi.org/10.1007%2Fjhep07%282014%29079>.
- [113] Richard D. Ball et al. “Parton distributions for the LHC run II”. In: *Journal of High Energy Physics* 2015.4 (Apr. 2015). DOI: 10.1007/jhep04(2015)040. URL: <https://doi.org/10.1007%2Fjhep04%282015%29040>.
- [114] Stefano Frixione et al. “Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations”. In: *Journal of High Energy Physics* 2007.04 (Apr. 2007), p. 081. DOI: 10.1088/1126-6708/2007/04/081. URL: <https://dx.doi.org/10.1088/1126-6708/2007/04/081>.
- [115] Pierre Artoisenet et al. “Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations”. In: *Journal of High Energy Physics* 2013.3 (Mar. 2013), p. 15. ISSN: 1029-8479. DOI: 10.1007/JHEP03(2013)015. URL: [https://doi.org/10.1007/JHEP03\(2013\)015](https://doi.org/10.1007/JHEP03(2013)015).
- [116] Torbjörn Sjöstrand et al. “An introduction to PYTHIA 8.2”. In: *Computer Physics Communications* 191 (June 2015), pp. 159–177. DOI: 10.1016/j.cpc.2015.01.024. URL: <https://doi.org/10.1016%2Fj.cpc.2015.01.024>.
- [117] Johannes Bellm et al. “Herwig 7.0/Herwig 3.0 release note”. In: *The European Physical Journal C* 76.4 (Apr. 2016). DOI: 10.1140/epjc/s10052-016-4018-8. URL: <https://doi.org/10.1140%2Fepjc%2Fs10052-016-4018-8>.
- [118] Manuel Bähr et al. “Herwig physics and manual”. In: *The European Physical Journal C* 58.4 (Nov. 2008), pp. 639–707. DOI: 10.1140/epjc/s10052-008-0798-9. URL: <https://doi.org/10.1140%2Fepjc%2Fs10052-008-0798-9>.
- [119] *ATLAS Pythia 8 tunes to 7 TeV data*. Tech. rep. Geneva: CERN, 2014. URL: <https://cds.cern.ch/record/1966419>.
- [120] T Gleisberg et al. “Event generation with SHERPA 1.1”. In: *Journal of High Energy Physics* 2009.02 (Feb. 2009), pp. 007–007. DOI: 10.1088/1126-6708/2009/02/007. URL: <https://doi.org/10.1088%2F1126-6708%2F2009%2F02%2F007>.
- [121] Michał Czakon and Alexander Mitov. “Top++: A program for the calculation of the top-pair cross-section at hadron colliders”. In: *Computer Physics Communications* 185.11 (Nov. 2014), pp. 2930–2938. DOI: 10.1016/j.cpc.2014.06.021. URL: <https://doi.org/10.1016%2Fj.cpc.2014.06.021>.

- [122] M. Beneke et al. “Hadronic top-quark pair production with NNLL threshold resummation”. In: *Nuclear Physics B* 855.3 (Feb. 2012), pp. 695–741. DOI: 10.1016/j.nuclphysb.2011.10.021. URL: <https://doi.org/10.1016%2Fj.nuclphysb.2011.10.021>.
- [123] ATLAS CMS Collaboration. *NNLO+NNLL top-quark-pair cross sections*. 2013. URL: <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO> (visited on 24/07/2023).
- [124] Rikkert Frederix and Ioannis Tsinikos. “On improving NLO merging for ttW production”. In: *Journal of High Energy Physics* 2021.11 (Nov. 2021), p. 29. ISSN: 1029-8479. DOI: 10.1007/JHEP11(2021)029. URL: [https://doi.org/10.1007/JHEP11\(2021\)029](https://doi.org/10.1007/JHEP11(2021)029).
- [125] Paul Jackson and Christopher Rogan. “Recursive jigsaw reconstruction: HEP event analysis in the presence of kinematic and combinatoric ambiguities”. In: *Phys. Rev. D* 96 (11 Dec. 2017), p. 112007. DOI: 10.1103/PhysRevD.96.112007. URL: <https://link.aps.org/doi/10.1103/PhysRevD.96.112007>.
- [126] Christopher Rogan. *RestFrames Homepage*. 2020. URL: <https://restframes.com> (visited on 02/08/2023).
- [127] Michael Feindt. *A Neural Bayesian Estimator for Conditional Probability Densities*. 2004. arXiv: physics/0402093 [physics.data-an].
- [128] M. Feindt and U. Kerzel. “The NeuroBayes neural network package”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 559.1 (2006). Proceedings of the X International Workshop on Advanced Computing and Analysis Techniques in Physics Research, pp. 190–194. ISSN: 0168-9002. DOI: <https://doi.org/10.1016/j.nima.2005.11.166>. URL: <https://www.sciencedirect.com/science/article/pii/S0168900205022679>.
- [129] <phi-t> Physics Information Technologies. *The NeuroBayes User’s Guide*. 2010. URL: <https://manualzz.com/doc/4120778/the-neurobayes-r-user-s-guide---neurobayes.de---phi-t> (visited on 03/08/2023).
- [130] Richard H. Byrd et al. “A Limited Memory Algorithm for Bound Constrained Optimization”. In: *SIAM Journal on Scientific Computing* 16.5 (1995), pp. 1190–1208. DOI: 10.1137/0916069. eprint: <https://doi.org/10.1137/0916069>. URL: <https://doi.org/10.1137/0916069>.
- [131] Yann A. LeCun et al. “Efficient BackProp”. In: *Neural Networks: Tricks of the Trade: Second Edition*. Ed. by Grégoire Montavon, Geneviève B. Orr and Klaus-Robert Müller. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 9–48. ISBN: 978-3-642-35289-8. DOI: 10.1007/978-3-642-35289-8_3. URL: https://doi.org/10.1007/978-3-642-35289-8_3.

