Department of Physics and Astronomy

Heidelberg University

Master thesis

in Physics

submitted by

Nikita Schmal

born in Heilbronn, Germany

2023

Staying on Top of Likelihood Analyses

Exploring various approaches of using likelihoods
for global SMEFT analyses in the Top sector

This Master thesis has been carried out by Nikita Schmal

at the

Institut für Theoretische Physik

under the supervision of

Prof. Dr. Tilman Plehn

Abstract - Exploring various approaches of using likelihoods for global SMEFT analyses in the Top sector:

While the Standard Model is able to describe numerous different measurements to an impressive extent, there are many observations that it still can not explain. In this thesis, global SMEFT analyses are performed to put constraints on new physics. The role of the likelihood, in particular, is studied in great detail.
For the first time, likelihoods published by ATLAS are used to implement new measurements into SFitter. Using these, one is able to validate approximations made in previous global analyses, showing that they already lead to very good results. Furthermore, the effect of profiling and marginalization of the likelihood in the global Top fit was studied, where the importance of theory uncertainties in the Top sector was shown. The additional implementation of a new $t\bar{t}$ differential cross section significantly improved the constraints on many SMEFT operators.
Finally, a combined fit using data from the Top, Higgs and Di-Boson datasets was performed using SFitter for the first time. It was shown that one finds stronger constraints for operators directly contributing to both Higgs and Top measurements.

Zusammenfassung - Untersuchung verschiedener Ansätze Likelihoods für globale SMEFT Analysen im Top Sektor zu verwenden:

Während das Standard Model eine Vielzahl verschiedener Experimente ausgezeichnet gut beschreiben kann, gibt es immer noch viele physikalische Beobachtungen, die es nicht erklären kann. In dieser Arbeit werden globale SMEFT Analysen durchgeführt, um Anzeichen neuer Physik einzuschränken. Vor allem die Rolle der Likelihood wird in diesem Kontext genauer analysiert.
Zum ersten Mal werden Likelihoods, die von ATLAS veröffentlicht wurden, verwendet, um neue Messung in SFitter zu implementieren. Mit diesen kann bestätigt werden, dass Näherungen, die in vorherigen globalen SMEFT analysen verwendet wurden, zu sehr guten Ergebnissen führen. Außerdem wurde der Effekt der Verwendung von profiling und marginalisierung der Likelihood analysiert. Es wurde gezeigt, dass vor allem die Theorieunsicherheiten im Top Sektor sehr wichtig sind. Daraufhin wurde ein neuer differentieller $t\bar{t}$ Wirkungsquerschnitt in SFitter implementiert, wodurch mehrere SMEFT Operatoren um einiges stärker eingeschränkt wurden.
Zum Schluss wurde zum ersten Mal ein kombinierter SMEFT fit mit Top, Higgs und Di-Boson Daten in SFitter durchgeführt. Man findet, dass die Operatoren, die gleichzeitig zu Top und Higgs Messungen beitragen, stärker eingeschränkt werden.

Contents

1

Introduction

It has been more than a decade since the discovery of the Higgs Boson at the Large Hadron Collider (LHC), the final experimental piece of evidence needed to complete the Standard Model (SM) [1,2,3]. It took almost 5 decades for this to be achieved after the proposal of the Higgs mechanism back in 1964 [4,5]. To this day, this marks one of the greatest discoveries in the field of particle physics. Ever since, more and more data has been taken at the LHC and the SM has proven to describe observations with impressive accuracy for a vast amount of different measurements. There are, however, some notable observations which the SM alone cannot explain.

The experimental discovery of neutrino oscillations [6,7,8], for example, proved that neutrinos actually do have a mass, a property which the framework of the SM cannot accommodate. Furthermore, there is the hierarchy problem [9] and the non-vanishing CP violating phase of QCD [10]. Going to larger, cosmological, scales, one finds oneself unable to explain Dark Matter [11,12] and the matter anti-matter asymmetry of the universe [13,14]. Consequently, the SM is far from being able to give a complete picture of all physics phenomena currently observed.

Although, various theoretical models have been studied to find explanations for all of these phenomena, conclusive results have yet to be found. In this thesis, a model agnostic Effective Field Theory (EFT) approach is used to determine constraints on physics beyond the Standard Model. For this purpose, the EFT extension of the SM, the Standard Model Effective Field Theory (SMEFT), is used. The SMEFT extends the SM via the addition of higher dimensional operators allowing for the study of new physics, under the assumption that the scale of these new physics is at energies much higher than those probed by current experiments. The data from measurements at the LHC is then used to determine constraints on the so-called Wilson coefficients (WC) which parametrize the effect of these operators.

Due to the large amount of data produced nowadays, an essential component of any such analysis is the statistical framework used to perform these fits. The tool of choice for the following analyses is SFitter, a tool which has been used for a variety of global SMEFT fits [15,16], with a strong focus on a comprehensive treatment of uncertainties and their correlations. The central object of any SFitter analysis is the likelihood, which gives a measure of the agreement between data and theory, incorporating the effect of uncertainties by introducing nuisance parameters (NP). Recently, the SFitter framework has been updated and different ways of handling these NPs have been implemented, allowing for statistical analyses using either profiling or Bayesian marginalization methods [16].

This thesis will be split into three distinct parts. The first part is concerned with the analysis of likelihoods published by the ATLAS Top working group [17,18]. Their properties, the information gained from them and how they can be used to implement new measurements into the SFitter dataset, are studied. Finally, the effect of these new measurements on the constraints of WCs from a global SMEFT fit in the Top sector, is analyzed.

The second part of the thesis is centered on the new Bayesian marginalization methods implemented into SFitter in Ref. [16]. An update to a previous global analysis in the Top sector [15] is provided, using the newly implemented marginalization techniques, checking whether there are any differences to be found in the constraints coming from the profiled or marginalized likelihoods. Afterward, the Top dataset from Ref. [15] will be updated by adding a new high energy distribution of the $t\bar{t}$ differential cross section in the leptons+jets channel, by the CMS collaboration [19]. The implementation of this distribution for two different observables, $m_{t\bar{t}}$ and $p_T(t_h)$, is studied to assess the influence that the choice of observable has on the implementation of the measurement and its constraints on the WCs.

Finally, the thesis will conclude with a combination of datasets from two previous global analyses using SFitter. The first is the global SMEFT analysis in the Top sector from Ref. [20] and the second is a global SMEFT analysis using Higgs, Di-Boson and Electroweak precision data from Ref. [16]. It will be the first fit using SFitter combining all Top, Higgs and Di-boson sectors and allows one to explore the interplay between these.

2

Standard Model Effective Field Theory (SMEFT)

This chapter will serve as a quick introduction to Effective Field Theories (EFTs), the assumptions they are based on, as well as the motivation behind their application to searches for physics beyond the Standard Model (SM). The majority of the theory is based on Ref. [21,22] and I refer to these for a more detailed introduction. We begin with a historical look at EFTs in particle physics and then extend the discussion to applications of the EFT formalism to searches for physics beyond the SM.

2.1 Effective Field Theories (EFTs)

As already mentioned in the Introduction, there is a multitude of different experimental observations which our current theoretical framework, the SM, cannot explain. On top of this, we have now entered an era of particle physics in which one has to look for the faintest signals for physics beyond the SM in the huge amount of data generated at the Large Hadron Collider (LHC). To deal with both of these problems, it is necessary to have a robust and well understood framework which is able to capture these faint signals.

Naturally, this is not the first time that physicists were not able to explain certain physical phenomena using the established theoretical frameworks of their time. Looking back to the 20th century, the physics community was puzzled by the beta decay

$$n \to p + e^- + \bar{\nu}_e \,. \tag{2.1}$$

The observed energy spectra were much wider than those one would expect from a 3-body decay. This led to the proposition that an additional particle had to be emitted in a beta decay carrying some energy with it and with that leading to a wider energy spectrum.

Following this, Enrico Fermi formulated a theory to describe this process, proposing a point-like 4 fermion interaction, which was already able to describe the weak interaction to an impressive extent for the time. From a modern point of view, however, it is very well known that the weak interaction is actually mediated by a very heavy W-Boson, although it would take another 50 years for this particle to actually be directly observed in experiment at the Super Proton Synchrotron.

To better understand how this approach can be extended to current problems, one can first take a closer look at the beta decay to find out why exactly Fermis 4-point interaction led to such good results at the time.

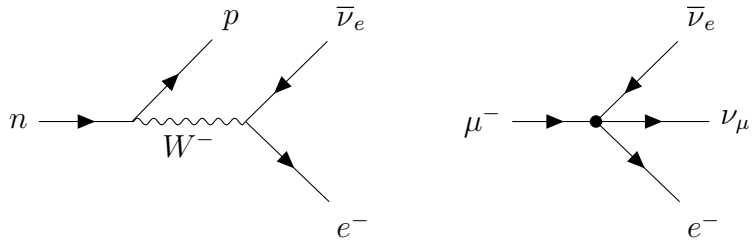


Figure 2.1: Left: Feynman diagram for leading order beta decay of a neutron into a proton. Right: Feynman diagram for μ^--decay via a 4 fermion interaction.

On the left side of Figure 2.1 the Feynman diagram for the beta decay is shown. Applying the usual Feynman rules, one can easily calculate the matrix element of this process

$$\mathcal{M}_{SM} = \left[\bar{\psi}_p \gamma^\mu \left(1-\gamma_5\right) \psi_n\right] \frac{\left(g_{\mu\nu} - p_\mu p_\nu\right)}{p^2 - m_W^2} \left[\bar{\psi}_e \gamma^\nu \left(1-\gamma_5\right) \psi_{\bar{\nu}}\right] , \tag{2.2}$$

where one now makes the key assumption that $m_W^2 \gg p^2$ which gives

$$\mathcal{M}_{SM} = \frac{g^2}{8m_W^2} \left[\bar{\psi}_p \gamma^\mu \left(1-\gamma_5\right) \psi_n\right] \left[\psi_e \gamma_\mu \left(1-\gamma_5\right) \bar{\psi}_{\bar{\nu}}\right] . \tag{2.3}$$

Doing this one essentially removes the degree of freedom introduced by the W-Boson effectively leading to a 4-point vertex as one can see on the right-hand side of Figure 2.1 just like the one proposed by Fermi.

Naturally one can also start with the 4-point interaction which gives

$$\mathcal{M}_{\text{Fermi}} = \frac{G_F}{\sqrt{2}} \left[\bar{\psi}_p \gamma^\mu \left(1-\gamma_5\right) \psi_n\right] \left[\psi_e \gamma_\mu \left(1-\gamma_5\right) \bar{\psi}_{\bar{\nu}}\right] \tag{2.4}$$

with Fermi constant G_F as it was introduced in the initial theory proposed by Fermi. Comparing the two expression one finds immediately that

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8m_W^2} \tag{2.5}$$

a procedure known as matching which already shows nicely why Fermis theory was able to explain the weak interaction so well.

However, if one were to now simply perform some dimensional analysis using this result, one finds that

$$\sigma(p \to n + e^- + \bar{\nu_e}) \propto G_F^2 E_{CMS}^2 \tag{2.6}$$

naturally leading to unitarity issues at very large energies. This is a consequence of the initial assumptions that $m_W^2 \gg p^2$ which naturally holds for lower energies, especially since it is now known that $m_W = 80.377$ GeV [23], but breaks down once one reaches energy scales close to m_W.

This leads to one of the most essential concepts within any EFT framework, the so-called

separation of scales, where one assumes that any kind of new physics takes place at some energy scale Λ assumed to be much larger than the energies probed in experiments $\Lambda \gg E$. This allows one to integrate out the heavier degrees of freedom of a theory as long as this is still valid. In the case of the weak interaction, this scale simply corresponds to m_W allowing the W-Boson to be integrated out as shown before. This is why Fermis description of the weak interaction held up so well at the time, since none of the experiments reached energies high enough to violate the assumption of a separation of scales.

2.2 EFT extension of the Standard Model

Before applying EFT concepts to the SM a basic understanding of its most important properties is necessary and as such a quick summary of them is provided here.
Its most defining properties include its symmetries

$$\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times U(1)_Y \tag{2.7}$$

and its field content, given in the following table.

Field	$SU(3)_C$	$SU(2)_L$	$U(1)_Y$	Mass Dimension
$Q_L = U_L, D_L$	3	2	1/6	3/2
U_R	3	1	2/3	3/2
D_R	3	1	-1/3	3/2
$L_L = l_L, \nu_L$	1	2	-1/2	3/2
$L_R = l_R$	1	1	-1	3/2
$B_{\mu\nu}$	1	1	0	2
$W^I_{\mu\nu}$	1	3	0	2
$G^A_{\mu\nu}$	8	1	0	2
ϕ	1	2	1/2	1

Table 2.1: List of the Standard Model field content and their respective charges.

Here $\{U_i, D_i\}$ are the up- and down-type quarks respectively and the L_i the different leptons with the index denoting their right- or left-handedness. On top of that one has the gauge fields B, W, G describing the force mediating particles and finally the Higgs field denoted as ϕ to give a consistent generation of the particle masses within the model. The SM Lagrangian is constructed using these fields, the gamma matrices $\{\gamma_\mu, \gamma_5\}$ and the covariant derivative, here defined as

$$\mathcal{D}_\mu = \partial_\mu + \frac{ig'}{2} B_\mu + \frac{ig}{2} \sigma_a W^a_\mu + \frac{ig_S}{2} \lambda_\alpha G^\alpha_\mu \,. \tag{2.8}$$

One can, however, not simply combine any of these without restrictions. To make sure that the theory is renormalizable, i.e. only a finite number of counter-terms is required to absorb any divergences from the theory, only operators up to mass dimension 4 can be included in the final Lagrangian.

Conceptionally it is now easy to extend the SM via an EFT approach. To do this, one simply defines the new physics scale Λ and adds higher-dimensional operators to the SM Lagrangian suppressed by corresponding factors of Λ.

$$\mathcal{L}_{SMEFT} = \mathcal{L}_{SM} + \sum_{d=5}^{n} \frac{C_i^{(d)}}{\Lambda^{d-4}} O_i^{(d)} \,. \tag{2.9}$$

Here the different $C_i^{(d)}$ are the Wilson coefficients (WCs) and the $O_i^{(d)}$ are all operators of dimension d one can construct from the field content of the theory while still respecting its fundamental symmetries. Naturally, this tower of operators gives rise to a huge amount of different possible contributions [24] which can't all be reasonably implemented within an analysis. Because of this, one needs to restrict the operators considered to those actually of interest to their specific analysis. Furthermore, the introduction of additional flavor symmetries helps to reduce their number even further. Finally, one restrict the analysis to only include operators u to a certain dimension, often chosen to only include contributions up to dimension six.

While this work is mostly interested in the SMEFT, which reproduces the complete SM at energies far below Λ, different EFTs can be defined depending on the scales one is interested in. Working below the weak scale, for example, reduces the symmetry of the model to a $SU(3) \times U(1)$ containing all SM fermions, apart from the top quark, from which one can define the so-called low-energy effective field theory (LEFT) [25]

$$\mathcal{L}_{LEFT} = \mathcal{L}_{\text{QCD+QED}} + \mathcal{L}_{\not{L}}^{(3)} + \sum_{d\geq 5} L_i^{(d)} \mathcal{O}_i^{(d)} \tag{2.10}$$

with majorana neutrino mass terms $\mathcal{L}_{\not{L}}^{(3)}$ and higher order LEFT operators $\mathcal{O}_i^{(d)}$. Other examples of EFTs are chiral perturbation theory [26] used to describe low-energy QCD mechanics, where the usual perturbative methods break down, based on its observed chiral symmetry for massless fermions, or Higgs Effective Field Theories [27] which allow for the study of a more general Higgs sector by dropping the assumption of a $SU(2)$ Higgs doublet.

Dimension 5 operators

At dimension 5 there is only a single SMEFT operator which can be constructed using these fields [28]. It is sometimes called the Weinberg operator [29] and defined as

$$\mathcal{O}_W = c_W \frac{(\bar{l}_p l_p)(\phi^\dagger \phi)}{\Lambda}. \tag{2.11}$$

This operator is not only interesting from a physics perspective, since after electroweak symmetry breaking it leads to a fermionic mass term for neutrinos, but also at a conceptual level for EFTs. Looking closer at this operator, one sees that it violates the lepton flavor symmetry of the SM, although it was previously established that an EFT should respect the symmetries of its underlying theory. This is where one needs to distinguish between the fundamental symmetries of a theory and accidental ones.

The fundamental symmetries of the SM are the previously mentioned $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times U(1)_Y$, while both lepton and baryon number conservation are only accidental symmetries and do not necessarily need to hold in an EFT in general. One can still decide to impose additional restriction via symmetries of this kind, and will do exactly this, neglecting any dimension 5 contributions in the following.

Dimension 6 operators and beyond

Going from dimension 5 to 6 already increases the amount of possible operators significantly to a total of 2499 different possible operators [24]. To make sure that the total number of operators stays at a reasonable number to perform precise analyses in a sufficiently fast manner, it is necessary to reduce the allowed operators via additional constraints. The simplest way of doing so is the introduction of additional symmetry assumptions or simply the omission of any operator which does not affect the physical observables one is interested in. The choice of these is highly dependent on the analysis considered, which is why they will be discussed more in the following section when the different datasets considered are introduced.

Going even further in the expansion to dimension 7, these operators can once again be neglected under the constraint that the lepton and baryon symmetry should hold even in the SMEFT. The inclusion of dimension 8 operators is where more detailed considerations are required. Due to the very fast-growing number of possible operators, one runs into the same problems one encountered at dimension 6 already. In this analysis, and many others, the SMEFT expansion is therefore truncated at dimension 6. This is motivated by the fact that all higher order contributions are also supressed by higher orders of Λ from which one naively expects only small contributions compared to those coming from dimension 6 operators. This does not necessarily have to actually be the case, however, and there have been several dedicated studies on this [30]. From these studies it was shown that there are some processes where one expects stronger contributions to come from dimension 8 operators as well as certain UV scenarious which lead to large dimension eight terms. None of the measurements included in the datasets used here include such processes.

Although the SMEFT expansion has now been truncated at dimension 6, more consideration needs to be put into at what level this is performed: at the level of the amplitude or the observable. Since one needs to compute squared matrix elements to compute the results at the level of the observable, one receives contributions linear and quadratic in these operators suppressed by, in the case of dimension six, Λ^2 and Λ^4 respectively. This does, however, also lead to additional interference terms between the dimension eight operators and the SM, giving an expression for the cross section of the form

$$\sigma = \sigma_{SM} + \frac{c_6}{\Lambda^2}\sigma_6 + \frac{c_6^2}{\Lambda^4}\sigma_{6\times 6} + \frac{c_8}{\Lambda^4}\sigma_8 + \mathcal{O}(\Lambda^5)\,. \tag{2.12}$$

Here the c_d are the Wilson coefficients of an operator with dimension d, while σ_d and $\sigma_{d\times d}$ describe the corresponding linear and quadratic contributions to the cross section, respectively.

As one can see the linear dimension 8 terms also scale as Λ^{-4}, so one could now ask whether one should truncate this expansion at the level of the Lagrangian and not at the level of the observable to allow for contributions from these dimensions 8 operators interfering with the SM. For SFitter analyses, the SMEFT expansion is truncated at the level of the Lagrangian, removing any contributions from dimension 8 operators and above from the observables.

2.2.1 The Top sector

The data used for the following analysis comes from different kinds of measurements of processes involving the top quark at the LHC. Due to this, it is necessary to understand how such processes come to be. To illustrate this further one can first consider the main production channels of a $t\bar{t}$ pair at the LHC shown in Figure 2.2. On the very left, one has the production of a $t\bar{t}$-pair from the annihilation of two quarks, while the other two show examples for gluon gluon fusion (ggF). Since the LHC operates at very high energies when colliding protons, top-pair production at the LHC is dominated by gluon-gluon fusion, making up approximately 90% of the total production rate [23].

As has already been mentioned, it is necessary to introduce additional restrictions on the SMEFT operators considered, to limit the analysis to a reasonable number of them. The easiest way to do this is via the introduction of additional symmetries. Following Ref. [15], an additional $U(2)_q \times U(2)_u \times U(2)_d$ flavor symmetry among quarks of the first and second generation is imposed. To separate the different generations the following notation is used

$$\begin{aligned} q_i &= \left(u_L^i, d_L^i\right) \qquad u_i = u_R^i, d_i = d_R^i \qquad i = 1, 2\,; \\ Q &= (t_L, b_L) \qquad\quad t = t_R, b = b_R\,. \end{aligned} \tag{2.13}$$

While this additional symmetry is not strictly necessary, it is well motivated by the fact that the different top quark observables considered cannot distinguish between the different light quarks of the same generation.

With this, the total number of operators one can construct already decreases by a large amount. The final set of operators included in the global fit is given in Table 2.2 separated into three different categories which will be discussed in more detail later.

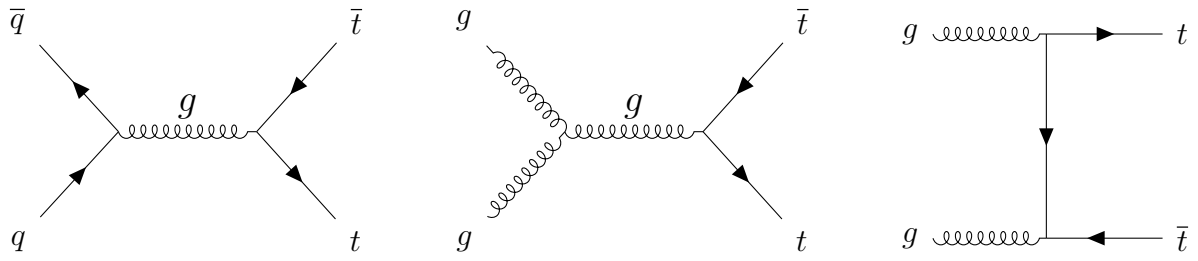


Figure 2.2: Dominant production channels for $t\bar{t}$ at the LHC. Left: quark antiquark annihilation, Middle and Right: gluon gluon fusion (ggF).

Operator	Definition	Operator	Definition
$O_{Qq}^{3,8}$	$(\bar{Q}\gamma_\mu T^A \tau^I Q)(\bar{q}_i \gamma^\mu T^A \tau^I q_i)$	$O_{Qq}^{3,1}$	$(\bar{Q}\gamma_\mu \tau^I Q)(\bar{q}_i \gamma^\mu \tau^I q_i)$
$O_{Qq}^{1,8}$	$(\bar{Q}\gamma_\mu T^A Q)(\bar{q}_i \gamma^\mu T^A q_i)$	$O_{Qq}^{1,1}$	$(\bar{Q}\gamma_\mu Q)(\bar{q}_i \gamma^\mu q_i)$
O_{tu}^{8}	$(\bar{t}\gamma_\mu T^A t)(\bar{u}_i \gamma^\mu T^A u_i)$	O_{tu}^{1}	$(\bar{t}\gamma_\mu t)(\bar{u}_i \gamma^\mu u_i)$
O_{tu}^{8}	$(\bar{t}\gamma_\mu T^A t)(\bar{d}_i \gamma^\mu T^A d_i)$	O_{tu}^{1}	$(\bar{t}\gamma_\mu t)(\bar{d}_i \gamma^\mu d_i)$
O_{Qu}^{8}	$(\bar{Q}\gamma^\mu T^A Q)(\bar{u}_i \gamma_\mu T^A u_i)$	O_{Qu}^{1}	$(\bar{Q}\gamma^\mu Q)(\bar{u}_i \gamma_\mu u_i)$
O_{Qd}^{8}	$(\bar{Q}\gamma^\mu T^A Q)(\bar{d}_i \gamma_\mu T^A d_i)$	O_{Qd}^{1}	$(\bar{Q}\gamma^\mu Q)(\bar{d}_i \gamma_\mu d_i)$
O_{tq}^{8}	$(\bar{q}_i \gamma^\mu T^A q_i)(\bar{t}\gamma_\mu T^A t)$	O_{tq}^{1}	$(\bar{q}_i \gamma^\mu q_i)(\bar{t}\gamma_\mu t)$
$O_{\phi Q}^{3}$	$(\phi^\dagger i \overleftrightarrow{D}_\mu^I \phi)(\bar{Q}\gamma^\mu \tau^I Q)$	$O_{\phi Q}^{1}$	$(\phi^\dagger i \overleftrightarrow{D}_\mu \phi)(\bar{Q}\gamma^\mu Q)$
$O_{\phi tb}$	$(\tilde{\phi}^\dagger i D_\mu \phi)(\bar{t}\gamma^\mu b)$	$O_{\phi t}$	$(\phi^\dagger i \overleftrightarrow{D}_\mu \phi)(\bar{t}\gamma_\mu t)$
O_{bW}	$(\bar{Q}\sigma^{\mu\nu} b)\tau^I \phi W_{\mu\nu}^I$	O_{tW}	$(\bar{Q}\sigma^{\mu\nu} t)\tau^I \tilde{\phi} W_{\mu\nu}^I$
O_{tB}	$(\bar{Q}\sigma^{\mu\nu} t)\tilde{\phi} B_{\mu\nu}$	O_{tG}	$(\bar{Q}\sigma^{\mu\nu} T^A t)\tilde{\phi} G_{\mu\nu}^A$

Table 2.2: List of dimension 6 operators included in the global Top fit, following [15]. Operators contributing to $t\bar{t}$ production are shaded in gray.

To fully define the set of operators included in the fit additional relations between top couplings to gauge bosons were used

$$\begin{aligned} C_{\phi Q}^{-} &= C_{\phi Q}^{1} - C_{\phi Q}^{3} \qquad & C_{tZ} &= -s_w C_{tB} + c_w C_{tW} \\ C_{\phi Q}^{+} &= C_{\phi Q}^{1} + C_{\phi Q}^{3} \qquad & C_{tA} &= c_w C_{tB} + s_w C_{tW}\,, \end{aligned} \tag{2.14}$$

allowing the degrees of freedom (dof) to be chosen freely, which in this case are chosen to be $C_{\phi Q}^{3}, C_{\phi Q}^{-}, C_{tW}$ and C_{tZ}. This gives the additional benefit that one can choose those dof that enter into the physical processes. In this case these would be $t\bar{t}Z$ production for $C_{\phi Q}^{-}$ and C_{tZ}, the tbW vertex for C_{tW} and both of these for $C_{\phi Q}^{3}$.

To fully understand the different operators considered in the final global analysis it is necessary to look at the dataset included in the fit since it only make sense to include operators which can actually be constrained using the considered data. All the operators from Table 2.2 contribute to different processes involving the top quark. They are split into three different categories within the table. The first group can be found in the top row, listing all four fermion operators with a LL or RR chiral structure, while those in the second row are the four fermion operators with either LR or RL chiral structure. The row at the very bottom shows the remaining operators, all of which describe the interactions of Bosons with heavy quarks. In addition to this, the operators contributing to $t\bar{t}$ pair production are shaded in gray since they will be the ones discussed in most detail in the following and will be the most important to understand for the rest of this work.

The effect of the four fermion operators on the production of a $t\bar{t}$ pair can be easily

understood when one considers the Feynman diagram on the left side of Figure 2.2 for quark antiquark production. In analogy to the weak beta decay where one replaced the W-boson as the mediator via an effective four fermion vertex, one can use the exact same logic to remove the gluon from this diagram and replace it with a four fermion interaction. Clearly any of the four fermion operators now affect this vertex and as such can lead to visible effects in $t\bar{t}$ observables. The final operator with significant contributions to top pair production is $\mathcal{O}_{tG}$. The effect of this operator can be easily seen when considering any of the Feynman diagrams in Figure 2.2. It induces corrections any vertex which involves a $t\bar{t}g$ coupling, of which one can find multiple already within the Feynman diagrams shown here.

From these considerations, anyone familiar with SMEFT operators might be asking why no corrections to the triple-gluon vertex are considered. It is possible to construct an additional dimension 6 operator via

$$\mathcal{O}_G = f_{abc} G_\mu^{a\nu} G_\nu^{b\rho} G_\rho^{c\mu} \tag{2.15}$$

which would affect the triple gluon vertex in the Feynman diagram shown in the middle of Figure 2.2. Since top production at the LHC is dominated by ggF one can expect significant contributions to come from any operator affecting the gluon. Previous work on the constraints of this operator in the context of multi-jet production [31,32] has shown that the constraints coming from these are much stronger than those one would get from $t\bar{t}$ production. Following these results and the recommendation from Ref. [15] any contributions from this coefficient to $t\bar{t}$ are set to zero.

While this covers all operators contributing to $t\bar{t}$ production, all other operators need to be constrained by other measurements. For this purpose the dataset considered does not only include $t\bar{t}$ but additional measurements for the production of a $t\bar{t}$ pair in association with a W- or Z-Boson, both s- and t-channel production of a single top as well as additional top decay observables and charge asymmetries. Processes describing the associated production of a W-Boson such as tW are included, but they are not expected to give better constraints than any of the other measurements in the dataset. On the other hand, the associated production of a Z via tZ production is expected to have a sizable sensitivity to $\mathcal{O}_{\phi Q}^-$ and $\mathcal{O}_{tZ}$.

The $t\bar{t}W$ measurements are added to distinguish between different four fermion operators since just $t\bar{t}$ production alone cannot tell the different contributions apart and the additional W-Boson in the initial state only allows for left-handed operators in the initial state to contribute, giving a probe to differentiate between different four fermion operators. For similar reasons, the $t\bar{t}Z$ measurements also allow for stronger discrimination between the four fermion operators while also constraining additional operators such as $\mathcal{O}_{\phi Q}^-, \mathcal{O}_{\phi t}$ and $\mathcal{O}_{tZ}$. Continuing with the single top measurements, they are mostly affected by $\mathcal{O}_{Qq}^{3,1}, \mathcal{O}_{\phi Q}^3, \mathcal{O}_{tW}$ already at leading order and $\mathcal{O}_{Qq}^{3,8}, \mathcal{O}_{\phi tb}$ and $\mathcal{O}_{bW}$ when additional quadratic terms are included.

The top decay observables considered in this dataset are the helicity fractions F_L, F_0 which describe the decay of the Top via $t \to bW$. The two observables distinguish between the

decay into a W-boson with negative (L) or zero (0) helicity, respectively. This gives additional constraining power to $\mathcal{O}_{Qq}^{3,1}, \mathcal{O}_{\phi Q}^{3}$ and especially $\mathcal{O}_{tW}$. They are very sensitive to $\mathcal{O}_{tW}$ and are the reason why the tW measurements are not expected to provide any better constraints. The final measurements of charge asymmetries within top rapidities helps to further distinguish the four fermion operators from one another. The exact list of measurements included in the fit and a more detailed explanation of their impact on the different WCs can once again be found in Ref. [15].

Kinematically enhanced operators

While the previous section showed how one can find which operators contribute to which process, their contributions relative to one another have not been discussed. Taking, once again, top pair production as an example, the different operators considered are the four fermion operators and $\mathcal{O}_{tG}$. Since the three-point vertices affected by $\mathcal{O}_{tG}$ are very different from the four-point vertices of the four fermion operators, they clearly contribute differently to the cross section. The different contributions from these operators and their interferences, as well as their kinematic behavior, has been discussed in much detail in Ref. [15], on which the following brief summary will be based.

The most important properties needed to understand the results of the following analysis is that $\mathcal{O}_{tG}$ mainly contributes to ggF processes, while the sensitivity to four fermion operators comes from quark antiquark annihilation. As has already been mentioned, about 90% of top pair production at the LHC is due to ggF making $t\bar{t}$ much more sensitive to $\mathcal{O}_{tG}$ than the four fermion operators. On the other hand, the contributions from four-fermion operators is kinematically enhanced, i.e. their contribution grows at larger energies making the tails of $t\bar{t}$ sensitive probes of these. In the dataset considered here, both total and differential cross section measurements for top pair production are included. From this one can now see that the total cross sections mainly constrain $\mathcal{O}_{tG}$ while the distributions help constrain the four fermion operators. This is especially true for normalized distributions since in those the contributions from $\mathcal{O}_{tG}$ cancel mostly, while those from the four fermion operators do not due to their kinematic enhancement.

Going global

The previous sections gave a brief summary of the different processes in the Top dataset considered, as well as their expected effect on the different WCs. Something that has not been addressed yet, however, is the reason for performing a global fit to all of these measurements simultaneously. Technically, it is possible to perform fits to only a single WC, for example. Most new physics model lead to more than just a single new operator, however, which is why fits such as those are rarely of practical use. Furthermore, fits to just a single coefficient spoil the basis independence of the fit, since a single operator in one basis can be expressed by a linear combination of multiple operators in another.

One can also use subsets of the full dataset. For example, one can fit only $t\bar{t}$ measurements to constrain C_{tG} and the four fermion operators. These operators are, however, not only constrained by $t\bar{t}$ measurements alone. Both $t\bar{t}Z$ and $t\bar{t}W$ clearly also constrain all the same operators that $t\bar{t}$ does. On top of those, one finds that $t\bar{t}W$ also constrains $\mathcal{O}_{tW}$, but it is known that this one is more constrained by the top decay. One can immediately

see that most of the WCs are constrained not by just a single process.

To get a complete picture of all physics in the Top sector and allow for the interplay of these processes when constraining all the different WCs a proper global fit is required. One of the biggest strengths of global fits is that even the measurements that do not directly affect certain WCs are able to constrain these indirectly through correlations. Furthermore, it can help improve constraints, since certain processes can constrain WCs in different ways than other processes. In Ref. [15] it was shown, for example, that the simultaneous fit of charge-symmetric and charge-assymmetric observables allows one to study constraints on both vector-like and axial-vector-like top couplings, which is able to remove flat directions from constraints on certain four fermion operators with different chiral structures.

The global SMEFT analysis does not need to stop at the level of the Top sector, however. In Ref. [16] the datasets from Higgs, Di-Boson and electroweak measurements have been combined to form a larger global SMEFT fit, for example. One of the measurements included in that Higgs dataset is $t\bar{t}H$ production. Just like for $t\bar{t}Z$ and $t\bar{t}Z$ this process is also able to constrain the same WCs $t\bar{t}$ does. Clearly one expects to find correlations between these measurements, which is why a combination of all of these datasets could prove interesting.

2.2.2 The Higgs, Di-Boson and electroweak sector

For this reason, the final part of this thesis will be concerned with the analysis of a combined global fit of both the Top dataset from Ref. [15] and the dataset from Ref. [16], which consists of Higgs and Di-Boson measurements as well as electroweak precision observables (EWPOs) from the Large Electron-Positron Collider (LEP). With this, the number of different operators considered increases significantly to account for the large number of different observables, with a full list of the additional dimension 6 operators given in Table 2.3. The final additional free parameter that was included in the fits was the branching ratio ($\mathrm{BR_{in}}$) of the Higgs into invisible final states $\mathrm{BR_{inv}}$. This was done to allow for the possible Higgs decay into a dark matter agent.

As one can see this nearly doubles the number of operators considered since only a single operator, $\mathcal{O}_{tG}$, has been considered in the analysis of both of these sectors. Unlike the operators for the Top sector, these are not given in the Warsaw basis, but are based on the HISZ basis [33] with slight modifications. This is most easy to see by comparing the different definitions of $\mathcal{O}_{tG}$ which, while similar, differ slightly. This will be important later, since the proper combination of the two sectors will require both of the sets of operators to be given in the same basis.

In Figure 2.3 the processes for which one would expect SMEFT contributions from both of the different sectors are shown. On the left side of this figure, one can see the Feynman diagram for the production of a Higgs boson, where one clearly see the $t\bar{t}g$ vertices which can be affected by $\mathcal{O}_{tG}$ just like before. Similarly, the production of a $t\bar{t}$ pair in association with a Higgs also includes additional gluon-fermion vertices, providing additional constraints to $\mathcal{O}_{tG}$.

Operator	Definition	Operator	Definition
O_{GG}	$\phi^\dagger\phi G^A_{\mu\nu}G^{A,\mu\nu}$	O_{WW}	$\phi^\dagger\hat{W}_{\mu\nu}\hat{W}^{\mu\nu}\phi$
O_{BB}	$\phi^\dagger\hat{B}_{\mu\nu}\hat{B}^{\mu\nu}\phi$	O_{BW}	$\phi^\dagger\hat{B}_{\mu\nu}\hat{W}^{\mu\nu}\phi$
O_B	$(D_\mu\phi)^\dagger\hat{B}^{\mu\nu}(D_\nu\phi)$	O_W	$(D_\mu\phi)^\dagger\hat{W}^{\mu\nu}(D_\nu\phi)$
O_{3W}	$\mathrm{Tr}(\hat{W}_{\mu\nu}\hat{W}^{\nu\rho}\hat{W}^{\mu}_{\rho})$		
$O_{\phi 1}$	$(D_\mu\phi)^\dagger\phi\phi^\dagger(D_\nu\phi)$	$O_{\phi 2}$	$\frac{1}{2}\partial^\mu(\phi^\dagger\phi)\partial_\mu(\phi^\dagger\phi)$
$O_{e\phi,22}$	$(\phi^\dagger\phi)\bar{l}_2\phi e_2$	$O_{e\phi,33}$	$(\phi^\dagger\phi)\bar{l}_3\phi e_3$
$O_{d\phi,33}$	$(\phi^\dagger\phi)\bar{q}_3\phi d_3$	$O_{u\phi,33}$	$(\phi^\dagger\phi)\bar{q}_3\tilde{\phi}u_3$
O_{4L}	$(\bar{l}_1\gamma_\mu l_2)(\bar{l}_2\gamma^\mu l_1)$	$O^{(1)}_{\phi e}$	$(\phi^\dagger i\overleftrightarrow{D}_\mu\phi)(\bar{e}_i\gamma^\mu e_j)\delta^{ij}$
$O^{(1)}_{\phi d}$	$(\phi^\dagger i\overleftrightarrow{D}_\mu\phi)(\bar{d}_i\gamma^\mu d_j)\delta^{ij}$	$O^{(1)}_{\phi u}$	$(\phi^\dagger i\overleftrightarrow{D}_\mu\phi)(\bar{u}_i\gamma^\mu u_j)\delta^{ij}$
$O^{(1)}_{\phi Q}$	$(\phi^\dagger i\overleftrightarrow{D}_\mu\phi)(\bar{q}_i\gamma^\mu q_j)\delta^{ij}$	$O^{(3)}_{\phi Q}$	$(\phi^\dagger i\overleftrightarrow{D}^I_\mu\phi)(\bar{q}_i\gamma^\mu\tau^I q_j)\delta^{ij}$
O_{tG}	$ig_s(\bar{Q}_3\sigma^{\mu\nu}T^A u_{R,3})\tilde{\phi}G^A_{\mu\nu}$		

Table 2.3: List of dimension 6 operators included in the global Higgs, Di-Boson and EWPO fit, following Ref. [16].

Since $\mathcal{O}_{tG}$ is already very strongly constrained by the Top fit, its effect was implemented into the global fit in Ref. [16] by implementing a prior for $\mathcal{O}_{tG}$ which follows the profile likelihood of $\mathcal{O}_{tG}$ from the full global Top fit from Ref. [15]. Since both datasets will be combined later, this prior can be removed and the cross-talk of $\mathcal{O}_{tG}$ within the Higgs and Top sectors can be examined. How this combination was carried out and what additional assumptions had to be made will be discussed fully in Chapter 5.

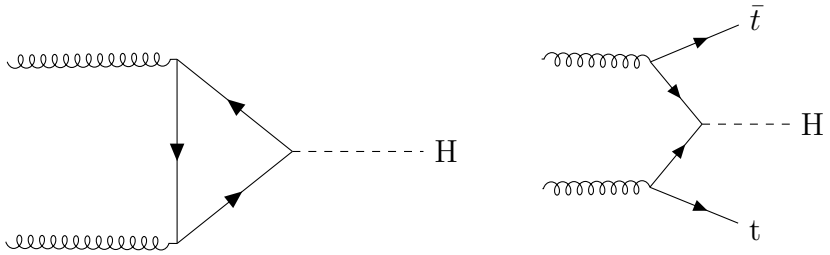


Figure 2.3: Feynman diagrams for processes which can affect both operators in the Top sector and those from the Higgs. Left: Higgs production via gluon-gluon fusion (ggF) Right: $t\bar{t}H$ production.

3

Likelihoods and SFitter

SFitter is a tool developed for statistical analyses, constructing likelihoods with a subsequent comprehensive treatment of its uncertainties using both profiling and marginalization methods. It has been used for a variety of different analyses, such as SMEFT analyses of Higgs and electroweak diboson measurements [34,35] and SMEFT analyses in the top sector [15]. This chapter begins with a quick introduction to the most essential statistical concepts needed to understand the following analyses. After that, the construction of the likelihood within SFitter, its main assumptions and the different frameworks used for either profiling or marginalization, are discussed. The final section of this chapter deals with likelihoods separate from those used in SFitter, shifting the focus to likelihoods published by the ATLAS group at CERN.

3.1 Statistics and Likelihoods

Prior to any discussion of SFitter and its inner workings, one first needs an understanding of the basics of a statistical analysis, especially in the context of high energy physics at the LHC. The following discussion will be mostly based on Ref. [36,37,38] and provide definitions for the concepts and terminology required to fully understand the following interpretation and implementation of experimental data.

The Likelihood

Consider an experiment which provides us with a result x. Repeating this experiment multiple times will give more and more different results, from which one can then construct a function $f(x)$ which assigns a probability to each of the results x one has obtained. Naturally, this needs to fulfill

$$1 = \int \mathrm{d}x \, f(x) \tag{3.1}$$

and be positive for all allowed values of x. This object is what will be referred to as the probability density function (PDF) from now on. Furthermore, one finds that usually this function is dependent on some parameters α describing the theory considered, i.e. $f(x|\alpha)$ read as 'f of x given α'.

In the context of the following analysis, these PDFs can be e.g. the result of an experiment at the LHC which counts signals meeting a certain selection criterion. Each criterion defines what is called a channel at the LHC where one measures an observable x_C. For now only a single channel will be considered and the additional index suppressed. It will be reintroduced later when the generalization to multiple channels is performed. In such an experiment there is naturally not just a single event but a large amount of

different events x_i which one can combine into a large dataset $D = \{x_1, ..., x_N\}$. Just like for the case of a single event, one can then assign a probability to this dataset $P(D|\alpha)$, ignoring for now the exact form this expression takes. With this, an object expressing the probability to obtain a certain dataset D for a given theory described by the previously mentioned model parameters α, has been constructed.

The final goal of this thesis is to look for physics not described by the current theory. Due to this one is interested not in the probability that the dataset agrees with the current theory, but the probability that the theory agrees with experiment for a fixed dataset D. To do so one defines the likelihood, the central object of this analysis

$$\mathcal{L}(\alpha) = \mathcal{L}(\alpha|D) = P(D|\alpha)\,. \tag{3.2}$$

It gives a measure for the agreement of the model parameters α for a fixed dataset D. Here it is important to realize that the likelihood is, in fact, not a probability density. As such it is not normalized to unity, i.e.

$$\int \mathrm{d}\alpha\, \mathcal{L}(\alpha) \neq 1\,. \tag{3.3}$$

Since the likelihood is the foundation of this entire analysis, its proper construction is essential. This is one of the main jobs of SFitter, discussed in detail in Section 3.2, where one finds that there are two well established methods to work with these likelihoods.

At the start of this chapter, PDFs were defined as the result of a large number of experiments. This is considered to be a frequentist way of interpreting PDFs, since these assign probabilities according to the relative frequency of their occurence. Another way to look at these probabilities is the Bayesian approach, derived from Bayes theorem

$$P(\alpha|D) = \frac{\mathcal{L}(\alpha)\pi(\alpha)}{P(D)} \tag{3.4}$$

which relates the probability density $P(\alpha|D)$, called the posterior, to the likelihood $\mathcal{L}(\alpha)$. In this case, however, there is an additional probability $\pi(\alpha)$, known as the prior, incorporating our knowledge of the model parameters α prior to the experiment. The term $P(D)$ is usually called the evidence and ensures that the posterior is normalized. For the purposes of this work it can be neglected, since the normalization does not affect the maxima of a function.

From a Bayesian perspective, these probabilities describe a degree of belief in a certain event, which allows one to assign probabilities to even those events with no consistent frequentist interpretation. Based on these two perspectives on statistics there are different ways of obtaining constraints from the likelihood. Neither of the two is conceptionally wrong, one just needs to make sure to properly interpret the results depending on the framework one chooses. What exactly is meant by these constraints and how SFitter is used to obtain them from the likelihood will be discussed in the following.

3.2 Likelihood construction and Nuisance Parameters

All measurements at the LHC essentially count the number of events in a certain channel. Due to this, each measurement can be interpreted as an individual counting experiment modeled by a Poisson distribution

$$\text{Pois}(d|p) = \text{Pois}(d|p(\alpha_n, \theta_i, b)) \,, \tag{3.5}$$

with measured and expected number of events d and p, respectively. Here p is a function of the different model parameters α_n, the background b and the nuisance parameters (NP) θ_i.

Independent of the different statistical approaches following later, the exclusive likelihoods can be constructed via

$$\mathcal{L}_{excl} = \text{Pois}(d|p(\alpha_n, \theta_i, b))\text{Pois}(b_{CR}|b\,k) \prod_i \mathcal{C}(\theta_i, \sigma_i) \,. \tag{3.6}$$

It is made up of the product of the Poisson contributions coming from the measurements of both data d and background b. The first Poisson contribution gives the likelihood for the measured data d as a function of its expected value $p(\alpha_n, \theta_i, b)$. The second Poisson contributions constrains the background b, taking into account the process of determining the background in a control region b_{CR} via dedicated measurements and its subsequent interpolation into the signal region using an interpolation factor k.

On top of this one needs to add additional constraints on the NPs via $\mathcal{C}(\theta_i, \sigma_i)$. One can easily recognize that from a Bayesian perspective, these are simply the priors for the different NPs in our likelihood. In the frequentist sense, they would represent additional constraints coming from some kind of auxiliary measurement. These constraints $\mathcal{C}(\theta, \sigma)$ are different depending on the type of uncertainty considered. The three different kinds of uncertainties implemented in SFitter are:

1) Systematic uncertainties
The systematic uncertainties considered in SFitter usually consist of uncertainties on values such as e.g. the luminosity or jet related uncertainties coming from the jet energy resolution and scale. These are usually determined via dedicated measurements consisting of a large number of events to allow for an accurate determination of these uncertainties.

As a consequence of this large number of events, one expects these parameters to follow a Gaussian distribution due to the central limit theorem. This is exactly why they are modeled as such using

$$\mathcal{L}(\mu, \sigma|x) = \mathcal{N}(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \,. \tag{3.7}$$

In the end this is still an assumption, however, and they could technically follow a different distribution. The validity of this approximation will be studied in more detail later when discussing likelihoods published by ATLAS in Section 3.3.

2) Statistical uncertainties
These describe the intrinsic statistical uncertainty of a measurement which is naturally described by a simple Poisson distribution,

$$\mathcal{L}(\nu|n) = \text{Pois}(n|\nu) = \frac{\nu^n e^{-\nu}}{n!}, \quad \nu > 0\,. \tag{3.8}$$

although for a very large number of events they can also be described by a Gaussian as a consequence of the central limit theorem.

3) Theory uncertainties
And finally for theory uncertainties one has flat constraints,

$$\mathcal{L}(\mu, \sigma|x) = \mathcal{F}(x|\mu, \sigma) = \frac{1}{2\sigma}\Theta\left[x - (\mu - \sigma)\right]\Theta\left[(\mu + \sigma) - x\right]\,. \tag{3.9}$$

The reasoning behind this can be easily understood by taking the scale uncertainty as an example. These uncertainties are usually determined by calculating the number of expected events at varied scales μ. Using these, one can define an interval of width $2\sigma_{\text{th}}$ in which one expects the true value to lie, although there is no reason for any of the different values to be more likely than any other. On top of this, it is also known that the true value is somewhere close to the central value, and it would be unreasonable for theory errors to allow for any value too far from this central value. Due to this, there needs to be some value at which the PDF for the theory uncertainty drops to 0.

Using the same logic for other theory uncertainties, one finds that a uniform distribution following Equation 3.9 allows exactly for such a behavior. Whether this is a reasonable assumption or not will also be discussed later when studying the differences between marginalization and profiling methods.

Profiling the likelihood

With this, it was shown how the exclusive likelihood is constructed on a surface level. The next thing to do now is to remove the different NPs θ_i since they are not physically interesting and this is where the two approaches start to differ. The traditional way SFitter dealt with these was via profiling methods which simply set the NPs to their best-fit values, i.e. those which maximize the likelihood

$$\mathcal{L}_{\text{prof}}(\alpha) = \max_{\theta} \mathcal{L}_{\text{excl}}(\alpha, \theta)\,. \tag{3.10}$$

Although conceptionally easy to write down, the actual application to the exclusive likelihood defined in Equation 3.6 requires much more involved calculations. The following is a brief summary of the most essential components of the profiled likelihood in SFitter, mostly based on Ref. [39,40] where a more detailed look at the likelihood profiling and its application can be found.

If the exclusive likelihood had been made up of just a single type of distribution the profiling would, while not necessarily trivial, be rather straightforward. Taking for example

the product of two gaussians with respective variances σ_1 and σ_2 one has

$$\mathcal{L}_{\text{prof}} \propto \max_{\theta} e^{-\theta^2/(2\sigma_1^2)} e^{-(d-\theta)^2/(2\sigma_2^2)} \tag{3.11}$$

$$= e^{-d^2/(2\sigma^2)} \tag{3.12}$$

where $\sigma^2 = \sigma_1^2 + \sigma_2^2$. One can see that the result is once again a Gaussian, with the variance as the quadratic sum of the individual ones.
For theory uncertainties one gets

$$\mathcal{L}_{\text{prof}} = \max_{\theta} \Theta[\theta - (p - \sigma_1)]\Theta[(p + \sigma_1) - \theta]\Theta[d - (\theta - \sigma_2)]\Theta[(\theta + \sigma_2) - d] \tag{3.13}$$

$$= \max_{\theta \in [p-\sigma_1, p+\sigma_1]} \Theta[d - (\theta - \sigma_2)]\Theta[(\theta + \sigma_2) - d] \tag{3.14}$$

$$= \Theta[d - p + \sigma_1 + \sigma_2]\Theta[-d + p + \sigma_1 + \sigma_2]\,. \tag{3.15}$$

Once again the final result has the same form as the input, although this time one can see that the uncertainties are actually added linearly as opposed to the quadratic Gaussian errors.

In the previous section it was shown that the exclusive likelihood is a product of a certain number of Poisson, Gaussian and flat distributions. Naturally one can also combine both a flat and Gaussian distribution with the same width σ

$$\mathcal{L}_{\text{prof}} = \max_{\theta} \Theta(p + \sigma_{\text{theo}} - \theta)\Theta(\theta - p - \sigma_{\text{theo}}) e^{-(d-\theta)^2/(2\sigma^2)} \tag{3.16}$$

which, after taking the logarithm, gives

$$\sqrt{-2\,\mathcal{L}_{\text{prof}}} = \begin{cases} (p + \sigma_{\text{theo}} - d)/(\sqrt{2}\sigma_{\text{syst}}) & d < p - \sigma_{\text{theo}} \\ 1 & d \in [p - \sigma_{\text{theo}}, p + \sigma_{\text{theo}}] \\ (p - \sigma_{\text{theo}} - d)/(\sqrt{2}\sigma_{\text{syst}}) & d > p + \sigma_{\text{theo}}\,. \end{cases} \tag{3.17}$$

Unfortunately the combination of a Poisson distribution with a Gaussian one is not trivial and one needs to either compute it numerically or approximate it via

$$\frac{1}{\log \mathcal{L}} = \frac{1}{\log \mathcal{L}_{\text{Poiss}}} + \frac{1}{\log \mathcal{L}_{\text{Gauss}}} \tag{3.18}$$

$$\approx \frac{1}{\log \mathcal{L}_{\text{Poiss},d}} + \frac{1}{\log \mathcal{L}_{\text{Poiss,b}}} + \frac{1}{\log \mathcal{L}_{\text{Gauss}}} \tag{3.19}$$

which once again is a heuristic formula motivated by the Gaussian case where this approximation becomes exact. The Poisson likelihood is given by

$$\log \mathcal{L}_{\text{Poiss}} = \max_{\tilde{b}} \left[(d - \tilde{s} - \tilde{b})\log(\tilde{s} + \tilde{b}) + (b - \tilde{b})\log(\tilde{b}) + \log\left(\frac{(\tilde{s} + \tilde{b})!}{d!} \frac{\tilde{b}!}{b!} \right) \right] \tag{3.20}$$

where $d = s + b$ and s, b describe the number of events measured for signal and background respectively and those variables with a tilde describe the number of the corresponding expected number of events. In the second line the Poisson likelihood is split into sep-

arate contributions from background and data, since computing this expression for the likelihood would be numerically inefficient

$$\log \mathcal{L}_{\text{Poiss,d}} = \left[(d - (\tilde{s} + b)) \log(\tilde{s} + b) + \log \left(\frac{(\tilde{s} + b)!}{d!} \right) \right] , \tag{3.21}$$

$$\log \mathcal{L}_{\text{Poiss,b}} = \left[(b - (d - \tilde{s})) \log(d - \tilde{s}) + \log \left(\frac{(d - \tilde{s})!}{b!} \right) \right] . \tag{3.22}$$

For large amounts of data the Poisson distributions can be approximated as a Gaussian and one finds that Equation 3.18 becomes exact.

One final concept needed to understand the frequentist framework is the concept of a test statistic, which on its most general level is defined as any function $t : T(D) \to \mathbb{R}$ that maps the data to a single real number. One very common choice for the test statistic is based on the profile likelihood, following Ref. [41], the most important being the likelihood ratio,

$$t = \frac{\mathcal{L}(H_1|D)}{\mathcal{L}(H_0|D)} = \frac{\mathcal{L}(\mu, \hat{\hat{\theta}})}{\mathcal{L}(\hat{\mu}, \hat{\theta})} \tag{3.23}$$

where the numerator describes the likelihood with maximized parameters $\hat{\hat{\theta}}$ for the specific given value of μ, while the denominator gives the global maximum likelihood with the most likely parameters $\hat{\theta}$ for the most likely value $\hat{\mu}$.

The choice of this test statistic is motivated by the Neyman-Pearson lemma [42] which states that, in the absence of nuisance parameters, the likelihood ratio is the most powerful test statistic i.e. it has the lowest probability to accept the null hypothesis when the alternate is true. One usually considers the so call negative log likelihood (NLL) defined as

$$t_{NLL} = -2 \log \frac{\mathcal{L}(\mu, \hat{\hat{\theta}})}{\mathcal{L}(\hat{\mu}, \hat{\theta})} \equiv \chi^2 . \tag{3.24}$$

There is a couple of reasons for this. First of all, it helps with the numerical analysis since the products in the likelihood turn into simple sums after taking the logarithm, furthermore it is easier to find the minima of the NLL rather than maximize the likelihood. Finally, according to Wilks theorem [41,43] this distribution follows a χ^2 distribution for large enough statistics, which will make it easier to define a confidence interval later in Section 3.2. For this reason, the NLL will be defined to be a generalized χ^2 from here on, which will be made use of later.

Marginalizing the likelihood

Although the majority of studies using SFitter have been performed via profiling methods, its first analysis using marginalization techniques has been published in Ref. [16]. The following will be a brief introduction to the computation of the marginalized likelihood within SFitter, mostly based on said paper.

Beginning, once again, from the exact same exclusive likelihood one can now eliminate the NPs by marginalizing over them i.e. they are integrated out instead of simply being set

to one specific best-fit value. As a result of this, one now needs to compute convolutions of different products of distributions. In the case of two Gaussian one finds that, just like in the profiling case, their convolution once again results in another Gaussian with its variance as the quadratic sum of the individual variances. In this case the profiling and marginalization agree, but this will change significantly when non-Gaussian distributions are considered.

To allow for a direct comparison, one can look at the exact same combinations as in the profiling case. Considering the product of two uniform distributions, one finds

$$\begin{aligned}
\mathcal{L}_{\text{marg}} &= \int \mathrm{d}\theta \, \frac{1}{4\sigma_1\sigma_2} \Theta\left[\theta - (p-\sigma_1)\right] \Theta\left[(p+\sigma_1) - \theta\right] \Theta\left[d - (\theta-\sigma_2)\right] \Theta\left[(\theta+\sigma_2) - d\right] \\
&= \int_{p-\sigma_1}^{p+\sigma_1} \mathrm{d}\theta \, \frac{1}{2\sigma_2} \Theta\left[(d-\theta) + \sigma_2\right] \Theta\left[\sigma_2 - (d-\theta)\right] \\
&= \frac{1}{4\sigma_1\sigma_2} \begin{cases} d - p + \sigma_1 + \sigma_2 & \text{for } p - \sigma_1 - \sigma_2 < d < p - \sigma_1 + \sigma_2 \\ 2\sigma_2 & \text{for } p - \sigma_1 + \sigma_2 < d < p + \sigma_1 - \sigma_2 \\ -d + p + \sigma_1 + \sigma_2 & \text{for } p + \sigma_1 - \sigma_2 < d < p + \sigma_1 + \sigma_2 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

in case that $\sigma_1 > \sigma_2$, although the expression for the opposite can be derived by just exchanging σ_1 with σ_2. Once again one finds that the uncertainties add linearly for theory uncertainties just like in the profiled case.

Continuing with the combination of multiple distributions, one finds for the convolution of a uniform and Gaussian

$$\begin{aligned}
\mathcal{L}_{\text{marg}} &= \frac{1}{2\sigma_{\text{theo}}} \int \mathrm{d}\theta \exp\left[-\frac{(d-\theta)^2}{2\sigma_{\text{syst}}^2}\right] \Theta[\theta - (p-\sigma_{\text{theo}})]\Theta[(p+\sigma_{\text{theo}}) - \theta] && (3.25) \\
&= \frac{1}{4\sigma_{\text{theo}}} \left[\text{erf}\left(\frac{p+\sigma_{\text{theo}}-d}{\sqrt{2}\sigma_{\text{syst}}}\right) - \text{erf}\left(\frac{p-\sigma_{\text{theo}}-d}{\sqrt{2}\sigma_{\text{syst}}}\right)\right] && (3.26)
\end{aligned}$$

where one simply adjusts the borders of the integral according to the uniform distribution and makes use of the following expression for the error function

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{d}t \, e^{-t^2} \,. \tag{3.27}$$

Comparing this to Equation 3.16 one can see that the expression look very different, making a direct comparison difficult. Because of this, the results for both the profiling and marginalization treatment are shown in Figure 3.1, taken directly from Ref. [16].

Looking at the left plot of Figure 3.1 one can see the behavior of the likelihood after either marginalization (blue) or profiling (green). Since the profiled results follow the so-called RFit prescription [44] they are labeled as such here. One finds that they are both maximal and flat in the ranges of the flat uncertainties and drop at larger values for the data. The main difference between the two methods is how rapidly this happens. In the profiling case, this is simply described by a Gaussian for values larger than the

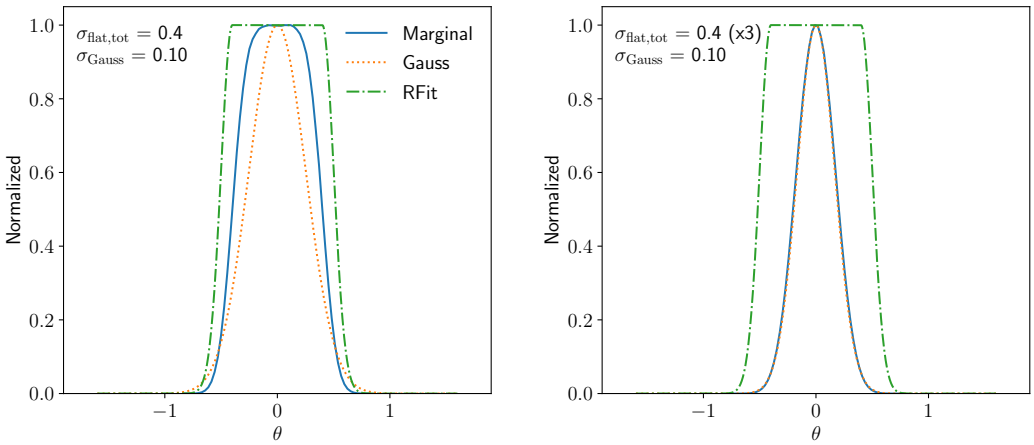


Figure 3.1: Profiled and marginalized likelihoods from the convolution of a Gaussian with either one (left) or three (right) theory uncertainties, taken from Ref. [16].

theory uncertainties, while the marginalized results drop to 0 much more rapidly due to the behavior of the error function. This result can be generalized to more than a single flat uncertainty, with the corresponding expression given in the Appendix. The behavior of this more general expression has already been studied in Ref. [16] with the most important results shown in Figure 3.1.

The right plot now shows the effect that a larger number of theory uncertainties has on the final distribution after performing the convolution. For the easy visualization of the differences the Gaussian one would obtain from adding the half-widths in quadrature, has also been plotted. In the left plot, one can still clearly see differences between the Gaussian and the final marginalized result, while the plot on the right shows that both the Gaussian and marginalized distribution agree very well. This can be understood as a consequence of the central limit theorem, which ensures that the distribution converges to a Gaussian with each further convolution. From this, one can already see clearly that a small number of theory uncertainties already suffices to get very Gaussian results.

The most general form of the likelihood after integrating out the background can be written as

$$\begin{aligned}\mathcal{L}_{\text{marg}} = &\int \mathrm{d}\theta_{tot}\, \mathcal{P}(d|s,\theta_{tot},b,k) \\ &\times \int \prod_j \mathrm{d}\theta_{theo,j}\, \mathcal{F}_{0,\sigma_j}(\theta_{theo,j}) \int \prod_{i \mathrm{d}\theta_{syst,i}} \mathcal{N}(\theta_{syst,i}|0,\sigma_i)\,,\end{aligned} \tag{3.28}$$

where the function $\mathcal{P}(d|s+\theta_{tot}, b_{SR}, k)$ in the first line is described by the Poisson-Gamma model, which results from the convolution of the two Poisson contributions from Equation 3.6. The analytic expression for this function can be found in the Appendix. To perform the marginalization on the second line of Equation 3.28 one makes use of the results for the convolutions of Gaussian and flat distributions from the previous section. The differences between marginalization and profiling will be studied further in Chapter 5, when the results from the global fit are discussed.

Generalization to multiple measurements

Ultimately one wants the likelihood to be generalized to multiple channels which can be easily achieved by multiplying the corresponding likelihood for each channel

$$\mathcal{L}_{\text{excl,full}} = \prod_c \text{Pois}(d_c|p_c)\text{Pois}(b_{CR_c}|b_c\, k_c) \prod_i \mathcal{C}(\theta_{i,c}, \sigma_{i,c})\,. \tag{3.29}$$

This does require the channels to be statistically independent which is usually the case and as such one does not need to worry too much about this, although one should still always keep that in mind. Another point one now needs to take into account is that since more than just a single channel is considered, it is also necessary to decide on the treatment of correlations between systematic and theory uncertainties of these measurements.