- [132] Glen Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *The European Physical Journal C* 71.2 (Feb. 2011). DOI: 10.1140/epjc/s10052-011-1554-0. URL: <https://doi.org/10.1140/2Fepjc%2Fs10052-011-1554-0>.
- [133] Kyle Cranmer et al. *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*. Tech. rep. New York: New York U., 2012. URL: <https://cds.cern.ch/record/1456844>.
- [134] G. Cowan. *Statistical Data Analysis*. Oxford science publications. Clarendon Press, 1998. ISBN: 9780198501558. URL: https://www.sherrytowers.com/cowan_statistical_data_analysis.pdf.
- [135] Jerzy Neyman, Egon Sharpe Pearson and Karl Pearson. “IX. On the problem of the most efficient tests of statistical hypotheses”. In: *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 231.694-706 (1933), pp. 289–337. DOI: 10.1098/rsta.1933.0009. eprint: <https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.1933.0009>. URL: <https://royalsocietypublishing.org/doi/abs/10.1098/rsta.1933.0009>.
- [136] A L Read. “Presentation of search results: the CLs technique”. In: *Journal of Physics G: Nuclear and Particle Physics* 28.10 (Sept. 2002), p. 2693. DOI: 10.1088/0954-3899/28/10/313. URL: <https://dx.doi.org/10.1088/0954-3899/28/10/313>.
- [137] F. James, Y. Perrin and L. Lyons, eds. *Workshop on confidence limits, CERN, Geneva, Switzerland, 17-18 Jan 2000: Proceedings*. CERN Yellow Reports: Conference Proceedings. May 2000. DOI: 10.5170/CERN-2000-005.
- [138] Michele Pinamonti et al. *TRExFitter Gitlab*. 2023. URL: <https://gitlab.cern.ch/TRExStats/TRExFitter> (visited on 11/08/2023).
- [139] Michele Pinamonti et al. *TRExFitter Documentation*. 2023. URL: <https://trexfitter-docs.web.cern.ch/trexfitter-docs/> (visited on 11/08/2023).
- [140] G. et al. Aad. “Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{13}$ TeV with the ATLAS detector”. In: *The European Physical Journal C* 81.8 (Aug. 2021), p. 689. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-021-09402-3. URL: <https://doi.org/10.1140/epjc/s10052-021-09402-3>.
- [141] *Studies on the improvement of the matching uncertainty definition in top-quark processes simulated with Powheg+Pythia 8*. Tech. rep. Geneva: CERN, 2023. URL: <https://cds.cern.ch/record/2872787>.