As has already been discussed thoroughly, there are systematic, statistical and theory uncertainties in SFitter which do not only differ in their respective choice of distribution, but also in the way they are correlated. In the case of systematic uncertainties, it is easy to see that one should expect the same type of uncertainty to be correlated between measurements, since constraints of the NPs used to describe them are usually based on the same secondary measurements. Since there are no values for these kinds of correlations provided by the experiments though, an assumption has to be made by picking a certain value for these. Within SFitter this leads to fully correlated systematic uncertainties between measurements of the same experiment, encoded in the covariance matrices. Here, it is important to note that the luminosity is correlated not just between measurements but also experiments, since it is the same for all experiments at the LHC.

Because of these correlations between the different systematics, the Gaussian constraints are replaced with the product of one Gaussian for each channel with the corresponding covariance matrix encoding the correlations between these uncertainties. Due to the marginalization, the expression for the systematics simplifies to a single N-dimensional Gaussian with the total covariance matrix as the sum of the individual covariance matrices, i.e.

$$\mathcal{N}(\theta_{syst,i}|0,\sigma_i) \longrightarrow \mathcal{N}(\vec{\theta}_{syst,i}|\vec{0},\Sigma_i)\,.$$

Due to the statistical independence of the different measurements the statistical uncertainties are assumed to be uncorrelated while the same usually applies to theory uncertainties as well. The constraints on theory and statistical uncertainties can easily be generalized by taking the product over each of the different channels by using

$$\begin{aligned}
\text{Pois}(d|p)\text{Pois}(b_{CR}|b\,k) &\longrightarrow \prod_c \text{Pois}(d_c|p_c)\text{Pois}(b_{CR_c}|b_c\,k_c)\,,\\
\mathcal{F}(\theta_{theo,j}|0,\sigma_j) &\longrightarrow \prod_c \mathcal{F}(\theta_{theo,cj}|\vec{0},\sigma_{cj})\,.
\end{aligned}$$

In the case of theory uncertainties, some consideration on implementing correlations have been made, which will be discussed further in Section 4.1.

Concerning correlations

The next step is to incorporate these correlations into the previously constructed likelihood. In the following, two different methods to do this will be discussed. The first one is applicable to both the profiled and marginalized likelihood, while the second one can only be applied to the latter.

For the more general treatment, SFitter makes use of the fact that any such correlation can be written in terms of the correlation matrix with 1 in its diagonal entries and

$$C_{ij} = \frac{\sum_{\text{syst}} \sigma_{i,syst}\sigma_{j,syst} \cdot \rho_{i,j,syst}}{\sigma_{i,exp}\sigma_{j,exp}} \qquad \text{with} \qquad \sigma^2_{i,exp} = \sum_{\text{syst}} \sigma^2_{i,syst} + \sum_{\text{pois}} \sigma^2_{i,pois}\,. \tag{3.30}$$

in the off-diagonals for channels i, j. This correlation matrix encodes the correlations of different uncertainties within SFitter. This way all systematic uncertainties in the same groups given in Table 8.2 can be correlated by picking an arbitrary value for $\rho_{i,j}$. In SFitter the assumption is made that all these systematics are fully correlated between measurements in the same experiment, to be more specific $\rho_{i,j} = 0.99$ is chosen, to ensure the invertibility of the final correlation matrix.

To construct the final likelihood from this, including both uncertainties and correlations, one uses the definition of the generalized likelihood from Equation 3.24 from which one can compute the χ^2 including all correlations via

$$\chi^2 = \vec{\chi}^T C^{-1} \vec{\chi}\,. \tag{3.31}$$

Laplace method

While the previous treatment of correlations can be applied in the exact same manner for the profiled and marginalized likelihoods, the following describes a different treatment making use of the marginalization. Denoting any one of the analytic expressions derived in the previous section as $\mathcal{B}$, the result after integrating out all NPs but those describing the systematic uncertainties is

$$\mathcal{L}_{marg} = \int \prod_i \mathrm{d}^{\mathrm{N}}\theta_{syst,i} \prod_c \mathcal{B}(d_c|b_{CR,c}, s_c + \sum_i \theta_{syst,ci}) \prod_i \mathcal{N}(\vec{\theta}_{\text{syst,i}}|\vec{0}, \Sigma_{\vec{\theta}_{\text{syst,i}}}) \tag{3.32}$$

$$= \int \mathrm{d}^{\mathrm{N}}\theta_{\text{syst,tot}} \prod_c \mathcal{B}(d_c|b_{CR,c}, s_c + \theta_{\text{syst,tot},c}) \mathcal{N}(\vec{\theta}_{\text{syst,tot}}|\vec{0}, \Sigma_{\vec{\theta}_{\text{syst,tot}}}) \tag{3.33}$$

where in the second line one makes use of the fact that the convolution of a multivariate Gaussian is once again another multivariate with $\theta_{syst,tot} = \sum_i \theta_{syst,i}$ and $\Sigma_{tot} = \sum_i \Sigma_i$. The multivariate Gaussian is parameterized by the covariance matrix Σ_i with off-diagonal entries describing the correlations between the different uncertainties.

Since Equation 3.32 describes very high dimensional integrals, it requires specialized techniques to efficiently compute these numerically. In the following, one makes use of

the Laplace method. To do this consider the following integral

$$\int \mathrm{d^n}x e^{\log f(x)} \approx \int \mathrm{d^n}x \exp\left(\log f(x_0) + B(x-x_0) - \frac{1}{2}F_{ij}(x_0)(x-x_0)_i(x-x_0)_j\right) \quad (3.34)$$

where

$$B = \frac{\partial}{\partial x}\log f(x_0) \qquad \text{and} \qquad F_{ij}(x_0) = -\frac{\partial^2}{\partial x_i \partial x_j}\log(x_0)\,. \quad (3.35)$$

Here the linear term is kept for now, although one would expect it to disappear at the maximum x_0, for reasons discussed later.

This is a Gaussian integral and one can simply write the solution as

$$\int \mathrm{d^n}x\; f(x) \approx f(x_0)\exp\left(\frac{1}{2}B^T(x_0)F^{-1}(x_0)B(x_0)\right)\sqrt{\frac{(2\pi)^n}{\det\, F(x_0)}}\,. \quad (3.36)$$

Applying this to Equation 3.32 one gets

$$\mathcal{L}_{\mathrm{marg}} = \prod_c \mathcal{B}(d_c|B_{CR,c}, s_c + \theta_{\mathrm{max},c})\mathcal{N}(\vec{\theta}_{\mathrm{max}}|\vec{0}, \Sigma_{\mathrm{syst,tot}})\exp\left(\frac{1}{2}B^T F^{-1} B\right)\sqrt{\frac{(2\pi)^n}{\det\, F(\vec{\theta}_{\mathrm{max}})}}\,.$$

It now remains to compute

$$\begin{aligned}\vec{\theta}_{\mathrm{max}} &= \max_\theta \prod_c \mathcal{B}(d_c|B_{CR,c}, s_c + \theta_c)\mathcal{N}(\vec{\theta}|\vec{0}, \Sigma_{\mathrm{max}}) \\ &\approx \left(\mathbb{1} + \Sigma^{-1}_{\mathrm{syst,tot}}\Sigma_{\mathrm{exp}}\right)^{-1}\left(\vec{d} - \vec{b}_{SR} - \vec{s}\right)\,.\end{aligned} \quad (3.37)$$

where $\Sigma_{\mathrm{exp}} = \mathrm{diag}(\sigma_{\mathrm{exp}})$ with σ_{exp} from Equation 3.30. Here, only an approximate result is given, although it becomes exact in the Gaussian case.

This is the reason why it was not possible to neglect the linear contribution in Equation 3.34 since this would require the position of the maximum to be known exactly. Using Equation 3.37 one can compute $B(\vec{\theta}_{\mathrm{max}})$ and $F(\vec{\theta}_{\mathrm{max}})$ using Equation 3.35. With this, the full correlations can be included in the likelihood using the Laplace method.

Determining the limits

One can now consider the case in which, after taking care of the NPs and performing a fit, one has a higher-dimensional likelihood which now only depends on the WCs $C_1, ..., C_x$. To determine the frequentist confidence intervals or bayesian credible intervals, it is now first necessary to reduce this likelihood to be just a function of the single WC one is currently interested in.

This is where once again the different approaches used previously lead to different ways to determine the constraints from the likelihood. Beginning, once again, with the method used in most of the previous SFitter analysis using profiling methods one simply sets the

values of the WC one is currently not interested in to those which maximize the likelihood

$$\mathcal{L}(C) = \max_{C_x} \mathcal{L}_{excl}(C, C_x)\,. \tag{3.38}$$

Since the proper construction of confidence intervals from these is far from trivial, different methods to do so exist, such as the one proposed by Feldman and Cousins [45]. This is, however, very complicated and inefficient, which is why one usually makes use of the previously mentioned property that for a large enough set of data the NLL from Equation 3.24 follows a χ^2 distribution.

Now the confidence intervals can easily be determined since the method to extract different confidence intervals from such a distribution is very well known and one simply needs to find its intersections with specific values of χ^2 depending on the number of degrees of freedom. For a single dof one finds that the 95% confidence interval can be found at $\chi^2 = 3.841$. This means that for this value of χ^2, the p-value, i.e. the probability for the value of the parameter to lie outside this range, reaches $p = 0.05$. Similarly, one finds the corresponding 68% intervals for $\chi^2 = 0.989$.

A different approach can be used when the marginalization treatment is applied. Just like for the nuisance parameters one can now integrate all remaining WCs out, i.e.

$$\mathcal{L}(C) = \int \mathrm{d}C_x\, \mathcal{L}_{excl}(C, C_x)\,. \tag{3.39}$$

From the marginalized likelihood the posterior can easily be computed using Bayes theorem. The definition of the different credible intervals is then rather straightforward. One simply needs to find the point of maximum posterior and integrates around it until the value of the integral corresponds to 68% or 95% of the total integral to get the σ and 2σ intervals, respectively.

Markov Chain Monte Carlo

From the previous section, one has now defined a procedure to extract constraints from either the profiled or marginalized likelihoods. As such, it is easy to see that for the determination of precise constraints on the values of the WCs, a precise description of the likelihood is required. In Section 2.2.1 the number of different operators one tries to constrain with a single fit was listed. From this one can see that the likelihoods considered usually have around 20 dimensions and one needs to find a sufficiently efficient and accurate method to map these.

The high dimensionality of the likelihood makes regular Monte Carlo methods not suited to our needs. Due to this, we make use of Markov Chain Monte Carlo methods which are based on the concept of a Markov process. The essential property defining such a process is that for all events X_t,

$$f(X_t = i | X_1 = x_1, X_2 = x_2, ..., X_{t-1} = x_{t-1}) = f(X_t = i | X_{t-1} = x_{t-1})\,, \tag{3.40}$$

i.e. every subsequent event is only dependent on the previous one. To make the connection

to this analysis, one can consider $\mathcal{L}(d|\alpha_i)$, which describes an initial point of the likelihood. One now proposes a new point α_{i+1} at which one can then once again compute the likelihood $\mathcal{L}(d|\alpha_{i+1})$. This point is chosen via a proposal function $g(\alpha_{i+1}|\alpha_i)$, which naturally needs to satisfy the Markov property from Equation 3.40. The Metropolis algorithm states that for a symmetric proposal g the new point α_{i+1} is accepted with probability [46]

$$p = \min\left\{1, \frac{\mathcal{L}(d|\alpha_{i+1})}{\mathcal{L}(d|\alpha_i)}\right\} . \tag{3.41}$$

If the point is accepted, one repeats this process using a new proposal now centered around the newly accepted point α_{i+1} replacing every instance of α_i with it. Otherwise, the point is rejected and one repeats the process with a new proposed point around α_i.

One can see that the Metropolis algorithm is highly dependent on the choice of the proposal distribution and one can show that so is the speed of convergence. It can, however, be shown that independent of the choice of proposal, this procedure converges to the same target distribution. Due to this, a suitable choice of the proposal function, while not necessarily required for proper convergence, is the deciding factor when it comes to speed and efficiency of the algorithm.

One can usually see whether the chosen proposal is a good one by checking the final acceptance rate for different proposals. If one finds that the acceptance rate is very high, i.e. almost all points are accepted, this simply means that almost every newly proposed point has a similar likelihood value to the previous. Considering the nature of the algorithm, this suggests that every newly proposed point is too close to the initial one. On the other hand, very low acceptance rates also do not lead to great results. This is easy to see since once the majority of new points are rejected one hardly moves from the current point or, in the worst case, gets stuck completely. Both of these issues can usually be fixed by varying the variance of the given proposal distribution. In the following analyses a Gaussian is used as the proposal function, although a choice of a Breit-Wigner is also possible, where the variance is adjusted such that the final acceptance rate in the fits is around 30% to 50%.

A different issue arises from the fact that the point at which one of these Markov chains is initialized can technically be arbitrarily far from the actual likelihood region one is interested in and it can take many steps for the chain to reach the region one actually wants to map. Since these points are very likely to describe the likelihood badly, the usual approach is to start a Markov chain, give it a certain number of steps to initialize and discards these points before continuing on with the rest of the chain. Naturally, one does not want this to be too long since it would lead to an unnecessarily longer fit. In SFitter this 'burn-in' phase makes up 10% of the total length of the final Markov chain after the points are discarded.

3.3 Published likelihoods

After the previous section provided a thorough explanation of likelihoods within the context of SFitter, the following will deal with likelihoods from a different perspective. For

years, the publication of full likelihoods as the main object for a proper interpretation of experimental results has been requested [47,48], since they provide the full mathematical description of the analysis. This led to further efforts in this regard, with the ATLAS collaboration providing the likelihoods for a non-insignificant number of their analyses.

The following section will be dedicated entirely to the analysis of two of these likelihoods, published by the ATLAS Top working group [17,18]. First, the HistFactory format they are published in will be discussed, following Ref. [49,50], after which the general tools to extract the most vital information from them are introduced. The final implementation of the corresponding measurements into the SFitter framework will be discussed further in Chapter 5.

The HistFactory format

To allow for a unified approach to the analysis of these likelihoods, it is necessary to have all of them be constructed based on the same basic template. The HistFactory format is exactly one such template which constructs, in its most general form, PDFs of the following form

$$\mathcal{L}(n_{cb}, a_\chi|\eta,\chi) = \prod_{c\in\text{channels}} \prod_{b\in\text{bins}} \text{Pois}(n_{cb}|\nu_{cb}(\eta,\chi)) \prod_{\chi\in\vec{\chi}} \mathcal{C}_\chi(a_\chi|\chi)\,. \tag{3.42}$$

Once again, similar to the likelihood within SFitter, one finds that the likelihood is simply the product of Poisson distributions where n_{cb} and ν_{cb} are the measured and expected number of events in bin b of channel c, respectively. Then there are the unconstrained parameters of interest (POI) which are denoted as η and additional parameters χ constrained by the terms $\mathcal{C}_\chi(a_\chi|\chi)$ where a_χ describes the auxiliary data constraining these.

Initially these likelihoods have only been implemented into the ROOT framework using input files in the XML format to construct them. The following, however, will be based on a purely python based implementation of these likelihoods via a module called pyhf [51,52]. It allows for easy construction of the likelihood based on a simple input file in the ubiquitous JSON format, giving easily readable models for both human and machine while also easily parsable using any programming language of choice due to its use in a large range of different fields.

To construct the likelihood from the input given in the JSON file one needs to calculate the event rates via

$$\nu_{cb} = \sum_{s\in\text{samples}} \left(\prod_{\kappa\in\vec{\kappa}} \kappa_{scb}\right) \left(\nu^0_{scb} + \sum_{\Delta\in\vec{\Delta}} \Delta_{scb}\right) \tag{3.43}$$

where one has the nominal event rate ν_{cb} and the different multiplicative $\vec{\kappa}$ and additive modifiers $\vec{\Delta}$ considered in the likelihood. These modifiers correspond to the previously introduced parameters χ and are used to describe the effect of uncertainties affecting the measured event rate ν_{cb}. The type of modifier and the constraints on its corresponding NP depends on the type of uncertainty considered, with the most common ones given

Description	Modification	Constraint Term c
Luminosity ('lumi')	$\kappa_{scb} = \lambda$	$\mathcal{N}(l = \lambda_0 \vert \lambda, \sigma_\lambda)$
Normalization unc. ('normsys')	$\kappa_{scb} = g_p(\alpha \vert \kappa_{scb,\alpha=\pm1})$	$\mathcal{N}(a = 0 \vert \alpha, \sigma = 1)$
Correlated Shape ('histosys')	$\Delta_{scb} = f_p(\alpha \vert \Delta_{scb,\alpha=\pm1})$	$\mathcal{N}(a = 0 \vert \alpha, \sigma = 1)$
MC Stat. ('staterror')	$\kappa_{scb} = \gamma_b$	$\prod_b \mathcal{N}(a_{\gamma_b} = 1 \vert \gamma_b, \delta_b)$
Uncorrelated Shape ('shapesys')	$\kappa_{scb} = \gamma_b$	$\prod_b \mathrm{Pois}(\sigma_b^{-2} \vert \sigma_b^{-2} \gamma_b)$
Normalization ('normfactor')	$\kappa_{scb} = \mu_b$	

Table 3.1: List of different modifiers considered in the construction of the HistFactory likelihoods, adapted with slight modifications from Ref. [50]

in Table 3.1. Here the index c denotes the different channels, describing disjoint binned distributions, s the samples, describing the physics process, and b the different bins.

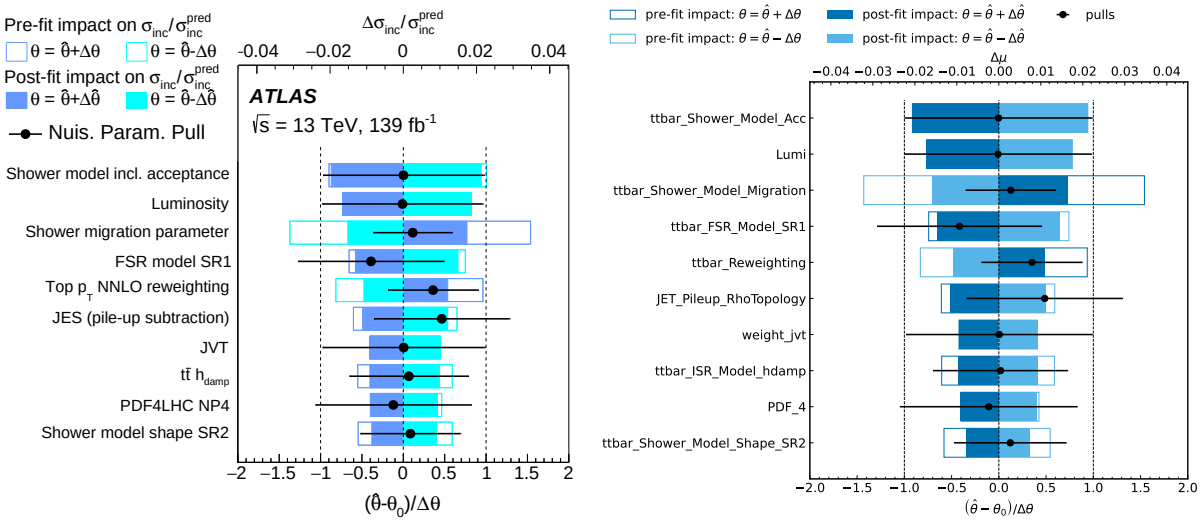
The most common such NP is the uncertainty on the luminosity, since it affects any measurement taken at the LHC and leads to a simple rescaling of the number of events, as can be seen in the table. Much more interesting are the NPs describing normalization uncertainties and correlated shape uncertainties, since they are the most common types found in the likelihoods and the implementation of their contributions is not nearly as trivial. Both of them are constrained by a Gaussian which can, once again, be understood as the results of a secondary measurement, which, with high enough statistics, approach a Gaussian. The actual modification of these modifiers is denoted here as $\kappa_{scb} = g_p(\alpha \vert \kappa_{scb,\alpha=\pm1})$ and $\Delta_{scb} = f_p(\alpha \vert \Delta_{scb,\alpha=\pm1})$ for normalization uncertainties and correlated shape uncertainties, respectively. Taking the normalization uncertainties as an example, for now, it is necessary to know that the only values given in the JSON file to fully describe the likelihood are $\kappa_{scb,\alpha=\pm1}$. This means that the effect of these NPs is only given for two specific values $\alpha = \pm1$, providing $g_p(\alpha \vert \kappa_{scb,\alpha=\pm1})$ in the JSON file and $g_p(\alpha = 0) = 1$ i.e. one recovers the nominal prediction at $\alpha = 0$.

The function $g_p(\alpha)$ describes the different interpolation strategies which allow one to calculate the values of κ_{scb} for any value of α. There are multiple choices for this interpolation, discussed in much detail in Ref. [49], using a piecewise exponential interpolation for normalization uncertainties or piecewise linear interpolation in the case of correlated shape uncertainties i.e. f_p. All the other modifiers are simple multiplicative corrections for each bin, where, once again for a more detailed discussion we refer to Ref. [49].

Using the values for these modifiers and the nominal rates ν_{cb} provided in the published likelihoods, one can then perform fits to experimental data. For this, a fit, using the profiling methods introduced in the previous section, can be performed to obtain the best fit value and uncertainties of the POI. In the case of these published likelihoods these correspond to the signal strength of the considered process which can be used to compute the cross section of this process.

Analyzing published likelihoods

Now that the general structure of these likelihoods has been established, the next step is to perform the actual analysis. As previously mentioned, this analysis is mostly con-



(a) Figure taken from Ref. [17]. (b) Values reproduced using pyhf.

Figure 3.2: Comparison of pulls and impact of individual NPs on the $t\bar{t}$ signal strength.

cerned with two different likelihoods. The first one describes the measurements of the total cross section for $t\bar{t}$ production at 13 TeV in the leptons + jets channel [17], while the second one describes the measurement of the cross section of $t\bar{t}$ production with an associated Z-Boson [18].

The first thing one can now do is check the validity of the provided likelihoods by reproducing some of the final results given directly in the papers. The most common results shown which one can reproduce are the so-called pull plots which show the best-fit points of each NP, i.e. the value at which they maximize the likelihood, and their pulls, which are defined as

$$\text{pull} = \frac{\hat{\theta} - \theta_0}{\Delta\theta}\,. \tag{3.44}$$

Here, the hatted variables describe the previously mentioned best-fit values and θ_0 the respective values before the fit. Finally, they are normalized to their respective pre-fit uncertainties $\Delta\theta$. Figure 3.2 shows the results for the $t\bar{t}$ likelihood while those for $t\bar{t}Z$ are given in the Appendix, since both practically work the exact same way. In addition to the pulls, it also displays the impact of different NPs on the POI, which in this case is the $t\bar{t}$ cross section. It can be determined by fixing the considered NP to its best-fit value $\hat{\theta}$, shifted by its prefit (postfit) uncertainties $\pm\theta(\pm\hat{\theta})$. They are ordered with decreasing post-fit impact and only the 10 most significant NPs are shown here, since both likelihoods have NPs numbering around 100 to 200.

As one can see, both results agree very well for both the pulls and their impact on the total cross section with only some barely visible insignificant differences. The plots were created by performing a profile likelihood fit via pyhf using the minuit optimization algorithm [53].

The visualization was done using code based on another python package called cabi-

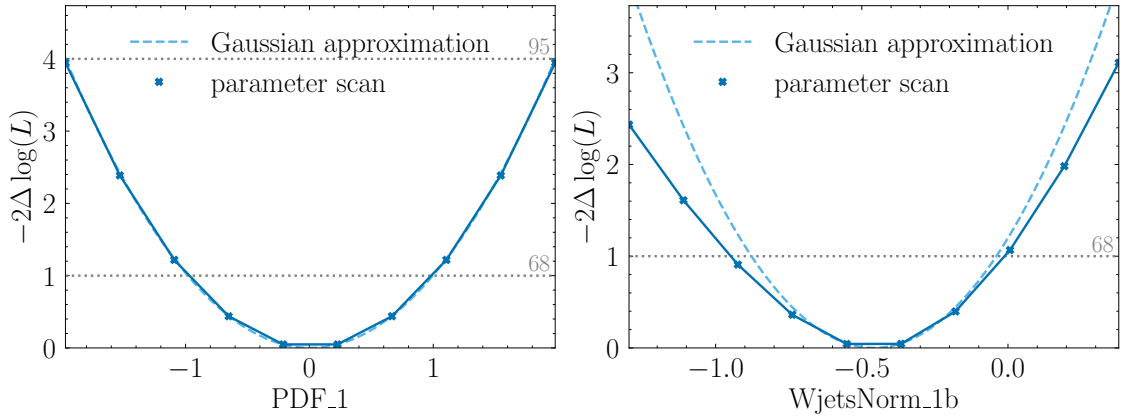


Figure 3.3: Profiled likelihoods for different individual NPs. Shown are examples for of a NP well described by a Gaussian (left) and a NP behaving unlike a Gaussian (right).

netry [54], another python package making use of the HistFactory likelihoods by interfacing with pyhf. Beside the visualization of these pull plots, cabinetry also allows for the easy analysis of the likelihood when varying a NP, allowing for an easy check of the shape of the likelihood for only considering this one NP. Once again, there are more than a hundred different NPs one could consider, which is why the NPs in Figure 3.3 only show the two most extreme cases for illustrative purposes.

On the left of Figure 3.3 one can see the behavior of the likelihood when varying one of the NPs used to determine the uncertainties from parton density functions[1] while the plot on the right shows the behavior for one of the NPs describing the W-jet normalization. One can see the very Gaussian behavior in the plot on the left, while the right plot clearly shows deviations from the Gaussian approximation, especially for negative values of the NP. To see why any of this is of interest to us, one needs to go back to one of the basic assumptions SFitter makes when constructing a likelihood. In Equation 3.7 the prior assumption was made that all systematic uncertainties in the likelihood can be described using Gaussian constraints[2]. Important to note now is that almost all the different NPs included in the likelihood exhibit a Gaussian behavior similar to the NP on the left side of Figure 3.3 while only a handful are actually non-Gaussian at all, and even if they are, the differences are very minor. These very small non-gaussianities of the individual NPs already shows that the Gaussian approximation used for systematic uncertainties in SFitter is well justified.

The final aspect of the likelihood discussed here is correlations. As has already been made clear in the previous sections, the comprehensive analysis using correlations is one of the key aspects of the SFitter likelihood. One of the advantages of using minuit for the optimization is that it does not only provide the bestfit values and uncertainties but also correlations between the different NPs. Since there are too many different NPs to display the entire correlation matrix reasonably, Figure 3.4 shows only those NPs where

[1] Often also abbreviated as 'PDF', this will not be done in this text, however, to prevent confusion with probability densities.

[2] Or priors in the Bayesian case, if you will.

at least one correlation is greater than 0.4. One finds that out of the total 178 nuisance parameters, only 12 show correlations stronger than 0.4, most of which are either uncertainties coming from the different modeling choices or are related to jets. If one were to decrease the minimal correlation in Figure 3.4 from 0.4 to 0.3 one finds that the number of NPs with correlations at least as strong as those more than double from 12 to 28. Interesting is that even now, the majority of highly correlated NPs are still almost only model and jet related. The only other strong correlations one finds apart from these is an anti-correlation between the measured $t\bar{t}$ cross-section and the luminosity of about -0.35.

There is, however, the caveat that minuit uses the inverse of the Hessian matrix to determine the uncertainties of the different NPs and can only return symmetric uncertainties. If one wants more precise and possibly asymmetric uncertainties one needs to use the MINOS algorithm since it also takes non linearities into account. How exactly these likelihoods are used in the final fit and how the uncertainties are extracted from them will be discussed in much greater detail in Chapter 5, when both measurements are added to the dataset and their effect on the constraints of different WCs are studied.

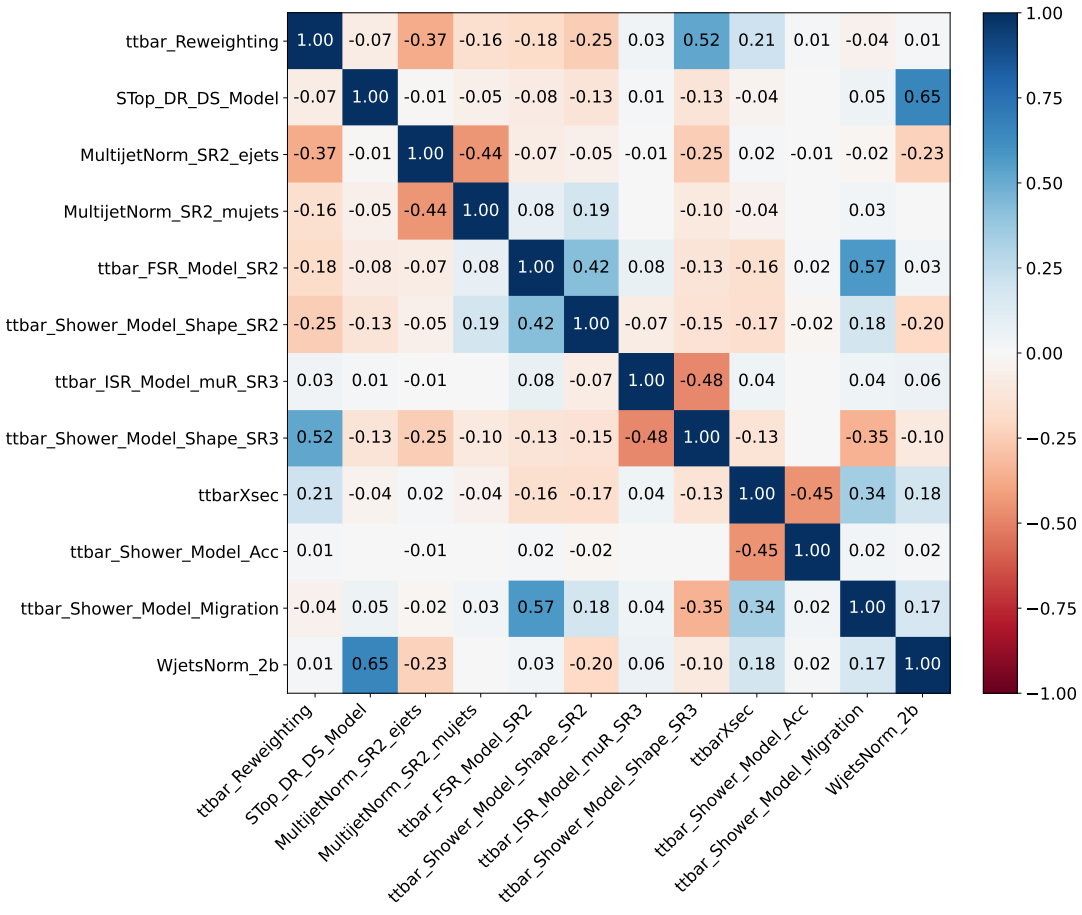


Figure 3.4: Correlations of different NPs with at least one correlation greater than 0.4 reproduced using pyhf using the likelihood of the $t\bar{t}$ production measurement published by ATLAS in Ref. [17].

4

Updating the dataset

With the methods to construct the likelihood covered in theory, the next step now is the actual implementation of a new measurement into SFitter. The following chapter will be split into two parts. First, the implementation of the experimental data and the corresponding uncertainties will be explained and the methods to compute the corresponding theoretical prediction are shown. The measurement considered here is the $t\bar{t}$ differential cross section measurement at 13 TeV in the leptons + jets channel with an integrated luminosity of 137 fb^{-1} by the CMS collaboration at the LHC [19].

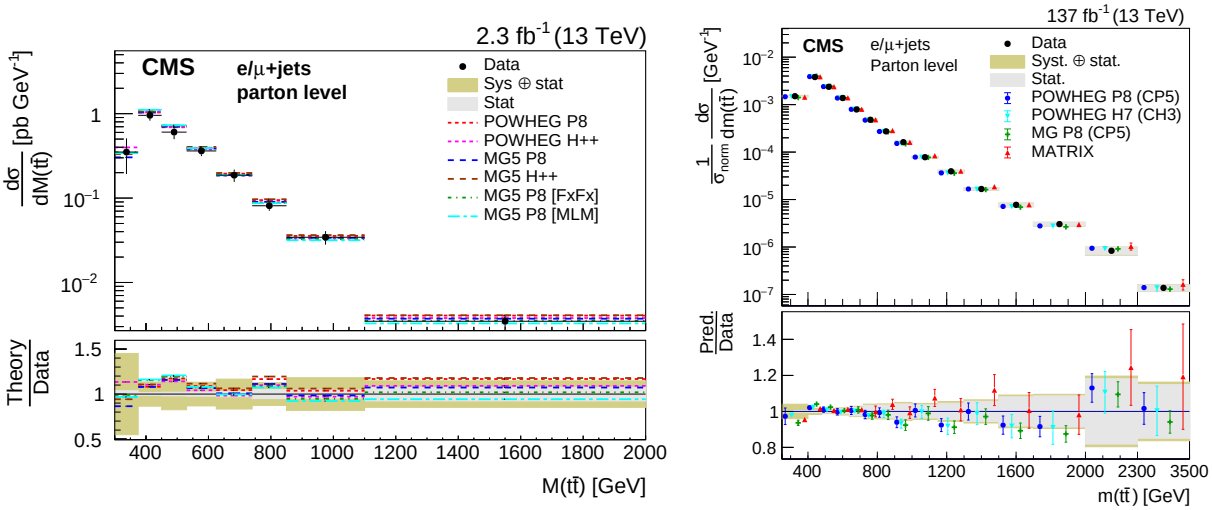


Figure 4.1: Left: Previously implemented measurement of the differential $t\bar{t}$ cross section taken directly from Ref. [55]. Right: Newly implemented measurement which replaces the one shown on the left, taken directly from Ref. [19].

Looking at all the data provided in Ref. [19] one finds that there are many different observables from which one can choose. For the purpose of this analysis, two different differential cross sections were chosen to be implemented, analyzed and compared to see which one will eventually be added into the final dataset.
The first is the differential cross section measurement as a function of the invariant mass $m_{t\bar{t}}$ of the $t\bar{t}$-pair, shown on the right of Figure 4.1. There are two main reasons for this choice. Firstly, this measurement is an update to a previously implemented measurement in the dataset, shown on the left of Figure 4.1, of a $t\bar{t}$ cross section measurement in the lepton+jets channel with an integrated luminosity of 2.3 fb^{-1} [55]. Choosing the exact same observable as before makes a direct comparison of the two much easier. Another reason is that one usually finds the predictions for the invariant mass to be easier to calculate, as one will see later in Chapter 5 when the implementation of the second observable, p_T, is discussed.

To be more specific, the second observable of choice is the transverse momentum $p_T(t_h)$ of the hadronically decaying top quark. This is, just like it was for $m_{t\bar{t}}$, motivated by a previous measurement implemented into SFitter. Similar to before, this was a measurement of the $t\bar{t}$ cross section in the lepton+jets channel but this time with an integrated luminosity of 35.8 fb^{-1} [56]. Because of this, it is essential to take the independence of these measurements into consideration. Clearly, one only wants to implement independent measurements into the dataset, since this is necessary for the proper likelihood construction. In the case of these two old measurements from Ref. [55] and Ref. [56] both were taken in two different data taking periods, making the two independent of one another. However, this is not the case anymore once the most recent measurements from Ref. [19] are included, as these measurements come from data taken over the entire Run II of the LHC. Regardless of which observable is chosen to be implemented into SFitter in the end, both of these two measurements will have to be removed from the dataset to ensure a fully independent dataset.

To understand what kind of improvement one expects from updating the dataset, one can take the $m_{t\bar{t}}$ distribution as an example. There are two main differences that one can already easily see when comparing the two plots shown here. First, the increase from an integrated luminosity of $2.3\,\text{fb}^{-1}$ to one of $137\,\text{fb}^{-1}$ gives much better statistics, reducing the statistical uncertainties by a significant amount. Similarly, the systematic uncertainties of the experiment have also improved significantly, most easily seen in the ratios of data and prediction. Here, the old measurement clearly shows that systematic uncertainties contribute significantly for almost all bins while the updated one is clearly statistics dominated, except for the very first low energy bins. Another key difference is the binning and kinematic ranges covered by these measurements. The previous measurement only included a total of eight different bins going up to energies of $m_{t\bar{t}} = 2000\,\text{GeV}$ while the updated measurement almost doubles the number of bins with a total of 15 while also increasing the covered energy range up to $m_{t\bar{t}} = 3500\,\text{GeV}$. This higher reach in energies is one of the main reasons for updating these measurements since, as was discussed back in Section 2, the four fermion operators are enhanced at higher energies. For this reason, higher energy bins are much more sensitive to these contributions and are crucial when it comes to constraining these four fermion operators.

Also important to note is that the new measurements and the previous ones are both implemented using their normalized distributions. Figure 4.1 shows the actual newly implemented data on the right side, although the data on the left only shows the not yet normalized data points, since no additional figures were provided for these in Ref. [55]. There are a few reasons to use these normalized distributions, with the simplest one being the reduction of systematic uncertainties. The best example to understand this is the uncertainty on the luminosity, since it always affects all measurements at the LHC in the same way. Because of this the effect of the luminosity on both the measured value in each bin and the total cross section they are normalized with cancel after taking the ratio. In a very similar way, this can also lead to the cancellation of higher order SMEFT contributions if they do not exhibit any kinematic behavior.

This is, for example, the case for C_{tG} in $t\bar{t}$ measurements, like the ones considered here, which is why the constraints on C_{tG} are not expected to improve significantly when only

normalized distributions are considered. Especially since the Top dataset also includes measurements of the total $t\bar{t}$ cross section, which will constrain C_{tG} separately.

The following will give a detailed step-by-step guide on how to implement a new measurement into SFitter, taking the invariant mass distribution as an example. Starting with the implementation of the measured data and then moving on to the simulation of theory predictions.

4.1 Experimental data

The extraction of the experimental data is made easy since the data of all the different measurements is available on the corresponding HEPData entry, making this step trivial. The total systematic and statistical uncertainties for each bin one can easily find in the HEPdata tables for each of the different observables provided. For SFitter the total systematics are not enough, however, since the correlation of different systematics is one of the core characteristics of its likelihood construction. Because of this, it is now necessary to take a more detailed look at the systematics. Typically, these are provided in the form of tables for each analysis, with the table for the analysis currently being implemented shown in Table 4.2.

Source	Uncertainty [pb]	[%]
Jet energy	11	1.38
Branching fraction	8.8	1.11
Lepton	7.8	0.98
NNLO	7.6	0.96
b tagging	7.0	0.88
Sim. event count	6.5	0.82
Background	6.1	0.77
CR model	5.5	0.69
Jet energy resolution	3.4	0.43
Scales μ_R, μ_F	3.2	0.41
Initial-state PS scale	3.2	0.40
Final-state PS scale	2.7	0.34
Subjet energy	2.4	0.31
b mistagging	2.2	0.28
UE tune	2.2	0.27
m_t	2.1	0.26
PDF	1.9	0.25
h_{damp}	1.5	0.19
L1 trigger	0.5	0.07
Pileup	0.4	0.05
Total syst.	21	2.66
Total stat.	0.6	0.07
Int. luminosity	14	1.75

Uncertainty group	Impact [%]
Theory uncertaintes	
TopMass	0.26
UnderlyingEvent	0.27
PDF	0.25
Scales	0.41
Statistical uncertainties	
MC	0.82
stat	0.07
Systematic uncertainties	
Jets	1.48
bTagging	0.92
PSscale	0.53
Background	0.77
Leptons	0.98

Figure 4.2: Left: List of systematic uncertainties for the CMS $t\bar{t}$ measurement, taken directly from Table 1 in Ref. [19]. Right: Groups of uncertainties, for this measurement, implemented into SFitter.

The provided systematic uncertainties on the left side of the table already give a comprehensive breakdown of the different uncertainties into a number of smaller categories. The actual number of systematics implemented in the final fit is reduced even further, since only the normalized measurement is implemented. The actual list of systematics implemented is shown in the second table in Figure 4.2 where the different uncertainties in a group, such as Jets, combines all jet related uncertainties by adding them in quadrature.

With this the systematic uncertainties on the total cross section have been found and in the case of a rate measurement one would now be able to implement this into SFitter. For distributions, such as this one, one final extra step is required. To properly account for the uncertainties, one needs to assign each of these to every bin separately. To do this, the same procedure used for the other previously implemented measurements is used. For this one simply takes the total systematic uncertainty for each bin, provided in the HEPdata tables, and distributes these according to the relative contributions

$$\omega_b = \frac{\Delta_{\text{tot},b}}{\sqrt{\sum_j \Delta_j \Delta_j}} \tag{4.1}$$

where $\Delta_{\text{tot},b}$ is the total systematic uncertainty in the respective bin and Δ_j the individual systematics from the previous table.

DataPrep

The final step necessary before the implementation of the experimental data is complete is to ensure that the correct systematic uncertainties are correlated with one another. Important to note is that SFitter is set up in such a way that any measurement is written in the following way

$$\text{MEASUREMENT\_NAME}_i = \sigma_{meas,i} \pm \Delta_{syst,i} \pm \Delta_{stat,i} \pm \Delta_{hat,i} \,,$$

listing the different uncertainties of the measurement after the measured value. Essential to this is now that the order of the different uncertainties is very important, since only those uncertainties in the exact same position in the data card are correlated in the final likelihood.

To allow for an easy creation of inputs such as these, a specialized tool called DataPrep, was developed for Ref. [15] which ensures the proper ordering of any such uncertainties independent of the initial inputs. This allows for a more flexible and easily generalizable input of the different uncertainties. In addition to this, it also provides a way to correlate the theoretical uncertainties of measurements describing the exact same process, i.e. those which measure the same observable of the same process and at the same energy. For the dataset from Ref. [15] the measurements averaged were, for example, the $t\bar{t}$ total cross section measurements from both ATLAS and CMS measured in multiple different channels such as lepton+jets or $e + \mu$.

It averages measurements of processes in such a way that the more constraining measurements are weighted more strongly. After averaging multiple measurements, one is left with only a single datapoint one implements into SFitter. This only requires a single

theory prediction and due to this all theory uncertainties are correlated for these measurements. Since none of the newly implemented measurements are weighted in such a way, this will not be explained in great detail here.

In the case of the implementation of a top measurement similar to those already included in the dataset, this did not require any changes to the code and it could be used as is. There have, however, been some slight modifications made to allow it to handle inputs of Higgs measurements on their own, or even both Higgs and Top data at the same time. This will be more important later for the combined fit of multiple different datasets in Section 5.3.

4.2 Predictions from Theory

To compute the likelihood using newly implemented data, it is now necessary to generate the theoretical predictions for this process. If one were to implement the data as it has been measured by the detector, this would require one to simulate the physics process using an event generator such as MadGraph5_aMC@NLO [57,58] and perform the showering of the resulting particles using additional software such as Pythia [59] or Herwig [60]. Finally, these generated showers would be measured using a detector which, as one would expect, is not a perfect tool and as such can only resolve events up to a certain resolution. This introduces a smearing into the final measured data, which one would simulate using dedicated frameworks such as Delphes [61].

Going through this whole chain does not only require the use of multiple different pieces of software, which does not only increase the amount of time required to generate predictions significantly but can also introduce additional uncertainties since e.g. the choice of showering algorithm can lead to different results. The good news is that all the data from Ref. [19] has been unfolded and given on the parton level. This means that all detector and showering effects have been removed from the measured data to give only information on the hard scattering process at parton level. Due to this, only the very first step of this chain is actually required when implementing these measurements, making this process notably more straightforward. In the following there will first be a basic overview of MadGraphs most important features, mostly following Ref. [62], to gain an understanding of how the cross section of different SM processes are computed.

Event simulation

As has already been mentioned, MadGraph relies on the generation of matrix elements to compute cross sections, decay widths or generate events. These can either be directly used for an analysis, or showering algorithms such as Pythia or Herwig can be applied to continue the detector simulation chain. This process of matrix element generation can be summed up in a few steps. After defining a process, such as, for example, in this case the production of a $t\bar{t}$-pair from a proton-proton collision, MadGraph generates all possible allowed Feynman diagram for this process. After all allowed diagrams are generated, it uses the ALOHA [63] library to compute the corresponding helicity wave functions for the different parts of the diagrams, in the end returning the full amplitude for each diagram. Taking the sum over all the different amplitudes and squaring the

result finally gives the matrix element for the entire process.

The formula to go from these matrix elements to the cross sections is conceptionally simple,

$$\sigma \propto \int \mathrm{d}\Phi(n)\, |\mathcal{M}|^2 \tag{4.2}$$

where one can immediately see that it is necessary to integrate the matrix elements over the entire phase space Φ. This is far from trivial, and large amounts of work and care has been put into the development of efficient and accurate algorithms to perform this task. The general idea is to employ Monte Carlo techniques to numerically compute these integrals.

For this thesis, the most important step of the computation of any process using MadGraph, is the generation of all allowed Feynman diagrams for the given process. All the information necessary for MadGraph to decide which Feynman diagrams are valid are encoded within the model used. These models are given in the Universal FeynRules Output (UFO) format [64,65]. Any model can be defined by their particle content, a few parameters such as their masses and coupling and finally the interactions between these particles. FeynRules uses these inputs to construct the corresponding Lagrangian, from which the different vertices and propagators are determined.

One such UFO model is SMEFTatNLO [66], implementing a long list of different dimension six SMEFT operators into MadGraph by extending the SM and makes it possible to compute SMEFT prediction up to NLO. The model mostly focuses on interactions involving a top quark and imposes a $U(2)_q \times U(2)_u \times U(3)_d$ flavor symmetry, very similar to the one imposed on the Top dataset in Section 2. More details on this model and its uses can be found in Ref. [66]. For the purpose of this study, it is only important to know that all the operators given in Table 2.2 are implemented in this model, which means that the contributions of all SMEFT operators affecting the $t\bar{t}$ production cross section can be determined using just this model.

Standard Model predictions

Before one can concern themselves with any higher order corrections to the SM prediction, it is first necessary to ensure that the SM itself can already be determined to sufficiently high accuracy. This can be easily validated using the data provided in Ref. [19] since these also give the ratio of data and prediction for different event generators for all different observables.

The predictions were computed using MadGraph5_aMC@NLO (v.3.5.0), simulating the process up to next-to-leading order (NLO) accuracy. Unlike the predictions given in the paper, however, no additional showering algorithms such as Pythia were used. The choice of renormalization μ_R and factorization scale μ_F follows Ref. [19] and is chosen to be half of the sum of the transverse masses of the top and anti-top

$$\mu_R = \mu_F = \frac{1}{2}\left(m_T(t) + m_T(\bar{t})\right) . \tag{4.3}$$

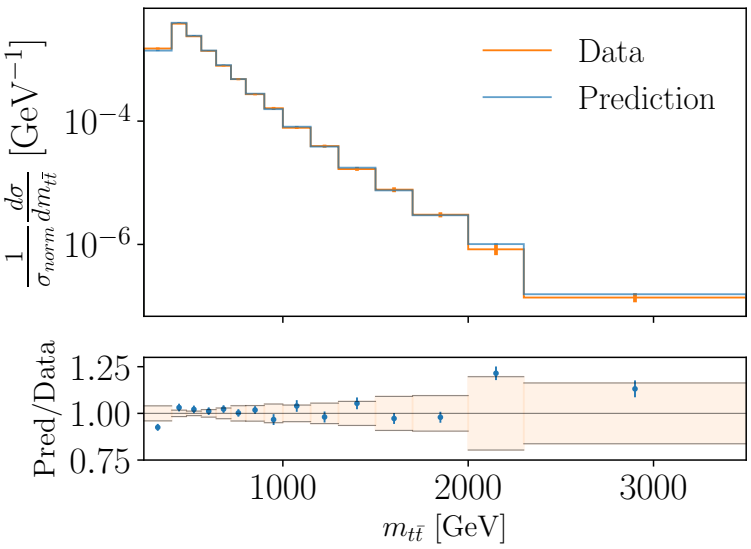


Figure 4.3: Comparison of the measured data from Ref. [19] (orange) and the values reproduced using the MadGraph event generator (blue). The corresponding ratio of prediction and data is shown in the lower plot, with the orange band describing the total uncertainty for the data.

To stay consistent with the parton density functions used in the previously implemented measurements the NNPDF3.1 NLO set was used at first [67]. This lead to inconsistent contributions from SMEFT operators in the tails of the distributions, where NNPDF3.1 has been shown to have issues with positivity. Because of this, the final predictions used the more recent NNPDF4.0 NNLO set of parton distribution functions, where these problems have been fixed [68].

Figure 4.3 shows both the data extracted directly from the paper and the corresponding SM prediction from MadGraph, with the ratio of these given directly below. The error bands in the lower plot depict the sum of statistical and systematic uncertainties provided by the experiment, while those for the predicitons come solely from MadGraph. Clearly they show very good agreement for bins 2 through 13 showing mostly minor deviations very similar to the deviations between data and theory in the paper [19]. On the other hand, the prediction in the first bin slightly underestimates the cross section compared to the data. Taking the uncertainties into account these differences are not too significant and one can see that compared to the comparison of data and theory in the paper, the deviations show a very similar pattern. Similarly, one finds that the last two bins seem to overestimate the cross section compared to the measured data, especially for the second to last bin. This does, however, once again agree with the results from Ref. [19].

Going to next-to-next-to-leading order (NNLO)

All previous calculations were performed using MadGraph_aMC@NLO which, as the name suggests, allows for the computation of predictions up to NLO accuracy. To account for the most precise measurements at the LHC, however, higher accuracy results of at least NNLO are necessary to gain predictions of sufficient accuracy. Previously implemented

measurements in Ref. [15] made use of fastNLO tables [69] to compute these so-called k-factors used to go from NLO to NNLO predictions and are simply defined as the ratio

$$k = \frac{\sigma_{NNLO}}{\sigma_{NLO}} \,. \tag{4.4}$$

For this analysis a recently published API called HighTea [70] is used to compute the NNLO level predictions for $t\bar{t}$ production for both $m_{t\bar{t}}$ and $p_T(t_h)$. HighTea provides a way to compute NNLO level predictions by making use of precomputed events stored in an online database, allowing for an easy and fast computation of these predictions without the needs of time intensive event generation.

After computing these NNLO prediction they are normalized to the total cross section of $\sigma_{t\bar{t}} = 832\,\text{pb}$, following again Ref. [19]. The results from this are used to calculate the k-factors for each bin which are subsequently implemented into SFitter to give the NNLO level predictions. These k-factors apply to the NLO SM predictions previously computed, meaning that the NNLO accuracy only holds for the SM prediction while the SMEFT contributions are calculated to NLO in QCD.

Additional theory uncertainties

While most of the uncertainties implemented into SFitter come directly from the experimental papers, there are a few uncertainties from the predictions computed here that need to be taken into account. All of these come from certain choices which have to be made before computing the predictions. They include the choice of renormalization and factorization scale, the parton density functions used, as well as the choice of event generator in general.

The way the scale uncertainties are currently determined is by varying the scales μ_R, μ_F by a factor of 2 in both directions which gives a maximal and minimal value for the cross section from which the uncertainty can easily be extracted as the width between these two values. In SFitter the scale uncertainty on the SM prediction is included as an overall theory uncertainty on the observable, which means that the scale uncertainty can easily be determined by computing the SM predictions in Figure 4.3 at different scales. MadGraph makes this easy since this computation at different scales has already been implemented and automated, which allows the easy determination of these uncertainties by simply telling it the different scales one wants to consider.

In a very similar way, the choice of the parton distribution functions used has an effect on the predictions. For the prediction calculated here, the NNPDF40_nnlo_as_0118 PDF set with $\alpha_S = 0.118$ is used. The uncertainties coming from these can, very similar to the scale uncertainty, be easily obtained using the functions built into MadGraph which computes the predictions for the 100 different members of the pdf set.

Finally, there are the uncertainties coming from the choice of the event generators used. For all predictions used in SFitter only MadGraph is used, while one can see from Figure 4.1 that the theory predictions provided by the different events generators Powheg, MadGraph and Matrix all give results that differ slightly. One could now determine

these uncertainties by reproducing the exact same predictions using the other event generators. To save the time of doing this, the information provided by CMS in Figure 4.1 will be used to estimate the effect of the different generators on the predictions. To get a conservative uncertainty estimate from these, the envelope of the uncertainties is used as a final theory uncertainty in SFitter. All of these uncertainties are added as theory uncertainties on top of those already implemented in Section 4.1.

SMEFT operator contributions

The process to determine the contributions of different operators to the cross section is technically quite similar to the SM, where the main difference comes from the underlying model used. Using the SMEFTatNLO UFO model, one can tell MadGraph which operators should be considered while generating a process. This replaces the rules for the construction of legal Feynman diagrams in the first step, such that all legal Feynman diagrams involving the considered operators are also generated. To determine the exact contributions from the WCs one sets the one currently considered to a specific value and performs the exact same computation as in the SM case. Similarly, this gives a result for the cross section which one can compare to those from the SM to determine the contribution of this WC at its current value. Repeating this process multiple times for different values of the WC, one can plot the contributions to the cross sections as a function of the value of that WC.

In Chapter 2 it was decided that only SMEFT contributions up to dimension six at the level of the Lagrangian will be included in the predictions. This means the SMEFT predictions using multiple different WCs C_i can be written as

$$\sigma_{\text{SMEFT}} = \sigma_{\text{SM}} + \frac{C_i}{\Lambda^2}\sigma_{\text{lin},i} + \frac{C_i^2}{\Lambda^4}\sigma_{\text{quad},i} + \frac{C_i C_j}{\Lambda^4}\sigma_{\text{Interf},ij}\,. \tag{4.5}$$

Neglecting the interference terms $\sigma_{\text{Interf},ij}$ for now, the following section will be dedicated completely to the linear and quadratic contributions.

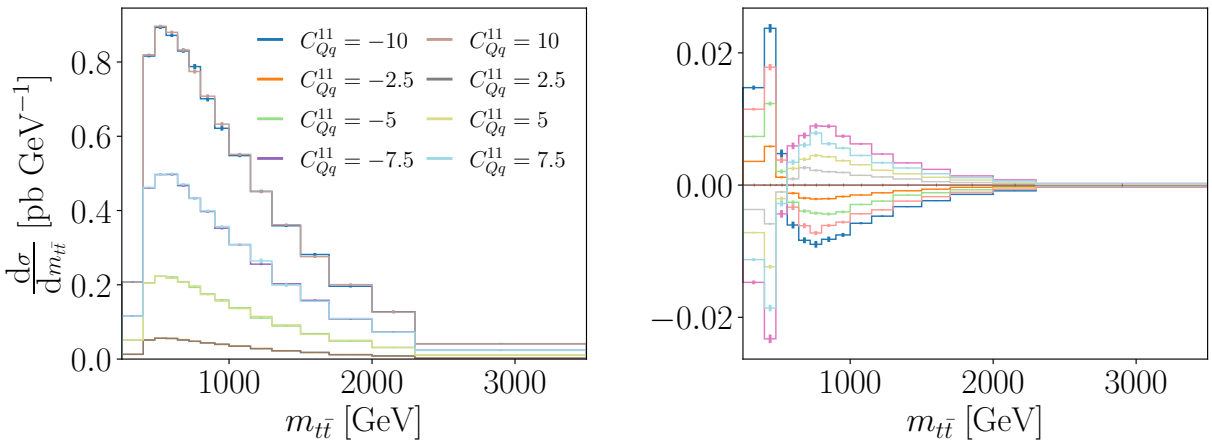


Figure 4.4: Histogram of SMEFT contributions from the C_{Qq}^{11} Wilson coefficient for both quadratic (left) and linear (right) contributions.

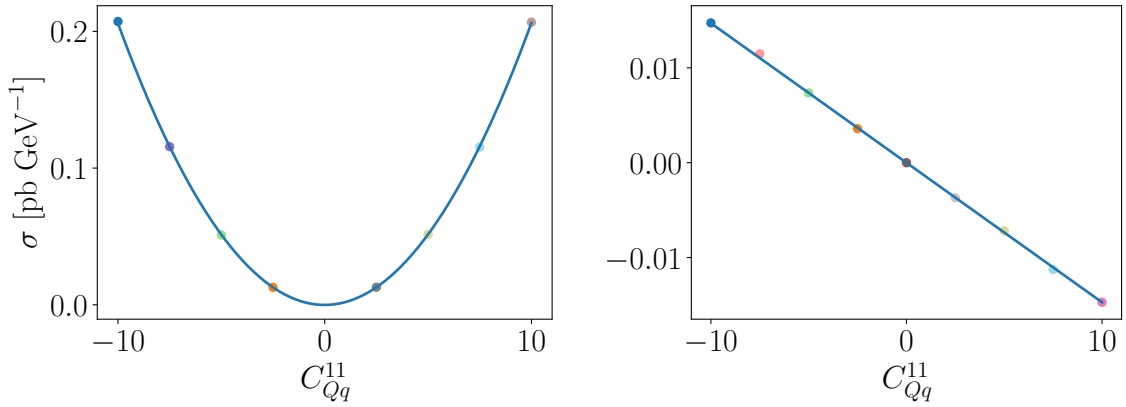


Figure 4.5: Quadratic (left) and linear (right) SMEFT contributions of C_{Qq}^{11} to the first bin of the $m_{t\bar{t}}$ distribution, shown in Figure 4.4 as a function of C_{Qq}^{11}.

To implement these into SFitter, the MadGraph analysis was set up in such a way such that the SM, linear SMEFT and quadratic SMEFT contributions are separated and put into different histograms. As an example, the plots for the contributions from C_{Qq}^{11} to $t\bar{t}$ production is shown in Figure 4.4 for several values of C_{Qq}^{11}. The plot on the left shows the quadratic contributions from the operator, while the one on the right shows the linear ones. To validate the linear and quadratic behavior of these contributions the plots in Figure 4.4 show the contributions from the C_{Qq}^{11} operator to the $t\bar{t}$ cross section, this time showing the behavior of its contributions in a single bin for different values of C_{Qq}^{11}. Both plots very clearly confirm the quadratic and linear behavior of the contributions.

The predictions for the cross section including both linear and quadratic SMEFT contributions are implemented in SFitter in such a way that the cross section in each bin can be easily computed using the following expression

$$\sigma_{\text{SMEFT}} = \left[1 + \kappa_{1,d} C + \kappa_{2,d} C^2\right] \sigma_{\text{SM}} \tag{4.6}$$

where the coefficients $\kappa_{1,d}$ and $\kappa_{2,d}$ are simply the σ_{lin} and σ_{quad} from Equation 4.5, respectively, divided by their SM prediction σ_{SM}. These coefficients are extracted via fits of the corresponding polynomial to the SMEFT contributions shown in Figure 4.5 after normalizing them to their SM predictions.

With this, the SMEFT contributions to have been determined, the final step now is to properly take the normalization into account. From Equation 4.6 the contributions to each bin are known and the subsequent normalization is easily performed by simply dividing each bin by the total cross section. The total cross section is calculated by taking the sum over the contributions of each bin, meaning that the total cross section also depends on the higher order SMEFT contributions. Because of this, the extraction of linear and quadratic terms from these is not trivial and requires a simple approximation. Assume for now that, for ease of notation, there is only a single WC contributing which will be denoted as C. From Equation 4.6 the normalized cross section for a bin can be

written as

$$\sigma_{\text{norm,SMEFT}} = \frac{\sigma}{\sigma_{tot}} = \frac{\left[1 + \kappa_{1,d}C + \kappa_{2,d}C^2\right]}{\left[1 + \kappa_{1,t}C + \kappa_{2,t}C^2\right]}\sigma_{\text{norm,SM}} \tag{4.7}$$

$$\approx \left[1 + \kappa_{1,d}C + \kappa_{2,d}C^2\right]\left[1 - \kappa_{1,t}C + (\kappa_{1,t}^2 - \kappa_{2,t})C^2\right]\sigma_{\text{norm,SM}} \tag{4.8}$$

where in the second line a Taylor expansion of the denominator was used since the contributions coming from the SMEFT operators are expected to be small. By ordering this in powers of C one finds

$$\frac{\sigma_{\text{norm}}}{\sigma_{\text{norm,SM}}} \approx 1 + (\kappa_{1,d} - \kappa_{1,t})C + \left[\kappa_{1,t}^2 - \kappa_{2,t} + \kappa_{2,d} - \kappa_{1,d}\kappa_{1,t}\right]C^2 + \mathcal{O}(C^3) \tag{4.9}$$

$$\equiv 1 + a_{\text{norm},b}C + c_{\text{norm},b}C^2 + \mathcal{O}(C^3)\,. \tag{4.10}$$

This defines the linear a_{norm} and quadratic c_{norm} contributions to the normalized cross section as functions of the coefficients $\kappa_{1,d}$ and $\kappa_{2,d}$ which were already extracted from the polynomial fits shown in Figure 4.5. The only additional information required are the linear and quadratic contributions to the total cross section $\kappa_{1,t}, \kappa_{2,t}$ which can be determined just as easily as the contributions to the individuals bins by performing a fit using the total cross section and not for individual bins.

SMEFT interference

Equation 4.6 already showed that there are additional interference terms between the differnt dimension six SMEFT operators that contribute to our observables. This means that besides the linear and quadratic contributions from the previous section, there are a lot of additional interference terms which would need to be determined to completely determine the predictions. To do this one used, once again, the SMEFTatNLO UFO model, but this time, instead of only setting a single operator to be different from zero, the two operators one wants to interfere are activated simultaneously. One of the two WCs is set to a fixed value different from zero and, after doing so, one follows the exact same steps as described in the previous section to determine the SMEFT contributions.

Previously, it was explained that MadGraph separates the linear SMEFT contribution from the quadratic ones. To be more specific, it separates contributions which scale differently with Λ. Because of this, the linear contributions which scale with Λ^{-2} are separated from the quadratic ones which scale with Λ^{-4}. Now that interference terms are included in MadGraph the same will happen for the interference terms. From Equation 4.6 we know, however, that these interference terms also scale with Λ^{-4} which is why MadGraph cannot separate the quadratic SMEFT contributions from those coming from the interference terms. This means that to properly extract the contributions coming from the interference terms, it is necessary to remove the quadratic contributions coming from the individual operators first. Fortunately, these have been computed already in the previous section, which means that one can simply remove their effect by subtracting them. After doing this, one is left with only contributions from the interference, with two examples of this shown in Figure 4.6.