- [142] Nuno Castro et al. *Search for flavor-changing neutral-current couplings between the top quark and the Z boson with LHC Run2 proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. Geneva: CERN, 2021. URL: <https://cds.cern.ch/record/2777010>.
- [143] CERN. *CERN Yellow Reports: Monographs, Vol 2 (2017): Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector*. en. 2017. DOI: 10.23731/CYRM-2017-002. URL: <https://e-publishing.cern.ch/index.php/CYRM/issue/view/32>.
- [144] *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*. Tech. rep. Geneva: CERN, 2021. URL: <https://cds.cern.ch/record/2758863>.
- [145] *Measurement of the total and differential cross-sections of $t\bar{t}W$ production in pp collisions at 13 TeV with the ATLAS detector*. Tech. rep. Geneva: CERN, 2023. URL: <https://cds.cern.ch/record/2855337>.
- [146] ATLAS collaboration. “Search for flavor-changing neutral currents in top quark decays $t \rightarrow Hc$ and $t \rightarrow Hu$ in multilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Physical Review D* 98.3 (Aug. 2018). DOI: 10.1103/physrevd.98.032002. URL: <https://doi.org/10.1103/PhysRevD.98.032002>.
- [147] ATLAS collaboration. “Search for top-quark decays $t \rightarrow Hq$ with 36 fb⁻¹ of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Journal of High Energy Physics* 2019.5 (May 2019). DOI: 10.1007/jhep05(2019)123. URL: [https://doi.org/10.1007/JHEP05\(2019\)123](https://doi.org/10.1007/JHEP05(2019)123).
- [148] A. Tumasyan et al. “Search for Flavor-Changing Neutral Current Interactions of the Top Quark and Higgs Boson in Final States with Two Photons in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV”. In: *Physical Review Letters* 129.3 (July 2022). DOI: 10.1103/physrevlett.129.032001. URL: <https://doi.org/10.1103/PhysRevLett.129.032001>.
- [149] *Search for flavour-changing neutral current interactions of the top quark and the Higgs boson in events with a pair of τ -leptons in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. Geneva: CERN, 2022. URL: <http://cds.cern.ch/record/2805360>.
- [150] *Search for flavor-changing neutral current interactions of the top quark and Higgs boson in proton-proton collisions at $\sqrt{s} = 13$ TeV*. Tech. rep. Geneva: CERN, 2023. URL: <https://cds.cern.ch/record/2872677>.

Danksagung

Auch wenn auf der Titelseite nur mein Name steht, so wäre diese Arbeit wohl kaum ohne die Hilfe von vielen anderen möglich gewesen. Seit der Bachelorarbeit wurde ich von Prof. Dr. Wolfgang Wagner exzellent betreut. Sein fachkundiges Wissen und verständliches Erklären waren immer enorm wertvoll, wenn man Rat brauchte. Sei es runds ums Thema Teilchenphysik oder welche Stauden er für den Garten empfehlen kann! Auch Joshua und Marvin gilt mein besonderer Dank. Joshua hat mir seit der Oberstufe aber insbesondere im Studium immer weiter geholfen, wenn etwas unklar war. Ohne ihn hätte mein Studium wohl ein paar Semester länger gedauert. Marvin war der beste Analysepartner, den man sich vorstellen kann: Immer gut informiert um komplizierte Aspekte zu besprechen und überaus kompetent um die Analyse alleine vorran zu bringen, sodass ich mich in den letzten Monaten komplett aufs Schreiben dieser Arbeit konzentrieren konnte.

Auch möchte ich praktisch dem ganzen Fachbereich Teilchenphysik an der Uni Wuppertal danken, denn der wissenschaftliche und freundschaftliche Austausch war über nahezu 8 Jahre sehr wertvoll. Dazu zählen Dominic, Frank, Jens, Volker, Shayma, Maren, Anna, Johanna, Lukas, Axel, und alle die ich in der Aufzählung vergessen habe. Ohne das Clusterteam wäre sowieso nix passiert, denn ohne die Instandhaltung und den Support des Rechenclusters hätte ich meine MC/data events von Hand auszählen müssen.

Dann möchte ich auch den anderen Mitgliedern meiner Prüfungskommission (neben Wolfgang) danken: Andrea, Dirk und Francesco. Es kostet viel Zeit nahezu 200 Seiten eines komplizierten Themas durchzulesen und ihr habt diese Aufgabe freiwillig übernommen um meine Promotion zu ermöglichen.

Zu guter letzt hat mich auch meine Familie immer unterstützt. Ich konnte lange Zeit Zuhause wohnen bleiben, wodurch ich zB keinen Nebenjob annehmen musste ich mich so vollständig auf das Studium konzentrieren konnte. Und sobald ich ausgezogen war, hatte mein Bruder immer Zeit um mir bei Elektrik in der Wohnung zu helfen.

Danke