In the left plot of Figure 4.6 the contributions from the interference between C_{Qq}^{11} and C_{Qq}^{31}

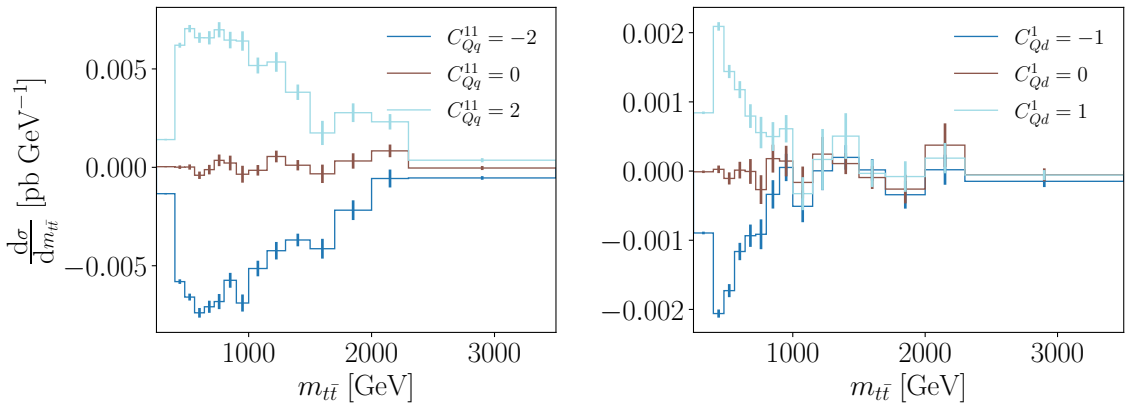


Figure 4.6: Contributions from interference between C_{Qq}^{11} and C_{Qq}^{31} (left) and C_{td}^{1} with C_{Qd}^{1} (right) for $t\bar{t}$ production. For both, the WC not being varied is fixed to 1.

is shown as an example. One can see very clearly the linear behavior of the contributions from interference, which is what one expects since one of the two WCs contributing to this interference is fixed and only the other one varies. The interference between C_{Qq}^{11} and C_{Qq}^{31} was chosen as an example here since these contributions are very clear and because of this the following steps of extracting them via a linear fit is straightforward. A problem arises when considering some of the other interference terms, however. To get a better understanding of this, the exact same procedure is repeated, for instance, for C_{td}^{1} and C_{Qd}^{1}. Comparing the results for this, shown on the right of Figure 4.6, one of the major issues with intereference terms becomes apparent. Although the contributions coming from interference are large enough compared to their uncertainties in the first bins, contributions to the tails are so small that a proper extraction of them is almost impossible.

Numerous attempts were made to reduce the noise in these interference terms. The first one was to request a higher accuracy from MadGraph of up to 0.01%, unfortunately this lead to computation times of approximately three days. On top of that, this also did not improve results significantly either, with the plot on the right side of Figure 4.6 already showing the results from such a high accuracy computation. The problem we found here is that as mentioned previously, the interference terms are computed simultaneously with the quadratic contributions from the individual operators. Since these quadratic contributions are clearly dominant compared to the interferences, MadGraph will focus on minimizing the uncertainty of these first. Due to this it takes a very long time to increase the accuracy of these interference terms and even then there was no guarantee that it would ever improve enough to properly extract the interferences.

After this, the value of the fixed WC coefficient was optimized in such a way that the contributions from interference is maximized compared to the quadratic ones. This was done by analyzing the relative contributions of the individual WC using the data of just the linear and quadratic contributions alone. Still, the interferences were too noisy to extract proper contribution from these. For this reason, it was decided that in the interest of time all following analyses involving the newly implemented measurement, will neglect

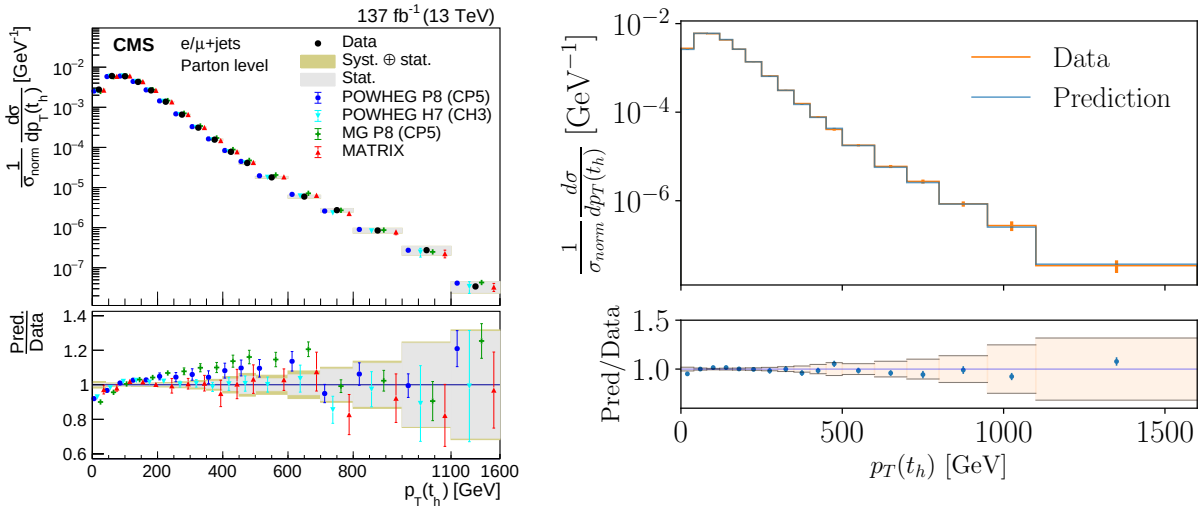


Figure 4.7: Left: Differential cross section as a function of the transverse momentum of the hadronically decaying top, as given in Ref. [19]. Right: Reproduced predictions using MadGraph compared to the data given on the left.

contributions from interference terms.

Implementing the $p_T(t_h)$ distribution

The procedure to implement the $p_T(t_h)$ distribution, which for convenience will also be referred to as the p_T distribution, is completely analogous to that of the $m_{t\bar{t}}$ distribution, meaning that all the previous discussions also apply to it. Starting again with the SM predictions, they are first compared to data in Figure 4.7.

Looking at the data provided in Ref. [19], one can already see that the predictions obtained from the different monte carlo generators already show much larger differences compared to m_{tt} in Figure 4.3. The 11th bin in particular shows a significant deviation from data for the MadGraph predictions, which are naturally the values of interest here, since the same generator is used for the SFitter predictions. There is, however, also an increased statistical uncertainty in this range as well, as such the generators are still consistent with one another and with the data within uncertainties. Comparing this with the reproduced results on the right of Figure 4.3 one finds that they show deviations similar to those in Ref. [19]. For both one finds that there is an underfluctuation in the first bin and the 11th bin shows again stronger deviations compared to the others. Taking the uncertainty on the data, shown as an orange band here again, into account they do still agree with the measurement very well.

Just like in the previous implementation of $m_{t\bar{t}}$ these predictions are only computed up to NLO accuracy since this is the limit of MadGraph . Once again, the k-factors used to go from NLO to NNLO are computed using the HighTea library. The determination of SMEFT contribution is also completely analogous to $m_{t\bar{t}}$ which is why this will not be repeated here. The interferences were just as noisy for this distribution still, which again means that they will be neglected for all fits involving this measurement as well.

5

Results

Now that the theoretical and technical foundation of this work are set, we can begin looking at the applications of these concepts, as well as analyze and discuss the results. This chapter is split into 3 different parts. First, the implementation of 2 different published likelihoods by ATLAS are discussed, including how they were implemented and how it changes our fit. After this, the differences between marginalization and profiling methods are analyzed using only data in the Top sector. Finally, a combined fit using the dataset from both the previous Top fit and the Higgs, Di-Boson and EWPO dataset is performed.

5.1 Implementation of published likelihoods

In Section 3.3 the general properties and information one can gain from the likelihoods published by ATLAS [17,18] has been studied. In the following, the use of these likelihoods in the context of a SFitter analysis will be examined. To this end, a comparison between the traditional implementation methods of SFitter with the new one using these likelihoods is made. Since these are updated measurements of the total cross sections for both $t\bar{t}$ and $t\bar{t}Z$, no additional theory predictions have to be computed, since those previously implemented into SFitter can be used again. The same applies to theory uncertainties coming from these simulations, simply using the same uncertainties coming from the choice of scale, parton densities and MC statistics from before.

Profiling published likelihoods

The traditional way of implementing the data from a new measurement into SFitter has already been discussed in detail in Section 4.1. Because of this, the main focus of this section will be on the values and results extracted from the published likelihoods. Extracting the measured cross sections follows essentially the same procedure for both implementation methods since regardless of whether one considers the full likelihood or just the measured data with systematics and statistical uncertainties, both are given in the form of a simple JSON file. Similar to 4.1 the main difference comes from the extraction of the different uncertainties.

From the discussion in Section 3.3 it was shown that the python package pyhf can be used to determine the effect of individual NPs on POI. Taking now the likelihood published for the measurement of the total cross section of $t\bar{t}Z$ as an example, the corresponding POI is the signal strength for this process from which the total cross section can be determined. Since there are, however, a total of 229 NPs in the full likelihood provided, it would be numerically inefficient, and at some point for a global analysis unfeasible,

Uncertainty	**Reproduced** $\frac{\Delta\sigma_{t\bar{t}Z}}{\sigma_{t\bar{t}Z}}$ [%]	**Paper** $\frac{\Delta\sigma_{t\bar{t}Z}}{\sigma_{t\bar{t}Z}}$ [%]
ttZ parton shower	3.1	3.1
tWZ modeling	2.9	2.9
b-tagging	2.9	2.9
WZ/ZZ + jets modeling	2.7	2.8
tZq modeling	2.6	2.6
Lepton	2.3	2.3
Luminosity	2.2	2.2
Jets + E_T^{miss}	2.1	2.1
Fake leptons	2.1	2.1
$t\bar{t}Z$ ISR	1.7	1.6
$t\bar{t}Z\mu_F$ and μ_r scales	0.9	0.9
Other backgrounds	0.8	0.7
Pile-up	0.7	0.7
$t\bar{t}Z$ PDF	0.2	0.2
Stat	5.2	5.2

Table 5.1: Comparison of different systematic uncertainties taken directly from Table 7 in Ref. [18] and those from the corresponding published likelihood using pyhf.

to include all of these individually. Considering, furthermore, that eventually SFitter correlates only those systematics within the same category, gives even more reason to reduce their number by grouping them corresponding to the usual SFitter uncertainty types listed in Table 8.1 in the Appendix.

To be able to do this, it is first necessary to know which of the NPs correspond to what kind of uncertainty. Fortunately, pyhf also gives labels for each and every one of the different nuisance parameters. In the case of the $t\bar{t}Z$ likelihood for example one has NPs such as 'BTag_B','BTag_C','BTag_light' where one can easily tell that these belong into the bTagging category. Similarly, most of the other NPs are labeled in such a way that all of those corresponding to jet uncertainties have 'JET' in their name or 'Fake' for anything related to fake leptons. Not all of them make it quite that easy though, since there is, for example, one nuisance parameter labeled 'xsec_other' and another one 'xsec_Other'. This makes it difficult to simply put them into the correct category, since just the explanations of the uncertainties given in the paper are not sufficient to clear this up. Because of this, some of the NPs have to be assigned either by choosing a category which looks like it would fit best, or by assigning it to one of the categories which show discrepancies between the reproduced results and those given in the paper.

Since going through every single NP can become very tedious for likelihoods with hundreds of NPs one would naturally want to automate this process. Unfortunately, the unclear labeling of some NPs within the same likelihood already makes this difficult. Another problem arises when trying to generalize this to different published likelihoods. This discussion only considered the likelihood published for the $t\bar{t}Z$ measurement up until now. Moving on to the $t\bar{t}$ likelihood now, one also needs to assign all of its NPs to the same groups of uncertainties from Table 8.1. Doing this, one finds that the naming

convention for the NPs do not appear to be consistent between different measurements. For example, the $t\bar{t}Z$ measurement labels bTagging related NPs with the key word 'BTag' while the $t\bar{t}$ measurement uses 'bTagSF'. Additional data from the experiment on exactly which NP describes what uncertainty and a more consistent naming scheme for all published likelihoods would be required to automate this process and allow for much a larger number of different published likelihoods to be worked with at once. For all the reasons listed above, every single NP was assigned to its group by hand here.

In Table 5.1 the systematics for the $t\bar{t}Z$ measurements, taken directly from Ref. [18], are compared to those one gets using pyhf and the corresponding published likelihood. To compute these values, one first performs a profiled likelihood fit to determine the best-fit values of each NP, which gives results such as those given in Figure 3.2. To compute the values given in the table, one performs another fit, where the NPs of the group considered are fixed to the best-fit values from this first fit. The uncertainty can be computed as the quadratic difference between the uncertainty on the POI from the full fit and the one from this fixed fit. As one can see, almost all the different uncertainties agree perfectly, with only a few of them being off by 0.1%. These differences can now be either attributed to numerical precision or possibly some incorrectly grouped NPs, since with just the given information, there is no way to be certain. Regardless, these values already agree well enough that any of the differences will not have a great effect on the results after implementing them into SFitter.

After this the uncertainties required to implement a measurement into SFitter are known, so the next step is to move on to correlations. In Section 3.3 it has already been shown that similar to the uncertainties, one also knows the correlations between the different nuisance parameters within a measurement, now that the full likelihood is provided. Usually there is no information on these correlations within a measurement, which is why SFitter assumes the uncertainties within a single measurement to be uncorrelated. This is why the next step in this study is to see whether one can implement these correlations into SFitter and to see what difference this would make.

In the previous part of the study, it has already been shown that the number of NPs implemented can be decreased by assigning them to different, more general, groups. Due to this, the correlation matrix in Figure 3.4 naturally also decreases in size to the correlations between just those categories. Figure 5.1 shows exactly those correlations, using the likelihood of the $t\bar{t}$ measurement to allow for a direct comparison with the results from Figure 3.4. One can easily see that basically all correlations are negligibly small, most values being in the range of ± 0.01. This is expected, since, as was seen in Section 3.3, the strongest correlations were found for NPs describing different modeling choices and jet related uncertainties. Within SFitter most of the NPs for the model choices now go into 'BkgTTBAR' while the jet related ones are absorbed into a larger 'Jets' uncertainty.

Clearly, most of the correlations are now absorbed into the larger groups, since almost all correlated NPs are now in the same uncertainty group. At the same time one can still see the effect of the stronger correlations between the different jet related NPs and the modeling NPs in the correlation between 'BkgTTBAR' and 'Jets' which, while at a value of -0.07 is still not very significant, is much larger than any of the others.

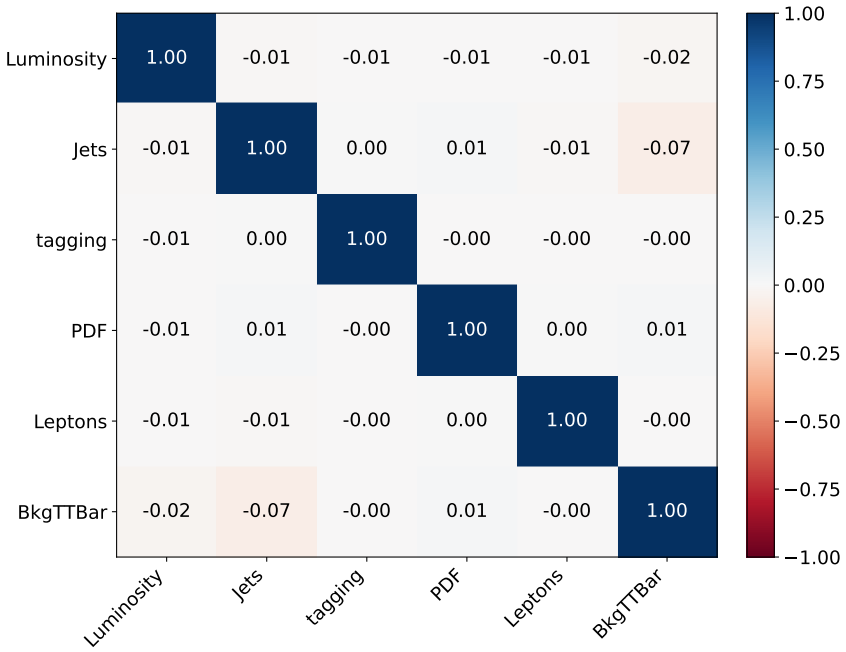


Figure 5.1: Correlations of final uncertainty categories considered in SFitter for the published likelihood of the $t\bar{t}$ measurement in Ref. [17].

This shows very clearly that all correlations between the uncertainties implemented into SFitter are negligibly small. Because of this, implementing any of these correlations into the fit would not lead to differences in the final results. This validates the assumption currently used in SFitter where all correlations between uncertainties within a measurement are neglected. This also means that, since the groups of uncertainties are the exact same for all other measurements previously used in the fit in Ref. [16], this holds for all of those as well. Furthermore, it also allows for a more flexible treatment of the NPs, which can be useful for future analyses where there is still more freedom in the modeling since one would not need to adapt new measurements to be compatible with older ones. For now, there is only a small number of likelihoods available publicly which is why a complete global fit only based on these is not feasible, it will be important to keep this flexibility in mind for future projects.

Constraints from new measurements

While the majority of the study of published likelihoods was dedicated to the likelihoods themselves, the main purpose of this analysis is to check the effect these measurements have on the constraints of different WCs. For this purpose, it is first necessary to ensure that the measurements implemented using the likelihoods lead to reasonable results in such a SMEFT fit. Choosing the $t\bar{t}$ total cross section measurement again, only C_{tG} will be included in the following fit, since this is the only operator contributing to $gg \to t\bar{t}$ at leading order and as such it is the most sensitive one to total cross section measurements and would show differences best.

Because of this, the dataset is restricted to only measurements of total $t\bar{t}$ cross sections

Experiment @ Energy (channel)	stat	syst	theo	total	New?
ATLAS @ 8 TeV (lj) [71]	0.3%	3.5%	25%	25.25%	
CMS @ 8TeV ($e\mu$) [72]	0.6%	3.8%	20%	20.37%	
CMS @ 13TeV (lj) [73]	0.5%	3.8%	16.7%	17.13%	
CMS @ 13TeV (ll) [74]	1.3%	3.7%	18.1%	18.5%	
ATLAS @ 13TeV ($e\mu$) [75]	1.2%	4.2%	16.6%	17.17%	
ATLAS @ 13TeV (lj) [17]	0.05%	4.6%	19.62%	20.15%	Added

Table 5.2: List of total uncertainties for the different $t\bar{t}$ total rate measurements of the Top dataset in SFitter.

already implemented in SFitter. Table 5.2 shows all the different $t\bar{t}$ rate measurements considered in the following fit and the corresponding total statistical, systematic and theory uncertainties. From this, one can already see that the theory uncertainties are clearly dominating, making up more than 95% of the total uncertainty for most. Because of this, all fits and the following comparisons will be performed while neglecting theory uncertainties, since otherwise all differences are completely washed out. Disregarding the theory uncertainties from now, one can immediately see the effect of the higher luminosity in the total statistical uncertainties of the different measurements while the systematics are very similar for all of them, with the new ATLAS measurement having slightly larger systematic uncertainties than any of the previous measurements.

Both curves in Figure 5.2 show the results of a profiled fit to C_{tG} after the new $t\bar{t}$ measurement was included, the difference between the two is that the blue curve shows the results from the implementation using the likelihood while the orange one is from the traditional approach using the data directly from the paper.

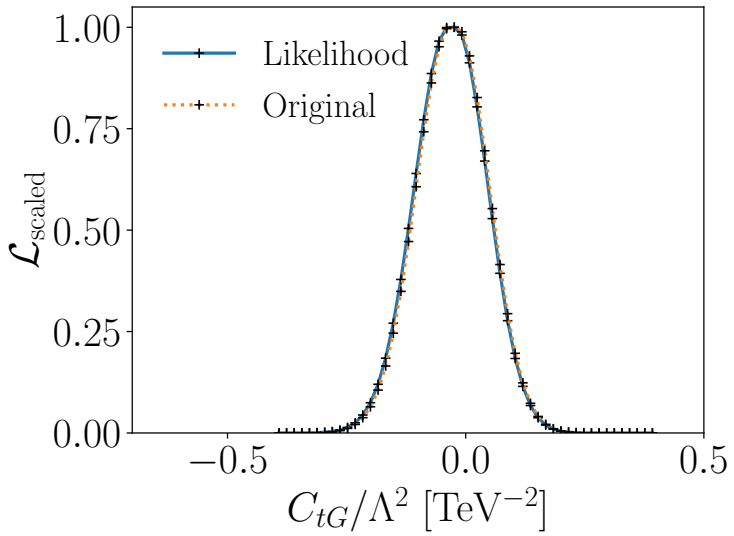


Figure 5.2: Results of a fit to only C_{tG}, comparing the implementation using the published likelihood (blue) with the implementation directly from the paper(orange).

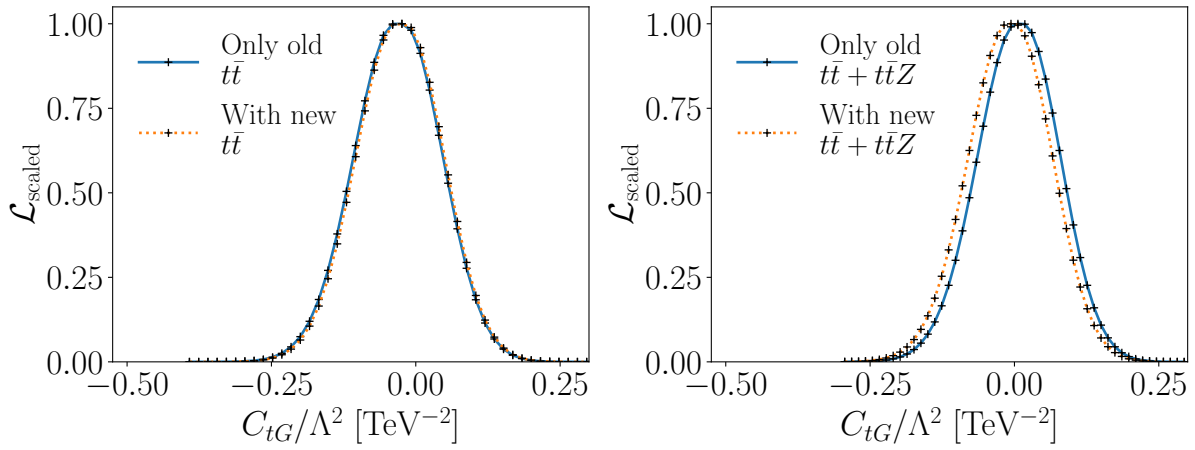


Figure 5.3: Left: Profiled fit to only C_{tG} using the measurement from Table 5.2 with (blue) and without (orange) the new $t\bar{t}$ measurement. Right: Same as left, but including the $t\bar{t}Z$ measurement from Table 5.3. Only including the old measurements (orange) or both new $t\bar{t}$ and $t\bar{t}Z$ measurements (blue).

As one can see, both methods lead to the exact same result, verifying the validity of the implementation using the published likelihood. The same kind of fit was also performed for the $t\bar{t}Z$ measurements, this time using the measurements listed in Table 5.3. For $t\bar{t}Z$ the exact same agreement is found, which can be easily explained when one considers the uncertainties from Table 5.1. Both implementations are based on the same measured data and theory predictions, which means that the only differences in constraining power come from these uncertainties which, as one could already see, are practically the same.

With this, it was shown that implementing the measurement using the likelihood leads to the same constraints in the very end. The next topic of interest is the effect of the new $t\bar{t}$ and $t\bar{t}Z$ measurements on the constraints of WCs. For this reason, the results shown in Figure 5.2 are compared to the results from a fit to the exact same dataset from Table 5.2 without the new $t\bar{t}$ measurement implemented using the likelihood. The results from this fit can be seen on the left side of Figure 5.3. One can very clearly see that the constraints do not change after the new $t\bar{t}$ measurement is included. To understand this, one only needs to consider the uncertainties of the measurements listed in Table 5.2. Although the statistics have improved significantly for the new $t\bar{t}$ measurement one can also see that the systematic uncertainties are larger for this measurement, which means that its constraining power is not expected to be stronger than any of the others in Table 5.2.

Now that the $t\bar{t}$ measurement has been shown not to have an effect on the final constraints, the $t\bar{t}Z$ measurements will now be included in the fit. Naturally, the $t\bar{t}Z$ measurements already included in the previous global fit from Ref. [15] should be included in the fit too. All $t\bar{t}Z$ measurements considered and their corresponding uncertainties are listed in Table 5.3. Important to note is that the new $t\bar{t}Z$ measurement from ATLAS is an update to an older one shown in the table, which means that instead of just adding it to the dataset, it replaces the previous one.

The results for this fit are shown on the right side of Figure 5.3 and one can see that

Experiment @ Energy (channel)	stat	syst	theo	total	New?
CMS @ 13TeV (multi lept.) [76]	5.1%	6.2%	13.4%	15.6%	
ATLAS @ 13TeV (multi lept.) [77]	10.6%	9.1%	13.4%	19.4%	Removed
ATLAS @ 13TeV (multi lept.) [18]	5%	7.6%	13.4%	15.2%	Added

Table 5.3: List of total uncertainties for the different $t\bar{t}Z$ total rate measurements of the Top dataset in SFitter.

there is a visible shift in the constraints of C_{tG}. This one can understand by checking once again the different uncertainties on the measurements, this time for the $t\bar{t}Z$ measurements. Comparing the different uncertainties in Table 5.3, it is easy to see that the new ATLAS measurement has uncertainties which are quite a bit smaller than those of the previous one, although if one compares them to the values of the CMS measurement one finds that their total uncertainties are very similar in size. From this one can see that previously the greater constraints came from the CMS measurements, while the new ATLAS is only similarly constraining. This explains why one can see slight shifts in the likelihoods in the final fit result in Figure 5.2.

The actual constraints are not changed too much, since for the 68% CL regions one has

$$C_{tG}/\Lambda^2 \in [-0.08, 0.07]/\text{TeV}^2 \longrightarrow C_{tG}/\Lambda^2 \in [-0.07, 0.08]/\text{TeV}^2\,, \tag{5.1}$$

and for the 95% CLs one has

$$C_{tG}/\Lambda^2 \in [-0.15, 0.14]/\text{TeV}^2 \longrightarrow C_{tG}/\Lambda^2 \in [-0.13, 0.14]/\text{TeV}^2\,. \tag{5.2}$$

To briefly summarize the most important results of this section, one found that the published likelihoods allow for a more flexible treatment of different nuisance parameters, which are in the end adapted to the general SFitter format, giving very similar results for both methods. After this, the effect of the new measurements on the constraints of the most relevant WC C_{tG} was studied. It was shown that there was only a very small effect on the constraints, indicating that the total rate measurements used in Ref. [15] already display a very high constraining power.

This concludes the studies on published likelihoods done for this thesis, there is, however, still some more work that can be done concerning these likelihoods. The measurements implemented here were both total rate measurements, for the future the implementation of a differential cross section using these likelihoods can be of interest. In Section 4.1 the systematic uncertainties were determined by weighting them according to the total systematic uncertainties. Using the full likelihood it should now be possible to compute the systematics for each bin, leading to better determination of these uncertainties. Furthermore, as will be shown in the following section, differential measurements will also help constrain a larger number of WCs.

5.2 Marginalization in the Top sector

In previous analyses, the SFitter framework has been used to study the effect of different statistical treatments on the constraints on WCs using the Higgs, EWPO and Di-Boson dataset [16]. This part of the thesis wants to expand on this, updating the previous study in the Top sector where profiling methods were used [15].

5.2.1 Analyzing the old Dataset

Before studying the dataset with the newly implemented measurement, the previous dataset used in the global Top fit of Ref. [15], introduced in Section 2, will be analyzed without any changes to allow for a direct comparison between these results and those from the previous paper. This allows for a study of the differences arising from the statistical treatment alone, while also allowing for an easier analysis of the effects of the new measurement on the constraints later.

Before a proper comparison between the marginalization and profiling methods could be performed, it was first necessary to ensure that the profiled results from the previous global Top fit in Ref. [15] were properly reproduced using the dataset within SFitter. To do so, a benchmarking fit was performed first, using the profiled likelihood, where it was found that the results agree completely with those given in Ref. [15]. The corresponding constraints are shown Figure 5.4 in blue, with the results from the exact same fit using marginalization shown in orange. The differences between the two methods were explained in much detail in Section 3, where, to list the most important differences, the treatment of NPs, their correlations and the determination of constraints differ.

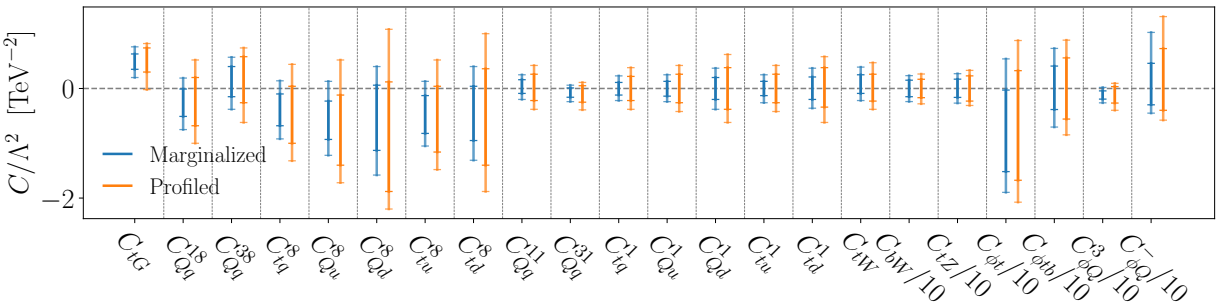


Figure 5.4: 68% and 95% CLs of all Wilson coefficients from the global Top fit for both profiling and marginalization.

One can clearly see that the marginalized results, shown here in blue, are much more constrained for almost all WCs. To properly understand this and for the differences between the two methods to become even clearer, it is useful to look at the distributions of the one dimensional profiled likelihoods and posterior probabilities for the individual WCs. To go from the marginalized likelihood to the posterior, the priors for each WC are assumed to be wide and flat for all constraints shown in the following. This minimizes the effect of the prior on the final results and also allows one to check for differences in the two methods by comparing the profiled and marginalized likelihoods, like the ones plotted in Figure 5.5.

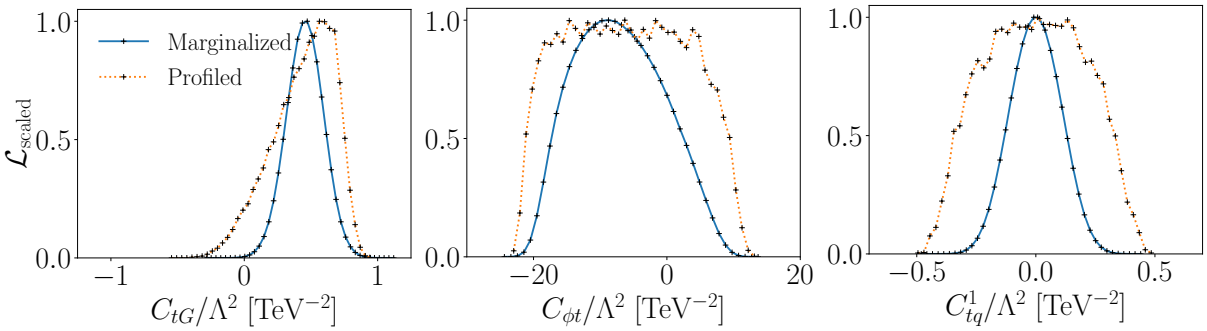


Figure 5.5: Comparison of marginalized and profiled results of the individual likelihoods for a few Wilson coefficients.

Looking at these plots, one finds that most of the profiled results show very flat likelihoods for a large range of values of the WCs, while the results after marginalization appear much more Gaussian with a distinctive peak. As a result, the marginalized posteriors are not as wide and the constraints on the coefficients become much tighter than they were before. This is the case for all the different WCs whose constraints are improved, not just for those shown here. To understand this, one needs to once again go back to the discussion of the two different treatments in SFitter from Section 3. It was already shown in that section that the convolution of a single Gaussian uncertainty with multiple flat uncertainties leads to very Gaussian results after marginalization, while the profiled results stay flat and follow, what is sometimes called the RFit prescription [44]. This clearly shows that the theory uncertainties have a much larger impact on the results in the profiled case, while for the marginalized results they do not.

The plots in Figure 5.5 show very clearly that the greatest effect these uncertainties have on the results are the flat regions in the profiled results, which disappear after marginalization. This does raise the question, whether a uniform distribution for the theory uncertainties is a reasonable assumption, since, if one were to perform the exact same fit with Gaussian theory uncertainties, one would expect these flat regions to disappear. To confirm this and show that this is not just an assumption, the previous fit was repeated, this time modeling all theory uncertainties as Gaussian and, which is the crucial difference here, profiling the NPs. Figure 5.6 shows the results from this fit and it becomes immediately apparent that there are no visible differences between the results obtained from either profiling or marginalization methods.

Consider that, after seeing the results in Figure 5.6, one would rather model the theory uncertainties as Gaussian since this gives results much more similar for the two statistical treatments. Since the distributions used to describe these uncertainties is only an assumption, one is technically free to do just that. Conversely, in the case of systematic uncertainties, where the constraint terms on the NPs can be understood as the results of some kind of auxiliary measurement, there is a statistical interpretation for these constraints. Finding an intuitive way to interpret and motivate the constraints on theory uncertainties is much more difficult, however. For this reason, one cannot claim any choice to be more reasonable than another.

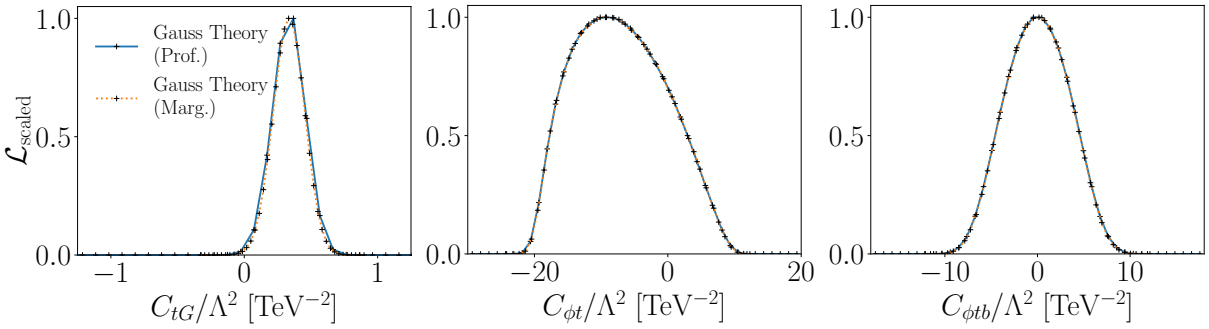


Figure 5.6: Profiled (blue) and marginalized (orange) likelihoods for Gaussian theory uncertainties.

In the end, this is a choice that has to be made at some point, and the reasoning for choosing them to be flat within SFitter has already been sufficiently explained in Section 3.2. Naturally, this does not mean that absolutely any kind of prior/constraint should be used, since the final results still needs to lead to realistic physics. The rest of this work will continue modeling the theory uncertainties as flat, although, unless otherwise specified, all results following this will come from marginalized fits.

With this, it was shown that profiling and marginalization agree with one another in the Gaussian case. What does not immediately become clear from this, is the effect that the modeling of theory uncertainties has when applying the marginalization to both of these. For this, the exact same fit was repeated, setting this time all theory uncertainties to be uncorrelated Gaussians. A few of the WCs affected the most by this are shown in Figure 5.7 and one finds that the greatest differences are found for C_{tG}, C_{tW}, C_{bW} and $C^3_{\phi Q}$. All the other WCs not listed here do not exhibit any sizable differences between the two fits. This shows that the choice of flat uncertainties does still have an effect on the results. Looking at them more closely, one finds that the constraints themselves do not get any stronger, mostly resulting in slight shifts of the peak closer to the SM.

To further emphasize the importance of theory uncertainties in the Top sector, another fit is performed, this time completely removing any of the theory uncertainties from the fit. Figure 5.8 shows the constraints on all WCs where one can immediately see a few strong differences between the two.

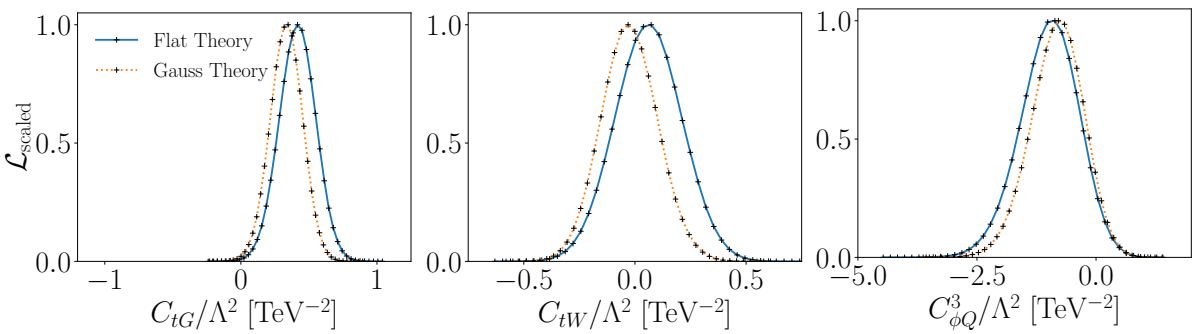


Figure 5.7: Results from marginalized fits for either Gaussian or Flat theory uncertainties.

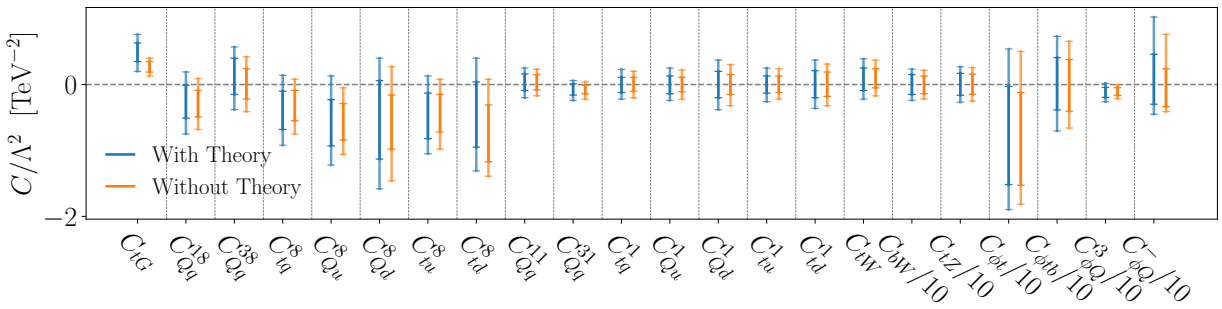


Figure 5.8: 68% and 95% constraints on the WCs in the Top sector for a marginalized global fit with full or no theory uncertainties.

Almost all WCs show significantly improved constraints, which is a result of the smaller uncertainties on the measurements, which gives them stronger constraining power. For the four-fermion octets, for example, one can see that the removal of theory uncertainties leads to constraints that exclude the SM much more strongly. Especially in the case of C_{td}^{8} one finds that there is a strong pull away from the SM which can usually be accommodated by large enough theory uncertainties.

The important role of theory uncertainties in the global Top fit has already been hinted at in Section 5.1 during the implementation of the total rate measurements using likelihoods. There they had to be neglected completely to even allow a proper comparison for some of the measurements. The usual total systematic uncertainty of a measurement was shown to be around the order of 5% while the theory uncertainties go up to $20 - 25\%$, clearly much more significant than the systematic, not to mention the statistical ones. These large theory uncertainties on the $t\bar{t}$ total rate measurements also explain why the constraints of C_{tG} improve so much after they are removed.

This is in stark contrast to the previous study using the Higgs, Di-Boson and EWPO dataset in Ref. [16], where it was shown that removing theory uncertainties from the fit did not have a significant impact on the results. It can, however, be easily explained due to the difference in size of the theory uncertainties of the two different datasets. For Higgs measurements they varied in the range of about $5 - 10\%$, already much smaller and more comparable to the size of its systematic uncertainties. In the case of the Higgs, higher accuracy predictions are much easier to compute since it does not interact strongly and as such one does not run into the problem of computing complicated QCD corrections, whereas any prediction involving the top quark requires the simulation of QCD processes for which the determination of higher order effects can lead to highly uncertain predictions.

5.2.2 Analyzing the new Dataset

The goal of this section is the implementation of the new differential cross section measurements from CMS, as explained in Section 4, into the dataset used for the final global fit. It will be structured in such a way that, first, the $m_{t\bar{t}}$ distribution will be analyzed. After that, the exact same studies will be repeated for the $p_T(t_h)$ distribution to allow for an easy comparison of the two, in the final part of this section.

The $m_{t\bar{t}}$ distribution

Before beginning with the global fit, a study of each distribution on its own can be of interest. Naturally, this means that not all of the operators listed in Table 2.2 can be constrained from just this $t\bar{t}$ measurement alone, since it is only affected by $\mathcal{O}_{tG}$ and the different four fermion operators. Out of the total 15 different operators contributing to $t\bar{t}$ pair production, only a subset of these operators will be considered for now. Since $t\bar{t}$ measurements on their own can not distinguish between any of the different four fermion operators, a total of three different WCs will be considered in the following fits. These are C_{tG} and one of the color octet four fermion operators C_{tq}^{8} and the corresponding color singlet C_{tq}^{1}. After confirming that these individual distributions give reasonable results on their own, one can move on and implement them into the global fit.

Beginning with the m_{tt} distribution, it was discusssed in Chapter 4 that it is an update to a previously implemented measurement published in Ref. [55]. Because of this, the $m_{t\bar{t}}$ distribution implemented in Chapter 4 will be referred to as the 'new' measurement, while the one from Ref. [55] used in the previous global Top fit, will be referred to as the 'old' one from now. Since this means that these measurements are not independent, the new one will replace the old one in the future. Due to this, the constraints one gets from the old measurement will serve as a good benchmark to validate the implementation of the new measurement. For this purpose, one fit to the three WCs listed above is performed for only the old and new measurements each. For both fits the NPs are marginalized and the resulting two-dimensional posteriors are shown in Figure 5.9.

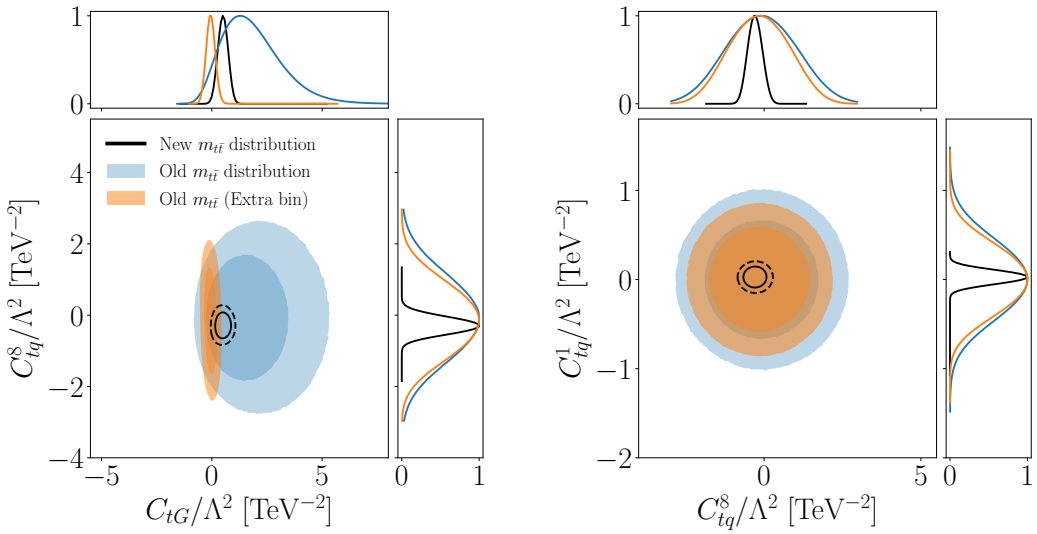


Figure 5.9: Two-dimensional marginalized constraints from a three-dimensional fit to the $m_{t\bar{t}}$ distribution from either Ref. [19] (black) or Ref. [55] (blue). Constraints after reimplementing the first bin from Ref. [55] are shown in orange.

The much stronger constraints, shown here in black, come from the fit to the newly implemented $m_{t\bar{t}}$ distributions, while the contour in blue shows the results from the fit to the old $m_{t\bar{t}}$ distribution. In both plots one can see very clearly that the constraints on all three coefficients C_{tq}^{1}, C_{tq}^{8} and C_{tG}, improve significantly. While stronger constraints

for these are expected due to the higher precision and wider energy range of the new measurement, the improvement for C_{tG} is very large, considering that normalized distributions are not expected to have too much of an effect on these. To explain this, one needs to understand the orange contour shown in the plot, for which it is necessary to explain the normalization of data in SFitter.

For every single distribution in the dataset the normalized data, as it is provided by the experimental papers, is implemented directly into SFitter. The key to understanding the differences in the strength of the constraints from the previous plot lies in how SFitter deals with this normalized data. It takes every single bin of these measurements and normalizes them again, by taking the sum over all the bins of the distribution included in the fit, and divides each of them by this total value. This means, that if the full normalized distribution is considered, this adds up to exactly 1 and the data does not change at all. This is the case for almost all the different distributions used in the previous fit, except for the old measurement being replaced here. In the case of this measurement only 7 out of the 8 total bins were implemented, which means that when normalizing the data the total adds up to less than 1, leading to larger values in each bin after once again normalizing the data. These larger values of the cross section require larger values of C_{tG} for the predictions to match the data, consequently leading to weaker constraints of C_{tG} as seen in Figure 5.9.

The bin that was removed from the old $m_{t\bar{t}}$ distribution was the very first one, which is one of the bins that contributes the most to the total cross section and as such also to the normalization. Because of this, it should lead to stronger constraints on C_{tG} when included in the fit. To check the effect that removing this bin has on the constraints, the first bin of the old measurement is reimplemented into SFitter using the data provided in Ref. [55]. The results from the fit to this distribution including all eight bins are shown in Figure 5.9 in orange. The effect of reimplementing the first bin can easily be seen by comparing the blue constraints with those in orange. Immediately, one finds that the constraints on C_{tG} change drastically, lowering the range of its allowed values. On the other hand, the constraints on C_{tq}^1 and C_{tq}^8 only show very slight improvements of their constraints, since their constraints are mostly driven by the high energy bins.

Moving on to the comparison with the constraints from the new measurement in black, one can see that the constraints agree much better once every bin is included in the fit. The improvements are much closer to what one expects from a normalized distribution. The constraints on C_{tG} now show almost no change, although there is still a slight pull away from the SM for the new measurement. Furthermore, one can also see that the constraints on the four fermion operators still show very significant improvements. This also follows our expectations since, as has been mentioned several times now, most of the constraining power for these comes from the tails of the distributions.

Whether one includes the additional bin to the old $m_{t\bar{t}}$ distribution does not matter for any of the future fits, however. In all following fits including the old $m_{t\bar{t}}$ distribution there is a large number of additional $t\bar{t}$ distributions. It was found that once all other $t\bar{t}$ measurements are included in the fit, the old $m_{t\bar{t}}$ distribution does not have any effect on the constraints at all. This means that regardless of whether the extra bin is included,

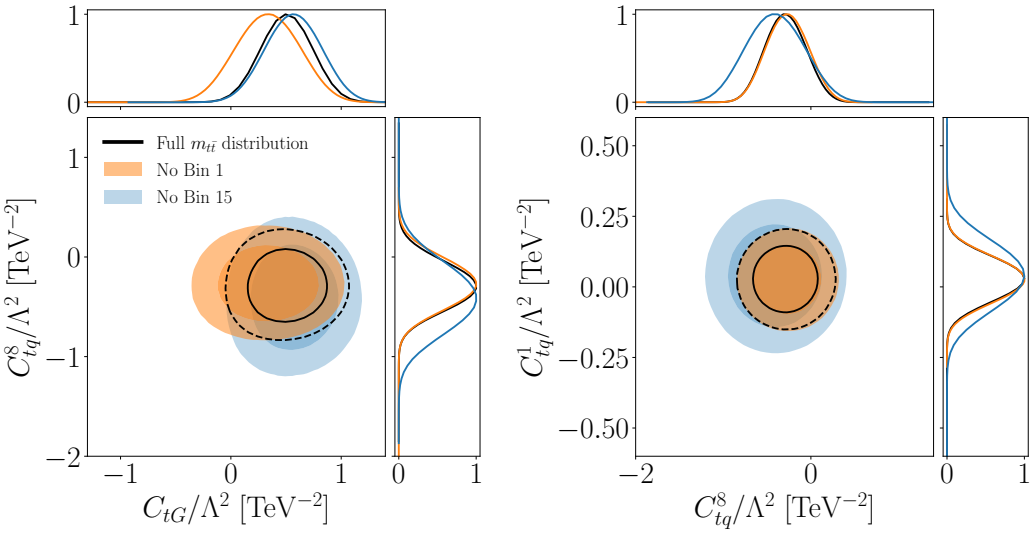


Figure 5.10: Two-dimensional marginalized constraints from a three-dimensional fit to only the newly implemented differential $t\bar{t}$ measurement. Different colored contours shows constraints after removing the bins given in the legend.

there is no difference in the final results. For this reason, all fits which include the old $m_{t\bar{t}}$ from now will not include the additional bin for simplicity sake.

With the benchmarking fit done, it was shown that the new $m_{t\bar{t}}$ distribution gives reasonable constraints for a normalized $t\bar{t}$ distribution. It only slightly affects the constraints of C_{tG} while giving much stronger constraints for the four fermion operators. Considering the strength of these constraints in this fit suggests that one can expect significant differences in the global fit as well. Before looking at this, however, a more detailed look at where these constraints come from is done by performing the exact same fits again, only this time removing certain bins from the new $m_{t\bar{t}}$ distribution beforehand. The results of a few such fits can be found in Figure 5.10.

Here, the black outlines show the constraints from the full fit, while the different colored contours show the constraints after removing different bins from the fit. Starting first with the contours shown in orange, which show the constraints after removing the very first bin from the distribution. Comparing this to the full distribution in black, one can see that there is a slight pull of C_{tG} closer to the SM, while also allowing for much smaller values of C_{tG}. This comes from the fact that there is an underfluctuation in the first bin of the distribution, as one can see in Figure 4.3. To accommodate this, larger values of C_{tG} are required, which, in turn, means that after removing this bin smaller values of C_{tG} are allowed. On the other hand, the removal of the final bin, shown here in blue, only leads to very small changes in the constraints of C_{tG}. Compared to this, the four fermion operators show the exact opposite behavior. The removal of the first bin does not change the constraints whatsoever, while removing the final bin leads to weaker constraints for both of them, once again confirming their kinematically enhanced nature.

Although Figure 5.10 only shows the effect of removing the first and last bins, respectively, the same analysis was also performed for all other bins. The results from these

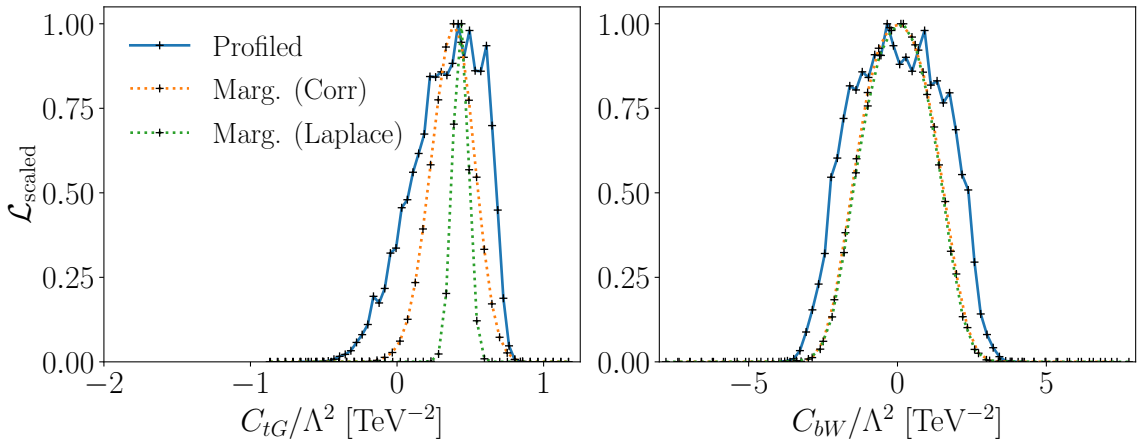


Figure 5.11: Comparison of the one dimensional profiled (blue) and marginalized likelihoods using either the correlation matrix (orange) or Laplace method (green) to include correlations.

fit are not shown here, however, since they show no significant deviations from the full distribution. This shows that the constraints on the four fermion operators is driven completely by the final bin of the distribution, while the first bin constrains C_{tG} the most. All of this confirms that the newly implemented measurement is well behaved and the final step is to include the new distribution into the complete dataset and perform a global fit of the entire space of operators.

The global fit is performed using the entire dataset from Ref. [15] after removing both the old $m_{t\bar{t}}$ and $p_T(t_h)$ distributions and adding this new $m_{t\bar{t}}$ distribution. This dataset will be referred to as the 'new' dataset in the following, similarly the dataset from the global Top fit in Ref. [15] will be referred to as the 'old' dataset. Before taking a look at the updated constraints of all WCs, it is necessary to look at the results of a profiled and marginalized likelihood fit to the new dataset of a few of the most interesting WCs in Figure 5.11. Unlike any previous study, these plots also show the results when using the different methods of introducing correlations between the measurements. In green are correlations included using the Laplace method from Section 3.2 and in orange those using the correlation matrix method, labeled as 'Corr'. No specific comparisons between these two methods were shown before, since their results always agreed for all fits discussed up to this point.

Only the two most extreme cases for the two different methods are shown here. Comparing the two correlation methods, one finds that in the right plot the constraints agree completely for both, while the plot on the left the constraints using the Laplace method show very large differences with a strong pull away from the SM. Clearly the new measurements has some sort of effect on these correlations, since none of this has been observed for any other fit before this. Taking this into account, the new measurement was studied in more detail to find the possible cause for these differences. Unfortunately, it was not yet possible to draw any definitive conclusions. This could either be an interesting property of the newly implemented measurement or a simple numerical problem whose cause has yet to be found. Either way, this will need to be studied further in the near future.

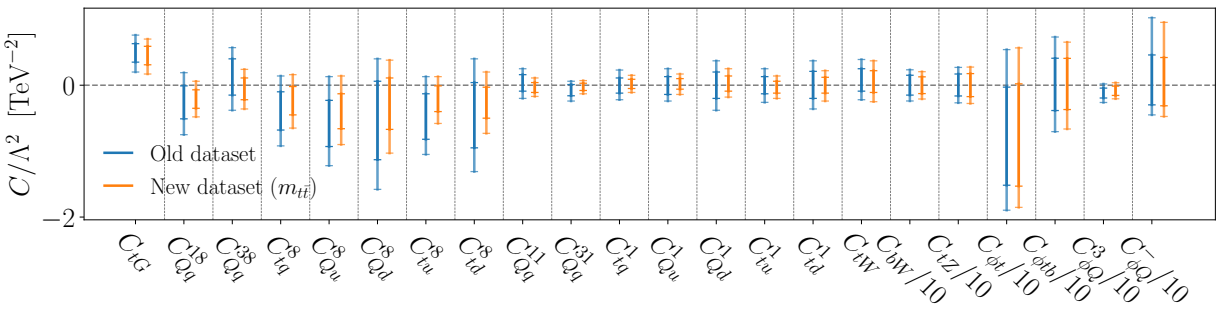


Figure 5.12: 68% and 95% CLs from a marginalized global SMEFT fit in the Top sector for the dataset in Ref. [15] (blue) and the new one including the $m_{t\bar{t}}$ distribution (orange).

For now, all future fits will use the correlation matrix method, unless specified otherwise, since the results obtained from these are much more conservative. At the same time, the improvement of the constraints between the profiled and the marginalized fit using the correlation matrix in Figure 5.11 look much more like those seen in Figure 5.4 and Figure 5.5.

Using the results from the global fit with the correlation matrix method, corresponding to the orange curves shown in Figure 5.11, their constraints are compared to those from the fit to the old dataset in Figure 5.12. Again, both fits make use of the marginalization of NPs. After the analysis of the new $m_{t\bar{t}}$ distributions on its own, the greatest differences one expects to see are in the four fermion operators. Looking at the results confirms exactly those expectations. All of these operators show much stronger constraints compared to before for both the singlets and octets.

Naturally, just like in the old dataset, the constraints on the color singlets are still stronger than those of the octets since the constraints from $t\bar{t}$ measurement scale with different color factors. Since the $t\bar{t}$ measurements alone does not distinguish between the different four fermions operators, they also show very similar improvements for all four fermion operators. The constraints for C_{tG} only display a very small shift, easily within the one sigma range and as such not significant. This is exactly what one expects, since C_{tG} is mostly constrained by the total rate measurements already in the dataset.

All other operators in the fit do not have any direct contributions to the $t\bar{t}$ measurement, meaning that one would not expect these to be constrained at all from these, if only $t\bar{t}$ measurements were included in the fit. Due to the global nature of this fit, however, an interplay between the WCs affecting different processes in the dataset can lead to correlations between WCs, which help further constrain these. One such coefficient not directly affected by the $t\bar{t}$ measurement but still showing improvements is $C_{\phi Q}^{3}$. This is a nice example of the strength of a global fit which allows these WCs to be affected due to correlations. Naturally, the correlations for both the old and new dataset were studied as well to see what kind of differences one can find with some handpicked examples shown in Figure 5.13.

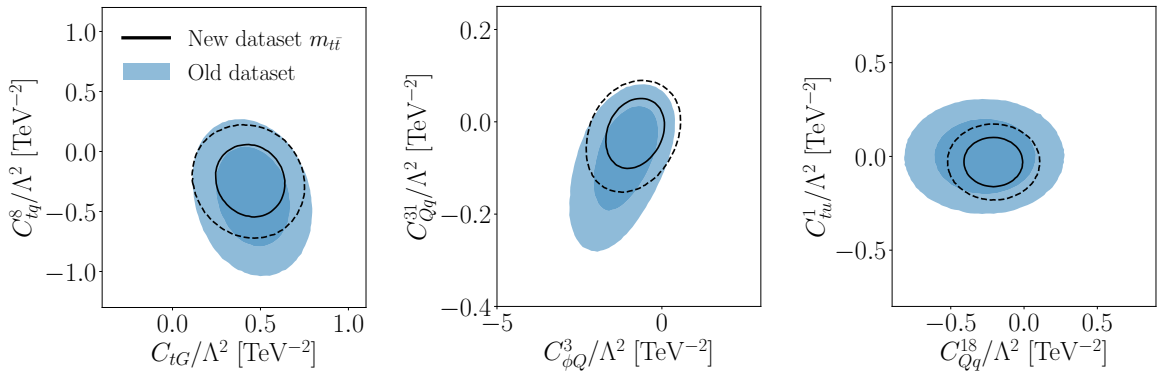


Figure 5.13: Comparison of the correlations of a few WCs from the marginalized global Top fit using the old dataset (blue) and the new one including the $m_{t\bar{t}}$ distribution (black).

The three different plots were chosen to show the differences for WCs with anti-correlations, correlations and no correlation at all. Once again, the black outlines show the results from the new dataset, while those from the old one are given in blue. Just like before, these plots make the stronger constraining power of the new dataset apparent. At the same time, one can also see that the correlations between the different WCs do not change significantly. This is not the case just for the coefficients shown here, all other two-dimensional contours show very similar results, not deviating a lot from any of the results from the fit to the old dataset.

The $p_T(t_h)$ distribution

To allow for a direct comparison of the different observables implemented, the same studies as those for $m_{t\bar{t}}$ will now be performed again for the $p_T(t_h)$ distribution, starting with the fit to determine the most constraining bins of the distribution. In Figure 5.14 the constraints from a fit to the full p_T distribution are shown compared to constraints after removing either the 6th or 16th bin from the measurement. Only the constraints from removing these bins are shown, since all other bins do not affect the constraints in any significant way. Both lead to very strong pulls, either away from the SM in the case of the 16th, shown here in blue, or pulls much closer towards the SM, shown in orange, when removing the sixth bin.

To understand where these strong pulls come from, one simply needs to take a look at the χ^2 values for these bins after setting all WCs to be exactly 0, i.e. the SM predictions. In Table 5.4 one can clearly see that there is already a very strong tension between the SM predictions at NNLO and the measured data from the Ref. [19], especially for the third, sixth and 16th bins. This tension between prediction and data can easily explain why removing the sixth bin from the distribution leads to constraints much more consistent with the SM. All other bins not listed in the table show very good agreement with the SM, with χ^2 values much smaller than 1.

Clearly the systematic uncertainties are too small to explain the discrepancies one can see in Table 5.4 and the theory uncertainties, although larger by about one order of magnitude, are still too small compared to the differences between data and prediction, to

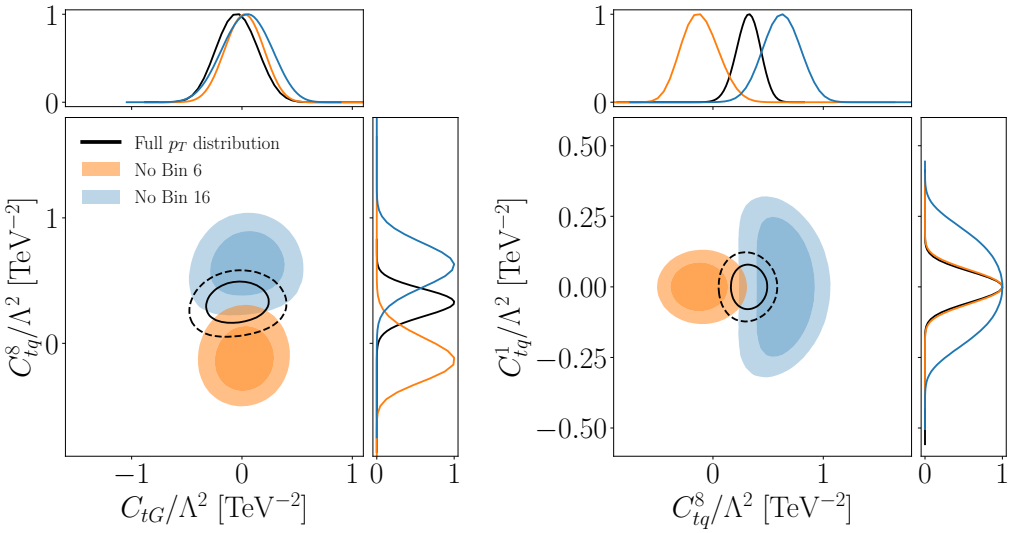


Figure 5.14: Marginalized constraints from a 3 operator fit to only the new $p_T(t_h)$ distribution after removing the bins specified in the legend.

do so either. The tension between the SM predictions and the data was not present in Section 4.2 when the prediction calculated using MadGraph were compared to those from Ref. [19]. This means that the source of this discrepancy is the calculation of the k factors for the NNLO predictions. Any discrepancies coming from this calculation are usually covered by the uncertainty from the choice of event generator, where the envelope of the MC generator uncertainties was used to get a conservative estimate of these. Regardless, the χ^2 values suggest that these first bins are not very well described and as such can not be included in the fit like this.

There are two different ways of fixing this issue. One can either include an additional theory uncertainty coming from the NNLO predictions, or simply remove the badly modeled bins from the fit. For the purpose of this analysis, it was decided to remove these bins from the fit completely. This was done for multiple different reasons. First, an additional theory uncertainty introduced to account for the tension between the NNLO SM predictions would have to be quite sizable. This would, in turn, already lower the constraining power coming from these bins dramatically, which is why one would not expect the constraints to be affected much by these bins anymore.

Bin# [Range]	Data	Prediction	σ_{syst}	σ_{theo}	χ^2
3 [80-120 GeV]	5.937e-03	5.357e-03	3.924e-05	2.253e-04	135.21
4 [120-160 GeV]	4.291e-03	4.580e-03	3.631e-05	1.896e-04	14.908
5 [160-200 GeV]	2.639e-03	2.837e-03	2.529e-05	1.312e-04	13.626
6 [200-250 GeV]	1.376e-03	1.118e-03	1.648e-05	1.189-04	168.46
16 [950-1100 GeV]	2.745e-07	5.318e-07	6.768e-08	9.933e-08	5.6418
17 [1100-1600 GeV]	3.422e-08	2.235e-08	1.083e-08	1.477e-08	0

Table 5.4: χ^2 values for SM predictions at NNLO and data taken from Ref. [19]. Corresponding theory and systematic uncertainties are also listed.

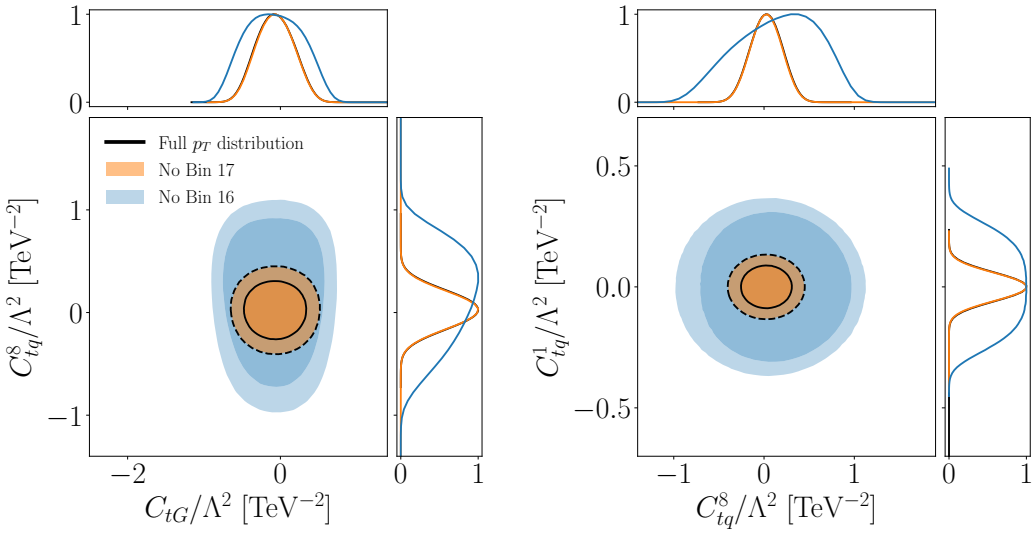


Figure 5.15: Marginalized constraints from a 3 operators fit to only the new $p_T(t_h)$ measurement after removing all the first six bins and the additional bins given in the legend.

The second reason is that kinematic distributions like this mostly constrain the kinematically enhanced operators, which means that the constraints one is interested in should only come from the last few bins for the most part. For all the reasons stated above, the first six bins of the p_T distribution will be removed from all following fits.

Having decided to remove the first six bins, one can repeat the fit of only C_{tG}, C_{tq}^8 and C_{tq}^1 to check which bin drives the constraints now that the number of bins has been reduced this much. Figure 5.15 shows the results from this fit as well as those from fits after removing the last and second to last bin, respectively. Starting with the constraints from a fit after the final bin was removed, shown in orange, one finds that those constraints agree perfectly with those from the full p_T distributions shown in black. Moving on to the results from the fit without the 16th bin, shown in blue, they once again show a very strong effect on the constraints. Just like for the $m_{t\bar{t}}$ distribution before, one finds that there is a slight improvement in the constraints of C_{tG} with a much more significant improvement for both four fermion operators. This clearly shows that it is the 16th bin that is driving the fit. To understand why it is the second to last bin driving the constraints this time, a quick look at Table 5.4 can help. Comparing the relative size of the uncertainties in the final bin with those in the 16th, one finds that the uncertainties in the final bin are much larger. Due to this, the constraining power of the final bin is reduced and the fit is mostly driven by the 16th bin. All other bins were also checked and have shown not to affect these constraints, consistent with their χ^2 values which are all much smaller than one.

With this, one can now move on to the implementation of the new p_T distribution into the full dataset from Ref. [15]. Once again, the dataset is updated by removing both the old $m_{t\bar{t}}$ and $p_T(t_h)$ distributions and replacing it with this new one. The complete global SMEFT fit can be repeated using this new dataset. Just like for the fit of the $m_{\bar{t}t}$ distribution the fit is performed, using the marginalization methods implementing correlations between the systematics using the correlations matrix method. The resulting constraints

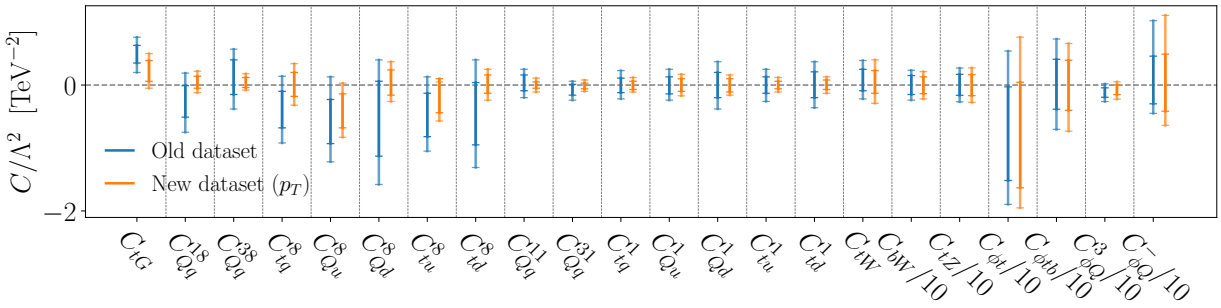


Figure 5.16: 68% and 95% CLs from a marginalized global fit to the old dataset from Ref. [15] and the updated dataset including the p_T distribution.

on the different WCs can be seen in Figure 5.16.

Starting, once more, with the constraints on the four fermion operators, they show, just as one would expect, much stronger constraints compared to before. All of their constraints are also still consistent with the SM. For now, the strength of the constraints relative to those from the previous $m_{t\bar{t}}$ distribution will not be discussed further, leaving this for the following section. In the case of C_{tG} one finds for the 95% CLs that

$$C_{tG}/\Lambda^2 \in [0.20, 0.76]/\text{TeV}^2 \longrightarrow C_{tG}/\Lambda^2 \in [-0.05, 0.50]/\text{TeV}^2\,, \tag{5.3}$$

In other words the constraints do not get stronger, but they are pulled much closer to the SM. This can be seen as a consequence of the removal of the first 6 bins of the distribution, considering the pull away from the SM one saw in Figure 5.14. For all other coefficients there are mostly small differences and one finds that there are actually weaker constraints for $C_{\phi t}$ and $C_{\phi Q}^{-}$ in particular. Since these are not affected by the new p_T distribution directly, studying the correlations between WCs can help understand this.

The final results discussed here show the correlations between the different WCs and can be found in Figure 5.17. For the first time, one can immediately see differences between the correlations of the WCs shown. In the left plot one can very clearly see the pulls of C_{tG} closer towards the SM, at the same time the correlations between C_{tG} and C_{tq}^{8} are reduced. One can see that they do still show an anti-correlation, which looks especially small since constraints from the new dataset shown in black are not to scale with those blue contours from the old dataset. The plot in the middle shows the exact same behavior for $C_{\phi Q}^{3}$ and C_{Qq}^{31}, showing weaker correlations between them after the p_T distribution is included in the dataset. The final plot shows, as an example, the effect for two WCs where the correlations do not change. In general one finds that for a large number of different WCs their correlations decrease. These smaller correlations consequently imply less interplay between the different WCs reducing the constraining power in the fit. This explains why the constraints on both $C_{\phi t}$ and $C_{\phi Q}^{-}$ become weaker after the new p_T distribution is included in the fit.

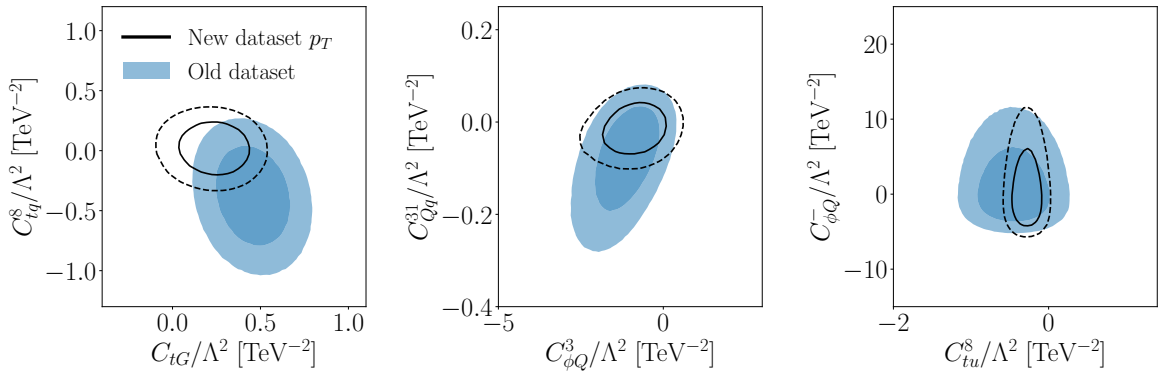


Figure 5.17: Correlations between a few of the WCs from the marginalized global SMEFT fit using the new dataset inlcuding the p_T distribution.

Comparing $m_{t\bar{t}}$ and $p_T(t_h)$

Now that the two different observables have been studied in great detail separately, the following section will provide the most important results by comparing them. Before doing so, it is necessary to make sure that both distributions are compared on the same footing, however. In the previous section the full $m_{t\bar{t}}$ distribution, i.e. all available bins, were implemented into the dataset, while for the p_T distribution the first six bins had to be removed. Throughout this thesis, it has already been shown multiple times that the energy range covered by these distributions very significantly impacts the results one obtains. Because of this, the following will give a comparison of full global fits using either the $m_{t\bar{t}}$ or p_T distributions after removing the first 6 bins for both of these distributions.

Figure 5.18 shows the final results from both different datasets. In blue the results from the global fit including the $m_{t\bar{t}}$ distribution is shown, while those from $p_T(t_h)$ are given in orange. Starting with C_{tG} this time, one finds that there is a slighlty stronger bound on the positive values of C_{tG} coming from the p_T distribution. For $m_{t\bar{t}}$ one finds that after removing the first 6 bins from the fit, its constraints also move closer to the SM, just like the values from p_T did. This can be explained by the underfluctuation in the first bin, one can see in Figure 4.1, just like it did for p_T. Continuing with the four fermion operators, the constraints coming from the p_T distribution are much stronger than those from $m_{t\bar{t}}$. As before, one can see that to compensate for the larger values of C_{tG} for $m_{t\bar{t}}$ there is a shift of all four fermion operators towards smaller values. The remaining WCs do not show any differences, which is to be expected since they did not show significant changes from the fit to the old dataset.

Looking at these results, one finds that using p_T as the observable leads to constraints quite a bit stronger than $m_{t\bar{t}}$. When it comes to trustworthyness of the final results, however, those from $m_{t\bar{t}}$ come out on top. Comparing the amount of issues one ran into when implementing either p_T or $m_{t\bar{t}}$ the latter proved to be much easier to implement. The main issues arising for p_T came from the determination of the NNLO k-factors which introduced very strong discrepancies between the SM predictions and the measured data. Compared to this, the k-factors for $m_{t\bar{t}}$ were much easier to compute and did not introduce any additional problems.

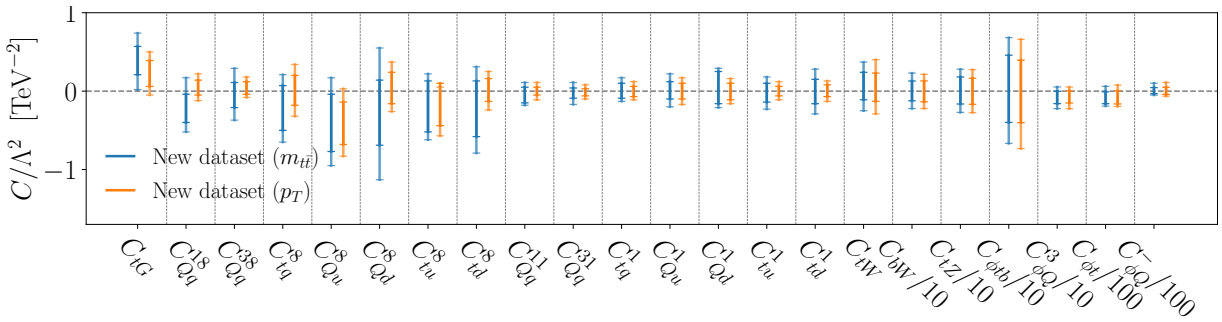


Figure 5.18: 68% and 95% CLs from a marginalized global fit to the new dataset, including the $m_{t\bar{t}}$ (blue) or p_T distributions (orange). The first six bins were removed for both of the distributions to cover similar kinematics ranges.

The conclusions to draw from this, is that while choosing p_T as the observable can lead to stronger constraints, it is much harder to get a handle on the corresponding predictions which, in turn, leads to more issues when implementing the distribution. Because of this, more work would need to be done to validate the implementation of the p_T distribution. On the other hand, the results from the $m_{t\bar{t}}$ distribution, while not quite as constraining, come from a much cleaner implementation where no troublesome bins had to be removed. For this reason, the $m_{t\bar{t}}$ distribution can be added to the dataset in its current state, while p_T distribution still requires a bit more care first.

5.3 Combined global fit

After the previous two parts of this thesis have been dedicated to studies of only the Top dataset, the final part of this thesis will be concerned with the combination of this Top dataset with data from the Higgs, Di-Boson and EWPO dataset. For the Top sector, the same dataset given in Ref. [15] is once again used, excluding the newly implemented $t\bar{t}$ measurement, while for the latter the complete dataset of Ref. [16] is used. Naively, one might expect the combination of these datasets to be trivial. Why not simply copy the data input from the one dataset into the other and just tell SFitter to fit all operators from both Table 2.2 and Table 2.3? The following will explain exactly why this is not possible, what issues one runs into and how these are dealt with for this analysis.

Reconciling Higgs and Top

To combine the two different datasets there are three distinct parts within SFitter that need to be adjusted. Firstly, there is the data from the experimental measurements that need to be correlated properly. Then there is the model considered, i.e. the list of operators constrained in the fit. Finally, there are the predictions one needs to look at, for reasons explained later.

We begin by properly including the correlations between the measurements of the different datasets within SFitter. In Section 4.1, DataPrep was introduced, a tool which automatically created inputs in the SFitter format while also making sure that the correct systematics are correlated. It was used to create the data input for the Top dataset

from Ref. [15] but has never been used for the fits using the Higgs, Di-Boson and EWPO dataset before. This meant that a few adjustments had to be made to DataPrep to make sure that the correct systematics are still correlated. To this end, it was necessary to consider the, up until now, separate tables of systematics used for the two different fits, given in the Appendix in Table 8.1 and Table 8.2.

Since none of the SFitter inputs used for the global Higgs, Di-Boson and EWPO data were created using DataPrep before, none of the inputs required to create them using it existed yet. This would mean that if one wanted to reproduce the same data used for those via DataPrep, new input files in the DataPrep format would be required for each and every single measurement. Because this would be more work than necessary, DataPrep was adjusted in such a way that it can create input files for SFitter compatible with the ones previously used for the Higgs, Di-Boson and EWPO fit. Doing this ensured that the systematics both datasets have in common are properly correlated, while all others stay independent of one another.

With this, the issue of the measurements and their correlations has already been taken care of. Continuing, the next step will be concerned with the combination of the two different sets of operators constrained in the fits. In Table 2.3 all the different operators considered in the previous Higgs, Di-Boson and EWPO analyses have already been listed. Naively, one might want to simply add these operators to those from the Top fit given in Table 2.2. This is where one small detail, which was already hinted at in Section 2.2.2, comes into play. The operators considered in the global Top fit are given in the Warsaw basis, while those for the Higgs, Di-Boson and EWPO dataset are given in the so-called HISZ basis. Naturally, one should not use two separate bases in the same fit, which is why it is necessary to express one of these bases in terms of the other. Since the Warsaw basis is the most commonly used SMEFT basis used nowadays, the rotation from the HISZ basis to the Warsaw basis was implemented into SFitter. This was done using expressions already derived in Ref. [15] which can also be found in the Appendix.

With this, the issue of the different bases has already been taken care of. There is, however, one final problem that had to be dealt with, before a first fit could finally be performed. As has already been mentioned, when the operators for the two different sectors were introduced, the number of different operators considered was reduced via additional flavor symmetry assumptions. In the case of the Top fit, this was a $U(2)_q \times U(2)_u \times U(2)_d$ flavor symmetry, while in the Higgs case a universal flavor symmetry was assumed. Clearly a direct combination of the two sectors without any additional considerations does not make any sense like this since some of the predictions previously calculated depend on these assumptions.

Within the dataset considered, the only measurements actually affected by these assumptions are those for the EWPOs. One is now left with a choice. One can either recompute the predictions for these, using the same flavor assumptions as in the Top fit, which would be far from trivial, or one can remove these measurements from the dataset. Since the largest correlations are expected between the Top and Higgs sectors, it was decided to remove all EWPO measurements from the dataset in the following analysis. This way, it is possible to first focus on the combination of Higgs and Top. Removing these leads to

some of the operators not being nearly as constrained anymore, especially those which had almost exclusively been constrained by EWPO data. The WCs affected by this are $C_{\phi D}, C_{ll}$ and $C_{\phi e}$, and have been removed from the fit completely for this reason.
The recalculation of the EWPO predictions using the flavor symmetry assumptions from the Top sector and their reimplementation into the dataset are left to a future work.

The combined fit

After all of these initial problems were solved, a first profiled fit of the combined dataset could now finally be performed. Since there is a total of 38 different WCs included in the fit, the resulting constraints will be split into two separate parts. In Figure 5.19 the constraints on the WCs which previously only entered the Top fit are shown. The blue constraints correspond to a global fit to only the Top dataset, while the orange ones give the constraints for the full combined fit of Higgs, Di-Boson and Top.

Comparing the constraints from the combined fit with those from just the Top dataset, it is quite clear to see that almost all the different WCs show at most very small, and definitely not significant, differences. The only coefficients which show visible changes are C_{tG}, where a very slight shift to higher values can be observed, and the four-fermion octets with a slight shift towards more negative values. The differences for the four-fermion operators can be understood as a result of the change in C_{tG}. Since all of these operators contribute to $t\bar{t}$ production, they are correlated and any changes in one can affect the other. Since C_{tG} is pulled to larger values, this needs to somehow be accommodated by the other coefficients, which leads to smaller or even negative values for the four-fermion operators. The only other strong difference can be found in $C^-_{\phi Q}$ which shows a slight shift towards more positive values, very similar to C_{tG}. Since $C^-_{\phi Q}$ is mostly constrained by $t\bar{t}Z$ which also constrains all the same coefficients as $t\bar{t}$, this is the result of the interplay of the $t\bar{t}$ and $t\bar{t}Z$ measurements where $C^-_{\phi Q}$ adjusts according to the changes of C_{tG} and the other four fermion operators.

The constraints for all remaining WCs for the Higgs, and Di-Boson fit are shown in Figure 5.20. Again the orange constraints correspond to the results from a fit to the combined dataset while the blue ones are those from the fit to the Higgs and Di-Boson data.

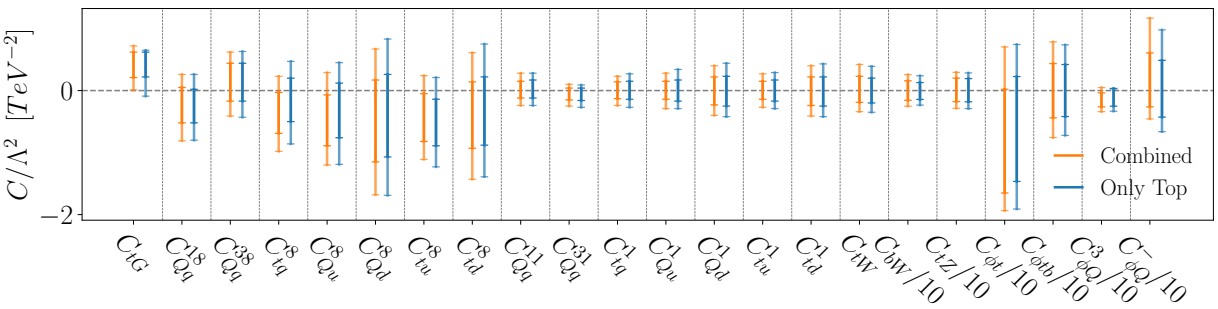


Figure 5.19: Comparisons of 95% and 68% CLs from the results of a combined fit with the constraints from only the Top dataset.

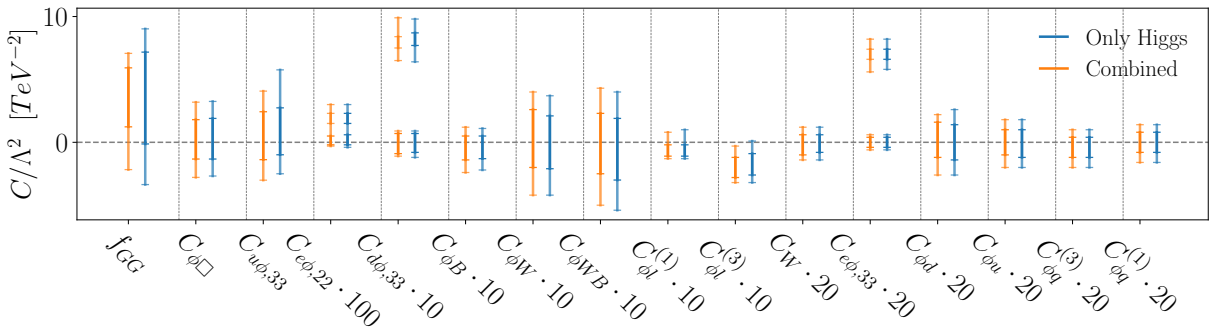


Figure 5.20: Comparison of 95% and 68% CLs from the results of a combined fit (orange) and the Higgs and Di-Boson dataset alone (blue).

This is the first time that a fit to the Higgs and Di-Boson data in SFitter was performed using the Warsaw basis. All constraints are given for the WCs in the Warsaw basis, except for f_{GG}. This is done to make it easier to visualize the constraints on the WCs side-by-side, since the numerical values of f_{GG} are much closer to those of the other WCs in the Warsaw basis, compared to its corresponding operator $C_{\phi G}$. This can be done since these operators only differ by a rescaling factor

$$C_{\phi G} = -\frac{\alpha_s}{8\pi} f_{GG} \,. \tag{5.4}$$

Because of this the shape of the constraints is not affected and the comparison between the two datasets still holds.

The only parameter not specifically shown in this plot is the branching ratio of the Higgs to some invisible final states $\mathrm{BR}_{\mathrm{inv}}$, which is why its constraints will be listed separately:

$$\begin{aligned} 68\%\ \mathrm{CL}: &\ \ \mathrm{BR}_{\mathrm{inv}} \in [0, 4.64]\% \rightarrow \mathrm{BR}_{\mathrm{inv}} \in [0, 5.3]\% \\ 95\%\ \mathrm{CL}: &\ \ \mathrm{BR}_{\mathrm{inv}} \in [0, 9.44]\% \rightarrow \mathrm{BR}_{\mathrm{inv}} \in [0, 9.5]\% \,. \end{aligned} \tag{5.5}$$

Beginning with comparison of the WCs shown in Figure 5.20 one can see that the majority of the WCs show no changes in their constraints. The only two WCs affected by the combination are f_{GG} and $C_{u\phi,33}$ which both show stronger constraints. The differences in $C_{u\phi,33}$ can be easily understood, considering its contributions to $t\bar{t}H$ production. Clearly $C_{u\phi,33}$ is one of the few operators contributing to processes affected by operators from both the Top and Higgs sector. The improvement in the constraints of $C_{u\phi,33}$ shows exactly why a global analysis leads to the best constraints for SMEFT fits.

Furthermore, both C_{tG} and f_{GG} are constrained by $t\bar{t}$, $t\bar{t}H$ production and also contribute to higgs productions via gluon fusion. Because of this, a strong correlation between the two is expected and this can also be seen in the left plot of Figure 5.21. This clearly shows the very strong correlations between these two operators, which explains why a combination of the datasets leads to such an improvement in the constraints of f_{GG}. As a comparison, the right side of Figure 5.21 shows the correlations from the Higgs, Di-Boson and EWPO fit from Ref. [16]. Clearly, the correlations between the two were already very large for this fit. To explain why the correlations were already this strong in this fit, it is necessary to once again talk about how the effect of C_{tG} was implemented into that fit.

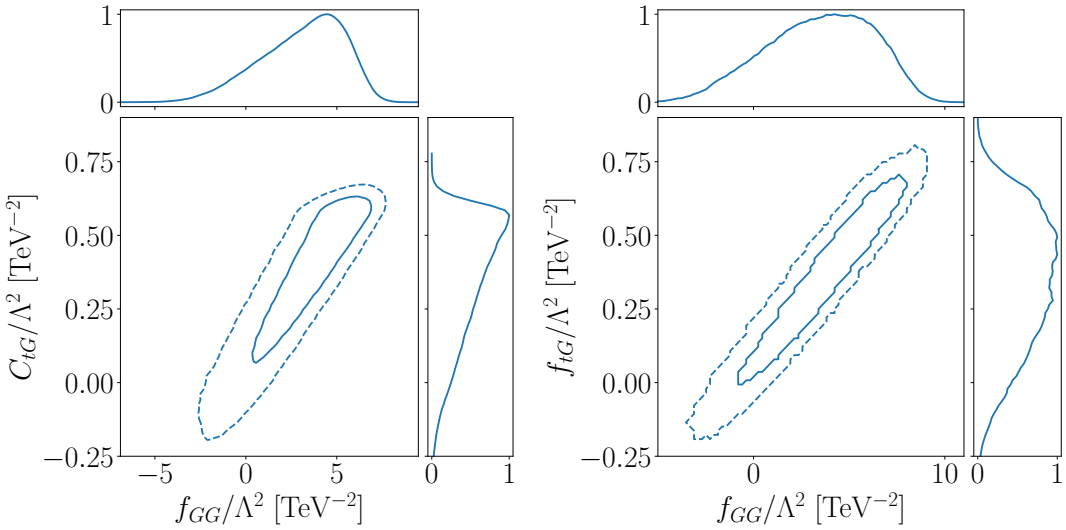


Figure 5.21: Two-dimensional marginalized results comparing the correlations between C_{tG} and f_{GG} from the combined fit (left) and the fit only using Higgs, Di-Boson and EWPO data (right), taken from Ref. [16].

In Section 2.2.2 it was already discussed that since the dataset of the previous Higgs, Di-Boson and EWPO fit included both $t\bar{t}H$ production processes and the production of the Higgs via gluon fusion, C_{tG} can not be neglected. To take this into account, a prior for C_{tG} is introduced as an additional constraint. This way C_{tG} was not a free parameter in the fit but constrained to values approximately following the constraints from the global Top fit in Ref. [15]. The right plot in Figure 5.21 shows that in this fit, the correlations between f_{GG} and C_{tG} have been shown to be significant already. These correlations are a result of the prior which approximated the constraints on CtG from the global Top fit of Ref. [15]. In the combined fit this prior is removed which means that the approximation is dropped and the correlations can be reassessed. The results from Figure 5.21 validates the approximation from Ref. [16].

Concluding, it was shown that the combined fit of using the data from the Top, Higgs and Di-boson dataset leads to the same constraints for most of the WCs in the fit. One finds that there are improvements for WCs such as $C_{u\phi,33}, C_{tG}$ and f_{GG} which all contribute to processes involving both the top quark and the Higgs. Furthermore, it was shown that correlations between f_{GG} and C_{tG} were already approximated really well in the previous Higgs, Di-Boson and EWPO fit by implementing the prior for C_{tG}.

6

Summary and Conclusions

The main goal of this thesis was the study of physics beyond the Standard Model by constraining WCs via global SMEFT analyses. To fully understand this, an introduction to the SMEFT and the higher dimension operators considered in the following fits was given. After this, SFitter was introduced, giving a summary of its most important properties and how it is used to determine the constraints on these WCs using either profiling or marginalization methods.

Before beginning with the global analyses, however, new measurements were studied and implemented into the Top dataset from Ref. [15]. We began with the implementation of total rate measurements for $t\bar{t}$ and $t\bar{t}Z$, for which likelihoods in the HistFactory format were published. It was shown that they allow key plots from experimental papers to be easily reproduced, using a python library called pyhf. Following that, the likelihoods were used to validate the assumption of Gaussian systematic uncertainties by studying its NPs, while also confirming that one can neglect the correlations between uncertainties within a single measurement.

After this, these published likelihoods were used to implement measurements into SFitter for the first time. Here, the total cross section for both $t\bar{t}$ and $t\bar{t}Z$ were added to the dataset from the previous global Top fit [15] and it was shown that the constraints on WCs do not get any stronger. This is due to the size of the systematic uncertainties of the new measurements, which are of similar size to those from the measurements already in the fit and as such do not lead to stronger constraints. This shows that the constraining power of the total rate measurements from Ref. [15] is already strong. This suggests that smaller systematic uncertainties are necessary to improve the constraining power coming from these total rate measurements. It is first necessary to improve the accuracy of the theory predictions, however, since it was seen that theory uncertainties are clearly dominant for total rate measurements, which means that more accurate experiments do not directly lead to better constraints unless the predictions improve as well.

Both measurements from published likelihoods studied in this thesis were total rate measurements. For a future work, it will be interesting to study the implementation of a differential cross section using these likelihoods. Using these, it should be possible to better determine the effect of the systematic uncertainties on each individual bin, since they were approximated for all measurements in Ref. [15]. Also, now that the ATLAS groups are starting to provide likelihoods for more of their measurements, giving access to more information from the experiment than before, it is the perfect opportunity for theorists and experimentalists to cooperate and continue this process towards extending the scope of information provided by experiments beyond these likelihoods published by ATLAS.

The next part of the thesis was dedicated to the study of constraints from a global SMEFT fit in the Top sector using profiled and marginalized likelihoods. It was shown that for flat theory uncertainties the profiled results give much wider flat likelihoods while the marginalization gave Gaussian likelihoods. This translates into stronger constraints for the marginalized likelihoods, once again displaying the importance of theory uncertainties in the Top sector.

Following this, a new differential $t\bar{t}$ measurements by CMS [19] was implemented to improve the constraints on C_{tG} and the four fermion operators. It was implemented choosing either $m_{t\bar{t}}$ or $p_T(t_h)$ as the observable. Starting with $m_{t\bar{t}}$ there were no issues when implementing this distribution and one found that the measurement lead to significantly stronger constraints for all four fermion operators while only very slightly affecting C_{tG}. Compared to this, the implementation of $p_T(t_h)$ distribution was more difficult, since the calculation of NNLO predictions lead to strong discrepancies between the measured data and SM predictions for some bins. This necessitated the removal of these troublesome bins, after which the same global fit as for $m_{t\bar{t}}$ was performed. For the fit using the $p_T(t_h)$ one observes an improvement for all four fermion operators, even more significant than for $m_{t\bar{t}}$ and a pull of C_{tG} closer to the SM.

Both distributions show strong constraining power for the four fermion operators coming from the tails of the distributions, while only p_T has a strong effect on C_{tG} due to the removal of a large number of its low energy bins. Taking these results into account, one finds that implementing some observables may be more constraining than others. One does, however, also need to consider that these can also prove much more difficult to include, requiring much more care before giving reasonable results. Regardless of whether one chooses $m_{t\bar{t}}$ or $p_T(t_h)$ as the observable, the SMEFT contributions from interference terms had to be neglected, since their contributions in high energy bins were too noisy to properly extract their effects. On top of that, we saw that using the Laplace method to introduce the correlations lead to significantly stronger constraints than expected. Both of these issues deserve more attention and will be of interest for future studies.

In the final part of this thesis, the measurements from the Top, Higgs and Di-Boson datasets were combined to perform a first combined fit using SFitter. It was shown that the combined fit lead to the same constraints for the majority of WCs considered in the fit. One does find improved constraints for the WCs directly affected by operators from both the Top and Higgs sector. Finally, it was shown that the implementation of a prior for C_{tG} to approximate its contributions in Ref. [16] already lead to very good results. With this, the first step towards a complete combined fit of the different datasets is done. The next step, left for future work, is to compute the predictions for the EWPOs using the flavor symmetry assumptions of the Top sector. After doing this the full combined fit of Top, Higgs, Di-Boson and EWPO can be performed.

Acknowledgments

First and foremost I have to thank Prof. Tilman Plehn for allowing me to work in his group, to learn from all those who worked within the group before me and most importantly gave me the chance to work on this thesis. I also thank Prof. Björn Malte Schäfter for immediately agreeing to be my second examiner for this thesis.

I thank Michel Luchmann and Emma Geoffray for gently guiding me into the incredible depths of SFitter while they were busy finding their way out. Furthermore, I want to thank Maeve Madigan, who not only always listened whenever I needed something, but also somehow managed to point me into a more productive direction regardless of the quality of what I had to share. In addition to this, I have to thank Luca Mantani for providing very much needed MadGraph guidance, which most definitely saved me from a few unnecessary headaches.

To everyone else in the group who was exposed to me for longer periods of time, I apologize, and thank you for having an open ear for questions[1] and discussions[2] before, during and after lunch.

Finally, I thank my family for supporting me not only throughout my studies, but my entire life, something that I surely did not make easy. They are the best.

[1] very physics related;
[2] or not.

7

Bibliography

[1] Georges Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett. B* 716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[2] Serguei Chatrchyan et al. "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC". In: *Phys. Lett. B* 716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[3] Georges Aad et al. "Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments". In: *Phys. Rev. Lett.* 114 (2015), p. 191803. DOI: 10.1103/PhysRevLett.114.191803. arXiv: 1503.07589 [hep-ex].

[4] Peter W. Higgs. "Broken symmetries, massless particles and gauge fields". In: *Phys. Lett.* 12 (1964), pp. 132–133. DOI: 10.1016/0031-9163(64)91136-9.

[5] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (1964). Ed. by J. C. Taylor, pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.

[6] Y. Fukuda et al. "Evidence for oscillation of atmospheric neutrinos". In: *Phys. Rev. Lett.* 81 (1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. arXiv: hep-ex/9807003.

[7] Q. R. Ahmad et al. "Measurement of the rate of $\nu_e + d \to p + p + e^-$ interactions produced by ^{8}B solar neutrinos at the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 87 (2001), p. 071301. DOI: 10.1103/PhysRevLett.87.071301. arXiv: nucl-ex/0106015.

[8] Q. R. Ahmad et al. "Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 89 (2002), p. 011301. DOI: 10.1103/PhysRevLett.89.011301. arXiv: nucl-ex/0204008.

[9] Stephen P. Martin. "A Supersymmetry primer". In: *Adv. Ser. Direct. High Energy Phys.* 18 (1998). Ed. by Gordon L. Kane, pp. 1–98. DOI: 10.1142/9789812839657_0001. arXiv: hep-ph/9709356.

[10] Jihn E. Kim and Gianpaolo Carosi. "Axions and the Strong CP Problem". In: *Rev. Mod. Phys.* 82 (2010). [Erratum: Rev.Mod.Phys. 91, 049902 (2019)], pp. 557–602. DOI: 10.1103/RevModPhys.82.557. arXiv: 0807.3125 [hep-ph].

[11] R. Adam et al. "Planck 2015 results. I. Overview of products and scientific results". In: *Astron. Astrophys.* 594 (2016), A1. DOI: 10.1051/0004-6361/201527101. arXiv: 1502.01582 [astro-ph.CO].

[12] P. A. R. Ade et al. "Planck 2015 results. XIII. Cosmological parameters". In: *Astron. Astrophys.* 594 (2016), A13. DOI: 10.1051/0004-6361/201525830. arXiv: 1502.01589 [astro-ph.CO].

[13] G. Steigman. "Observational tests of antimatter cosmologies". In: *Ann. Rev. Astron. Astrophys.* 14 (1976), pp. 339–372. DOI: 10.1146/annurev.aa.14.090176.002011.

[14] Edward W. Kolb and Michael S. Turner. *The Early Universe*. Vol. 69. 1990. ISBN: 978-0-201-62674-2. DOI: 10.1201/9780429492860.

[15] Ilaria Brivio et al. "O new physics, where art thou? A global search in the top sector". In: *JHEP* 02 (2020), p. 131. DOI: 10.1007/JHEP02(2020)131. arXiv: 1910.03606 [hep-ph].

[16] Ilaria Brivio et al. "To Profile or To Marginalize – A SMEFT Case Study". In: (Aug. 2022). arXiv: 2208.08454 [hep-ph].

[17] Georges Aad et al. "Measurement of the $t\bar{t}$ production cross-section in the lepton+jets channel at $\sqrt{s} = 13$ TeV with the ATLAS experiment". In: *Phys. Lett. B* 810 (2020), p. 135797. DOI: 10.1016/j.physletb.2020.135797. arXiv: 2006.13076 [hep-ex].

[18] Georges Aad et al. "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". In: *Eur. Phys. J. C* 81.8 (2021), p. 737. DOI: 10.1140/epjc/s10052-021-09439-4. arXiv: 2103.12603 [hep-ex].

[19] Armen Tumasyan et al. "Measurement of differential $t\bar{t}$ production cross sections in the full kinematic range using lepton+jets events from proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Phys. Rev. D* 104.9 (2021), p. 092013. DOI: 10.1103/PhysRevD.104.092013. arXiv: 2108.02803 [hep-ex].

[20] Ilaria Brivio et al. "From models to SMEFT and back?" In: *SciPost Phys.* 12.1 (2022), p. 036. DOI: 10.21468/SciPostPhys.12.1.036. arXiv: 2108.01094 [hep-ph].

[21] Ilaria Brivio and Michael Trott. "The Standard Model as an Effective Field Theory". In: *Phys. Rept.* 793 (2019), pp. 1–98. DOI: 10.1016/j.physrep.2018.11.002. arXiv: 1706.08945 [hep-ph].

[22] Aneesh V. Manohar. "Introduction to Effective Field Theories". In: (Apr. 2018). Ed. by Sacha Davidson et al. DOI: 10.1093/oso/9780198855743.003.0002. arXiv: 1804.05863 [hep-ph].

[23] R. L. Workman et al. "Review of Particle Physics". In: *PTEP* 2022 (2022), p. 083C01. DOI: 10.1093/ptep/ptac097.

[24] Brian Henning et al. "2, 84, 30, 993, 560, 15456, 11962, 261485, ...: Higher dimension operators in the SM EFT". In: *JHEP* 08 (2017). [Erratum: JHEP 09, 019 (2019)], p. 016. DOI: 10.1007/JHEP08(2017)016. arXiv: 1512.03433 [hep-ph].

[25] Wouter Dekens and Peter Stoffer. "Low-energy effective field theory below the electroweak scale: matching at one loop". In: *JHEP* 10 (2019). [Erratum: JHEP 11, 148 (2022)], p. 197. DOI: 10.1007/JHEP10(2019)197. arXiv: 1908.05295 [hep-ph].

[26] Stefan Scherer. "Introduction to chiral perturbation theory". In: *Adv. Nucl. Phys.* 27 (2003). Ed. by John W. Negele and E. W. Vogt, p. 277. arXiv: hep-ph/0210398.

[27] Claudius G. Krause. "Higgs Effective Field Theories - Systematics and Applications". PhD thesis. Munich U., 2016. DOI: 10.5282/edoc.19873. arXiv: 1610.08537 [hep-ph].

[28] B. Grzadkowski et al. "Dimension-Six Terms in the Standard Model Lagrangian". In: *JHEP* 10 (2010), p. 085. DOI: 10.1007/JHEP10(2010)085. arXiv: 1008.4884 [hep-ph].

[29] Steven Weinberg. "Baryon- and Lepton-Nonconserving Processes". In: *Phys. Rev. Lett.* 43 (21 1979), pp. 1566–1570. DOI: 10.1103/PhysRevLett.43.1566. URL: https://link.aps.org/doi/10.1103/PhysRevLett.43.1566.

[30] Sally Dawson, Samuel Homiller, and Matthew Sullivan. "Impact of dimension-eight SMEFT contributions: A case study". In: *Phys. Rev. D* 104.11 (2021), p. 115013. DOI: 10.1103/PhysRevD.104.115013. arXiv: 2110.06929 [hep-ph].

[31] Frank Krauss, Silvan Kuttimalai, and Tilman Plehn. "LHC multijet events as a probe for anomalous dimension-six gluon interactions". In: *Phys. Rev. D* 95.3 (2017), p. 035024. DOI: 10.1103/PhysRevD.95.035024. arXiv: 1611.00767 [hep-ph].

[32] Valentin Hirschi et al. "Constraining anomalous gluon self-interactions at the LHC: a reappraisal". In: *JHEP* 07 (2018), p. 093. DOI: 10.1007/JHEP07(2018)093. arXiv: 1806.04696 [hep-ph].

[33] Kaoru Hagiwara et al. "Low-energy effects of new interactions in the electroweak boson sector". In: *Phys. Rev. D* 48 (1993), pp. 2182–2203. DOI: 10.1103/PhysRevD.48.2182.

[34] Anja Butter et al. "The Gauge-Higgs Legacy of the LHC Run I". In: *JHEP* 07 (2016), p. 152. DOI: 10.1007/JHEP07(2016)152. arXiv: 1604.03105 [hep-ph].

[35] Anke Biekötter, Tyler Corbett, and Tilman Plehn. "The Gauge-Higgs Legacy of the LHC Run II". In: *SciPost Phys.* 6.6 (2019), p. 064. DOI: 10.21468/SciPostPhys.6.6.064. arXiv: 1812.07587 [hep-ph].

[36] Kyle Cranmer. "Practical Statistics for the LHC". In: *2011 European School of High-Energy Physics*. 2014, pp. 267–308. DOI: 10.5170/CERN-2014-003.267. arXiv: 1503.07622 [physics.data-an].

[37] Tilman Plehn. "Lectures on LHC Physics". In: *Lect. Notes Phys.* 844 (2012), pp. 1–193. DOI: 10.1007/978-3-642-24040-9. arXiv: 0910.4182 [hep-ph].

[38] Emma Geoffray. "Build from the SMEFT up - SMEFT global analyses as a bottom-up approach to constrain BSM physics". PhD thesis. Heidelberg University Library, U. Heidelberg (main), Heidelberg University, Dec. 2022. DOI: 10.11588/heidok.00032920.

[39] Anke Biekötter. "Know its limits: A global view on Higgs couplings at the LHC". PhD thesis. Heidelberg U., 2019. DOI: 10.11588/heidok.00026907.

[40] Anja Butter. "Global Fits for New Physics at the LHC and Beyond". PhD thesis. U. Heidelberg (main), 2017. DOI: 10.11588/heidok.00023710.

[41] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J. C* 71 (2011). [Erratum: Eur.Phys.J.C 73, 2501 (2013)], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [physics.data-an].

[42] Jerzy Neyman and E.S. Pearson. "On the Problem of the Most Efficient Tests of Statistical Hypotheses". In: *Philosophical Transactions of the Royal Society, A* 231 (Jan. 1933), pp. 289–337. DOI: 10.1098/rsta.1933.0009.

[43] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". In: *The Annals of Mathematical Statistics* 9.1 (1938), pp. 60 –62. DOI: 10.1214/aoms/1177732360. URL: https://doi.org/10.1214/aoms/1177732360.

[44] Andreas Hocker et al. "A New approach to a global fit of the CKM matrix". In: *Eur. Phys. J. C* 21 (2001), pp. 225–259. DOI: 10.1007/s100520100729. arXiv: hep-ph/0104062.

[45] Gary J. Feldman and Robert D. Cousins. "A Unified approach to the classical statistical analysis of small signals". In: *Phys. Rev. D* 57 (1998), pp. 3873–3889. DOI: 10.1103/PhysRevD.57.3873. arXiv: physics/9711021.

[46] Andrew Gelman. *Bayesian data analysis*. eng. Third edition. Texts in statistical science. Includes bibliographical references and indexes. - Description based on online resource; title from title page (viewed May 10, 2017). Boca Raton: CRC Press, 2014, 1 online resource (1 volume). ISBN: 978-1-4398-9822-2 and 1-4398-9822-7. URL: https://learning.oreilly.com/library/view/-/9781439898222/?ar.

[47] S. Kraml et al. "Searches for New Physics: Les Houches Recommendations for the Presentation of LHC Results". In: *Eur. Phys. J. C* 72 (2012), p. 1976. DOI: 10.1140/epjc/s10052-012-1976-3. arXiv: 1203.2489 [hep-ph].

[48] F. Boudjema et al. "On the presentation of the LHC Higgs Results". In: *Workshop on Likelihoods for the LHC Searches*. July 2013. arXiv: 1307.5865 [hep-ph].

[49] Kyle Cranmer et al. "HistFactory: A tool for creating statistical models for use with RooFit and RooStats". In: (June 2012).

[50] "Reproducing searches for new physics with the ATLAS experiment through publication of full statistical likelihoods". In: (2019).

[51] Lukas Heinrich, Matthew Feickert, and Giordon Stark. *pyhf: v0.7.2*. Version 0.7.2. https://github.com/scikit-hep/pyhf/releases/tag/v0.7.2. DOI: 10.5281/zenodo.1169739. URL: https://doi.org/10.5281/zenodo.1169739.

[52] Lukas Heinrich et al. "pyhf: pure-Python implementation of HistFactory statistical models". In: *Journal of Open Source Software* 6.58 (2021), p. 2823. DOI: 10.21105/joss.02823. URL: https://doi.org/10.21105/joss.02823.

[53] F. James. "MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1". In: (1994).

[54] Kyle Cranmer and Alexander Held. "Building and steering binned template fits with cabinetry". In: *EPJ Web Conf.* 251 (2021), p. 03067. DOI: 10.1051/epjconf/202125103067.

[55] Vardan Khachatryan et al. "Measurement of differential cross sections for top quark pair production using the lepton+jets final state in proton-proton collisions at 13 TeV". In: *Phys. Rev. D* 95.9 (2017), p. 092001. DOI: 10.1103/PhysRevD.95.092001. arXiv: 1610.04191 [hep-ex].

[56] Albert M Sirunyan et al. "Measurement of differential cross sections for the production of top quark pairs and of additional jets in lepton+jets events from pp collisions at $\sqrt{s} = 13$ TeV". In: *Phys. Rev. D* 97.11 (2018), p. 112003. DOI: 10.1103/PhysRevD.97.112003. arXiv: 1803.08856 [hep-ex].

[57] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

[58] R. Frederix et al. "The automation of next-to-leading order electroweak calculations". In: *JHEP* 07 (2018). [Erratum: JHEP 11, 085 (2021)], p. 185. DOI: 10.1007/JHEP11(2021)085. arXiv: 1804.10017 [hep-ph].

[59] Torbjörn Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Comput. Phys. Commun.* 191 (2015), pp. 159–177. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].

[60] Johannes Bellm et al. "Herwig 7.0/Herwig++ 3.0 release note". In: *Eur. Phys. J. C* 76.4 (2016), p. 196. DOI: 10.1140/epjc/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph].

[61] J. de Favereau et al. "DELPHES 3, A modular framework for fast simulation of a generic collider experiment". In: *JHEP* 02 (2014), p. 057. DOI: 10.1007/JHEP02(2014)057. arXiv: 1307.6346 [hep-ex].

[62] Johan Alwall et al. "MadGraph 5 : Going Beyond". In: *JHEP* 06 (2011), p. 128. DOI: 10.1007/JHEP06(2011)128. arXiv: 1106.0522 [hep-ph].

[63] Priscila de Aquino et al. "ALOHA: Automatic Libraries Of Helicity Amplitudes for Feynman Diagram Computations". In: *Comput. Phys. Commun.* 183 (2012), pp. 2254–2263. DOI: 10.1016/j.cpc.2012.05.004. arXiv: 1108.2041 [hep-ph].

[64] Celine Degrande et al. "UFO - The Universal FeynRules Output". In: *Comput. Phys. Commun.* 183 (2012), pp. 1201–1214. DOI: 10.1016/j.cpc.2012.01.022. arXiv: 1108.2040 [hep-ph].

[65] Adam Alloul et al. "FeynRules 2.0 - A complete toolbox for tree-level phenomenology". In: *Comput. Phys. Commun.* 185 (2014), pp. 2250–2300. DOI: 10.1016/j.cpc.2014.04.012. arXiv: 1310.1921 [hep-ph].

[66] Céline Degrande et al. "Automated one-loop computations in the standard model effective field theory". In: *Phys. Rev. D* 103.9 (2021), p. 096024. DOI: 10.1103/PhysRevD.103.096024. arXiv: 2008.11743 [hep-ph].

[67] Richard D. Ball et al. "Parton distributions from high-precision collider data". In: *Eur. Phys. J. C* 77.10 (2017), p. 663. DOI: 10.1140/epjc/s10052-017-5199-5. arXiv: 1706.00428 [hep-ph].

[68] Richard D. Ball et al. "The path to proton structure at 1% accuracy". In: *Eur. Phys. J. C* 82.5 (2022), p. 428. DOI: 10.1140/epjc/s10052-022-10328-7. arXiv: 2109.02653 [hep-ph].

[69] Michal Czakon, David Heymes, and Alexander Mitov. "fastNLO tables for NNLO top-quark pair differential distributions". In: (Apr. 2017). arXiv: 1704.08551 [hep-ph].

[70] Michał Czakon et al. "HighTEA: High energy Theory Event Analyser". In: (Apr. 2023). arXiv: 2304.05993 [hep-ph].

[71] Morad Aaboud et al. "Measurement of the inclusive and fiducial $t\bar{t}$ production cross-sections in the lepton+jets channel in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 78 (2018), p. 487. DOI: 10.1140/epjc/s10052-018-5904-z. arXiv: 1712.06857 [hep-ex].

[72] Vardan Khachatryan et al. "Measurement of the t-tbar production cross section in the e-mu channel in proton-proton collisions at sqrt(s) = 7 and 8 TeV". In: *JHEP* 08 (2016), p. 029. DOI: 10.1007/JHEP08(2016)029. arXiv: 1603.02303 [hep-ex].

[73] Albert M. Sirunyan et al. "Measurement of the $t\bar{t}$ production cross section using events with one lepton and at least one jet in pp collisions at $\sqrt{s} = 13$ TeV". In: *JHEP* 09 (2017), p. 051. DOI: 10.1007/JHEP09(2017)051. arXiv: 1701.06228 [hep-ex].

[74] Albert M Sirunyan et al. "Measurement of the $\mathrm{t\bar{t}}$ production cross section, the top quark mass, and the strong coupling constant using dilepton events in pp collisions at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 79.5 (2019), p. 368. DOI: 10.1140/epjc/s10052-019-6863-8. arXiv: 1812.10505 [hep-ex].

[75] Morad Aaboud et al. "Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b-tagged jets in pp collisions at $\sqrt{s}$=13 TeV with the ATLAS detector". In: *Phys. Lett. B* 761 (2016). [Erratum: Phys.Lett.B 772, 879–879 (2017)], pp. 136–157. DOI: 10.1016/j.physletb.2016.08.019. arXiv: 1606.02699 [hep-ex].

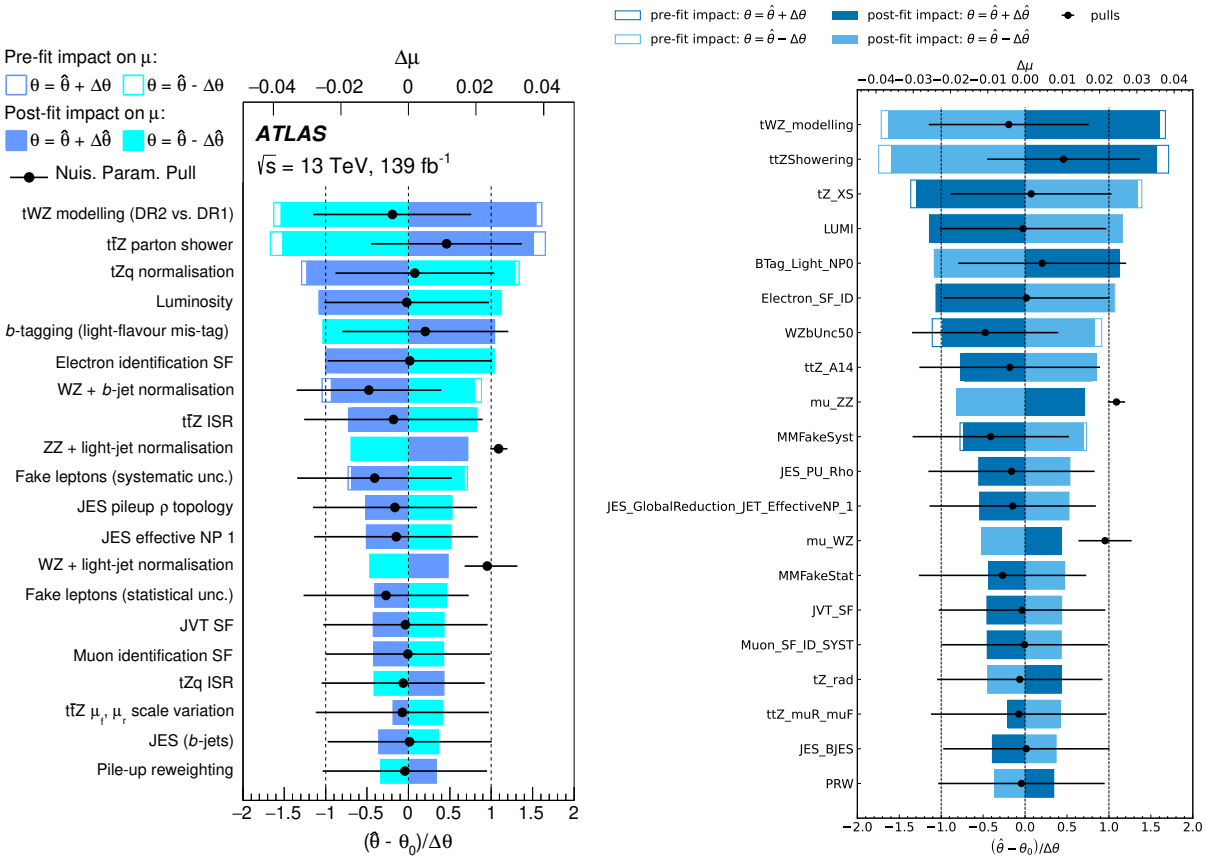
[76] Albert M Sirunyan et al. "Measurement of top quark pair production in association with a Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *JHEP* 03 (2020), p. 056. DOI: 10.1007/JHEP03(2020)056. arXiv: 1907.11270 [hep-ex].

[77] Morad Aaboud et al. "Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ production cross sections in multilepton final states using 3.2 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 77.1 (2017), p. 40. DOI: 10.1140/epjc/s10052-016-4574-y. arXiv: 1609.01599 [hep-ex].

8

Appendix

Published likelihood results for $t\bar{t}Z$



(a) Figure taken from Ref. [18].

(b) Reproduced values using pyhf.

Figure 8.1: Comparison of pulls and individual impact on the $t\bar{t}Z$ cross section.

Groups of uncertainties in the Top fit

Theory uncertainties	Statistical uncertainties	Systematic uncertainties
Color Reconnection	stat	Beam
Extrapolation	KinRec	Background (Separate for each channel)
Matching	MC	ETmis
NLOMatching	Modeling	Jets
PDF	bFragmentation	Leptons
PDFSim	Others	LightTagging
ScaleSim		Luminosity
Scales		Pileup
ScalesT		Trigger
Scheme		Tune
TopMass		bTagging
UnderlyingEvent		partonShower
hatPDFNLO		tTagging
		tauTagging

Table 8.1: List of different uncertainty groups considered within the Top sector in SFitter. Systematics in the same groups are assumed to be fully correlated between measurements.

Systematic uncertainties in SFitter

Order	Uncertainty	Order	Uncertainty
1/16	Luminosity (CMS/ATLAS)	9/24	BkgZZ4l (CMS/ATLAS)
2/17	Detector (CMS/ATLAS)	10/25	BkgHtGG (CMS/ATLAS)
3/18	LeptonReconstruction (CMS/ATLAS)	11/26	BkgHtTT (CMS/ATLAS)
4/19	PhotonReconstruction (CMS/ATLAS)	11/26	BkgHtTT (CMS/ATLAS)
5/20	bTagging (CMS/ATLAS)	12/27	BkgHtWW (CMS/ATLAS)
6/21	tauTagging (CMS/ATLAS)	13/28	BkgHtbb (CMS/ATLAS)
7/22	VBF (CMS/ATLAS)	14/29	WWprod 8 TeV (CMS/ATLAS)
8/23	LeptonIsolation (CMS/ATLAS)	15/30	WZprod 8 TeV (CMS/ATLAS)
31	WZprod 7 TeV (ATLAS)		

Table 8.2: List of different systematic uncertainties included in the Higgs,Di-Boson and EWPO fit and their ordering within SFitter.

Analytic expressions used in marginalization

This is a list of different analytic expression used in the marginalization of the likelihood.

Poisson-Gamma Model

$$\begin{aligned}\mathcal{P}(d|s, b_{SR}, k) &= \int d(bk)\text{Pois}(d|s+b)\text{Pois}(b_{SR}k|bk) \\ &= \int db \frac{1}{\Gamma(d+1)\Gamma(b_{SR}k+1)} e^{-s+b}(s+b)^d e^{-kb}(bk)^{b_{SR}k} \\ &= \sum_{i=0}^{d} \frac{\Gamma(i+kb_{SR}+1)}{\Gamma(i+1)\Gamma(b_{SR}k+1)} \left(\frac{k}{1+k}\right)^{kb_{SR}} \left(\frac{1}{1+k}\right)^{i+1} \text{Pois}(d-i|s)k\end{aligned}$$

Flat-Flat-Gauss convolution

$$\text{GFF}(d|p, \sigma_{\text{th},1}, \sigma_{\text{th},2}, \sigma_{\text{syst}}) = \int \text{d}\tilde{p}\, \mathcal{N}(d|\tilde{p}, \sigma_{\text{syst}}) p_2(\tilde{p}|p, \sigma_{\text{th},1}, \sigma_{\text{th},2})\,.$$

where p_2 is the results of a convolution of two flat distributions.
Computing the integral gives

$$\begin{aligned}\text{GFF}(d|p, \sigma_{\text{th},1}, \sigma_{\text{th},2}, \sigma_{\text{syst}}) = \frac{1}{4\sigma_{\text{th},1}\sigma_{\text{th},2}} &[F_1(b_1, b_2) - b_1 F_1(b_1, b_2) \\ &+ 2\sigma_{\text{th},2} F_1(b_2, b_3) + b_4 F_1(b_3, b_4) - F_2(b_3, b_4)]\end{aligned}$$

where

$$\begin{aligned} b_1 &= p - \sigma_{\text{th},1} - \sigma_{\text{th},2} \\ b_2 &= p - \sigma_{\text{th},1} + \sigma_{\text{th},2} \\ b_3 &= p + \sigma_{\text{th},1} - \sigma_{\text{th},2} \\ b_4 &= p + \sigma_{\text{th},1} + \sigma_{\text{th},2}\,, \end{aligned} \tag{8.1}$$

assuming that $\sigma_{\text{th},1} > \sigma_{\text{th},2}$.

Finally F_1 and F_2 are defined as

$$F_1(a, b; \mu, \sigma) = \frac{1}{2}\left[\text{erf}\left(\frac{b-\mu}{2\sigma}\right)\text{erf}\left(\frac{a-\mu}{2\sigma}\right)\right] \tag{8.2}$$

and

$$F_2(a, b; \mu, \sigma) = \sigma^2\left[\mathcal{N}(a|\mu, \sigma) - \mathcal{N}(b|\mu, \sigma)\right] + \mu F_1(a, b; \mu\sigma)\,. \tag{8.3}$$

The generalization to more than two theory uncertainties is then done by only taking the two largest theory uncertainties and adding the remaining ones in quadrature to the total Gaussian uncertainty. This approximation has been checked and has shown excellent agreement with numerical results.

Rotation from Warsaw to HISZ basis

The following is a list of the expressions needed to rotate from the Warsaw basis to the HISZ basis, as they are used in the combined fit in Chapter 5 and are taken from Ref. [38].

$$
\begin{aligned}
-\frac{\alpha_s}{8\pi} f_{GG} &= C_{\phi G} & f_{4L} &= \bar{C}'_{ll} \\
f_{3W} &= \frac{4}{g^3} C_W & f_{tG} &= -\frac{i}{g_s} C_{uG,33} \\
f_B &= \frac{8}{g'^2} \bar{C}^{(1)}_{\phi l} & f_W &= -\frac{8}{g^2} \bar{C}^{(3)}_{\phi l} \\
f_{\phi 1} &= C_{\phi D} + 4\bar{C}^{(1)}_{\phi l} \\
f_{\phi 2} &= -2C_{\phi\Box} - 2\bar{C}^{(1)}_{\phi l} + 6\bar{C}^{(3)}_{\phi l} \\
f_\phi &= C_\phi - 4\lambda \bar{C}^{(3)}_{\phi l} \\
\frac{m_\tau}{\nu} f_\tau &= C_{e\phi,33} - 2(Y_e)_{33} \bar{C}^{(3)}_{\phi l} & f^{(1)}_{\phi e} &= \bar{C}_{\phi e} - 2\bar{C}^{(1)}_{\phi l} \\
\frac{m_t}{\nu} f_t &= C_{u\phi,33} - 2(Y_u)_{33} \bar{C}^{(3)}_{\phi l} & f^{(1)}_{\phi u} &= \bar{C}_{\phi u} + \frac{4}{3}\bar{C}^{(1)}_{\phi l} \\
\frac{m_b}{\nu} f_b &= C_{d\phi,33} - 2(Y_d)_{33} \bar{C}^{(3)}_{\phi l} & f^{(1)}_{\phi d} &= \bar{C}_{\phi d} - \frac{2}{3}\bar{C}^{(1)}_{\phi l} \\
\frac{m_\mu}{\nu} f_\mu &= C_{e\phi,22} - 2(Y_e)_{22} \bar{C}^{(3)}_{\phi l} & f^{(1)}_{\phi Q} &= \bar{C}^{(1)}_{\phi q} + \frac{1}{3}\bar{C}^{(1)}_{\phi l} \\
f^{(3)}_{\phi Q} &= 4\left[\bar{C}^{(3)}_{\phi q} - \bar{C}^{(3)}_{\phi l}\right] \\
f_{BB} &= -\frac{4}{g'^2}\left[C_{\phi B} - \bar{C}^{(1)}_{\phi l}\right] \\
f_{WW} &= \frac{4}{g^2}\left[C_{\phi W} + \bar{C}^{(1)}_{\phi l}\right] \\
f_{BW} &= 4\left[-\frac{C_{\phi WB}}{gg'} - \frac{\bar{C}^{(3)}_{\phi l}}{g^2} + \frac{\bar{C}^{(1)}_{\phi l}}{g'^2}\right]
\end{aligned}
$$

Finally the quartic Higgs coupling gets redefined as

$$\lambda_{\text{HISZ}} = \lambda_{\text{Warsaw}} + \frac{4m_h^2}{\Lambda^2}\bar{C}^{(3)}_{\phi l}$$

Erklärung:

Ich versichere, dass ich diese Arbeit selbstständig verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den (Datum) 13.09.2023 ..... Schmaf .....